Keele UNIVERSITY

An investigation of factors that interact with the effect of sleep on false memory in the DRM paradigm

Zainab Mohumad A Alyobi

Thesis submitted for the degree of Doctor of Philosophy

June 2021

Keele University

Contents

List of Tables

List of Figures

An Open Science Framework Declaration

The raw data from all the experiments are available on the Open Science Framework (https://osf.io/vj2ec/).

Acknowledgments

I have always believed that success depends on personal effort, but, throughout my time as a PhD student, I have come to realise that this not true. I could not have completed this journey without the support of a number of amazing individuals. Here, I would like to express my sincere gratitude to all of them for making this thesis possible.

First of all, I would like to thank my supervisory team Dr. Sue Sherman, Dr. Helen Williams and Prof. Nicky Edelstyn for all their help throughout this journey. A special thank you goes to my lead supervisor Dr. Sue Sherman; her support and guidance were the candles that lit my way. I greatly appreciate both the many hours that she put aside to meet me and her constructive comments, which developed my thinking and strengthened my overall writing. I could not have asked for a better supervisor.

I would also like to thank my friends for being around, for being my second family; thank you for companying me during the most stressful time in my life.

My family deserves my wholehearted thanks for their continued support, unparalleled love and dedicated efforts, all of which contributed to the completion of this thesis. They had to listen to all my tales of woe throughout this journey, but, as usual, they provided me with the much-needed reassurance and encouraged me to pursue this degree. My beloved family, I owe you a debt of gratitude.

Last but not least, I am thankful for the situations that made me feel sorry for myself, because those moments have changed me, forced me to make big decisions, gave me the courage to begin the PhD journey and made me who I am now.

A very unique thank you goes to myself for being strong enough to deal with all the challenges throughout this journey and life in general. You have proven that you have courage, strength and fortitude, as you were committed to and working on this thesis during a period in which we were living the worst nightmare of COVID-19. Thank you for holding onto your dream and never giving up. It has been a long road of late nights and long mornings, but you made it!

Abstract

It is well established that sleep plays an important role in memory consolidation. However, there is a substantial body of evidence which suggests that sleep has the capacity to distort memory and facilitate the creation of false memories. Previous studies examining the effect of sleep on false memory within the Deese-Roediger-McDermott paradigm have reported mixed results, whereby sleep has been found to increase, decrease, and have no effect on false memory. Previous research has suggested multiple factors that could affect false memory formation within DRM. Therefore, this thesis reports on a series of experiments designed to study some of these factors to examine whether they influence the size of sleep effects on the formation of false memories within the DRM paradigm, namely emotionality of the lists, backward associative strength (BAS), list length and inter-item connectivity. This is in order to address inconsistencies in previous research and to increase our understanding of the factors which influence sleep-dependent consolidation of false memories within the DRM paradigm. Additionally, this thesis further differentiates between early and late sleepers to examine which sleep stage is involved in false memory consolidation. It also compares the total sleep time (TST) scored derived from actigraphy watches and sleep diaries to identify the congruence level between the two methods over five consecutive days. The results of the experiments reveal that sleep increases the level of false memory within the DRM paradigm when inter-item connectivity, BAS and list length are manipulated. However, this effect was not observed when emotional lists were used. The results also indicate that there is no significant effect of sleep stages on the formation of

false memory. Furthermore, the results from a comparison of the TST measures from the actigraphy watch and sleep diaries suggest that differentiation between the two methods varies on a day-to-day basis.

Experiments reported in this thesis are the first which have sought to investigate how sleep interacts with other factors to create false memory within the DRM paradigm. Specifically, the first two experiments experiment are the first that explored the effect of sleep and three types of emotional DRM lists. The third and fourth experiments are the first that investigated how sleep interact with BAS, inter-item connectivity and list length to effect false memory. The fifth experiment is the first that explored the connection between early and late sleep periods and an increase or decrease in false memory. This thesis offers novel findings and adds significantly to the existing literature on sleep and false memory.

Chapter 1: Literature Review

1.1 General Introduction to the Thesis

Sleep has been widely linked to memory in the research setting (Gais, Lucas, & Born, 2006). Indeed, there is a substantial body of evidence suggesting that sleep plays an important role in the consolidation of memory (Diekelmann & Born, 2010; Marshall & Born, 2007). During sleep, newly acquired memory traces are strengthened within neural circuits (synaptic consolidation) playing an important role in future memory retrieval (Diekelmann & Born, 2010). However, other processes also occur during sleep in relation to how memories are stored and arranged. For instance, memory traces may be redistributed to specific regions of the brain for longer periods of storage, depending on the nature of the memory trace (Born & Wilhelm, 2012). Memory traces may also be integrated with existing long-term memories, a process of consolidation, in order to associate memories and allow for more efficient cognitive processes of these memories (Born & Wilhelm, 2012). This process in particular is of interest in memory research, as active restructuring of memories may permit existing memories and experiences to distort or influence recently acquired memory traces, thus potentially leading to false memory formation (Inostroza & Born, 2013). When integrated with previous long-term memories, new memories may have qualitative differences from the original encoded memory trace (Inostroza & Born, 2013). Sleep may therefore be considered a factor in the generation of false memories, as sleep promotes this state of contextual integration for new memory traces. However, inconsistencies in previous research exploring whether sleep increases false memory or reduces it have emerged. Sleep has been found to increase (Diekelmann, Born, & Wagner, 2010; Payne et al., 2009), decrease (Fenn, Gallo, Margoliash, Roediger, & Nusbaum, 2009), and have no effect (Diekelmann, Landolt, Lahl, Born, & Wagner, 2008) on the creation of false memories. These inconsistencies in the results could be for a number of reasons, some of which are explored in this thesis.

The aim of this thesis is to explore the role of sleep by examining its impact on DRM word lists that vary in terms of the level of emotionality, density of interrelations between words in a thematically-related list, associative strength between list words and the critical lure and list length. A further aim is to explore whether sleep stages are involved in false memory. This review chapter will start by outlining the current literature on memory and the key theories that have been proposed to explain the phenomenon of false memory before reporting the current literature on sleep and sleep stages. It will then highlight the questions that this thesis addresses. The following chapter will present a review of the most commonly methods used to measure sleep. Chapters 3, 4, 5 and 6 will report series of experiments that explore some factors that may influence the size of the effect of sleep on DRM false memories. General discussion of the primary findings, limitations, further direction and implications will be outlined in Chapter 7.

1.2 Memory

1.2.1 Human memory system

The system of human memory is complex, multi-factorial in nature (Izawa & Ohta, 2014) and, despite the best efforts of sustained and wide-reaching research focus across recent decades, not fully understood at this point (Radvansky, 2017). In order to conceptualise and theorise on the workings of human memory, researchers have concentrated on two broad forms of human memory, classified as short-term memory and long-term memory (Norris, 2017). Short-term memory is a temporary storage system that is of paramount importance in a variety of informational processing tasks ranging from complex reasoning, all the way to speech comprehension (Baddeley, 2017). In contrast to short-term memory, the long-term system can store information in considerable volume, quantity and over a long period of time (Warrington, 2014). The long-term memory system comprises multiple components such as declarative (explicit) memory and non-declarative (implicit) memory (Salvato et al., 2016; Squire, 1998; see Figure 1). These frameworks can be further subdivided into more specific constructs. For example, in the case of non-declarative memory: skill learning, priming and classical conditioning (Squire, Knowlton, & Musen, 1993). On the other hand, declarative memory can be subdivided into two types: semantic memory and episodic memory (Kilman, Zekveld, Hällgren, & Rönnberg, 2015; Squire et al., 1993). Semantic memories are related to principles and facts and require little or no contextual background (Huth, De Heer, Griffiths, Theunissen, & Gallant, 2016). Episodic memory is contextually situated within places, dates, times and social settings (Burke et al., 2014). Episodic memory is used for events, sensations, emotions and places (Light, 2015); it is what laypeople may refer to or think of as ‘autobiographical’ memory (Moscovitch, Cabeza, Winocur, & Nadel, 2016). Human memory is not a perfect system and retrieval of episodic memory information is often inaccurate. The next section outlines a form of episodic memory distortions, false memory, to describe instances in which our memories fail to represent events as they actually happened.

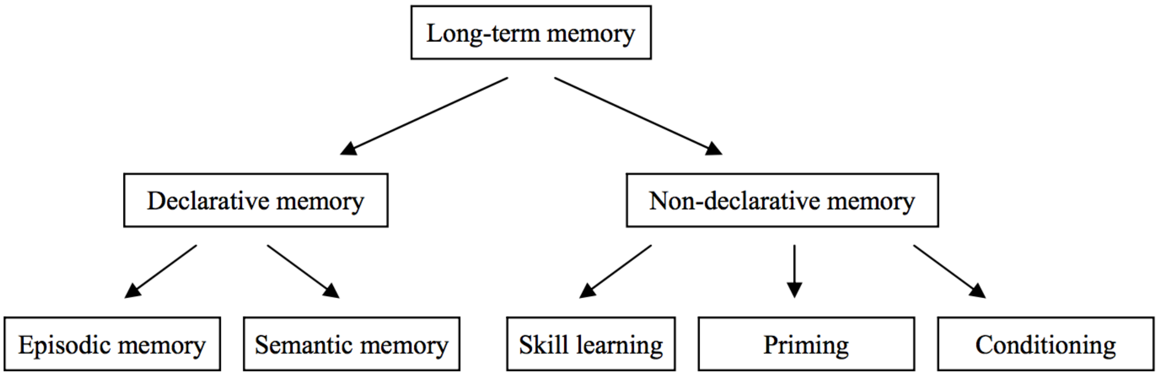


Figure 1.1 Multiple systems of long-term memory (Squire & Zola, 1996).

1.2.2 False Memory

Human memory is a fallible process, not necessarily reflecting a literal record of events as they have occurred or been experienced (Gallo, 2010). Initially, memories are formed through a process of encoding (learning), whereby information is actively stored in neural circuitry (Winters, Saksida, & Bussey, 2008). The next phase is consolidation, whereby memories are ‘processed’ and the strength of memory traces increases following initial encoding (Diekelmann & Born, 2010). Finally, the retrieval of learned material completes the memory cycle, wherein an individual is capable of recalling information or events that have occurred (Winters et al., 2008). At each of these phases, there is the potential for varying influences on memory accuracy and strength, which may influence how an individual remembers events in an accurate or inaccurate manner. One example of how memory may be prone to error is the concept of false memories. False memories are defined as recollections of events that never actually occurred (Brainerd & Reyna, 2005). False memories are typically strongly associated with actual events that have been encoded and individuals who have false memories are often highly confident that they have remembered the event correctly (Gallo, 2010). Because all memories are reconstructions of events rather than accurate recordings of events (Roediger, Watson, McDermott, & Gallo, 2001), it is possible for new memories to become mixed with old and for these memories to take on new forms and content which are errant in nature (Tononi & Cirelli, 2014). The genesis of false memories is a complex issue in contemporary research, although consideration of memory sub-processes may provide some insight into how false memories may occur. For instance, false memories can be generated by a form of data that has become corrupted in the encoding process due to systemic errors in a system that has evolved to be parsimonious rather than entirely accurate at times (Walker & Stickgold, 2004). Additionally, during the consolidation phase of memory, it may be the case that existing representations in the mind are associated with a new memory, leading to the restructuring of a memory and the development of false memories (Diekelmann & Born, 2010). Essentially, these memories would be false in that they are associated with previous knowledge and/or experiences, and abstract from the initial memory event that triggered the encoding and consolidation. An alternative view is that the retrieval of memories may be impaired in the case of false memory, leading to a false representation of what actually happened (Okado & Stark, 2003). Alterations in attention, working memory and executive functions may be associated with poor information retrieval in this context (Winters et al., 2008). Therefore, it is important to formulate experimental contexts in which the genesis of false memories and the factors associated with their formation can be examined in order to understand this phenomenon in greater detail (Winters et al., 2008).

1.2.2.1 Deese-Roediger-McDermott (DRM) paradigm

Assessing the role of sleep in false memory requires a clear and concise experimental design to ensure that false memories may be reliably defined and recorded. The Deese-Roediger-McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995) has been extensively used to assess false memory in experimental settings for over two decades. The DRM paradigm involves the presentation of word lists (e.g. sedan, automobile, vehicle, garage etc.), where words are associated with a critical word not presented in the list (e.g. car). During the memory test, participants reliably have false memories for the non-presented critical lure at a similar rate as other presented words.

In the original experiment by Deese (1959), 50 participants were presented with lists of 12 words, each with an associated critical lure. When asked to recall the word lists, recall of the critical lures was as high as 44%. The actual rate of false recall of the critical lure varied across lists and Deese reasoned that the likelihood of false recall was mediated by the ability of words in the lists to activate the critical lure. For example, Deese suggested that a list containing the word 'short' could activate the critical lure 'man' (backwards association), but a list containing the word 'man' was less likely to activate the critical lure 'short' (forwards association). Roediger and McDermott (1995) revived and popularised the DRM paradigm. In their study, two different experiments took place. The first of these experiments effectively replicated the findings of Deese (1959) on false recall. They used 6 of the original word lists that had produced the highest level of false recall and used both recall and recognition tests. In the recall test, participants were instructed to freely recall as many words as they could remember from the list they had heard previously. For the recognition test, they were presented with words, some of which had been previously presented, some of which were critical non-presented lures, some that were not related to the lists and some that were weakly related to the lists but hadn't been previously presented. Participants were asked to indicate on a four-point scale how confident they were that each word had appeared on the lists. False recall of critical lures occurred roughly 40% of the time, a result that was consistent with the original study. The critical non-presented lures were falsely recognised 58% of the time. In Roediger and McDermott's second experiment, they developed and expanded the lists that had been used in their first experiment and also used a Remember/ Know task (Tulving, 1985). They found that the rate of false recall was 55% and found that prior recall increased later false recognition, as 81% of the lures presented in the study and subsequently recalled were falsely recognised.

The DRM paradigm is a simple experimental design and has been used to study false memory in numerous research populations. For example, it has been used with children (Calado, Otgaar, & Muris, 2018; Khanna & Cortese, 2009; Metzger et al., 2008), adults (Calvillo & Parong, 2016; Otgaar, Moldoveanu, Wang, & Howe, 2017; Payne et al., 2009) and older adults (Akhtar, Howe, & Hoepstine, 2020; Howe & Akhtar, 2020). The DRM paradigm has also been employed to study the effect of different variables on false memory, for example; sleep (Darsaud et al., 2011; Fenn et al., 2009; Payne et al., 2009), mood (Ruci, Tomes, & Zelenski, 2009; Zhang, Gross, & Hayne, 2017), stress (Pardilla-Delgado, Alger, Cunningham, Kinealy, & Payne, 2016; Smeets, Jelicic, & Merckelbach, 2006a) and mindfulness (Baranski & Was, 2017; Rosenstreich, 2016; Sherman & Grange, 2020; Wilson, Mickes, Stolarz-Fantino, Evrard & Fantino, 2015). The DRM paradigm will be the method used for all the studies reported in this thesis.

1.2.2.2 False Memory Theories

Many theories have been developed to help explain the formation of false memory. This section outlines two core theories that have been extensively used to explain false memories created using the DRM paradigm. These are fuzzy trace theory, which explains false memories via a small set of principles that implement a single representational distinction, and activation monitoring theory in which individuals produce false memories because the activation spreads from seen items to unseen similar items.

1.2.2.3 Fuzzy Trace Theory

Fuzzy trace theory (FTT) was developed in (1990) by Brainerd and Reyna and provides a framework for psychologists to understand the functioning of 'gist' representations within learning, memory and reasoning. FTT distinguishes between these gist representations and verbatim representations across the memory storage process. According to this theory, there are two parallel memory traces formed in the mind at encoding. First, there is a verbatim trace which represents the details of the items being remembered and a gist representation which captures more of a general overview (Brainerd & Reyna, 2005). These gist memories may be nebulous patterns, impressions or essences and are traces that become stored in a parallel encoding process (Reyna, 2008). The result of this process, fuzzy trace theory posits, is the storage of multiple traces of memory in relation to the same stimulus (Reyna, 2008). The gist memories and verbatim memories (although stored simultaneously and in parallel) are functionally independent and processing one type of trace does not necessarily facilitate the processing of other types of trace (Brainerd & Reyna, 2005). However, it has been suggested that verbatim traces degrade faster than gist traces over time (Payne, Elie, Blackwell, & Neuschatz, 1996). Individuals' reliance on gist memory can often lead to memories becoming corrupted, fragmented or false in nature (Brainerd & Reyna, 2005). In the DRM paradigm, gist representation refers to the general semantic representation generated by the studied words but with an absence of specific details whereas the verbatim representation refers to the specificity of information generated by the seen words during encoding (Brainerd & Reyna, 2005). Researchers who have used the DRM paradigm and have interpreted the false memory phenomenon within this framework, suggest that gist representation is what underlies false memory while verbatim representations support correct memory (Lo, Sim, & Chee, 2014; Metzger et al., 2008; Payne et al., 2009).

1.2.2.3.1 Activation Monitoring Theory

According to the activation monitoring theory (AMT), a false memory is generated by the combination of the activation and monitoring process (Roediger et al., 2001). The activation component comes from Underwood's (1965) implicit associative response theory which argues that activation happens during the time that the words are presented and this activation spreads to related words that have not been presented, through a semantic network. The original work by Deese (1959) supports the idea of associative activation by highlighting the role of association strength. He demonstrated that lists with items that most likely to generate the unseen critical lure in subsequent tasks can be identified as being the most likely lists to produce false memory. Therefore, the stronger the association between the words in the lists and the critical lure, the greater the amount of activation for the lure.

However, monitoring of memories also has a role to play in false memory generation (Roediger et al., 2001). The concept of monitoring was derived from Johnson's source-monitoring framework. Johnson, Hashtroudi and Lindsay (1993) claimed that there are certain differences between events that have occurred and those that are generated internally. Memories for events that are actually experienced tend to have more details and context, whereas those for internally generated events are less likely to have specific details and tend to focus on more cognitive functions (Johnson et al., 1993). The aim of monitoring is to use these differences in order to distinguish between what is real and what is imagined. AMT proposes that the study of words associated with a critical lure leads to activation of the lure representation in semantic memory. During the encoding process (e.g. within a DRM paradigm study) a list of related words (e.g. truck, bus, train, automobile and vehicle) indirectly activates the non-presented lure (i.e. car) due to associations with these words. As a result of this process, at the retrieval stage false recollection is more likely to occur because individuals fail to distinguish between internally generated information and externally presented information (Sergi, Senese, Pisani, & Nigro, 2014). A strong activation may result in the presented information and unpresented information sharing similar features. In this instance, monitoring becomes difficult and source-monitoring errors may occur, thereby contributing to the likelihood of creating false memories.

1.2.2.4 Variables that Influence DRM False Memories

Given the wide range of literature that has reported false memory creation using the DRM paradigm, it has been suggested that false memory formation is mainly dependent on task-specific features (Newbury & Monaghan, 2019). The current section outlines some of the factors that found to influence the formation of false memories within the DRM paradigm.

1.2.2.4.1 Recall versus Recognition Testing

As outlined above, Roediger and McDermott (1995) conducted two experiments to replicate the findings of Deese (1959). Their participants completed a free recall test after the presentation of each list and following this, participants completed a recognition test. The authors found that participants recalled the critical lure at a rate of 40%; meanwhile, the critical lure was recognised with high confidence 58% of the time (Experiment 1). This suggests that recognition tests lead to a higher rate of false memory than free recall. Roediger and McDermott suggested that the false recognition originated during the study phase, due to an implicit associative response whereby the presented words activated their associated ones consciously or unconsciously. Moreover, the findings of McEvoy, Nelson and Komatsu (1999) support this difference in false memory magnitude between recognition and recall . It could be argued that recognition testing is more reflective of many real-world scenarios, as recounting events and experiences is often not entirely freeform and may involve prompts or interviewer guidance.

The debate around recall versus recognition testing has been noted in further DRM studies. Although the evidence suggests that recognition testing leads to higher false memory than recall testing, it is possible that the rate of false memory, in general, is not related to the nature of the tests themselves but the variables that are manipulated in each study. It has been suggested that after a period of sleep, participants are more likely to produce a false memory in a recall test than they are in a recognition test (Newbury & Monaghan, 2019).

This can be explained in the light of AMT. The theory proposes that monitoring cues are activated during the presentation of words in a recognition test which allow participants to suppress and reject related but un-presented words (Watson, McDermott, & Balota, 2004). Meanwhile, no monitoring cues are available in a recall test, therefore, a large number of associated words are activated. Sleep is thought to enhance source monitoring ability which helps participants to better reject unpresented similar words during recognition tests than during recall tests (Newbury & Monaghan, 2019).

Moreover, Oliver, Bays and Zabrucky (2016) followed a design similar to that of Roediger and McDermott (1995); they included an additional element by asking some participants to imagine the list items in their head. This was intended to address the issue of whether imagery may serve to lower false memory in tests. While Oliver et al. (2016) found that imagery seemed to lower false memory overall, specific test-type interactions highlighted some interesting results with implications for the recall versus recognition discussion. Specifically, semantic (imagery) lists tended to cause higher false recognition while phonological lists led to higher false recall. This suggests that the magnitude of false memories measured by recognition and recall tests depends on the nature of the study.

1.2.2.4.2 List length

The number of words in each list is one of the factors that has been suggested to influence the false memory effect within DRM (Newbury & Monaghan, 2019).

In the original study by Deese (1959), participants were asked to freely recall lists of 12 words and the frequency of false recall for critical lures was counted. In the second experiment by Roediger and McDermott (1995), the lists used in Deese's study were extended; they developed 24 lists with 15 words, for their first experiment, using both recognition and recall tests. They found that the extended set of lists led to increases in both false recall and false recognition rates compared to the 12-words list used in their first experiment, indicating that false memory within DRM could be related to the number of words in the lists. Activation monitoring theory explains this by postulating that the greater number of associate words in a list, the stronger the activation received by the lure and as a result, the larger the effect of false memory.

Robinson and Roediger (1997) investigated the effect of list length on DRM false recall and recognition. They presented their participants with 6 groups of 4 lists each; the list lengths were rotated between each group (0, 3, 6, 9, 12, 15 words). They found that the false recall probability significantly increased when the length of the list increased. In the 3-word list condition, the false recall probability was .03 and it increased to .31 for the 15-word list condition. The same pattern was found for false recognition. These findings were further supported by Sugrue, Strange and Hayne (2009). In this study, the researchers used 8 DRM lists containing either 7 or 14 words. These were presented to their participants, both children and adults, who were randomly assigned to one of the length conditions. They found that, across all ages, longer lists significantly increased false recall compared to shorter lists. Moreover, Swannell and Dewhurst (2012) followed a similar design to look at the effect of list length on false memory, they found that longer lists tended to increase instances of false memory. Similar findings on the effect of list length have been also reported in other studies (Gallo & Roediger, 2003; Libby & Neisser, 2001; Marsh & Bower, 2004; Watson, Balota, & Roediger III, 2003). The high rates of false memories when list length increased can be explained in the light of both FTT and AMT. According to FTT, increasing the number of associates systematically increases the level of false memory as stronger gist representations are formed when studying many associates. AMT suggests two processes that explain why participants tended to make more mistakes when given a longer DRM list. Firstly, the large number of words in a list might increase the likelihood of activating the critical lure, and therefore result in increased false memory effects. Certainly, when simply manipulating the length of the list, it has been indicated that the list's total associative strength is a better predictor of whether the critical lure will be remembered (Robinson & Roediger, 1997). When the list is longer then its total associative strength will be greater (Gallo, 2006). Secondly, more errors may occur with a longer DRM list as a shorter list allows participants to better utilise their source-monitoring capabilities. For instance, compared to a longer list, there is a lower cognitive load linked to retaining short-list items in the memory (Sugrue et al., 2009). Additionally, research indicates that source-monitoring errors will more probably happen when there are no defining features which allow the lure words to be rejected (Israel & Schacter, 1997; Smith, Hunt, & Gallagher, 2008). From this standpoint, a shorter list allows for increased distinctiveness of any list item which may allow participants to remember that even though they thought of the critical lure when the list was presented, it did not actually appear on the original list (Roediger et al., 2001; Sugrue et al., 2009).

1.2.2.4.3 The Number of Different Lists Learned

The number of studied lists may contribute to the increased level of false recall and recognition observed in the second experiment of Roediger and McDermott (1995) relative to their first experiment. In their second experiment, participants studied 16 lists (as opposed to 6 in their first experiment) suggesting that increasing the number of lists that participants are exposed to may play a role in increasing the false memory rate. With fewer lists, there is only a small number of critical lures to potentially recall or recognise; hence, if a higher number of lists is used then the level of false memory would be expected to increase. The recall data from Payne et al. (2009) and Pardilla-Delgado and Payne (2016) supports this idea. In Payne et al.'s (2009) research, a total of 8 DRM word lists were used. Pardilla-Delgado and Payne (2016) expanded the number of studied lists from 8 to 16. The study by Payne et al. shows that the mean false recall rate is approximately four critical words (M= 3.6 in exp. 1); however, the average false recall in the Pardilla-Delgado and Payne study was higher (M=4.2) Hence, increasing the number of lists may be one of the factors that led to the increase of false memory level in their study. The FFT explanation for why an increased number of learned lists are more likely to produce false memory is that individuals have to increasingly rely on their gist as a large amount of information (i.e. number of lists) has been presented to them. This can also be explained in the light of AMT; the difficulty of source monitoring increases with an increase in the number of lists, thereby increasing the chances of false memories. The spread of activation over a larger number of distinctive thematic lists might additionally increase the potential effects of generating false memories (Newbury & Monaghan, 2019).

1.2.2.4.4 Backward Associative Strength

Backward associative strength (BAS) is the link between list words and the critical lure. The part that associative strength plays in producing false memories can mainly be attributed to list items influencing backward associative links to the critical lure (Deese, 1959). Deese argued that BAS is the variable that can best explain this phenomenon. He discovered that the mean BAS of a list predicts the probability of that list eliciting false recall of the critical item. While BAS is the probability that list items will elicit critical items as a response in free association tasks, the mean BAS of a list is the average of the backward association between each word in a list and the critical item. Deese found that the probability of the intrusion of a critical item correlates highly with the study list's mean BAS ($r = +.87$). Therefore, the more likely list items are to evoke the critical item in free association tests, the greater the likelihood of false recall from those lists. McEvoy et al., (1999) explored this further and varied the degree of the associative strengths from the list items to the critical lure on each list. They discovered that higher BAS increases both false recognition and false recall. Moreover, Roediger et al. (2001) demonstrated that BAS, but not forward associative strength (FAS; the strength from critical lure to list words), could predict false recall and recognition of the lure, suggesting that BAS is an important factor in DRM false memories. There is a substantial body of research that supports these findings and confirms that BAS plays a considerable role in the production of false memory within the DRM (Arndt, 2012; Beato & Arndt, 2017; Gallo & Roediger, 2003; Hicks & Hancock, 2002; Hicks & Marsh, 1999; Howe, Wimmer, Gagnon, & Plumpton, 2009; Knott, Dewhurst, & Howe, 2012; Roediger et al., 2001). FTT suggests that gist extraction is easier in the high BAS lists as the associative strength from the list words to the critical lure is strong. It has been argued that BAS gives a metric of the degree to which activation spreads to a lure's representation when the associates are studied (Arndt, 2015). According to AMT, BAS influences false memory rates via spreading activation. This expectation stems from the claim that false recall and recognition of the lures, are increased by factors influencing the spread of activation from the study item to representation of the lure, for instance the strength of directional associations from the study item to the representation of the lure (i.e. BAS; Arndt, 2015).

1.2.2.4.5 Inter-item Connectivity

Inter-item connectivity refers to the degree of the connection strengths from each word in the list to every other word in the list and how likely each word is to generate the other words in an association task. It has been suggested that the level of associative 'connectedness' of the words on a list influences DRM false memory (Deese, 1959; McEvoy et al., 1999). For instance, the degree of density between the words in the DRM list seems to be a good predictor of the likelihood of false memory. This associative interconnectedness might also result in increases in the chance of source confusion leading to increased false memories (McEvoy et al., 1999). McEvoy et al. noted that when participants are presented with a list of words, the amount of connections between the words on the list seems to have an influence on the probability of recalling the unpresented critical lure. In McEvoy et al's study, the density of word associations and the strength between the critical lures and the list words were manipulated in each list. The probability of a false memory being produced in a free recall test was seen to increase as the connection strength between the list words and the critical items increased, yet it was inversely connected to the association density amongst the list words. Contrastingly, correct recall of the list words was positively related to association density. For recognition, the likelihood of both correct and false memory increased when the words in the lists had a denser interconnection. The difference in the effect of inter-item connectivity between false recall and false recognition was explained by the assumption that it was due to "the list cuing factor, the degree to which recalled words cue other list words at test" (McEvoy, et al 1999, p. 1185). This process occurs only in a recall test, but it is not involved in a recognition test.

According to the AMT, studying words associated with a lure leads to activation of the lure representation in semantic memory (Brainerd, Wright, Reyna & Mojardin, 2001). In comparison to lists with weaker interrelations, lists that have stronger interrelations of words have an increased opportunity to activate unpresented associate words resulting in higher levels of false memories. This is due to the fact that activation is spread amongst the associates within semantic memory (Newbury & Monaghan, 2019). Moreover, Knott et al. (2012) explored the roles of BAS and inter-item connectivity within DRM and category lists. They found that correct recall and recognition were higher when the lists were designed with high inter-item connectivity. False memories for both recall and recognition tests were found to be greater for lists with high BAS and low inter-item connectivity. The inconsistent findings in McEvoy et al's and Knott's studies with regard to false recognition could be due to the different list lengths used (15 words vs 10 words in each list respectively). As outlined above, longer lists produce more false memory than shorter lists; a list with high inter-item connectivity and more semantic association words could lead to greater activation of a critical lure.

1.2.2.4.6 Emotionality of the Lists

Previous research has indicated that emotionally charged information is more likely to be remember than neutral information (Adelman & Estes, 2013; Kensinger & Corkin, 2004; Wiesner et al., 2015) This effect of emotion is not limited to correct memory alone; it also influences false memory formation (Brainerd, Stein, Silveira, Rohenkohl, & Reyna, 2008; Pesta, Murphy, & Sanders, 2001). It has been suggested that the high level of associations between emotional words compared to neutral words can lead to a high level of false memory within the DRM paradigm (Buchanan, Etzel, Adolphs, & Tranel, 2006)

A number of studies have focused on negative compared to neutral word lists when investigating the effect of emotion on false memory within the DRM paradigm. They have found that negative lists produce more false memories than neutral lists (Howe, Candel, Otgaar, Malone, & Wimmer, 2010; Pesta et al., 2001; Shah & Knott, 2018; Sharkawy, Groth, Vetter, Beraldi, & Fast, 2008). This suggests that negatively valenced lists have stronger connection to their critical lures than neutral lists (Janiszewska, 2014), resulting in greater activation of these associated critical lures. The increase in the false recognition rate for emotional lists may be the result of emotionally charged lists generating a higher level of familiarity, indicating that participants have a reduced ability to use item-specific correct memories to inhibit false memories when presented with emotional lists.

A limited number of studies have compared both negative and positive lists to neutral lists to examine the effect of different types of emotional lists on false memory within DRM and the findings have been inconclusive. While Palmer and Dodson (2009) found that the false recall rate was significantly higher for neutral lists than for both negative and positive lists, the results of Dehon et al.'s (2010) study indicated that negative and positive lists produce a higher level of false recall and recognition than neutral lists. The latter findings might suggest that gist representations of negative and positive lists are more broader and robust relative to those of neutral lists (Goodman et al., 2011). However, Brainerd et al. (2008) drew different conclusions. They presented their participants with a total of 18 DRM lists (six negative, six positive and six neutral). Their results demonstrated that only negative word lists increased false recognition, and there was a reduction in false recognition for positive word lists in comparison to neutral lists. This can be explained by the suggestion that the semantic connections in negative lists are more salient than those with positive valence and that gist representations are strengthened more by negative valence than by positive valence (Bookbinder & Brainerd, 2016). Analysis of memory processes has suggested that two primary memory processes contribute to the effect of valence: “gist similarity and recollection to rejection” (i.e. being able to use verbatim traces to suppress critical lure acceptance; Yeh & Lu, 2019, p. 70). There is an increase in gist similarity weight when taking a recognition decision for negative critical lures in comparison to positive and neutral lures. That is, the lure gist trace and the corresponding list items increase in familiarity with negative lists. Yet, with the positive and neutral lists the recollection rejection incrementally increases in weight compared to the recognition decision (Yeh & Lu, 2019).

Furthermore, differences have been also identified between the test of recall and test of recognition when emotionality of the lists are investigated. The effect of emotional lists on DRM false memory has been found to be clearer in recognition tests compared to recall tests. Howe et al. (2010) explored the effect of emotional word lists on false memory within the DRM paradigm and found significant differences in false memory for negative vs. neutral lists. However, the differences varied between the recognition and recall tests. False recall decreased for negative lists compared to neutral lists. Contrastingly, false recognition increased for negative lists compared to neutral lists. The same pattern was also found in Sharkawy et al’s (2008) study. They found that false recognition of a critical lure was higher in negative word lists compared to neutral lists, whereas no difference has been found for false recall in negative and neutral lists. The increase in the false recognition rate for emotional lists may be the result of emotionally-charged lists generating a higher level of familiarity, indicating that participants had a reduced ability to use item-specific correct memories to inhibit false memories when presented with emotional lists. Moreover, it has been suggested that the effect of emotional stimuli on false recognition may occur in two ways: (1) by influences on recollection-based suppression of false memories, or (2) by influences on familiarity-based acceptance of false memory (Bookbinder & Brainerd, 2016).

1.3 Sleep

1.3.1 Role of Sleep in Memory Consolidation

It is well known that memory consolidation is largely sleep dependent and occurs primarily within segments of sleep (Stickgold, 2005). Jenkins and Dallenbach were among the first researchers to propose evidence from their experiments suggesting that sleep assists memory consolidation (Jenkins & Dallenbach, 1924). They performed systematic tests on learned nonsense syllable retention over time. Their findings show that memory performance was improved after a night of sleep compared to after an equal amount of wakefulness. Since then, the effect of sleep on memory consolidation has been reported in many studies, using different learning and retrieval types and different tasks of memory (Jan Born, Rasch, & Gais, 2006; Diekelmann, Wilhelm, & Born, 2009; Peigneux, Laureys, Delbeuck, & Maquet, 2001; Walker & Stickgold, 2006).

Furthermore, sleep has been confirmed to benefit both declarative and procedural memory in different types of task (Diekelmann et al., 2009; Marshall

& Born, 2007). For procedural memory consolidation, it has been evidenced that the learning of motor skills is dependent on post-learning sleep intervals. For example, Walker, Brakefield, Morgan, Hobson and Stickgold (2002) demonstrated that motor skills were more effectively learnt after a period of sleep than a period of wakefulness. Such findings have been supported in recent years by the work of King et al. (2017) who found that the effect of sleep is not limited to the learning of motor skills but other cerebral skills such as phonological rule learning are improved by the process of sleep. Also, other studies have found that sleep improves procedural skill performance using other tasks, such as the mirror tracing task (Plihal & Born, 1997) and visual texture discrimination task (Mednick, Nakayama, & Stickgold, 2003; Stickgold, Whidbee, Schirmer, Patel, & Hobson, 2000). On the other hand, sleep also plays an important role in the consolidation of declarative memories. One study which addresses this issue examined the effect that sleep had on both children (8-12 years) and adults in terms of their declarative-episodic memory (Wang, Weber, Zinke, Inostroza, & Born, 2018). Participants were allocated to either a sleep or a wake group and presented with 24 female faces for a later what-where-when memory task. The researchers found that in both age groups, episodic memory was improved after a period of sleep (Wang et al., 2018). A similar result was found by Gais et al. (2006) who observed that sleep supports the consolidation of declarative memories. In another study, van Dongen, Thielen, Takashima, Barth, and Fernández (2012) presented their participants with two sets of picture-location association and tested them after a period of sleep or wakefulness. They found that the sleep group showed better recall performance compared to the wake group. The recall of word pairs has been also found to be greater after a delay that includes a period of sleep, compared to a delay of the same period of wakefulness (Plihal & Born, 1997; Wilson, Baran, Pace-Schott, Ivry, & Spencer, 2012). Moreover, significant evidence exists to support the role of sleep in declarative memory recognition. Dumay and Gaskell (2007) found that after sleep, recognition of newly learned words was significantly increased. They suggested that novel information integrated with previous knowledge during sleep, and participants who stayed awake for a 12-hour period did not exhibit the same consolidation level. Similar findings were reported by Davis, Di Betta, Macdonald and Gaskell (2009), who also found that recognition of novel words increased after sleep compare to wake.

Interestingly, not all information is equally remembered after sleep. It has been suggested that sleep plays a selective part in memory consolidation, to the extent that memories are filtered to be selectively reactivated during sleep. Specifically, it has been discovered that sleep preferentially enhances memories for both positive and negative emotionally relevant information in comparison to non-emotional events (Rasch & Born, 2013; Walker & van Der Helm, 2009). It is believed that emotional memory enhancement occurs due to hippocampus and amygdala co-activation (Walker & van Der Helm, 2009), with increased amygdala activity correlating with increased emotional memory (McGaugh, 2004). Interestingly, it has been found that emotional content retention was significantly improved for emotional stories in comparison to neutral stories four years after they had been encoded, yet only when sleep followed learning rather than when participants remained awake (Wagner, Gais, & Born, 2001). Additionally, Hu, Stylos-Allan and Walker (2006) found that a full night of nocturnal sleep after learning increased retention of emotions in comparison to when participants remained in a wakeful state for a similar retention interval during the day.

Sleep, therefore plays an important role in the processing of a variety of memory and learning tasks due to the fact that the sleep process takes advantage of the plasticity of the brain and consolidates related neural pathways (Tononi & Cirelli, 2014). Memory consolidation during sleep can strengthen associations as well as qualitative changes in memory representations. Memory strengthening is expressed as resistance to interference from similar tasks, which is known as 'stabilisation' and improved performance ('enhancement'). As noted above, sleep has stabilising effects that have been observed in a variety of memory types.

1.3.2 Sleep stages

Sleep consists of multiple stages, and these stages are transient in nature and cyclical in the manner in which they are switch from one stage to the next (Meddis, 2018). There are generally thought to be five distinct phases of sleep, the first four of which are grouped and categorised as non-rapid eye movement (non-REM) and one of which is categorised as rapid eye movement sleep (REM). Collectively, these stages are referred to as the sleep cycle (Meddis, 2018; Figure 2). The first stage of the sleep cycle is light sleep, which is considered to be the transition phase from wakefulness to sleep and is thought to account for 5 and 10% of the total time spent sleeping (Dementienko et al., 2018). At this stage alpha brain levels demonstrate a frequency of approximately 8-13 Hz (Meddis, 2018). Stage 2 sleep accounts for approximately 40-50% of total sleep time (Dementienko et al, 2018) and is characterised by alpha waves of between 4-8 Hz. During this phase of sleep, brain waves slow down with occasions of sporadic bursts of rapid waves throughout (Dementienko et al, 2018). Stages 3 and 4 of the sleep cycle are almost indistinguishable from one another and between them represent approximately 20% of the total time spent sleeping (Rasch, Büchel, Gais, & Born, 2007). These stages are only distinguished from each other by the percentage of data wave activity, with a frequency oscillation of approximately 0 and 4 Hz (Meddis, 2018). As a result, these stages are known as slow wave sleep (SWS) (Rasch et al., 2007). Within these stages, all muscle and eye movement ceases and it is SWS that is thought to provide the most recuperative effects of the sleep process and is often though to define the quality of a night's sleep (Halász, Bódizs, Parrino, & Terzano, 2014). The final stage (stage 5) of the sleep cycle is known as REM sleep and it can last up to 60 minutes. REM sleep is a stage in which brain wave activity mimics that of an individual within a waking state at a rate of 15-30 Hz (Nishida, Pearsall, Buckner, & Walker, 2009). In this period, breathing and heart rate increase, eyes jerk rapidly, and the body loses some of its ability to regulate temperature (Meddis, 2018). It is estimated that REM sleep occurs on average between three and five times during a night of sleep (Nishida et al., 2009).

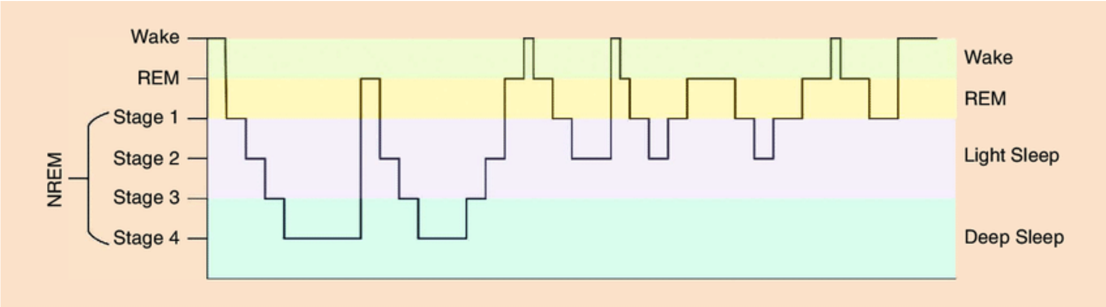


Figure 1.2 Sleep cycle over a night of sleep (Hong et al., 2019).

Specific sleep stages are required for memory consolidation in different memory systems. The role that sleep stages play in memory consolidation is highlighted in two main approaches. These are dual process theory and sequential hypothesis. The dual process theory presumes that the consolidation of various type of memory depends on the specific stage of sleep, whereby SWS sleep enhances the consolidation of declarative memory and REM sleep supports procedural memory consolidation (Gais & Born, 2004; Plihal & Born, 1999). The sequential hypothesis suggests that memory is benefitted by sleep through both REM and SWS sleep in cyclic succession. Originally, this hypothesis assumed that SWS was responsible for the weakening of traces of non-adaptive memory. However, REM sleep is proposed to restore any remaining traces (Ficca & Salzarulo, 2004; Giuditta et al., 1995). A comparison between early vs. late sleep is supported by the dual process hypothesis. For example, this regards a specific approach that compares the effects of retention intervals that take place in the first (SWS-rich) half of nocturnal sleep as well as the second half, which is REM-rich sleep. Declarative memory consolidation was found to be supported by early sleep (SWS-rich), such as word-pairs (Plihal & Born, 1997). This is also assumed to support consolidation for spatial relations (Plihal & Born, 1999). However, REM sleep is shown to be beneficial to memory types that are non-declarative, such as priming skills (Wagner, Hallschmid, Verleger, & Born, 2003) as well as mirror-tracing (Plihal & Born, 1997). However, this is not true for all results. Research shows that SWS also supports various non-declarative tasks, such as discrimination of visual texture (Gais, Plihal, Wagner, & Born, 2000) and rotation adaptation (Huber, Ghilardi, Massimini, & Tononi, 2004). It also seems that REM sleep benefits certain facets of declarative memory (Fogel, Smith, & Cote, 2007). This is particularly the case when using emotional materials (Nishida et al., 2009; Payne, Stickgold, Swanberg, & Kensinger, 2008). Therefore, the findings may mean that tasks used to measure memory are impure and involve more than one memory type. This concurs with the sequential hypothesis which proposes that sleep is beneficial for consolidating both declarative and non-declarative memories when REM and SWS sleep happen successively (Giuditta et al., 1995). The sequential hypothesis is primarily supported by research that introduces natural cyclic sequence disruptions of REM and SWS sleep when subjects are awakened from REM sleep (Ficca, Lombardo, Rossi, & Salzarulo, 2000; Ficca & Salzarulo, 2004). This method has weaknesses, as being awakened from REM sleep induces a stressful response on the subject, which could skew the results (Gais & Born, 2004). Various undisturbed sleep studies which have used correlation analysis have proposed that overnight performance gains on procedural visual discrimination tasks are greatest when SWS occurs in addition to REM sleep successively in post-learning sleep (Gais et al., 2000; Mednick et al., 2003).

Therefore, both SWS sleep and REM sleep might consolidate procedural and declarative memory. Each approach— sequential and dual process hypotheses - can be reconciled by proposing that the succession of both SWS and REM sleep can have benefits for both procedural and declarative memory. However, because declarative memory is integrative in nature (features bound from memories in various memory systems), it may gain increased benefits from consolidating a SWS-associated system. On the other hand, because procedural memories are specific and discrete, they may benefit more from the consolidation of REM sleep-associated (Diekelmann & Born, 2010).

Moreover, REM sleep especially is known to reinforce emotional memory (Tempesta, Socci, De Gennaro, & Ferrara, 2019) and numerous studies on memory and sleep stages link REM sleep to consolidating emotional memories (Groch, Wilhelm, Diekelmann, & Born, 2013; Groch, Zinke, Wilhelm, & Born, 2015; Nishida et al., 2009; Payne, Chambers, & Kensinger, 2012; Popa, Duvarci, Popescu, Léna, & Paré, 2010; Sopp, Michael, Weeß, & Mecklinger, 2017). Wagner et al. (2001) made a comparison of memory retention of neutral vs emotional texts in early nocturnal sleep periods with predominantly late REM-rich sleep and slow wave sleep (SWS) and periods. They found that late REM-rich sleep improves memory for emotional texts and the enhancing effects of sleep rich in REM in emotional material has significance compared to the effect of corresponding wake retention intervals. Other research has also indicated that late sleep rich in REM aids emotional memories, while early SWS-rich sleep does not affect emotional memory (Groch et al., 2013). Additionally, Nishida et al. (2009) demonstrated that napping after encoding neutral and emotional images led to increased consolidation of emotional rather than neutral stimuli; this was correlated with increases in REM sleep.

Carr and Nielsen (2015) discovered that increases in the priming of emotional words after sleeping leads to increases in REM sleep. They suggest that REM sleep causes increased integration and consolidation of emotional memory, in addition to increases in spreading semantic network activations toward other primed associates (Carr & Nielsen, 2015). This theory raises the question of whether REM sleep might additionally enhance false memories for emotional stimuli via an activation of a wider set of associations with the to-be-remembered stimuli or through an increase in gist-based memory representations.

1.3.3 The Role of Sleep in False Memories

As outlined previously, sleep is considered to be one of the primary mechanisms through which potential memories are consolidated (Plihal & Born, 1997). During the process of sleep, newly acquired memory traces are consolidated within the neural pathways that will later be used to retrieve such information (Diekelmann & Born, 2010). Sleep, however, is a complex and multi-factorial process, much as memory is. The consolidation of encoded information is not just carried out via the repetition of information along the relevant neural pathways. Consolidation also takes place in a different way during the sleep process (i.e. through schemas).

Many theories of memory formation and processing have emerged in recent years and these may explain how false memories can be generated in relation to sleep. For instance, it has been suggested that three distinct phases of memory processing occur during sleep: schema formation, schema integration and schema disintegration (Landmann et al., 2014). A schema can be defined as a cognitive framework that helps to organise information. Schema formation defines the way in which pieces of information may be combined to form one distinct schema (Landmann et al., 2014). For example, a young child identifies a cat after having seen different breeds of cats. During sleep, a process known as multi-item generalisation has been known to occur, which plays a vital role in the development of schema (Walker & Stickgold, 2010). Integration reflects the association of memories with existing knowledge and long-term memories, a process that is essential in complex learning such as language learning (Tononi & Cirelli, 2014). Integration of recent and remote (older or infrequently accessed) memories may occur and the incorporating of newer data may lead to the corruption of remote memories (Lewis & Durrant, 2011). Finally, the process of schema disintegration is considered to be the foundation of the creative process, where novel ideas and concepts are generated (Landmann et al., 2014). The formation and integration processes are particularly associated with the potential for false memory generation during sleep because these processes allow for the corruption of memory data based on existing associations and concepts (Landmann et al., 2014). Given the focus on the DRM paradigm in this thesis, this section will consider the role of sleep in false memory generation, based on studies utilised the DRM paradigm.

Sleep has been shown to be a time when a complex set of processes takes place that support memory consolidation as well as memory deterioration (Landmann et al., 2014). Accordingly, sleep may play a considerable role in the generation of false memories. Experimental studies into this phenomenon have investigated this process and the DRM testing procedure has been used with subjects who have subsequently remained awake or who have slept in order to explore the effect of sleep on false memory. Fenn et al. (2009) assessed the role of sleep in the formation of false memories in three experiments, using the DRM paradigm. Participants studied 16 DRM lists either at 10am (wake group) or 10pm (sleep group) and returned 12 hours later for the test phase. They found that sleep was associated with a reduction in false recognition within the DRM paradigm compared to an equal period of wakefulness. Interestingly, they found no difference between the sleep and wake groups in terms of how many studied words were recognised. The researchers explained their results by suggesting that sleep influences the accuracy of episodic memory, thereby reducing false memory recognition. They also suggested that the role of sleep was an important buffer against the formation of errant consolidated information. An additional control group was tested; participants in this group completed all the tasks in one session, either in the morning or evening. No significant differences were found between the morning and evening groups in terms of both correct and false recognition. Furthermore, a small study of 14 healthy older adults supported the findings of Fenn et al. (2009). Lo et al. (2014) demonstrated that there was a decrease in false memory recognition after a night of sleep compared to wakefulness while no difference was found between groups in terms of correct recognition of the studied words.

Therefore, sleep may reduce false memory as the above studies have suggested. It has been suggested that sleep helps to actively consolidate information which supports recollections and might increase source memory performance (Drosopoulos, Wagner, & Born, 2005; Johnson et al., 1993). With regard to the AMT, representations of memory can be accompanied by recollections of the detail of an item's context during presentation instead of a basic sense of familiarity. For instance, remembering specific elements of presentations (i.e. yellow print) might act as a cue for remembering that a specific word has been previously seen. Sleep might increase the perception of detail or make the memory stronger for an internal cognitive process which occurs during encoding, allowing greater efficiency and accuracy of retrieval (Fenn et al., 2009). However, the precise mechanism of the benefits of sleep on episodic memory remain unclear. It is possible that improvement in recollection might increase due to being able to strategically avoid any false memory during retrieval. Diekelmann et al. (2008) found that sleep deprivation – but not sleep – increased false memories within the DRM paradigm, suggesting that sleep itself may reduce false memory.

However, while the above studies found that sleep leads to a reduction in false memory within DRM, a number of studies have demonstrated the opposite relationship between sleep and DRM false memories. These suggest that sleep is more likely to increase the chances of false memories. A study by Payne et al. (2009) used the DRM paradigm and tested subsequent recall of false and correct memories following either a defined period of wakefulness or nocturnal sleep to explore the influence of sleep on false memory. Participants studied word lists at 9am (wake group) or 9 pm (sleep group) and were tested after 12h. Two additional control groups, who had a 20-minute delay between the study and test, were included separately to rule out the effects of circadian rhythm on the effects of false memory recall. Payne et al. (2009) found that there was an increase in both correct and false memory recall after a night of sleep compared to a day of wakefulness. In line with Fenn et al. (2009), they also found no significant differences between the morning and evening control group in terms of correct and false memory. In their third experiment, the authors explored the effect of a short nap (an average of 88 minutes) on false memory. They found that the nap group recalled more critical words than the wake group. This is consistent with the finding of false recall from the sleep and wake groups. Overall, the results from this study indicate that sleep, compared to wakefulness, may lead to an increased probability that false memories will be generated over time. This suggests that sleep helps to distort an original memory that has been stored and form it into a false memory. A similar result was found by Monaghan, Shaw, Ashworth-Lord and Newbury (2017), who also found that sleep enhances false recognition. Diekelmann et al. (2010) conducted an interesting study comparing three distinct sleep-wake states following the application of the DRM paradigm: nocturnal sleep, nocturnal sleep deprivation, and normal daytime wakefulness. The effects of each of these sleep-wake states were compared in order to highlight which state was most likely to be associated with false memories. Compared to memory performance after daytime wakefulness, both nocturnal sleep and sleep deprivation were associated with a significant increase in false memory (i.e. enhanced recall of non-studied words). However, this pattern was only seen in specific participants: those with a low general memory performance (high and low general memory performance groups were created using a median split on adjusted recall, see Chapter 4 for details). Therefore, sleep may induce false memories through variations in consolidation and retrieval of memory traces, but only where there is a weak initial memory encoding event, as seen in individuals with low general memory performance.

In theory, sleep could lead to an increase in false memories through a variety of mechanisms. One possible mechanism is related to the phenomenon that false memories are known to degrade at a slower rate than true memories, suggesting that true memories are more susceptible to being forgotten over time (Hardt, Nader, & Nadel, 2013). Additionally, it has been suggested that the ability to generalise past studied information is more likely after sleeping (Wagner, Gais, Haider, Verleger & Born, 2004). Likewise, when subjects are presented with DRM lists which they have to remember, they might be more likely to generalise past the particularities of the lists and remember the common themes in each list after sleep. FFT may explain the increase of false memory after sleep. Gist traces are responsible for the formation of false memories in FTT and sleep is believed to promote memory integration (Landmann et al., 2014). This may result in representations of the memory becoming more gist-based after sleep. Moreover, general associative memory is considered to be the basis for false memories (Roediger, Balota, & Watson, 2001). It has been suggested that sleep may be an important factor in this type of memory preservation and may be associated with false memories over the long term. Therefore, sleep may increase false memory as it supports memory integration from context-dependent hippocampal systems to neocortical, context-independent representations. This can occur because the number of associations are more widely activated, or because improved memory for the gist leads to a high probability of false memory following sleep (Landmann et al., 2014).

Whilst the findings from the above studies have shown either that sleep increases or decreases false memory, a study conducted by Diekelmann et al. (2008) had a different result. They conducted a study comparing three distinct sleep-wake states following the application of the DRM paradigm (nocturnal sleep, nocturnal sleep deprivation and normal daytime wakefulness). They found that sleep had no specific effect on false recognition when comparing individuals who had slept with those who had stayed awake during the day. They explained their results by saying that this may have been because of the memory test used (i.e. recognition). However, this is unlikely to be the reason because the effect of sleep on false memory was previously observed in studies using the recognition test (for example, Fenn et al., 2009; Lo et al., 2014).

Taking all the evidence together, it is difficult to draw firm conclusions regarding the role of sleep in false memory generation. There are studies to support a range of findings: that false memories may be reduced, increased or unaffected by sleep within the DRM paradigm. The disparity in these findings is likely to be due to the differences in the experimental design. As previously mentioned, there are a number of factors that affect false memory within the DRM paradigm and that may be responsible for these mixed results. For example, the variation in the type of tests used in the studies (i.e. recall or recognition), the length of the word lists, the emotionality of the lists and the level of inter-item connections or some interaction between these factors. There remain a number of areas of uncertainty with respect to the role that sleep plays in false memory formation. Most importantly, there are conflicting findings regarding whether sleep increases or reduces false memories. Hence, there is a need to further explore the influence of sleep on the generation of DRM false memories. Each of these factors need to be examined further in order to confirm the findings of previous research and to further explore the role of sleep in false memory.

To date, none of the above factors have been considered when looking to the effect of sleep on false memory within DRM paradigm. To our knowledge, McKeon, Pace-Schott, and Spencer's (2012) study is the only one that has explored the effect of sleep on emotional DRM false memory. In this study, participants were presented with neutral and negative word lists and assigned to either a sleep or wake group to recall the lists 12 hours later. The study found that false recall for emotional negative and neutral lures was higher after sleep. Moreover, sleep increased correct recall for neutral but not for negative list words. The results from the morning and evening control groups did not differ significantly in terms of correct and false recall but, when comparing all four groups (sleep, wake, am, and pm), in terms of both correct and false recall, the sleep, morning and evening groups, all exceeded the wake group. These findings support previous studies that suggest that sleep enhances gist memory and offers a wider spread of activation, but they do not support the effect of sleep on emotional memory (Lewis, Cairney, Manning & Critchley, 2011; Baran, Pace-Schott, Ericson & Spencer, 2012). However, although McKeon et al. (2012) examined the emotional content of DRM lists, they used only negative and neutral lists. There is a need to study all three (positive, negative and neutral) types of list to address whether sleep influences the way that emotional words generate false memory.

1.4 Objectives of the Research in this Thesis

The aim of this thesis is to examine the role of sleep in false memories and explore the factors that might interact with sleep and influence DRM false memory generation. The literature so far indicates the beneficial role of sleep in declarative memory; however, the role of sleep in false memory is unclear. Additionally, research investigating the specific influences of BAS, the effects of inter-item connectivity and emotionality on DRM false memories after sleep has yet to be fully explored. This research aims to obtain an increased understanding of the factors that influence sleep-dependent consolidation of false memories. It also provides an opportunity to confirm findings of previous research and to further explain the role of sleep on false memory and how different stimuli may influence the effect of sleep on false memory formation.

Firstly, to inform the design of the experiments investigating the role of sleep reported in this thesis, sleep assessments need to be evaluated in order to decide which sleep measures to use. In Chapter 2, I will present and evaluate some common tools used to measure sleep.

As outlined above, memories with emotional events particularly benefit from the effect of sleep on memory consolidation (Payne & Kensinger, 2010; Walker & van Der Helm, 2009). Hence, experiments 1 and 2 reported in chapters 3 and 4, investigated whether the selective enhancement of sleep for emotional memories can be expanded to false memories. It has been established that emotional words have higher level of associations compared to neutral words (Buchanan et al., 2006). Higher association between words gives rise to more false memories being created in a subsequent memory test. Although the only study to have examined the emotional content of DRM lists (McKeon et al., 2012) found no specific effect of sleep on false memories for negative and neutral lists, there is yet to be a study that takes all three valence conditions (positive, negative and neutral) into account. Therefore, I investigated whether variations in valence had a significant impact on the effect of sleep on false memories and if different emotional DRM lists respond differently with sleep. Also, the findings of a comparison between the actigraphy watch and sleep diary are reported in Chapter 3. This was to identify the extent to which there is congruence in total sleep time between the sleep diary and actigraphy watch on a day-by-day basis. This was done in order to check the validity of sleep diaries in measuring sleep duration for a single night. In Chapter 5, I investigate how the manipulation of the associative strength from the words in the list to the critical lure, the interrelations between words in the list and list length have been considered as important factors in the role of sleep in false memories creation. Studying lists that have strong associations between the words and the critical lure or lists that have strong associations between their words increases the probability of activating the critical lure that is similar in meaning.

Conversely the activation of the critical lure is more difficult to achieve in lists with weak semantic associations between list words and the critical lure and those that have less closely related words. Similarly, short lists have been shown to produce fewer false memories because the number of list items is not enough for sufficient semantic activation of the critical lure. Sleep has been found to support the spread of activation and consolidation of semantic connections. Therefore, sleep may offer the required additional activation and lead to enhanced false memories in low BAS /inter-item connectivity lists and shorter lists compared to a period of wakefulness. Hence, in Experiment 3, I manipulated the inter-item connectivity and BAS of the lists and in Experiment 4, I manipulated the length of the presented lists to examine how sleep will affect false memory for these manipulations. Lastly, the results from Experiment 3 prompted me to further examine which sleep stage is involved in false memory formation. It has been suggested that SWS impairs declarative memory consolidation and it is thought of as being important for sleep-dependent memory processing. As the time spent in SWS is 5 times longer during the 'early sleep' period (the first part of the night) when compared to the 'late sleep' period (the later part of the night), this perhaps points to some connection between this period and false memory. Moreover, it has been suggested that memory integration, rule and pattern extraction and generalisation occur mainly during SWS. Hence, in Chapter 6, early and late sleepers were recruited to investigate the role of SWS on false memory formation with manipulated levels of inter-item connectivity and BAS. Finally, Chapter 7 discusses the primary findings of the experiments reported in this thesis, outlines the limitations of the current study and suggests topics that need to be considered in future research. Specific literature and hypotheses relating to each experiment will be covered in the introduction to the relevant chapters.

The experimental works reported in this thesis will contribute to the current understanding of false memory formation. They will shed light on the factors that interact with sleep to give rise to this phenomenon.

Chapter 2: Methodology for the Assessment of Sleep

2.1 Introduction

As reviewed in the previous chapter, there are some questions surrounding the role of sleep on false memory. In order to conduct research in this area, it is important to choose an appropriate method to measure and analyse sleep quality and quantity. This chapter outlines the various methods used to assess sleep and considers their respective strengths and weaknesses.

2.2 Methods Used to Assess Sleep

With growing interest in studying the role of sleep in many areas, the issue of the most appropriate, cost effective and easiest ways to measure sleep has attracted considerable attention. Many researchers consider PSG to be the gold standard for sleep assessment (Marino et al., 2013), and polysomnogram tests such as polysomnography (PSG) typically measure a wide range of variables. However, this high performance comes at a cost – requiring large, laboratory-based equipment. Such equipment is not only expensive and bulky but can also mean that the participant sleeps in an unfamiliar and uncomfortable condition (Marino et al., 2013). First night effects frequently occur with PSG evaluations and can impair sleep in good healthy sleepers (Babson & Feldner, 2015). Moreover, some research into sleep does not actually need all the detailed information given by such equipment like eye movements, brain activity and breathing patterns during sleep. Also, some research needs to collect sleep data for an extended period of time (e.g., a week or more). Therefore, there is a range of alternative methods – both subjective and objective – which can be used to measure sleep, depending on the nature of the data and time span required. The decision about which sleep measure to use needs to be carefully considered because of the inherent difficulties with collecting data, and the associated costs. The following sections reviews the different approaches that commonly used as alternatives to PSG to measure sleep in psychological research, with a view to informing the methods that will be chosen for subsequent chapters.

2.2.1 Actigraphy Watch

There have been major advances in sleep monitoring as a result of the technological progress in the field of wrist actigraphy, which, in recent years, has become central both as a sleep assessment and sleep diagnosis tool (Sadeh, Dan, & Bar-Haim, 2011). Actigraphy is a method of monitoring movement for extended periods of time using a small watch-like device. An actigraphy watch monitors movement during a sleep period. An accelerometer is used for recording motion. Movements indicate wakefulness, whilst periods of stillness are more likely to correlate with being asleep. The actigraphy watch records and stores information on the activity in fixed or varying epoch lengths (e.g. 10 s, 20 s or 1 min). The recorded activity can then be transferred from the actigraphy to a computer for analysis of the data using algorithms. These algorithms are unique and can only be recognised by the software associated with the actigraphy watch (Ward, Lentz, Kieckhefer, & Landis, 2012).

The increase in the use of such devices in sleep research has overtaken that of PSG (Sadeh et al., 2011). They are considered to offer a reliable and valid means of conducting naturalistic sleep studies, and have been demonstrated to produce meaningful and crucially reliable data feedback within research settings (Rupp & Balkin, 2011). Furthermore, this method is significantly less expensive than PSG, and does not involve the discomfort of the PSG process for participants (Landry, Best, & Liu-Ambrose, 2015).

There is a variety of elements of sleep that actigraphy seeks to measure; total sleep time (TST), sleep latency (SL – how long it took to fall asleep), and wake after sleep onset (WASO – the extent to which sleep is disturbed after initially falling asleep). A number of studies have been conducted in order to understand the levels of reliability and validity that are achieved by the technology (Babin, Lee, Halko, Boudreau, & George, 1997; de Souza et al., 2003; Kanady, Drummond, & Mednick, 2011; Kosmadopoulos, Sargent, Darwent, Zhou, & Roach, 2014; Kushida et al., 2001; Sadeh, Hauri, Kripke, & Lavie, 1995). Fonseca et al. (2017) identified that in comparison to PSG, actigraphy watches tended to overestimate TST and underestimate WASO. However, other studies, such as that conducted by Vallieres and Morin (2003) have drawn different conclusions, specifically that total wake time was overestimated, and TST was underestimated. The latter study drew its sample from a population already suffering from sleep disturbances, which suggests that the suitability of actigraphy might be dependent on the nature of the condition which it is seeking to address.

However, a number of studies have concluded that the actigraphy sleep measurements were correlated with PSG. For instance, Tonetti, Pasquini, Fabbri, Belluzzi, and Natale compared the actigraphy watch with PSG in healthy adults, and found no difference between the actigraphy watch and PSG in terms of the accuracy of measurements for TST, WASO and SE. Marino et al. (2013) examined the performance of the actigraphy watch in comparison to PSG with 77 participants for multiple nights. They found a positive significant correlation between the actigraphy watch and PSG. They concluded that actigraphy watches are useful and accurate when measuring WASO and TST; consequently, they asserted that they are suited to use in future studies. A recent study by Brena et al. (2020) compared sleep measurements obtained via actigraphy and PSG. Sleep was recorded for 55 participants using both actigraphy and PSG for five nights. They found no significant difference between the actigraphy watch results and PSG in terms of measuring TST, suggesting that actigraphy is a reliable tool for measuring TST.

In Ancoli-Israel et al.'s (2003) review, the authors reported that the findings from validation studies show that the overall agreement between actigraphy and PSG is up to 93%. In their review, the important features of an actigraphy watch were outlined; it was demonstrated that it does extend continuous recording of nocturnal and daytime sleep periods, is cost-effective and non-intrusive. They also explained how actigraphy was first used only for research purposes and, after numerous reviews on how reliable the actigraphy watch is, professionals began to use it in clinical settings as well.

There are various actigraphy watch devices, and in recent years, the functions of such devices have extended to measure a variety of types of bio-feedback alongside movement. Different brands of actigraphy watches have been validated and used in many studies (Falck et al., 2019; Njamnshi et al., 2012; Tanum & Ringen, 2017; Zinkhan & Kantelhardt, 2016). There are a number of studies which have investigated the validity of different actigraphy brands by comparing them to PSG. One study by Weiss, Johnson, Berger and Redline (2010) compared three different actigraphy watches with PSG and reported on their performance. The watches were; the Actiwatch (Respironics, Pittsburgh, PA); Sleepwatch (Ambulatory Monitoring, Inc., Ardsley, NY) and the Actical (Respironics, Pittsburgh, PA). They found no differences between the three watches in terms of TST. Compared to PSG, the Sleepwatch and Actiwatch showed a stronger positive correlation compared to the correlation between PSG and Actical which was weaker. Overall, Actical was found to accurately measure the duration of sleep; however, in relation to PSG, its performances were weaker than the other two actigraphy watches.

Another study done by Mantua, Gravel and Spencer (2016) used five different brands of actigraphy devices in order to investigate their validity. The actigraphy devices were Basis Health Tracker, Misfit Shine, Fitbit Flex, Withings Pulse O2, and a research-based actigraph, Actiwatch Spectrum. This study compared different sleep measures recorded by the different actigraphy watches and PSG for sleep monitoring. Participants had to wear these different devices while they were sleeping in a way that did not affect their sleep characteristics (two watches were worn on one wrist and three on the other wrist). In terms of measuring TST, there were no major differences between the five actigraphy devices; the Misfit device was the only one close to showing a significant difference. However, all the devices had a significant and strong correlation with PSG for TST scores. The devices with the strongest correlation to PSG were the Actiwatch, Fitbit and Withings. In terms of sleep efficiency, Actiwatch and Fitbit did not differ from the PSG measure whereas Misfit and Withings overestimated SE compared to PSG and the Basis device underestimated SE. In terms of deep sleep, the scores from the Misfits and Withings were different from those of PSG, as they were found to overestimate deep sleep. In contrast, the Basis underestimated deep sleep when it was low and tended to overestimate it when deep sleep was high. In terms of light sleep, there was a significant difference between PSG and the devices, which all underestimated it. Overall, the research demonstrates that the actigraphy watch can be used as a sleep measuring device, but it depends on the measurements and application the researcher is interested in. The researchers noted that the devices are fine to measure TST and sometimes SE; however, their overall reliability is very low, and PSG should continue to be relied upon for research. They also demonstrated that actigraphy devices do not have sufficient information for accurate sleep staging in order to accurately distinguish between light sleep and deep sleep.

The choices for devices used by researchers in sleep studies are varied. Most of the available devices have many useful features for both the researcher and for the wearer but preference for one over another in research should be based on the nature of the study. For example, in one of the experiments reported in this thesis, participants were required to wear the watch for five consecutive days; therefore, the watch needed to meet certain criteria to ensure that it served the study purpose well. Hence, the researcher looked for a watch that was lightweight, waterproof, had an event marker – to enable the participant to mark when they went to bed and what time they woke up – and had a long-lasting battery to ensure that it recorded without any interruptions for charging the watch (see appendix T for a comparison between different brands of actigraphy watches).

The MotionWatch 8® from CamNTech Inc. (see Figure 3) met the study requirements, being only nine grams in weight, waterproof, having an in-built ambient light sensor and event marker and a replaceable battery that lasts for roughly three months. In addition, it has fast USB direct transfer to download the data and it comes with software (MotionWare) which permits the researcher to access and analyse sleep data. Moreover, CamNTech Inc. is based in Cambridge, UK, which made it easy for the researcher to contact the company in relation to any necessary maintenance or any other problem that arose. This also ensured faster delivery with no extra charge.

Figure 2.1 The MotionWatch 8® from CamNTech Inc.

This MotionWatch 8 has been used and validated in a variety of sleep studies in recent years. For one night of recording, Elbaz, Yauy, Metlaine, Martoni and Leger (2012) compared the sleep data from MotionWatch8 and PSG for 70 adults and they found that the MotionWatch 8® provided clear agreement with PSG and accurate measurement of TST, SE and SL. No clear agreement was found between the watch and PSG in term of number of awakenings and WASO, however these measurements were not relevant for the studies reported in this thesis. MotionWatch 8® has been used and reported on in over 100 studies in the past five years in different research areas (Fafrowicz et al., 2019; Falck et al., 2019; Falck, Best, Davis, & Liu-Ambrose, 2018; Leduc, Jones, Robineau, Piscione, & Lacome, 2019; Roveda et al., 2019; Yap, Rice-Lacy, Bei, & Wiley, 2018).

2.2.1.1 Hand Placement of the Device

In a study by Violani, Testa and Casagrande (1998), the actigraphy watch was placed on each wrist, and on the ankles to assess the changes in motor activity in different limbs. They found that the dominant wrist records greater activity than the non-dominant wrist and no differences were found between the two ankle placements (Violani et al., 1998). Sadeh, Sharkey and Carskadon (1994) performed a ‘twin-wrist’ actigraphy validation study together with PSG with 36 healthy participants who wore wrist actigraphy on both their dominant and non-dominant wrists. They discovered that activity levels measured on the dominant wrist were significantly greater than those for the non-dominant wrist. Similar results were found by Natale, Martoni, Esposito, Fabbri and Tonetti (2007) and

Bravi, Cohen, Martinelli, Gottard, and Minciacchi (2017) who concluded that the non-dominant hand is slightly less active than the dominant hand. However, the differences between the dominant and non-dominant hands seems to be related to the activity of the limbs during wakefulness not during sleep. Driller, O'Donnell and Tavares (2017) investigated the difference in sleep data between dominant and non-dominant hands using actigraphy watch and they found no significant differences in sleep measures between hands. They concluded that the accuracy of the sleep measures is not affected by the placement of the device and the actigraphy watch should be worn on the most comfortable hand. Moreover, placement difference assessments have shown that scoring algorithms (a method to calculate sleep data) lack sensitivity when placed on the wrist regardless of any significant differences in activity level (Middelkoop, Van Dam, Smilde-van den doel, & Van Dijk, 1997).

Furthermore, since the validation studies and most of the studies that used actigraphy watch have instructed their participants to use their non-dominant hand, it may better, for consistency, to ask the participants to wear the actigraphy watch on their non-dominant hand to minimise any difference in the results that may be derived from different watch placement.

2.2.2 Sleep Diaries

A sleep diary is a popular tool in sleep studies and forms the basis of understanding patients' subjective complaints. Sleep diaries permit the self-assessment of sleep and they are easy to use and only require a few minutes every day to fill in. Sleep self-monitoring using sleep diaries facilitates identification of nightly perceived metrics and usually includes the following estimation measurements: sleep onset latency (SOL), WASO, SE and a numerical estimation of TST (Miller, Kyle, Melehan, & Bartlett, 2015). Usually, the patient is requested to fill in their diary at the start of the day and record the previous night's sleep with 'approximations' to the closest five minutes (Miller et al., 2015). Compared to the gold standard PSG, Lichstein et al. (2006) found that healthy participants will normally overestimate their overall sleep time, and those who have insomnia will underestimate their sleep time. Similar results were also found by Maes et al. (2014). However, sleep diaries are considered to be the benchmark for subjective sleep assessment and crucial in assessing insomnia (Carney et al., 2012). Rogers, Caruso, and Aldrich (1993) attempted to discover if sleep diaries can be utilised to gather reliable data regarding home wake/sleep patterns across a period of 24-hours. A total of 50 participants (aged 18–66 years) completed sleep diaries while undertaking 24-hour ambulatory PSG monitoring (a more portable but still costly alternative to lab-based PSG). The percentage of agreement between the sleep diary data and PSG data was 87%. Specificity and sensitivity were also high (95.6% and 92.3%, respectively). The results suggested that sleep diaries are a reliable tool for the collection of sleep/wake pattern data.

Much research has examined the agreement between actigraphy sleep data and self-reported measures in particular populations, including women (Girschik, Fritschi, Heyworth, & Waters, 2012), adolescents (Arora, Broglia, Pushpakumar, Lodhi, & Taheri, 2013), and older people (McCrae et al., 2005).

Yet, these have led to inconsistent results in terms of overall agreement and showed differences in level of accuracy between methods. For example, research has discovered that a poor correlation exists between the two methods and that the sleep diary overestimates TST in comparison to actigraphy (Arora et al., 2013; Currie, Malhotra, & Clark, 2004; Landry et al., 2015). Contrastingly, in other research the agreement between the two methods ranged from moderate to good (Gonzalez, Tamminga, Tohen, & Suppes, 2013; Kölling, Endler, Ferrauti, Meyer, & Kellmann, 2016; Lockley, Skene, & Arendt, 1999). A study conducted by Campanini et al. (2017) assessed the agreement between actigraphy and sleep diaries on sleep parameter assessment in healthy adults (aged 45 ± 9 years). Participants completed sleep diaries and wore a wrist actigraphy for 7 days. Positive agreement was found between the actigraphy data and the sleep diary entries for total sleep duration, time spent in bed, sleep start time and wake-up time. The results suggested that each method is considered a reliable sleep assessment tool on its own. Moreover, Werner, Molinari, Guyer and Jenni (2008) concluded that actigraphy and sleep diaries can be used either for sleep start, sleep end, and assumed sleep, instead of variables linked to nocturnal wake time.

It has been suggested that sociodemographic elements, especially age and education level, might have an influence on differences between objective and subjective methods for the assessment of sleep data. (Cerin et al., 2016; Lauderdale, Knutson, Yan, Liu, & Rathouz, 2008). Moreover, a particular level of disagreement can be anticipated, since actigraphy can capture the absence or presence of movement, while self-reported sleep data is a complex variable and might be prejudiced by the perceptions of individuals (Campanini et al., 2017).

Given that there has been a high adoption of electronic devices in recent years, an electronic version of a sleep diary has been designed as one alternative option to a paper diary for individuals. Tonetti, Mingozzi and Natale (2016) conducted research to compare using paper and electronic sleep diaries. They reached the conclusion that there are similarities in their diagnostic power, and that an electronic sleep diary offers many positives compared to paper sleep diaries. The application can reduce the time taken for data entry and automatic scoring, as well as the avoidance of 'parking lot syndrome' (when patients complete a number of days retrospectively at once) and automatically recording the time of day when the entry was completed (e.g., straight after waking). This suggests that both methods – electronic and paper – provide useful sleep information that can be used in sleep assessment.

2.2.2.1 Number of Nights Needed to Measure Sleep Pattern via Actigraphy Watch and Sleep Diaries

Aggregating measures of sleep across a number of nights will stabilise individual differences and therefore reveal how other variables can be predicted. It has been suggested that "the best index of whether aggregation is useful is the reliability estimate of a single measurement item compared with the reliability of scores obtained by aggregating scores over multiple nights" (Sadeh & Acebo, 2002, p. 4).

Very few studies address the stability of actigraphy measurement over a number of nights for individual subjects. (Sadeh et al., 1995; Sadeh, Raviv, & Gruber, 2000). Acebo et al. (1999) suggested that reliability estimation should be undertaken for aggregated values for a period of one to seven nights. The results indicate that value reliability for five-night aggregation for sleep start time, overnight wake minutes, and SE was adequate (<0.70, coefficients greater than .70 indicate an adequate reliability, based on the criteria suggested by Rosenthal and Rosnow,1991, for measures of individual differences).

In healthy young women, Tworoger, Davis, Vitiello, Lentz, and McTiernan (2005) aimed to investigate factors related to sleep quality when using actigraphy with 77 participants for an average of 7 nights. They found that for reliable SE assessment using actigraphy, a minimum of three nights is recommended. They also suggested that data from multiple nights, including weeknights and weekend nights, are necessary in order to acquire reliable estimates of sleep pattern. Using adult participants, Knutson, Rathouz, Yan, Liu, and Lauderdale (2007) explored the intraindividual variability in sleep parameters using actigraphy watches for a period of one-year. They concluded that five to six nights are sufficient for a reliable assessment of sleep duration. This was also supported by Sadeh and Acebo (2002), who suggested in their review that measurements are recorded for a minimum of five nights, and ideally one week, to gather aggregated measurements which are reliable in terms of characterising individual subjects and which, therefore, have a greater chance of predicting other characteristics. They also suggested that the reliability of aggregated values over one or two nights was poor for most measures (Sadeh & Acebo, 2002).

Regarding the sleep diary, limited research has performed an evaluation of the amount of nightly sleep diary records required to reliably estimate sleep. Many elements have an effect on how many might needed, including the sleep measurement under assessment (i.e., sleep onset latency SOL, total sleep time, bedtime, wake time, etc.) and the sample population (i.e., healthy participants or individuals who have a sleep disorder). A commonality which these factors share is what is driving the difference in the number of nights required for reliability, i.e., intra-individual stability. Increased stability necessitates the aggregation of more data points to give better estimates of 'typical' sleep. Wohlgemuth, Edinger, Fins and Sullivan (1999) aimed to investigate the number of nights on average which are required to achieve reproducible, representative, and stable sleep parameter estimates. Both sleep diaries and PSG were used for 32 participants. They found that one week's worth of sleep diary records is required to afford the necessary stability. Similarly, Short, Arora, Gradisar, Taheri and Carskadon (2017) investigated the amount of sleep diary entries required to reliably estimate WASO, TST, SOL and time in bed across 12 weeks. In their study, a total of 1,766 participants from four different countries (United Kingdom, United States, Australia and Qatar) completed a sleep diary for a week. They suggested that a minimum of five nights of sleep diary entries should be completed when measuring time in bed, SOL, and TST. However, WASO was unreliably estimated using one week's worth of sleep diary entries. Among patients with insomnia, two weeks' worth of entries were required for estimating SOL, whilst reliable estimates of WASO required three weeks of sleep diary records (Wohlgemuth et al., 1999). This leads to the conclusion that the amount of nights' sleep which should be recorded is dependent upon the types of sleepers and the settings in which the recording takes place.

2.2.3 Questionnaires

A widely used tool for preliminarily evaluating sleep is a questionnaire (Ibáñez, Silva, & Cauli, 2018). A questionnaire is frequently the initial diagnostic test applied in primary care and provides a common (quantitative) measure about subjective sleep quality. Several sleep questionnaires have been validated using larger statistical analyses, and they represent a common tool which has been used in many sleep studies for many years (Miller et al., 2015). Previous studies have investigated the accuracy of the questionnaires (Jinmei, Rong, Xu, Yi, & Jiong, 2014; Pataka, Daskalopoulou, Kalamaras, Passa, & Argyropoulou, 2014; Silva, Goodwin, Vana, & Quan, 2016). All used PSG as the gold standard and found that questionnaire specificity varied from 50%-96% while sensitivity varied from 73.0%-97.6% (Ibáñez et al., 2018). Single- or multi-item sleep questions can also be used as a sleep measurement. However, a smaller amount of questions and a reduced answer range means that questionnaires are simpler to use, although this does mean that less information is collected (Ibáñez et al., 2018).

A recent study by Mallinson, Kamenetsky, Hagen, and Peppard (2019) compared self-reported sleep duration measured by sleep diaries and sleep questionnaires. They tested whether a sleep-related disorder is linked to differences between diaries and questionnaires regarding sleep duration. From a longitudinal cohort of 1,516 adults, 5,432 questionnaire-sleep diary pairs were compared for self-reported sleep duration. The questionnaire asked for details about sleep history. For example: “How many hours of sleep do you usually get during the night”. The primary measurements from the questionnaire data gave an average for overall sleep duration, average weekend sleep duration, and average weekday sleep duration. The results found a positive correlation between questionnaires and diaries entries of sleep duration. Suggesting that both sleep diaries and questionnaires can be used to compare the collection of self-reported sleep duration, although questionnaires can produce shorter responses than diaries for self-reported sleep duration.

There are numerous questionnaires available to enable participants to report on sleep duration, quality of sleep, sleep type, and daytime functioning (Shahid, Wilkinson, Marcu, & Shapiro, 2012). Here are some of the more popular and applicable validated questionnaires for assessing sleep and associated characteristics (Buysse, Ancoli-Israel, Edinger, Lichstein, & Morin, 2006; Miller et al., 2015).

2.2.3.1 The Pittsburgh Sleep Quality Index (PSQI; see Appendix U for a copy)

The PSQI questionnaire (Buysse, Reynolds, Monk, Berman, & Kupfer, 1989) is a popular self-report measure for assessing sleep quality, with more than 18900 citations as of May 2020. This is a general retrospective evaluation of sleep disturbance and sleep quality recorded over a period of one month. Participants provide scores on 19 factors. The first four factors are open questions, whilst 5-19 are questions rated on a four-point Likert scale. Each individual item score is used for deriving 7 component scores: daytime dysfunction, sleep efficiency, sleep medication used, sleep disturbance, sleep latency, subjective sleep quality, and sleep duration. (Buysse et al., 1989). Lastly, a total overall sleep quality score (from 0 to 21) is calculated by adding the seven component scores. To determine sleep quality a global PSQI score is used, from good to poor (> 5 = poor sleeper). The PSQI is suitable for use in patients and research to differentiate between poor and good sleepers. In the original PSQI validation study by Buysse et al. (1989), healthy, depressed and sleep disordered participants were tested to assess their sleep quality using PSQI. The questionnaire showed high sensitivity and specificity in distinguishing poor and good sleepers. Mollayeva et al. (2016) conducted a systematic review to evaluate the validity and reliability of PSQI. They suggested that the PSQI questionnaire is a reliable tool for identifying poor sleep quality. Moreover, the PSQI has previously been compared with actigraphy in non-clinical samples. Grandner, Kripke, Yoon and Youngstedt (2006) conducted a comparison of PSQI scores with seven days of actigraphy sleep data, depression assessment and concurrent sleep diary entries in 59 older and 53 younger adults. It was found that there was no significant correlation between PSQI scores and actigraphy data in either older or younger adults, however, PSQI scores were correlated with depression assessment scores and sleep diary. The results indicate that subjective measures differ in comparison with sleep actigraphy measures. They suggest that the PSQI doesn't reflective the actual sleep parameters but that it can be used for assessing sleep quality. It should be noted that Grandner et al. (2006) used actigraphy for seven days and the retrospective measure in PSQI – a global sleep quality estimate which spans the past month – and this may explain its limited agreement with actigraphy. Potentially, a longer actigraphy record would have correlated better with the PSQI scores.

2.2.3.2 The Epworth Sleepiness Scale (ESS; see Appendix V for a copy)

The ESS is used to grade a subjective estimation of extreme sleepiness during the day (Johns, 1992). The overall score (out of 24) is comprised of one factor – the tendency to fall asleep over the preceding two weeks. The scale requires participants to self-rate how likely they are to fall asleep across 8 situations (for instance, reading and sitting). Participants rate their likelihood of dozing in the eight settings on a scale of 0-3 (0 = '*would never doze*', 3 = '*high chance*'). A score over 16 indicates extreme sleepiness, which can be implicated in patients with idiopathic narcolepsy or hypersomnolence (Gallo, Foster, & Johnson, 2009).

The ESS has been validated in prior studies. For instance, Vignatelli et al. (2003) aimed to investigate the validity of the ESS against the multiple sleep latency test (MSLT, the gold standard for assessing excessive daytime sleepiness). They compared the results for the two tests conducted on 91 participants and found that ESS positively correlated with MSLT. In a study by Spira et al. (2012), the researchers investigated the validity and reliability of the ESS in a large cohort of 3,059 participants. The participants were required to complete the ESS and wear an actigraphy watch for 24h. The results showed that the correlation between ESS and the actigraphy results was statistically significant. Moreover, Lauderdale et al. (2006) investigated the stability of the scores from two questionnaires, PSQI and ESS, over a period of one year. More than 600 participants completed the questionnaires twice (one year apart), and a comparison of the scores for the two years for each individual showed that 85% of the participants had the same ESS score and 76% of the participants had the same PSQI classification. The authors suggested that the ESS and PSQI questionnaires are a stable measure of sleep, sleepiness and sleep quality.

2.2.3.3 The Morningness-Eveningness Questionnaire (MEQ; see Appendix W for a copy)

The MEQ (Horne & Östberg, 1976) score is correlated using circadian organisation core parameters of humans, for example, time of sleep (Carrier, Humphreys, & Petersen, 1997; Giannotti, Cortesi, Sebastiani, & Ottaviano, 2002; Laberge et al., 2000). Using MEQ in an inter-individual experiment allows the segregating of 'early sleepers' from 'late sleepers' (Baehr, Revelle, & Eastman, 2000; Bailey & Heitkemper, 2001). The MEQ consists of 19 questions which aim to determine when, in the daily temporal span, participants were most likely to be active. The majority of the questions are preferential, in that the participants are asked to indicate when they prefer to wake up or go to sleep. Each question is multiple choice, with the answers given a value. The total score ranges from 16-86, with a lower value corresponding to an evening type and a higher value to a morning type. The MEQ has been validated in previous studies; for example Natale et al. (2006) compared MEQ scores with sleep onset and wake-up time using actigraphy. A total sample of 110 students completed the questionnaire and wore an actigraphy for three nights. The authors found that each sleep-wake rhythm recorded by the actigraphy was significantly correlated with the MEQ score. This suggests that the MEQ is a useful tool for classifying individuals according to circadian type. A similar findings were also found by Thun et al. (2012), indicating that MEQ scores are strongly correlated with actigraphy parameters..

In summary, all the aforementioned methods for assessing sleep have their strengths and weaknesses. Therefore, they need to be used and modified according to particular requirements. For example, PSG is a methodology which offers the most complete and accurate sleep data. It permits precise identification of sleep stages. Yet, it is very costly, and requires special hardware. Thus, it can only be used for short periods of time, with a small number of participants and where the sleep is assessed outside the usual context (e.g., in a hospital). Conversely, actigraphy, sleep diaries and sleep questionnaires offer discretised approximations to actual quality of sleep. They are comparatively inexpensive and result in useful information about sleep, for example, sleep duration and pattern. The benefits of these approaches may also extend beyond the capability to provide measurements of sleep outside the context of a laboratory environment, and to do so for an extended period of time (Van De Water, Holmes, & Hurley, 2011). Because of the greatly reduced cost of such methods when compared to traditional approaches of sleep measurement (i.e. PSG), it has finally become possible for researchers in this field to engage with larger scale samples in an efficacious manner.

2.3 Conclusion

Previous research investigating the role of sleep on DRM false memory has utilised PSG (Diekelmann et al., 2008; Payne et al., 2009; Diekelmann et al., 2010; Lo et al., 2014) questionnaires (Darsaud et al., 2011; McKeon et al., 2012), sleep diaries (Fenn et al., 2009; McKeon et al., 2012) and actigraphy watches (Darsaud et al., 2011; Lo et al., 2014) for the assessment of participants' sleep for experimental purposes. However, based on the financial constraints of the project and on the evaluations of the alternatives to PSG outlined above, Experiment 1 used an actigraphy watch, specifically, MotionWatch 8® alongside sleep diaries and sleep questionnaires. The method of evaluation was adjusted for subsequent experiments and the reasoning for each change is outlined at the relevant points in the thesis.

Chapter 3: Experiment 1: The Role of Emotional Stimuli on Sleep and False Recognition

(Lab-based study)

3.1 Experiment 1: The Role of Negative, Positive and Neutral Stimuli on the Effect of Sleep on False Recognition

3.1.1 Introduction

Emotional information has been found to affect memory processing, and it has been suggested that memories that include emotional events persevere more than memories that have no emotional tone (Adelman & Estes, 2013; Kensinger & Corkin, 2004; Wiesner et al., 2015). It is well known that sleep significantly affects memory consolidation. A growing body of research has established that sleep is central to the consolidation of memory generally (Payne et al., 2009). However, memories with emotional events particularly benefit from the effect of sleep on memory consolidation (Payne & Kensinger, 2010; Walker & van Der Helm, 2009). Payne et al. (2008) presented their participants with negative and neutral scenes positioned on a neutral background. The encoding and retrieval phases for all participants were undertaken 12hrs apart and, during this time, participants who were divided into two groups experienced either a period of sleep or wakefulness. The wake group showed a reduction in memory recognition for the objects and backgrounds in the negative scenes, while the sleep group recognised more negative objects than neutral ones. This suggests that sleep may selectively preserved memory which includes emotional information.

Furthermore, the effect of emotion is also thought to influence the generation of false memory as the emotionality of the stimuli has been found to increase false memory relative to non-emotional stimuli (Gallo et al., 2009; Porter, Spencer, & Birt, 2003). As mentioned in Chapter 1, the DRM paradigm has been used in many studies to investigate the phenomenon of false memory and it has been found that emotionality of the lists is one of the factors that contribute to the increased level of false memory with DRM paradigm (Brainerd et al., 2008; Dehon et al., 2010; Howe et al., 2010; Sharkawy et al., 2008). It has been suggested that the effect of emotion on false memory within the DRM paradigm may be related to a high level of associations between the emotional words compared to neutral words (Buchanan et al., 2006). Dehon et al. (2010) found that the rates of false recognition and recall of critical lures relating to positive and negative lists were higher than for those relating to neutral lists, but correct memory for studied words did not differ between list types. Similar results were found by Bauer, Olheiser, Altarriba, and Landi (2009) who noted that false recall was highest for emotional lists over concrete and abstract neutral lists. Correct recall was higher for concrete lists than the other two lists. Another study was conducted by Zhang et al. (2017) to investigate the effect that emotional words have on false memory with induced moods. Zhang et al. (2017) assigned their participants to one of three different mood conditions (positive, negative, or neutral). Mood was induced using music combined with self-imagery. Participants were then presented with positive, negative, or neutral DRM word lists. They found that participants in the negative mood condition had a higher false recognition rate for negative lures than for the other two types of lures. The false recognition rate for participants in the positive and neutral mood conditions was high across all three types of lures. They also found a similar pattern of results for participants' true recognition of the word lists. They suggested that mood in a negative state selectively aids gist representation of the information with a negative tone.

Taking these results together, if the increases in false memory are associated with the emotionality of the information; and sleep selectively promotes false memories in the DRM paradigm for emotional information, then there should be a greater increase in false memories of emotional material after a period of sleep. To the author's knowledge, there is just one previous study which has used emotional DRM word lists to explore the effect of sleep on false memory. This study was conducted by McKeon et al., (2012), and they used neutral and negative word lists to investigate the false memory formation of emotion stimuli after a night of sleep. The study found that false recall for both negative and neutral lures were higher after sleep, while, correct recall was higher for neutral over negative lists. The authors suggested that it is still unclear whether sleep influences the way that emotional words generate false memory. The reasons for the findings in this study are uncertain; however, the observation that sleep enhanced correct recall for neutral words, rather than negative words, suggests that emotional context may lead to enhancement of gist memory at the expense of specific details, thereby reducing accuracy of recall during emotional word list procedures (Bookbinder & Brainerd, 2016).

Moreover, there was a study conducted by Howe et al. (2010, Exp. 3) which measured false memory for negative and neutral DRM lists after a one-week delay. The participants in the study were presented with a list of words before being asked to recall the words. They were then split into two groups. One group was given a recognition test immediately and the other group was asked to return a week later to complete theirs. The study found that correct recognition of the neutral and the negative word lists decreased over time. False recognition for the neutral lists did not change, but false recognition for the negative lures increased over time. It is important to note that this study did not explicitly measure sleep but, due to the one-week delay, sleep is likely to have had an impact on the findings.

There is evidence to suggest an interaction between sleep and emotional valence in false memory, so it is important to examine this further. Whilst several studies have examined whether or not sleep influences false memory (as reviewed in Chapter 1), none of these studies has included emotion as a factor in the design of their word lists (Darsaud et al., 2011; Diekelmann et al., 2010; Diekelmann et al., 2008; Fenn et al., 2009; Payne et al., 2009). As mentioned above, only one study has investigated sleep using emotional DRM lists (McKeon et al., 2012). But it only examined negative and neutral word lists, and there has yet to be a study that takes positive, negative and neutral lists into account. Moreover, McKeon et al. (2012) used lists that were matched for BAS but not for arousal. Therefore, the effect of sleep observed from the comparisons between neutral and emotional negative word lists could have resulted from the differences in arousal between the words in the list (Brainerd, Holliday, Reyna, Yang, & Toglia, 2010). Hence, this experiment aims to explore the possible effect that negative, positive and neutral DRM lists may have on effects of sleep on false memory. All lists are matched in term of both BAS and arousal.

Based on the findings from Sharkawy et al. (2008), it was decided to use a recognition test rather than the recall test used by McKeon et al. (2012). Sharkawy et al. (2008) presented their participants with negative and neutral DRM lists in a cross-sectional study. Participants completed both recall and recognition test. The researchers found that participants recognised more negative critical lures relative to neutral lures. There was no difference between negative and neutral false recall. A similar pattern was also observed on Howe et al. (2010) study. This suggests that the effect of emotion on false memory is more easily detected in the test of recognition rather than the test of recall.

This study is the first study to use three valenced types of DRM word lists with two being of an emotional nature (positive and negative) and the third being neutral to study the effect of sleep and emotion on DRM false recognition. Following the results from previous studies, the main hypotheses for this experiment are:

1. Correct memory will be greater for negative words relative to neutral words (Howe et al., 2010; McGaugh, 2004)
2. False memory will be higher for negative critical lures (Sharkawy et al., 2008) or higher for both negative and positive lures over neutral lures (Bauer et al., 2009; Dehon et al., 2010).

3. Sleep compared to wake will result in higher levels of false memory for negative and neutral critical lures (McKeon et al., 2012).

3.1.2 Method

3.1.2.1 Participants

A total of 94 native English-speaking students and staff from Keele University signed up for the study. Prior to data collection, exclusion criteria were established that led to the removal of a number of participants. Participants who had a score of 5 or greater in PSQI (n=4), those who had score greater than 9 in ESS (n=2), participants who slept less than 6 hours in the sleep group (n=2) and those who had a nap in wake group (n=1) were excluded from the study. Additionally, 14 participants withdrew from the study, and three more participants were excluded due to missing sleep data from the watch (it is unknown whether this was a technical issue or if the participants took the watch off). The total number of participants who successfully completed the experiment was 68 participants. They were aged 18-65 years (mean age= 21 years, standard deviation= 6.2) and there were 15 males, 52 females and 1 participant who preferred not to say. A sample size calculation was conducted using G*Power (version 3). A minimum of 68 participants was needed to find a medium effect size of sleep (.06) with power of .80 and an alpha value of .05 (see Appendix Q).

Participants were allocated to either the “wake” group or the “sleep group”. The “wake” group studied the DRM word lists at 9AM and returned for testing at

9PM that evening (n=34). Conversely, the “sleep” group studied the DRM word lists at 9PM and returned for testing at 9AM the next morning (n=34). Therefore, the “wake” group did not experience any sleep between the encoding and retrieval phases, while the “sleep” group had a night of sleep between encoding and retrieval phases. All groups were instructed to abstain from caffeine and alcohol two days before the encoding process (i.e. initiation of data collection) and until the experiment is over. Alcohol and caffeine have been shown to influence memory and may also disrupt sleep patterns (Capek & Guenther, 2009). Participants received either £10 payment or 1 hour of course credit as a reward for participating in this study.

3.1.2.2 Stimuli

The DRM word lists that were used in this study varied in their emotional content; they consisted of negative, positive and neutral word lists. An example of the negative word lists is ‘fear, fight, annoy, hate, hatred, wrath, fury, enrage, mad and rage’ which were related to the non-presented word ‘anger’. There were 12 lists in total (4 lists for each condition) with 10 words in each list. The full lists of the three types used are listed in Appendix (A). The word lists used at the encoding stage were taken from Zhang et al., (2017) and were matched on both arousal and BAS (see Table 3.1). The unrelated distractor words were taken from Brainerd et al. (2010) and Roediger et al. (2001).

Table 3.1 Means of BAS, arousal and valence for negative, positive and neutral lists

	BAS	Arousal	Valence
Negative Lists			
Anger	0.293	4.568	6.308
Sick	0.337	4.972	6.527
Lie	0.257	4.922	6.352
Sad	0.273	5.254	6.746
Positive Lists			
Beautiful	0.168	4.362	3.445
Nice	0.150	4.985	3.900
Love	0.277	4.577	3.640
Happy	0.315	4.530	3.550
Neutral Lists			
Foot	0.203	5.212	4.728
Car	0.353	5.175	4.667
Mountain	0.198	4.772	4.393
Music	0.350	4.758	3.947

3.1.2.3 Materials

In order to ensure that participants did not have existing deficits in sleep or memory, which may distort the findings of the assessment, all participants were asked to complete two questionnaires. Firstly, they filled out the Pittsburgh sleep quality assessment (PSQI), this is used to determine poor or good sleeping patterns by measuring components in seven different areas (see Chapter 2). The second questionnaire was The Epworth Sleepiness scale

(ESS). This is a test which determines how quickly one can fall asleep in varying scenarios ranging from standing to sitting or mid conversation. Both questionnaires have a minimum score that a participant needed to achieve in order to be deemed suitable for the experiment.

3.1.2.4 Apparatus

All participants were required to wear an actigraphy watch (Motion Watch8 from CamNtech ®) and complete a sleep diary for the five days prior to the experiment and throughout the study. This was to record their sleep performance during the experimental period.

3.1.2.5 Design

The experiment had two independent variables. The first variable ‘group’ had two levels: sleep and wake. The second independent variable ’valence’ had three levels: positive, negative, and neutral. There were two dependent variables: accurate response to presented words and false response to lures. The experiment comprised three distinct phases. In the first phase, the pre-study phase, participants were asked to complete two questionnaires in order to check their eligibility and they were given the equipment needed for the study. In the second phase, the study phase, participants were presented with lists of words and were asked to remember them for later memory task. In the third phase, the retrieval phase, participants performed a recognition task. The overview of the study and test design is presented in Figure 3.1

.

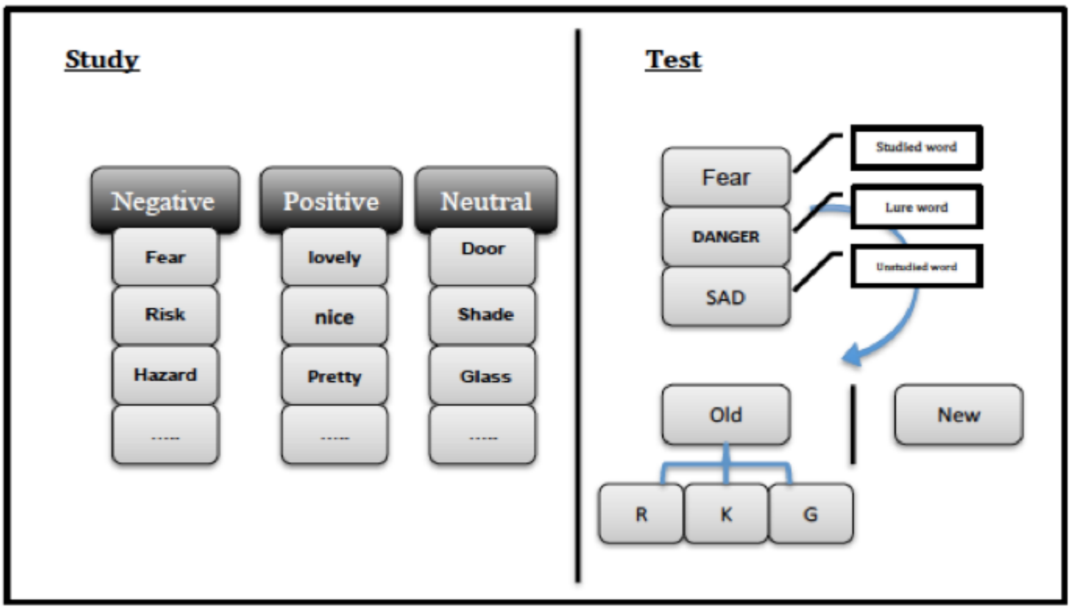


Figure 3.1 The study design is based around three-word list categories (negative, positive and neutral) and the overview of the testing process is presented.

3.1.2.6 Procedure

<u>Pre-study phase</u>

The pre-study phase was to discuss the nature of the study with the participants, establish baseline information, while permitting participants to become familiar with the equipment. This was done five days prior to the actual experiment. All participants had the purpose of the study explained to them clearly and were provided with written information sheet and informed consent regarding their desire to participate and have their data used for research purposes (see Appendix F). In this phase participants were required to complete the two questionnaires, the PSQI and ESS in order to check their eligibility for the study. They were also given the actigraphy watch to wear and a copy of the sleep diary that they need to complete. Participants were then asked to return 5 days later to study the word lists.

Encoding phase

Five days after the pre-study phase, the "wake" group came to the lab at 9AM to study the DRM word lists, while the "sleep" group came to study the word lists at 9PM. Participants were presented with the word lists via Qualtrics (an internet-based experimental software package) on a computer and participants heard the words through headphones. The word lists were presented orally by a digital female voice one at a time in a random order using Qualtrics with a delay of 10 seconds between lists and 2 seconds between words. There were 10 words in each list and they were presented in decreasing order of association strength to the critical lure. Lists were divided into three valence-specific blocks of four lists, the presentation order of word lists was random across participants. A phrase "NEW LIST" appeared on the screen before each new list. At the end of the word presentation participants were asked to return 12 hours later for the recognition test. Subjects in the sleep group were asked to sleep for at least 6hr after the encoding period. While subjects in the wake group were asked to refrain from having naps throughout the day after the encoding process.

Retrieval Phase

This was the final phase of the experiment. Participants were presented with 72 words in a random order. The 72 words comprised of 36 words that appeared in the initial study (studied words), 12 related un-presented words (critical lures), and 24 unrelated distractor words (filler words). Words presented in the recognition test are listed in Appendix (B).

All the words were presented via Qualtrics and all the instructions were provided on the first page before they started the test. Participants were asked to indicate if each word was old or new. A Remember/Know/Guess judgment (Tulving, 1985) followed each item that was identified as old. A 'Remember' response is one in which a person can consciously recollect seeing the item during the list presentation as a discrete event in their past. They may recall details of the event, such as any thoughts, feelings, or memories they experienced when they had the item presented to them, an association they formed with another item, or some aspect of the item's physical or auditory appearance. A 'Know' response is one in which a person can recognise the item because it feels familiar in this context, but they cannot recall its actual occurrence in the experiment. They recognise the item purely on the basis of a feeling of familiarity. A 'Guess' response is when there may be other items that they neither recollect nor recognise on the basis of familiarity, but which they cannot definitely reject. The instructions were based on Dewhurst and Anderson (1999). Debriefing information was provided at the end of this phase (Appendix J).

3.1.3 Statistical Analysis

Data was analysed using IBM SPSS (version 24). Independent t-tests were run to control for any potential differences in sleep scores (PSQI & ESS) between the two sleep/wake groups. Means and standard deviations were calculated for the study variables (see Table 3.2). Total correct recognition was the sum of correct recognition for Remember, Know and Guess judgments. Total false recognition of the critical lures was the sum of false recognition for Remember, Know and Guess judgments. A series of 2 x 3 mixed ANOVAs were conducted on overall correct recognition, overall false recognition, filler recognition and Remember and Know judgments with group (sleep, wake) as the between-participants variable and word valence (negative, positive, neutral) as the repeated measures variable. Bonferroni post hoc tests were used to follow up any significant main effect of word valence. Data are available on the Open Science Framework (https://osf.io/g9mdt/).

3.1.4 Results

PSQI and ESS scores as a function of group

There was no significant difference in usual sleep habits during the past month as measured by the Pittsburgh Sleep Quality Index (PSQI) when comparing participants in the Sleep group (*M* = 3.12, *SD* = 1.04) and participants in the Wake group (*M* = 2.97, *SD* = 2.97, *t*(66) = .587, *p* = .559). Similarly, no differences were found between the Sleep group (*M* = 4.68, *SD* = 2.34) and the Wake group (*M* = 4.26, *SD* = 1.76) regarding participants' sleepiness at the moment of participation as measured by the Epworth Sleepiness Scale (ESS), $t(66) = .818$, $p = .416$. Therefore, any group differences in recognition cannot be explained by sleep related factors.

3.1.4.1 Proportion of recognition responses

Mean proportions and (standard deviations) for correct and false recognition are presented see Table 3.2.

Table 3.2 Correct and false recognition as a function of word type, word valence and sleep sleep/wake groups.

	Sleep			Wake		
Word Valence	Negative	Positive	Neutral	Negative	Positive	Neutral
Studied words						
Remember	.38(.19)	.35(.19)	.34(.23)	.43(.19)	.31(.19)	.28(.19)
Know	.27(.17)	.27(.15)	.14(.13)	.25(.15)	.23(.13)	.16(.16)
Guess	.07(.12)	.09(.10)	.11(.12)	.08(.09)	.10(.10)	.10(.11)
Critical Lures						
Remember	.34(.32)	.37(.27)	.31(.29)	.37(.34)	.28(.32)	.28(.27)
Know	.27(.22)	.26(.24)	.25(.18)	.36(.29)	.30(.25)	.22(.23)
Guess	.12(.18)	.10(.19)	.11(.20)	.04(.10)	.10(.14)	.11(.18)
Filler words						
Remember	.08(.11)	.05(.12)	.02(.05)	.07(.11)	.03(.05)	.01(.04)
Know	.16(.13)	.07(.14)	.08(.19)	.13(.13)	.06(.10)	.07(.10)
Guess	.10(.09)	.07(.10)	.04(.07)	.09(.11)	.07(.11)	.09(.13)

Correct recognition of studied words

The proportion of times that participants recognised a studied word is shown in Figure 3.2 as a function of group and word valence. The 2 x 3 mixed ANOVA showed that there was no main effect of group on correct recognition rate $F(1, 66) = .362$, $p > .05$, $\eta^2_p = .005$. There was a significant main effect of word valence on correct recognition rate $F(2, 132) = 20.02$, $p < .001$, $\eta^2_p = .233$. A Bonferroni post-hoc test found that correct recognition rate was significantly higher for negative ($M = .75$) compared to positive ($M = .67$) and neutral words ($M = .57$) all $ps < .05$ and for positive vs. neutral words ($p < .05$). The interaction between group and word valence failed to reach significance $F(2, 132) = 1.798$, $p = .170$, $\eta^2_p = .027$.

There was a significant main effect of word valence on the Remember judgments of correct recognition $F(2, 132) = 7.089$, $p = .001$, $\eta^2_p = .097$. A Bonferroni post-hoc test found that recognition rate was higher for negative ($M = .41$) compared to positive ($M = .33$, $p < .05$) and neutral words ($M = .31$, $p < .05$).There was no main effect of valence on the Know judgments of correct recognition $F(1, 66) = .362$, $p > .05$, $\eta^2_p = .005$. There was no significant main effect of group for Remember $F(1, 66) = .211$, $p > .05$, $\eta^2_p = .003$ or Know responses $F(1, 66) = .201$, $p > .05$, $\eta^2_p = .003$. Similarly, no significant interactions were observed between group and word valence for the Remember $F(2, 132) = 2.173$, $p > .05$, $\eta^2_p = .032$ or Know $F(2, 132) = .950$, $p > .05$, $\eta^2_p = .014$ responses.

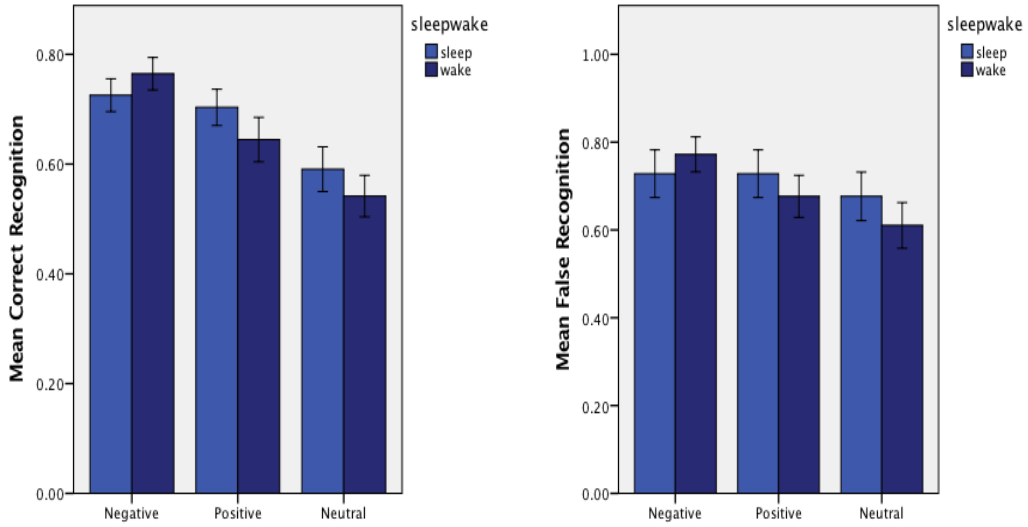


Figure 3.2 Correct recognition rate (left) and false recognition rates as a function of word valence and sleep/wake groups respectively (Error bars represent standard error).

False recognition of critical lures

The proportion of times that participants falsely recognised a critical lure is shown in Figure 3.2 as a function of group and word valence. The 2 x 3 mixed ANOVA showed that there was no main effect of group on false recognition rate $F(1, 66) = .362$, $p > .05$, $\eta^2_p = .005$. There was a significant main effect of word valence on false recognition rate $F(2, 132) = 3.456$, $p = .03$, $\eta^2_p = .050$. A Bonferroni post-hoc test found that false recognition rate was significantly higher for negative ($M = .75$) compared to neutral words ($M = .64$, $p < .05$). No other differences were significant. The interaction between group and word valence failed to reach significance $F(2, 132) = 1.798$, $p > .05$, $\eta^2_p = .027$. Next, three separate 2 x 3 mixed ANOVAs found no main effect of group on the Remember $F(1, 66) = .354$, $p > .05$, $\eta^2_p = .005$ or Know $F(1, 66) = .737$, $p > .05$, $\eta^2_p = .011$. There was no main effect of word valance on the Remember responses $F(2, 132) = .782$, $p > .05$, $\eta^2_p = .012$ and Know responses $F(2, 132) = 2.528$, $p > .05$, $\eta^2_p = .037$. Similarly, no significant interactions were observed between group and word valence for the Remember $F(2, 132) = .893$, $p > .05$, $\eta^2_p = .013$ and Know $F(2, 132) = 1.361$, $p > .05$, $\eta^2_p = .020$ responses.

Recognition of filler words

There were no significant main effects of group $F(1, 66) = .065$, $p > .05$, $\eta^2_p = .001$, valence $F(2, 132) = 28.097$, $p > .05$, $\eta^2_p = .299$ or interaction $F(2, 132) = 1.616$, $p > .05$, $\eta^2_p = .024$ on false recognition of filler words.

3.1.4.2 Signal Detection Analyses

For the sake of completeness and comparability, the data were further analysed using signal detection theory analysis (Green & Swets, 1966). Signal detection analyses were conducted for correct old responses to studied words and false old responses to the critical lure in order to correct participant response bias. This is important for studies manipulating emotion and using longer retention intervals because participants tend to make more false alarms to distractors for emotional stimuli and when retention periods are longer (Pixton, 2011). Signal detection parameters d' (discrimination ability) and C (response bias) were computed using the valenced list items/critical lures as hits and the correspondingly valenced filler items as false alarms. The analysis of d' allows us to examine the discriminability of the participant where higher values indicate better memory performance, while the analysis of C allows us to distinguish the decision criterion used by the participant (Hellenthal, Knott, Howe, Wilkinson &

Shah, 2019). Values above 0 represent more conservative bias and values of 0 indicate no decision bias. To avoid an infinite z value in computing the d′s, all hit and false-alarm rates were corrected by adding 0.5 to the frequency of hits or false alarms and dividing this adjusted frequency by N + 1 where N was the number of old or new trials (see Snodgrass & Corwin, 1988). Signal detection parameters d' and C were analysed separately for correct and false recognition using the same repeated measure ANOVAs as before. Any significant effect of word valence was examined using the Bonferroni correction for pairwise comparisons. The results of d' and C are summarised in Table 3.3.

Table 3.2 Signal detection of discriminability (d') and Criterion Bias(C) for studies words and critical lures

	Sleep						Wake					
	d'			C			d'			C		
	M	LB	UB	M	LB	UB	M	LB	UB	M	LB	UB
Studied words												
Negative	.95	.61	1.30	.00	-.19	.19	.99	.65	1.34	.09	-.09	.29
Positive	1.50	1.17	1.81	.25	.08	.43	1.40	1.07	1.72	.37	.20	.55
Neutral	1.38	1.05	1.70	.51	.36	.66	1.00	.68	1.32	.48	.33	.63
Critical Lures												
Negative	-.14	-.51	.24	.55	.38	.73	.24	-.13	.61	.47	.30	.65
Positive	.45	.06	.85	.77	.61	.93	.78	.38	1.16	.69	.52	.85
Neutral	.90	.57	1.25	.75	.56	.93	.74	.40	1.08	.61	.43	.80

Note: M = Mean; LB= Lower Bound; UB = Upper Bound for 95% confidence intervals.

Correct responses to studied words

For discrimination ability d', there was no main effect of group on correct responses $F(1, 66) = .718$, $p > .05$, $\eta^2_p = .011$.There was a significant main effect of word valence on correct responses $F(2, 132) = 5.344$, $p < .05$, $\eta^2_p = .075$. with better discrimination ability in the positive lists ($M = 1.44$) compared to negative ($M = .97$, $p < .05$). No other differences were significant (all *ps* > .05). The interaction between group and word valence was not significant $F(2, 132) = 1.798$, $p > .05$, $\eta^2_p = .016$.

For response bias *c*, there was no main effect of group on correct responses $F(1, 66) = .466$, $p > .05$, $\eta^2_p = .007$. There was a significant main effect of word valence $F(2, 132) = 17.77$, $p < .001$, $\eta^2_p = .212$, showing a relatively liberal bias in the negative ($M = .050$) compared to positive ($M = .311$) and neutral lists ($M = .496$) lists, all *ps* < .05 and for positive vs. neutral lists ($p < .05$). The interaction between group and word valence was not significant $F(2, 132) = .586$, $p > .05$, $\eta^2_p = .008$.

Correct Remember responses

For discrimination ability d', there was no main effect of group $F(1, 66) = .001$, $p > .05$, $\eta^2_p = .000$ or valence $F(2, 132) = 1.383$, $p > .05$, $\eta^2_p = .021$ on correct Remember responses. Similarly, no significant interaction was observed between group and word valence for the Remember judgment $F(2, 132) = 1.815$, $p > .05$, $\eta^2_p = .027$

For response bias *C*, there was no main effect of group on Remember responses $F(1, 66) = .684$, $p > .05$, $\eta^2_p = .010$. There was a significant main effect of word valence $F(2, 132) = 14.994$, $p < .001$, $\eta^2_p = .185$, showing a relatively liberal bias in the negative ($M = .796$) compared to positive ($M = .993$) and neutral lists ($M = 1.064$) all *ps* < .05. No other differences were significant (all *ps* > .05). The interaction between group and word valence failed to reach significance $F(2, 132) = 1.250$, $p > .05$, $\eta^2_p = .019$.

Correct Know responses

For discrimination ability d', there was no main effect of group $F(1, 66) = .42$, $p > .05$, $\eta^2_p = .001$. There was a significant main effect of word valence $F(2, 132) = 11.537$, $p < .001$, $\eta^2_p = .149$. with better discrimination ability in the positive lists ($M = .655$) compared to negative ($M = .364$) and neutral ($M = .216$) all *ps* < .05. No other differences were significant (all *ps* > .05). The interaction between group and word valence failed to reach significance $F(2, 132) = .706$, $p > .05$, $\eta^2_p = .011$.

For response bias *c*, there was no main effect of group on Know responses $F(1, 66) = .457$, $p > .05$, $\eta^2_p = .007$. There was a significant main effect of word valence $F(2, 132) = 20.280$, $p < .001$, $\eta^2_p = .235$, showing a relatively liberal bias in the negative ($M = .906$) compared to positive ($M = 1.058$) and neutral lists ($M = 1.220$), all *ps* < .05 and for positive vs. neutral lists ($p < .05$). The interaction between group and word valence failed to reach significance $F(2, 132) = .749$, $p > .05$, $\eta^2_p = .011$.

False responses to critical lures

For discrimination ability d', there was no main effect of group on false responses *F*(1, 66) = .907, $p > .05$, η^2_p = .014.There was a significant main effect of word valence on false responses *F*(2, 132) = 12.741, $p < .001$, η^2_p = .075. with better discrimination ability in the neutral lists (*M* = .825) compared to negative (*M* = .052) and for positive (*M* = .614) compared to negative lists, all *ps* < .05. The interaction between group and word valence failed to reach significance *F*(2, 132) = 1.771, $p > .05$, η^2_p = .026.

For response bias *c*, there was no main effect of group on false responses *F*(1, 66) = 1.747, $p > .05$, η^2_p = .026. There was a significant main effect of word valence *F*(2, 132) = 3.641, $p = .029$, η^2_p = .52, showing a relatively liberal bias in the negative (*M* = .512) compared to positive (*M* = .728, $p < .05$)). No other differences were significant (all *ps* > .05). The interaction between group and word valence failed to reach significance *F*(2, 132) = .059, $p > .05$, η^2_p = .001.

False Remember responses

For discrimination ability d', there was no main effect of group *F*(1, 66) = .837, $p > .05$, η^2_p = .013 or valence *F*(2, 132) = 1.074, $p > .05$, η^2_p = .016 on false Remember responses. Similarly, no significant interactions were observed between group and word valence for the Remember judgment *F*(2, 132) = 1.291, $p > .05$, η^2_p = .019

For response bias *C*, there was no main effect of group $F(1, 66) = 2.611$, $p > .05$, $\eta^2_p = .038$ or valence $F(2, 132) = 2.230$, $p > .05$, $\eta^2_p = .033$ on false Remember responses. The interaction between group and word valence failed to reach significance $F(2, 132) = 1.628$, $p > .05$, $\eta^2_p = .024$.

False Know responses

For discrimination ability d, there was no main effect of group $F(1, 66) = .482$, $p > .05$, $\eta^2_p = .007$. However, the analysis revealed that a significant main effect of word valence $F(2, 132) = 3.563$, $p = .031$, $\eta^2_p = .051$. with better discrimination ability in the positive lists ($M = .813$) compared to negative ($M = .535$, $p < .05$) . No other differences were significant (all *ps* > .05). The interaction between group and word valence failed to reach significance $F(2, 132) = .634$, $p > .05$, $\eta^2_p = .010$.

For response bias *C*, there was no main effect of group on Know responses $F(1, 66) = .014$, $p > .05$, $\eta^2_p = .000$. There was a significant main effect of word valence $F(2, 132) = 6.809$, $p < .05$, $\eta^2_p = .094$, showing a relatively liberal bias in the negative ($M = .823$) compared to positive ($M = .981$) and neutral lists ($M = 1.010$) all *ps* < .05. No other differences were significant (all *ps* > .05). The interaction between group and word valence failed to reach significance $F(2, 132) = .382$, $p > .05$, $\eta^2_p = .006$.

3.1.4.3 Methodological analysis

As mentioned in Chapter 2, PSG has been identified as the ‘gold standard’ within the field, however, actigraphy watches are commonly used as an objective alternative (Peltz, Rogge, & Sturge-Apple, 2018). It has been found that sleep data obtained from actigraphy are more accurate than data from a sleep diary when compared to PSG (Lichstein et al., 2006). In comparison to sleep diaries, actigraphy watches do not rely on an individual’s memory and allow for quantitative and qualitative measurements of sleep duration across many days. Most of the studies to date have engaged in a comparison between actigraphy and sleep diaries by evaluating overall TST measurements (i.e. for entire days) derived from the actigraphy and sleep diaries. Hence, sleep data from sleep diary and actigraphy watch collected for this experiment was used to identify in more detail, the extent to which there is congruence in TST scores on a day-by-day basis between actigraphy and sleep diaries in healthy young adults, over five consecutive days.

Sleep diary and Actigraphy watch results

A related t-test revealed a significant difference in the overall TST between the actigraphy watch (*M* = 490, *SD* = 54) and sleep diary (*M* = 507, *SD* = 58, $t(67) = -3.15$, $p < 0.01$). The means and standard deviations for the TST are presented in Table 3.4.

Table 3.4 The means (standard deviations) of total sleep time (TST) in minutes for actigraphy watch and sleep diary.

Day	Actigraphy	Sleep Diary
Day 1	488 (117)	497 (118)
Day 2	475 (102)	496 (91)
Day 3	515 (96)	552 (112)
Day 4	496 (100)	512 (112)
Day 5	456 (89)	485 (110)
Overall	490 (54)	507 (58)

However, related t-tests run on the daily scores reveals that the average difference between actigraphy and sleep diaries does not remain consistently different (see Figure 3.3). There was no significant difference between actigraphy and diary for TST on day one $t(67) = -.85$, $p = 0.394$, day two $t(67) = -1.84$, $p = 0.069$ and day four $t(67) = -1.38$, $p = 0.172$. There was a significant difference on day three $t(67) = -2.91$, $p < 0.01$ and day five $t(67) = -2.95$, $p < 0.01$ with the diary recording a greater mean TST than actigraphy, with mean reported differences of 37 min ($SD = 105$) and 28 min ($SD = 79$) respectively.

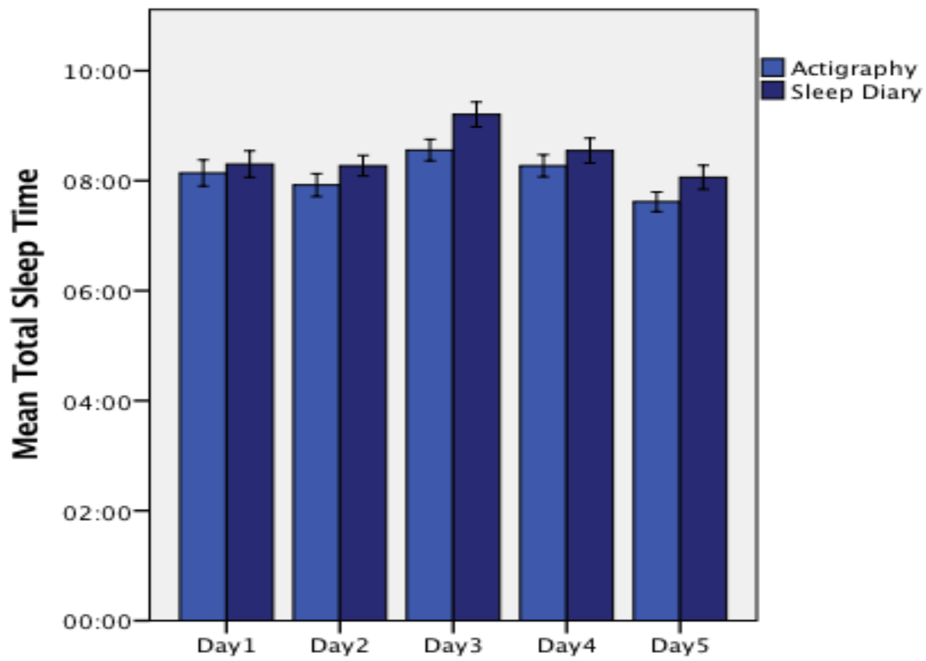


Figure 3.3 TST mean across the five days as assessed via Actigraphy and Sleep Diary (Error bars represent standard error).

The relationship between actigraphy and sleep diary was further explored with Pearson correlation coefficients. Results demonstrated a strong positive correlation between the TST reported by both methods for overall TST ($r = 0.69$, $p < 0.01$), day one TST ($r = 0.68$, $p < 0.01$) and day five TST ($r = 0.70$, $p < 0.01$); and a moderate positive correlation between the TST scores recorded on day two ($r = 0.53$, $p < 0.01$), day three ($r = 0.50$, $p < 0.01$) and day four, ($r = 0.57$, $p < 0.01$).

3.1.5 Discussion

Experiment 1 aimed to investigate the potential relationship between sleep and false recognition and valence on the formation of false memories. The findings suggest that word valence has a significant effect on level of correct recognition, which was highest for negative words than for positive and neutral words. For levels of false recognition of critical lures, it was found that there was a significant main effect of valence. The false recognition rate was significantly higher for negative lures compared to neutral lures. There was no difference in false recognition between negative and positive lures. Lastly, there was no significant effect of sleep on correct or false recognition.

The first finding, that valence has a significant effect on levels of correct recognition is in line with the findings of Brainerd et al. (2008). They used negative, positive and neutral DRM lists and found that true recognition was higher for negative lists over neutral and positive lists. A possible explanation for the increase of correct recognition for negative lists is that negative information appears to evoke more cognitive, behavioral and physiological activity and enhance the cognitive processing compared to positive and neutral information (Taylor, 1991). Thus, negative words would involve deeper encoding, which leads to increased accuracy on the memory task. This may also be explained as being the result of a subsequent decrease in true recognition rates for the positive and neutral conditions, as opposed to just an increase for negative stimuli as such. Moreover, this current finding supported Adelman and Estes (2013) suggestion that the emotionality of words increases the accuracy of the memory performance. Interestingly, this finding appears to conflict somewhat with the findings of Howe et al. (2010) who carried out a similar experiment and found that true recognition was higher for neutral than negative items. However, it is worth noting the key differences between the current study and that conducted by Howe et al. (2010) was that they used only negative and neutral word lists and had no positive valence lists unlike this experiment.

The finding that negative valence words were the most often correctly recognised, followed by the positive ones, then the neutral ones has been echoed in other studies (Inaba, Nomura, & Ohira, 2005). These findings are supported by signal detection analysis with participants appearing to adopt a more liberal response criterion for accepting the negative and positive studied words as old. Wider research has demonstrated some variation in explaining the effect of positive valence words on recognition; however, those studies also suggest that positive words might impair or improve recognition accordingly, depending on many circumstances. For example, Kuperman, Estes, Brysbaert and Warriner (2014) found that negative words are typically recognised more slowly than positive words, suggesting that more time may be taken over negative words, leading to more acceptance of studied negative words than positive ones. Domes, Heinrichs, Rimmele, Reichwald and Hautzinger (2004) implied that stress impairs positive word recognition, suggesting that individual stress levels may reduce positive word recognition below the level for negative word recognition. As for why neutral words are most poorly recognised, previous research demonstrates some form of effect of emotional valence over neutral words (Adelman & Estes, 2013); this might support the view that neutral words naturally exhibit less correct recognition by extension (Brainerd et al., 2010).

Another key finding of this study lends support to the view that emotional valence is a factor in false memory recognition. In the current study, emotional words were found to be the most potent in eliciting false recognition. This is consistent with other studies which found that participants highly recognised emotional lures compared to neutral (Brainerd et al., 2008; Howe et al., 2010; Sharkawy et al., 2008). Moreover, Gallo, Foster, Wong and Bennett (2010) found that false recollection of emotional pictures were greater relative to neutral pictures. This suggests that emotional materials induced attention bias, that resulted in an increase in recognition of emotional items regardless of prior exposure to these items. Fuzzy-trace theory may provide one possible explanation of how this phenomenon might occur. Fuzzy-trace theory argues that memory retrieval can come in two forms; verbatim (exact recollection of facts) or gist (broad recollection of the whole). In the DRM paradigm, gist memory recollection might explain how individuals are susceptible to critical lures. This is because the critical lures are semantically related to those words that appeared in the actual list; and a more comprehensive recollection of wider meaning causes people to falsely recognise these critical lures (Brainerd & Reyna, 2002). In a standard form of the DRM paradigm without any additional extraneous variables, this explanation might serve well in its own right. However, the role of emotion in the current study will need to be further addressed, in order to ascertain how emotional lists might result in false recognition. Bookbinder and Brainerd (2017) actually studied the relationship between emotional images and gist memory, finding that emotional images elevated both true and false memories. They explained this in terms of 'conjoint recognition modelling' whereby erroneous suppression of true recognition is reduced while the relative familiarity of both true and false memories is simultaneously enhanced. Also, it has been suggested that gist representations of emotional lists are stronger and broader than non-emotional lists (Goodman et al., 2011). The current findings suggest that emotional words enhance the similarity of unseen critical lures to the words in the lists. It also reduced the use of verbatim memory to reject the lures (Brainerd et al., 2008). However, signal detection analysis found that participants tended to adopt a more liberal response criterion for accepting negative critical lures as old compared to positive lures. It has been suggested that negative information attracts more attention than positive information and it is processed faster due to its relevance for survival (Hansen & Hansen, 1994). This suggestion points to an attentional bias for the negative information and this may explain the current findings.

As is also the case for correct recognition, neutral lures attracted the least amount of false recognition. A possible explanation for the increase in false recognition for emotional lures compared to neutral ones is that emotional word lists have been suggested to be more dense (i.e. have high inter-connectivity; Howe et al., 2010). This leads to a greater number of associations between the words in the lists which activate the lure to a greater extent, resulting in acceptance of more emotional critical lures than neutral ones.

The lack of impact which sleep had on both true and false recognition alike is a finding which at first seems counterintuitive, as one would assume –based on the literature- that a period of sleep would affect the memory. However, findings from other DRM studies on false memory recognition after sleep have been somewhat mixed. For instance, Lo et al. (2014) demonstrated that there was a decrease in false memory recognition after a night of sleep compared to wakefulness while there was no difference has been found between groups in correct recognition of studied words. Similarly, Fenn et al. (2009) found that sleep reduced the rates of false recognition in the DRM paradigm compared to an equal period of wakefulness. Conversely, Darsaud et al. (2011) found that sleep enhance both false and correct memory recognition.

The current findings are consistent with the findings from Diekelmann et al. (2008) study, in which no difference was observed between the sleep and wake group in terms of false recognition. They found that the increase in false memory was only observed in the sleep deprivation group. The lack of impact of sleep was also found for true recognition and this also supports the findings from the current study. However, one explanation for the absence of the effect of sleep in this study may relate to the test used (i.e. recognition test). Although the effect of sleep was observed in previous studies using the recognition test (Darsaud et al., 2011; Fenn et al., 2009; Lo et al., 2014), recent research has suggested that the recognition test is less sensitive to the effect of sleep than the recall test (Newbury & Monaghan, 2019). This suggestion of the difference between the test of recall and recognition tests within DRM have been explained by the AMT (Roediger et al., 2001). The theory proposes that monitoring cues are activated during the presentation of the words in the test of recognition which allow participants to suppress the related but un-presented words. Meanwhile, in the test of recall, the monitoring cues are unreachable, resulting in the activation of more associated words. This gives rise to more false memory in the recall test than the recognition test after sleep. As sleep has been found to enhance the abilities relating to source monitoring (Johnson et al., 1993) hence, the ability of reject the unseen words would be greater for the test of recognition compared to the recall test (Newbury & Monaghan, 2019). Furthermore, this may also may explain the inconsistency between our results and those of McKeon et al. (2012). Further research is needed using the recall test with the same experimental design as was used for this study to address the issue of the differences between the recognition and recall tests.

The findings of the TST measures derived from actigraphy watches and sleep diaries are consistent with those found by Short et al. (2012), Campanini et al. (2017) and Maich et al. (2018) that is actigraphy generally measured a lower overall TST relative to that recorded by sleep diary. Interestingly, the differences between actigraphy scores and diary scores are not consistently different. Across the five days, the results showed that there was no difference in TST scores between the actigraphy and diary on the first, second and fourth days. While a significant difference was observed on the third and fifth days. This suggests that there is not a consistent level of differentiation between the scores generated from actigraphy and sleep diaries, but instead variation on a day by day basis.

There are a number of factors which may contribute to the unevenness of these scores. The first of these is that there may be an element of ‘getting used to’ the idea of assessing sleep patterns – with a greater awareness of sleep being developed as the study develops. However, the fact that the diary results showed a differentiation from actigraphy measures on both third and fifth days precludes this assessment from the data given. An alternative explanation given by Lawrence and Muza (2018) is that respondents may simply fail to complete their diaries, and only do so on the final day before they are due to submit results. They state that this can lead “to inaccuracies due to problems with recall and a tendency to just put rough guesses to complete the document” (Lawrence & Muza, 2018, p.180). Interpreting our results in light of this suggestion, it is possible that the participants are more diligent at the start of the study, completing the sleep diary daily for the first couple of days, but then becoming less diligent, having to complete it retrospectively. If this is the case, caution is needed when using sleep diaries as participants may become less accurate in their record-keeping over time. Another possible explanation is that the present of actigraphy might have an impact on participants’ performance on the diary. In the first days of the study, they may feel very conscious that their sleep is being monitored, as a result they are more compliant in recording it, but over time that feeling fades. If this explanation is correct, then caution is needed relying solely on self-reported TST.

Despite these inconsistencies, the results showed a strong to moderate correlations in TST between the actigraphy and the sleep diary both overall and across each of the five days. Taken together, the results of this study suggest that the broad pattern of sleep is accurately recorded, but that the precise details are not recorded effectively and that the level of precision will fluctuate across even a 5-day recording span. If one assumes that the objective actigraphy watch is the more consistent of the two recording methods, the current findings suggest that the biggest problem with sleep diaries is not an inherent difficulty in the subjective assessment of sleep, but instead the fact that non-compliance and/or misrepresented data might undermine conclusions about the sleep patterns of respondents, especially where these conclusions rest on details such as TST. A modification to this method might be to send participants a prompt every morning throughout the diary recording period to record their sleep, possibly electronically.

An evaluation of the day to day fluctuations in consistency between TST recorded by actigraphy and sleep diaries over 5 days suggests that sleep diary is a reliable tool to measure sleep for one night as the smallest difference between sleep diary and actigraphy TST was found in day one with difference of only 9 minutes between the two methods. However, the current findings suggest that these fluctuations may increase over time which may have implications for the accuracy of longitudinal studies using sleep diaries. Further studies are required to confirm these findings and explore ways to overcome any fluctuations.

3.1.6 Comments on the Current Experiment

Conducting lab-based sleep research has been challenging. It was not easy to find participants who were willing to participate in such study, even for financial reward. This may be because they were required to commit for a longer time (i.e. five days) compared to other studies. Participants were also required to attend one of the sessions out of working hours (i.e. 9pm) and had to wear an actigraphy watch for the entire duration of the experiment without taking it off or switching it between hands. Furthermore, there were some challenges for the researcher as well, for example, I was required to notify university security every time that I was using the lab at night, so that they could let me access the building. Also, there were many participants who attended the first session and then withdrew without notifying me, resulting in my having to spend time waiting for them at the lab and then having to chase them to recover the watch, which was both time and effort consuming. Generally, lab-based sleep studies take a very long time to complete; therefore, it is not the best option for researchers who have a limited time frame (i.e. PhD researchers). Online data collection is a possible solution in such situation. The methodological analysis reported above suggest that sleep diary on its own is a valid tool to measure sleep for one night. This may help to answer the question of how reliable a sleep diary is if a researcher needs to use it alone to measure TST in an online study. Hence, this finding, in conjunction with the unsustainability of lab-based data collection led to my decision to conduct the rest of the experiments online to complete this thesis in the given time for the PhD.

Chapter 4: Experiments 2: The Role of Emotional Stimuli on Sleep and False Recall and Experiment 2A: The Effect of Time of Day on Emotional False Recall.

(Online studies)

4.1 Experiment 2: The Role of Negative, Positive and Neutral Stimuli on the Effect of Sleep on False Recall

4.1.1 Introduction

As noted in experiment 1, no impact of sleep on correct and false recognition was found. McKeon et al., (2012) studied the effect of sleep on false recall using negative and neutral DRM lists. They found that sleep increases false recall for negative and neutral lists while correct recall for neutral - but not negative - lists were higher after sleep compared to wakefulness. Hence, it was decided to replicate Experiment 1 but with the recall test instead of the recognition test in an attempt to recreate the effects of sleep that McKeon et al. (2012) found. Moreover, as previously mentioned; presentation of the words to participants during the recognition result in monitoring cues being activated. These can lead to the suppression of related un-presented words. As such, in a free recall test with no words presented to participants in the retrieval phase and which requires them to rely on their gist memory, the likelihood of the effect of sleep on false memory may increase. Newbury and Monaghan (2019) conducted a meta-analysis to examine the effect that sleep has on both correct and false memory within studies using the DRM paradigm. They found that the sleep effect was less prominent for the recognition test than the recall test and that the recall test leads to false acceptance of unseen lures to a greater extent than the recognition test. Thus, this experiment aims to explore the possible effect that emotional DRM lists may have on the effect of sleep on false recall. Based on the findings of McKeon et al. (2012), the main hypotheses were as follows:

1. Levels of false recall will be higher after sleep than after wakefulness.
2. Both negative and neutral lists will produce similar level of false recall.

4.1.2 Method

4.1.2.1 Participants

A total of 212 native English-speaking participants aged 18-65 years from the UK signed up to participate in the study via Prolific (www.prolific.ac). Prolific is an online platform created by Oxford and Sheffield Universities to assist in data collection. The website allows researchers to apply custom pre-screening in order to make the study available to specific participants (i.e. who fall within the study criteria). For example; only participants who aged over 18 and below 65 and those who were native English-speakers could access to the study. Participants received a payment of £3 for completing the two parts of the study. All participants were screened for self-reported sleep to ensure that participants did not have existing deficits in sleep. They were required to answer some questions about their sleep (e.g., how would they rate their sleep quality overall, how many hours they sleep per night and are they taking any medicine to help them sleep?). There were two groups in this study the 'sleep' and the 'wake' groups. The sleep group (n=34) studied the list at 9pm and was invited to take the test at 9am the next morning, while the wake group (n=34) studied the list at 9am and tested at 9pm in the same evening. Participants in the sleep group were asked to sleep for at least 6 hours after the study phase, while participants in the wake group were asked to refrain from having naps throughout the day after the encoding process. Participants who slept for less than 6 hours in the sleep group (n=6), and those who had a nap in the wake group (n=2) were excluded from the analysis. Participants who reported sleep problems (n=48) or were taking medication that may affect their sleep (n=4) were also excluded from the final analysis. Additionally, 83 participants failed to complete the second part of the experiment, so they were excluded as there was no recall data collected from them. The total number of participants who successfully completed the experiment was 69 participants (mean age= 36 years and standard deviation= 13.09; 21 males and 48 females). A sample size calculation was conducted using G*Power (version 3). A minimum of 68 participants was needed to find a medium effect size of sleep (.06) with power of .80 and an alpha value of .05 (see Appendix Q).

4.1.2.2 Stimuli

The lists used were the same negative, positive and neutral lists used in the Experiment 1 (see Appendix A).

4.1.2.3 Material and Apparatus

Participants were required to answer some questions about their sleep in order to check their eligibility for the study (see 4.1.2.1). They were also required to complete a sleep diary for one night. Because this was an online study, participants were able to use computers or laptops to complete the study. They were instructed to not use any other devices (i.e. smart phones) as some of the study features may not display on these devices.

4.1.2.4 Design

The experiment had two independent variables. The first variable 'group' had two levels: sleep and wake. The second independent variable 'valence' had three levels: positive, negative, and neutral. There were two dependent variables: accurate response to presented words and false response to lures.

The experiment comprised two phases. In the first phase, the study phase, participants were presented with lists of words and were asked to remember them. In the second phase, the retrieval phase, participants performed a free recall test in which they required to type as many words as they could remember from the study phase of the experiment.

4.1.2.5 Procedure

The data was collected using Prolific and the experiment was created using Qualtrics. The study was distributed at two different times, at 9am and 9pm. This is because the wake group was required to study the lists at 9am and the sleep group at 9pm. For this we used Prolific's prescreening filter to select participants living in the UK in order to make sure that the participants completed the study at these specific times. Information about the study (see Appendix G) and consent information was provided on the first page of the web-based experiment. Participants were asked to check one of two boxes: 'I have read the information on this page and I consent to take part in this study' or 'I do not wish to take part in this study' (see Appendix P). Checking the latter box exited the participant from the experiment. In the next page, they were required to answer some questions about their sleep before moving on to studying the lists. If the participants' answers for the sleep questions did not fit the criteria for the study, they were automatically exited from the experiment. Participants who satisfied the criteria were able to proceed and move to the study phase where they were presented with the word lists. The lists were presented auditorily by a digital female voice one at a time with a delay of 10 seconds between lists and 2 seconds between words. Lists were divided into three valence-specific blocks of four lists, the presentation order of word lists was random across participants. Twelve hours later participants were invited back - using their prolific ID- to complete the second part of the study (the test phase). In that part, they were required to type all the words they could remember from the lists they had heard previously. Before starting the test, they were informed that this part had no time limit. Debriefing information was provided at the end of this section (see Appendix K).

4.1.3 Statistical Analysis

The data were analysed using IBM SPSS (version 24). Two 2 (Group: sleep / wake) × 3 (Valence: positive/ negative/ neutral) mixed ANOVAs were conducted on correct recall rate and false recall rate. Bonferroni post-hoc tests were used to follow up any main effect of valence. Moreover, the number of intrusions was calculated for sleep/wake groups (see Table 4.1). Data is available on the Open Science Framework (https://osf.io/af8qg/).

4.1.4 Results

Intrusions

Table 4.1 shows the intrusions data, with participants in both groups making more unrelated intrusions as opposed to related intrusions. These intrusions were not analysed further.

Table 4.1 Mean of intrusions per person as a function of Sleep/Wake groups.

Intrusions	Sleep	Wake	Total
Related Intrusions	1	1.28	2.28
Unrelated Intrusions	1.85	3.08	4.93
Total Intrusions	2.85	4.36	7.21

Correct recall of studied words

The proportion of times that participants recalled a studied word is shown in Figure 4.1 as a function of group and word valence. The 2 x 3 mixed ANOVA showed that there was no main effect of sleep/wake group $F(1, 67) = 3.12$, $p > .05$, $\eta^2_p = .044$. However, there was a significant main effect of valence on correct recall $F(1.69, 113.50) = 5.76$, $p < .01$, $\eta^2_p = .079$. This was explained by the significantly higher recall for neutral trials (M =.17) compared to positive trials (M =.11, $p < .05$) found in the Bonferroni post-hoc test. No other differences were significant. The interaction between valence and sleep/wake group did not reach significance $F(2, 134) = 1.17$, $p > .05$, $\eta^2_p = .017$.

False recall of critical lures

The proportion of times that participants falsely recalled a critical lure is shown in Figure 4.1 as a function of group and word valence. The two-way ANOVA showed that there were no significant main effects of valence $F(1.98, 132.83) = 679$, $p < .01$, $\eta^2_p = .010$ or sleep/wake group $F(1, 67) = .91$, $p > .05$, $\eta^2_p = .013$ on false recall and no significant interaction between these variables $F(2, 134) = .660$, $p > .05$, $\eta^2_p = .010$.

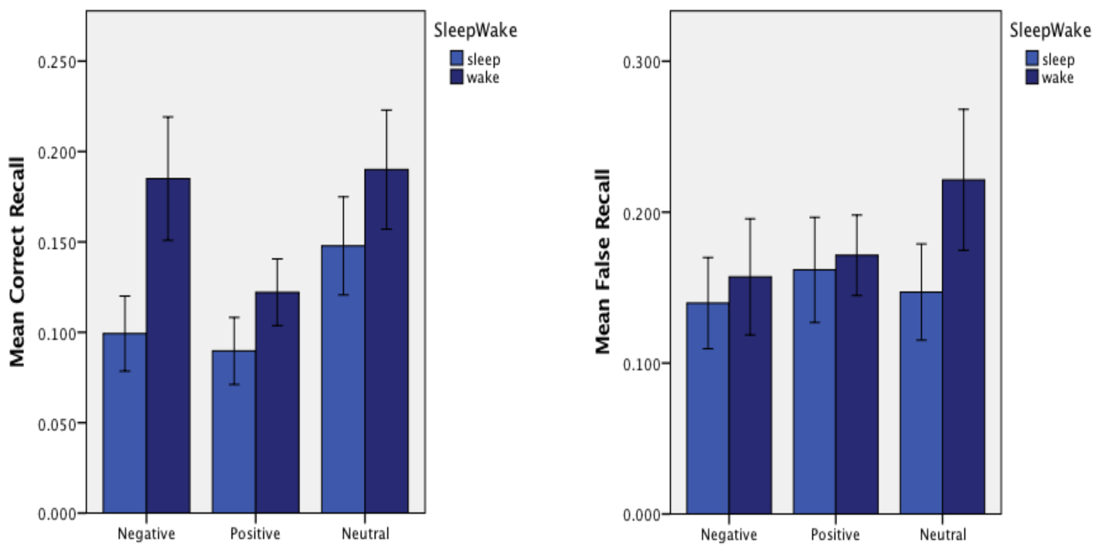


Figure 4.1 Correct (left) and false recall as a function of valence and sleep/wake groups (Error bars represent standard error).

4.1.5 Exploratory Analysis

As no difference has been found between sleep and wake group in correct and false memory. Diekelmann et al. (2010) also found no difference between sleep and wake groups, however they found that sleep was associated with a significant increase in false memory when they analysed participants by general memory performance. Thus, an exploratory analysis was conducted to see whether the participants in this experiment would show the same pattern.

The recalled words were categorized as follows: (1) studied words (words that were presented in the study phase); (2) critical lures (non-presented words that are associated with list items) and (3) intrusions (other words reported at recall which weren't presented in the learning phase). High and low general memory performance groups were created using a median split on adjusted recall (adjusted recall = the number of studied words minus intrusions). This resulted in sup-groups of n=14 and n=21 participants in the 'sleep' and 'wake' groups for the high memory performance respectively, and n=20 for low performance in the 'sleep' group and n=13 in the 'wake' group.

To test if there was a significant difference between high and low memory performance. A two 2 (Group: sleep-high / wake-high) ×3 (Valence: positive/ negative/ neutral) mixed ANOVAs were conducted on correct recall rate and false recall rate (see Figure 4.2). Another two 2 (Group: sleep-low / wake-low) ×3 (Valence: positive/ negative/ neutral) mixed ANOVAs were conducted on correct recall rate and false recall rate. Bonferroni post-hoc tests were used to follow up any main effect of valence. Means and standard deviations are found in Table 4.2.

Table 4.2 Correct/false recall as a function of valence and sleep/wake groups and high/low general memory performance groups

	Sleep/High			Wake/High		
Valence	Negative	Positive	Neutral	Negative	Positive	Neutral
Correct Recall	.19(.14)	.16(.13)	.26(.19)	.27(.22)	.16(.12)	.28(.21)
False Recall	.16(.19)	.14(.23)	.14(.19)	.21(.25)	.17(.16)	.27(.31)
	Sleep/Low			Wake/Low		
Valence	Negative	Positive	Neutral	Negative	Positive	Neutral
Correct Recall	.03(.04)	.04(.06)	.07(.06)	.05(.06)	.07(.06)	.06(.07)
False Recall	.13(.17)	.18(.18)	.15(.19)	.07(.15)	.18(.15)	.15(.20)

Note: standard deviations in parentheses.

High General Memory Performance

The proportion of times that participants recalled a studied word and falsely recalled a critical lure is shown in Figure 4.2 as a function of group and word valence. The 2 x 3 mixed ANOVA with the variables group and valence (sleep/high vs. wake high) on correct recall found no significant main effect of group, $F(1, 33) = 453$, $p > .05$, $\eta^2_p = .014$. There was a significant main effect of valence $F(1.72) = 5.37$, $p = .01$, $\eta^2_p = .140$. Bonferroni post-hoc test found that correct recall was greater for neutral trials ($M = .27$) compared to positive trials (M = .16, $p < .05$). The interaction between valence and group was also not significant $F(2, 66) = .83$, $p > .01$, $\eta^2_p = .024$. For false recall, the 2 x 3 mixed ANOVA showed that there were no main effects of valence $F(2, 66) = .626$, $p > .05$, $\eta^2_p = .019$ or group $F(1, 33) = 1.47$, $p > .05$, $\eta^2_p = .043$ and no significant interaction between these variables $F(2, 66) = .656$, $p > .05$, $\eta^2_p = .019$.

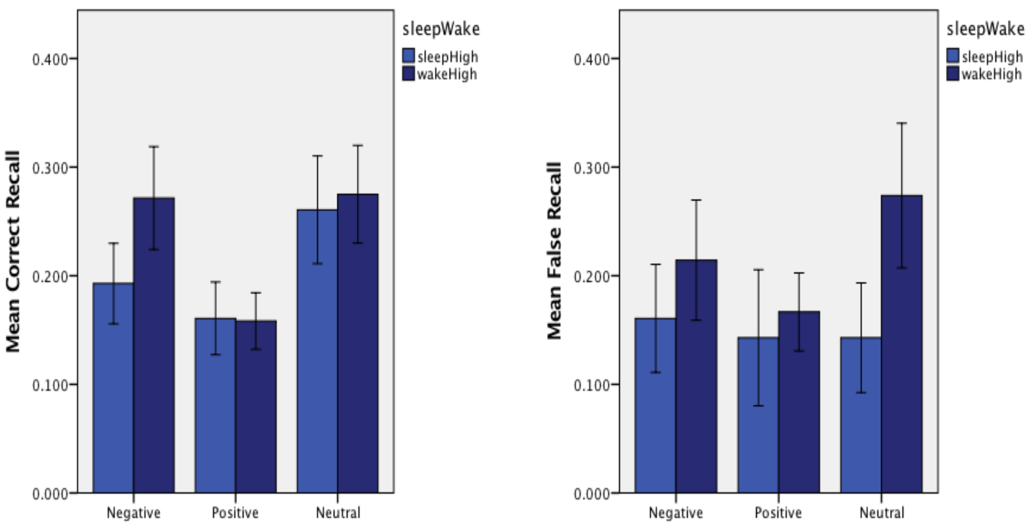


Figure 4.2 Correct recall rate (left) and false recall rates as a function of word valence and high-sleep/high-wake groups respectively (Error bars represent standard error).

Low General Memory Performance

The proportion of times that participants recalled a studied word and falsely recalled a critical lure is shown in Figure 4.3 as a function of group and word valence. The 2 x 3 mixed ANOVA showed no significant main effects of group $F(1, 33) = 1.66$, $p > .05$, $\eta^2_p = .049$ or valence $F(1.96, 62.85) = 1.10$, $p > .05$, $\eta^2_p = .033$ on correct recall and no significant interaction between these variables $F(2, 64) = .725$, $p > .05$, $\eta^2_p = .022$. For false recall, there were no main effects of group $F(1, 32) = .201$, $p > .05$, $\eta^2_p = .006$ or valence $F(2, 64) = 2.01$, $p > .05$, $\eta^2_p = .059$ and no significant interaction between these variables $F(2, 64) = .296$, $p > .05$, $\eta^2_p = .009$.

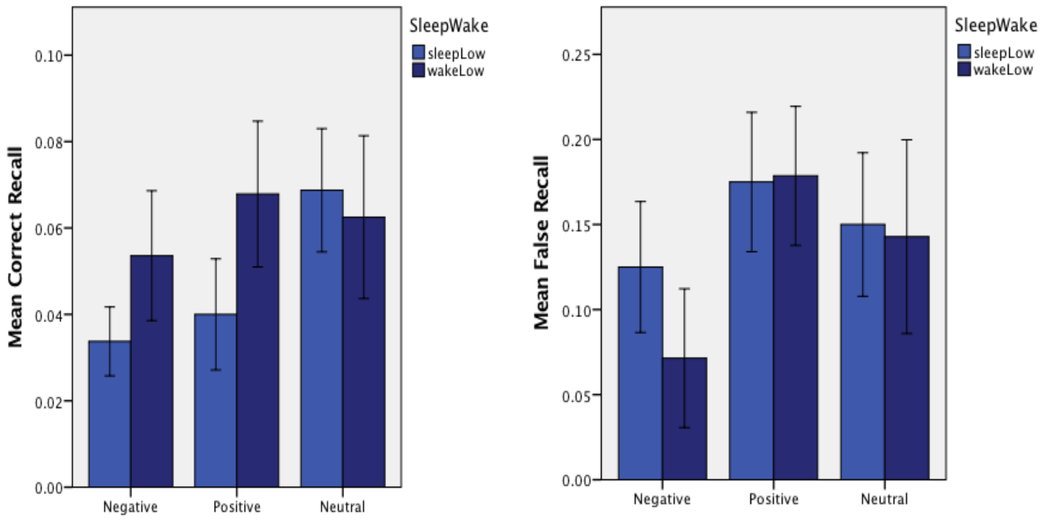


Figure 4.3 Correct recall rate (left) and false recall rates as a function of word valence and low-sleep/low-wake groups respectively (Error bars represent standard error).

4.1.6 Discussion

Like the Experiment 1, Experiment 2 aimed to explore the possible relationship between emotional valence, sleep, correct and false memory. However, this time a free recall test was used, as opposed to a recognition test. The results showed that valence affected correct recall. The neutral valence had the strongest effect on correct recall, compared to positive. However, no significant main effects of valence were found on false recall. The difference between the sleep and wake groups didn't reach significance for both correct and false recall. The finding that sleep has no effect on false memory is consistent with that of the first experiment.

As the key difference between the two experiments is that Experiment 2 was recall-based while Experiment 1 was recognition-based, it is important to focus on a wider body of research geared more towards recall. Findings in this area have varied and produced multiple possible explanations. In contrast to the current study Gomes, Brainerd and Stein (2013) found that positive and negative valence words aided recall more than neutral words. It is worth noting, however, that their experiment differentiated between two different forms of recall; recollective (that which draws upon prior experience for recall) and non-recollective (that which does not). In this instance, it was found that positive valence words improved recollective recall, while negative valence words improved non-recollective recall. As this was a dimension which the current study did not consider, it cannot be ruled out that a combination of factors might have resulted in differing degrees of recollective and non-recollective recall. This could have led to inconsistency between participants and brought the average recall scores for valence words down far enough to drive neutral words up by comparison However, our results were partially consistent with those of Howe et al. (2010). They found that the correct recall of neutral lists exceeded the recall of emotional lists. This result was also supported by the findings from McKeon et al. (2012) study. Across all groups, they found that neutral lists were preferentially recalled more effectively than negative lists. It has been argued that valence information yields more gist than information with no emotional tone (i.e. neutral lists). Accordingly, correct recall for neutral lists should be better than emotional negative and positive lists (Howe et al., 2010). Moreover, it has been suggested that negative information attracts more attention than positive information and it is processed faster due to its relevance for survival (Hansen & Hansen, 1994). This suggestion points to an attentional bias for the negative information. However, these explanations do not hold true, as the current experiment found no significant difference in correct recall between negative and neutral lists. Studies in age differences suggested that older adults show a greater tendency to remember positive information relative to negative and neutral information (Spaniol, Voss, & Grady, 2008). This means that younger adults, by contrast, may show a bias, favoring negative and neutral over positive stimuli. Hence, according to the age of the participants in this experiment, this could be a possible explanation for the current finding

However, the current findings are not entirely consistent with other studies of a similar nature that found that valence had an impact on levels of false recall. Howe et al. (2010) found false recall was highest for the neutral items, suggesting that there was a significant main effect of valence. Dehon et al. (2010) found a significant main effect of valence on false recall where negative lure produced the highest false recall. This trend was further echoed by Brainerd et al. (2010). However, the finding of Sharkawy et al. (2008) is consistent with our finding, as they also found no difference between emotional and neutral false recall. It has been suggested that within the DRM paradigm false recall level tends to be inversely related to correct recall level (Roediger et al., 2001) but this is not the case in this study. The higher rates of correct recall for neutral and negative word lists don’t contribute to false recall. The lack of significance found for valence on false recall may be related to the encoding mechanisms. It has been argued that learning distinctive information obstructs the encoding of relational information and the spread of activation to the unseen lure (Hege & Dodson, 2004; Palmer & Dodson, 2009). As two sets of lists out of three were emotional in our experiment, the absence of the valence effect might be a consequence of this encoding mechanism. (Hege & Dodson, 2004; Palmer & Dodson, 2009).

The finding that sleep had no impact on levels of false recall seems counterintuitive to current understanding and previous findings, since previous studies which used recall test found an effect of sleep on false recall (i.e. McKeon et al., 2012; Payne et al., 2009). However, Diekelmann et al. (2010) found in their study, that false recall was more likely after a period of sleep but this effect of sleep typically only occurred in individuals who performed lower on general memory tests. Therefore, we followed Diekelmann et al.'s (2010) strategy of grouping participants according to their memory performance. The analysis showed that there was no effect of valence on correct recall, suggesting that valence is more likely a more determining factor here. The next key finding pertained to levels of false recall. Interestingly, none of the variables had a significant main effect. In short, the results from the low and high-memory performance groups support the initial findings for the of sleep and wake groups.

However, it is possible that the lack of sleep and valence effects on false recall found in the current experiment could have arisen as a function of the study being conducted online. In order to verify whether this might be a factor and to replicate the control conditions of other sleep studies which check for the possible impact of circadian rhythms (i.e. Payne et al., 2009; McKeon et al., 2012), it was decided to conduct a control experiment in which time of day but not sleep was manipulated.

4.2 Experiment 2A: The Effect of Time of Day on Emotional False Recall

4.2.1 Introduction

Some previous DRM studies has included control conditions to check whether there is any effect of circadian rhythms on memory performance (e.g., Payne et al., 2009; McKeon et al., 2012). They found no difference between morning and evening groups in term of both correct and false memory. However, since Experiment 2 is the first study to explore the effect of sleep on the DRM paradigm using an online experiment it was important to conduct a similar control experiment. This was to establish baseline measures of memory retrieval after a short period of delay in order to rule out any influence of circadian rhythms. Another reason for conducting a control experiment was to ensure that the absence of the effect of sleep on false recall found in Experiment 2 was not due to online data collection.

4.2.2 Method

4.2.2.1 Participants

A total of 151 native English-speaking participants aged 18-65 years from the

UK signed up to participate in the study via Prolific. Participants received a payment of £2 for completing the two parts of the study. The same as Experiment 2 participants were screened for self-reported sleep to ensure they not have existing deficits in sleep. Participants who reported sleep problems (n=45) and those who were taking medication that may affect their sleep (n=5) were excluded from the final analysis. Additionally, 32 participants failed to complete the second part of the experiment, so they were also excluded from the analysis the total number of participants who successfully completed the experiment was 69 participants (mean age= 37 year and standard deviation= 11.13; 34 males and 35 females). There were two groups in this study the "sleep" and the "wake' groups. The sleep group (n=34) studied the list at 9pm while the wake group (n=35) studied the list at 9am. Both groups were invited back to take the test 20 minutes later. A sample size calculation was conducted using G*Power (version 3). A minimum of 68 participants was needed to find a medium effect size of sleep (.06) with power of .80 and an alpha value of .05 (see Appendix Q).

4.2.2.2 Stimuli and Materials

Stimuli and materials were the same as Experiment 2.

4.2.2.3 Design

The experiment had two independent variables. The first variable 'group' had two levels: sleep and wake. The second independent variable 'valence' had three levels: positive, negative, and neutral. There were two dependent variables: accurate response to presented words and false response to lures. The experiment again comprised two phases. The study and retrieval phases were exactly as reported in Experiment 2.

4.2.2.4 Procedure

The current experiment followed the same procedure as Experiment 2. The only difference was that the current experiment used a delay interval of 20 minutes instead of 12 hours.

4.2.3 Statistical analysis

The data was analysed using IBM SPSS (version 24). A 2 (Time of day: AM / PM) ×3 (Valence: positive/ negative/ neutral) mixed ANOVAs were conducted on correct recall rate. A separate ANOVA was run on the false recall data. Bonferroni post-hoc tests were used to follow up any main effect of valence. Means and standard deviations for the control data are reported in Table 4.3. The data are available on the Open Science Framework (https://osf.io/s2nxr/)

4.2.4 Results

Correct recall of studied words

The proportion of times that participants recalled a studied word is shown in Figure 4.4 as a function of group and word valence. The 2 x 3 mixed ANOVA showed no main effect of time of day group $F(1, 67) = .33$, $p > .05$, $\eta^2_p = .005$. There was a significant main effect of valence on correct recall $F(2, 134) = 5.53$, $p < .01$, $\eta^2_p = .076$. A Bonferroni post-hoc test showed that correct recall was significantly higher for neutral trials ($M = .19$) compared to positive trials ($M = .12$, $p < .01$). No other differences were significant. The interaction between valence and time of day group was not significant $F(2, 134) = 1.44$, $p > .05$, $\eta^2_p = .021$.

Table 4.3 Correct/false recall as a function of valence separately for time of day groups.

	AM			PM		
Valence	Negative	Positive	Neutral	Negative	Positive	Neutral
Correct Recall	.19(.21)	.12(.16)	.19(.18)	.14(.15)	.12(.11)	.19(.17)
False Recall	.17(.26)	.14(.18)	.16(.25)	.18(.20)	.19(.18)	.16(.23)

Note: standard deviations in parentheses.

False recall of critical lures

For false recall, the proportion of times that participants falsely recalled a critical lure is shown in Figure 4.4 as a function of group and word valence. The 2 x 3 mixed ANOVA showed that there were no main effects of time of day group $F(1, 67) = .311$, $p > .05$, $\eta^2_p = .005$ or valence $F(1.74, 121.37) = .068$, $p > .05$,

η^2_p = .001 and no significant interaction between these variables $F(2, 134)$ = .417, $p > .05$, η^2_p = .006.

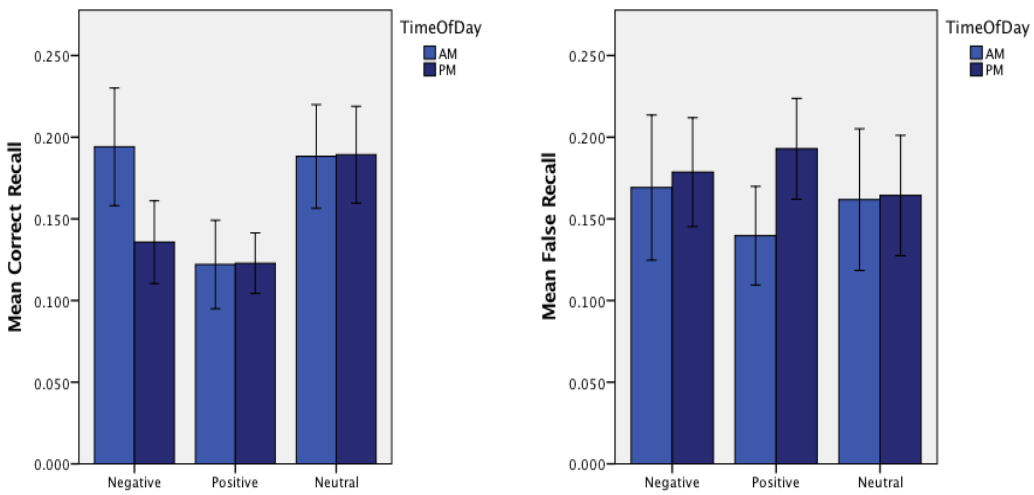


Figure 4.4 Correct (left) and false recall as a function of valence and Morning (AM) and evening (PM) groups (Error bars represent standard error).

4.2.5 Discussion

The findings from the control group showed that the time of day for the group had no significant effect on correct or false recall. These findings are consistent with those of Payne et al. (2009) and McKeon et al. (2012); they used the DRM paradigm and a similar procedure as the current study. They found that the level of correct and false memory was similar between participants tested in the morning and those who were tested in the evening. This suggests that there is no effect of circadian variation on correct and false memory generation and the control study reported here extends that finding to the online domain.

Also, the current results suggest that the findings of Experiment 2 are not due to issues related to online data collection or circadian rhythms. One possible explanation may be related to the accuracy of the participants in reporting their sleep duration during the night between the learning and test phase. They were instructed to adhere to the requirement to sleep for a minimum of 6 hours and while all participants included in the analysis reported a sleep period of between 6 and 9 hours, it is possible that participants failed to adhere to this criterion. Alternatively, it may be that there are individual differences in responding to emotional stimuli that are interfering with the findings. Or it may be that other factors are at play such as the factors investigated in the subsequent chapters.

4.2.6 Summary of Experiments 1 and 2 Results

Experiments 1 and 2 reported in this thesis explored the link between sleep and false memory formation through emotional DRM lists using a recognition test (in Experiment 1) and a free recall test (in Experiment 2). The results of these experiments showed that participants correctly and falsely recognised more negative words and tended to correctly recall more neutral words than emotional words. There was no effect of emotion on false recall. No effect of sleep was found in either experiment. The contradictory findings from the two experiments, with regard to the effect of emotion on false memory, support the belief that emotional critical lures are more likely to be accepted in a recognition test than a recall test (Howe et al., 2010; Sharkawy et al., 2008). The lack of an effect of sleep was an unexpected finding but this may be a result of many factors. For instance, the type of lists used may have affected the results; whilst the lists have been used previously in the study by Zhang et al. (2017), they were used to test a different factor other than sleep (i.e. mood). Another possible explanation is that it may be the case that sleep in general does not affect false memory formation but that a specific stage of sleep has the effect. It has been found in previous research that REM sleep enhances emotional memories (Groch et al., 2013; Nishida et al., 2009; Wagner et al., 2001). This maybe point to a possible link of REM sleep to the enhancement of emotional false memory.

However, the current findings suggest that sleep (or a period of consolidation) is not overtly affecting levels of false recall or recognition in this study. While it has been well established that sleep is a key to memory consolidation, the results suggest that false memories themselves are likely not to be formed during this time (Stickgold & Walker, 2007). There is some debate about this, as multiple studies support the notion that false memories are linked to some kind of an impairment in the consolidation process (Darsaud et al., 2011; Lo et al., 2014; Payne et al., 2009). However, other studies suggest that false memories are formed at the moment of encoding and later consolidated. Instead, the false memory may be the result of missing or misinterpreted details caused by acute stress, arousal, distraction, or the individual's own working memory capacity (Buchner, Rothermund, Wentura, & Mehl, 2004; Corson & Verrier, 2007; Smeets, Jelicic, & Merckelbach, 2006b).

Overall, while this study produced some intriguing findings which merit further research, it also served to highlight the complex nature of false memories and emotional valence.

Chapter 5: Experiments 3-4: Effect of Sleep, Inter-item Connectivity, BAS and list length on False Memory Formation Within DRM Paradigm.

(Online Studies)

5.1 Introduction

As mentioned in Chapter 1, previous studies using the DRM paradigm have resulted in inconsistencies when looking at whether sleep increases false memory or reduces it. Sleep has been found to increase (Darsaud et al., 2011; Diekelmann et al., 2010; Payne et al., 2009), decrease (Fenn et al., 2009) and have no effect on (Diekelmann et al., 2008) the creation of false memories. Newbury and Monaghan (2019) conducted a meta-analysis which pointed to some aspects of the lists that may contribute to the effect of sleep on false memory within the DRM paradigm, specifically emotionality of the lists, inter-item connectivity (the density levels of the relationships between words in a list that is semantically related), backwards associative strength (BAS) and list length. Experiments 1 and 2 reported in this thesis explored one of these factors (i.e. emotionality of the lists), but contrary to predictions the effect of sleep on false memory was not observed. In this chapter, it was decided to leave emotional lists behind and manipulate BAS and list length instead. In line with the findings of Newbury and Monaghan's meta-analysis (2019), it is predicted that there will be a significant effect of sleep on false memory which will interact with BAS and list length.

Spreading activation has been considered a potential explanation for the generation of false memories in DRM tests (Collins & Loftus, 1975). When participants are exposed to word lists, activation spreads through the semantic network to activate unseen words that are related in meaning to the presented words. According to activation monitoring theory (AMT), studying words associated with a lure leads to activation of the lure representation in semantic memory (Roediger et al., 2001). This activation increases the probability that errors will be made in association with that lure on subsequent memory tests. Hence, words that have stronger semantic associations with the lure word result in higher levels of false memories, than words with weaker associations. (Newbury & Monaghan, 2019). This theory elicits a question about the extent to which the role of sleep on false memory creation is affected by varying the BAS and inter-item connectivity of the words in a list.

There is consensus about the important role of the associative relationships between words as a key factor for false memory formation in the DRM paradigm (McEvoy et al., 1999; Roediger et al., 2001). This is also supported by the finding that the strength of the associative connection from the seen words to the unseen critical lure (i.e. BAS) is considered as a strong predictor of false memory (Roediger et al., 2001). That is, DRM lists with high BAS may activate the lure word as the list words are presented, more than DRM lists with low BAS. Knott, Dewhurst, and Howe (2012) explored the roles of BAS and inter-item connectivity within DRM and category lists. They found that false memories for both recall and recognition tests were greater for items with high

BAS. They suggested that the strong association between the critical lure and list words is the primary predictor of false memories. In terms of inter-item connectivity, they found that false recall and recognition were higher for low inter-item connectivity lists. They explain this by stating that strong relationships between list items allow for better encoding of studied words, making those words more distinguishable from the lure words. However, when McEvoy et al. (1999) manipulated the density of the word in each list, they found that both correct and false recognition increased when the word lists had a higher interconnection.

The disparity in McEvoy et al's and Knott's findings regarding the role of inter-item connectivity on false recognition could be related to the experimental designs of the two studies (i.e. different list lengths used; see Chapter 1 for more details). However, acording to the FTT, gist traces are responsible for the formation of false memories whereby gist representation refers to the general semantic representation generated by the studied words. From this standpoint, lists with stronger interrelations would produce high levels of false memory since extracting the theme word from closely related words is easier. Moreover, the AMT proposes that the study of words associated with a lure leads to activation of the lure representation in semantic memory (Brainerd, Wright, Reyna & Mojardin, 2001). Hence, strong association between the words can make the activation of the critical lure easier (or stronger). Meanwhile, in less closely related (sparser) DRM lists, more spreading activation would be required over semantic networks.

Moreover, such activation is also seen in long DRM lists, where activation of the lure word is facilitated by the large number of related words in the list. In an experiment to examine the effect of list length on false memory within the DRM paradigm, half of the participants in different age groups (children and adults) were randomly assigned to a 7-word list condition and the other half assigned to a 14-word list condition (Sugrue et al., 2009). The findings indicated that participants assigned to the longer lists recalled more critical lures than those assigned to the shorter lists, especially among adults. Sugrue et al. (2009) explained these results by stating that longer lists activate the critical lures more than shorter lists because critical lures in shorter lists received less activation from the small number of related words in the lists.

Interestingly, it has been suggested that sleep has a facilitative effect on memory when a task is difficult (Newbury & Monaghan, 2019). Sio, Monaghan and Ormerod (2013) gave their participants remote-associate tasks (RATs; a test where participants are given three words and asked to think of a fourth word that relates to each of the given words) to solve. They manipulated the task difficulty by varying the strength of the stimuli-answer associations. When participants returned to have a second attempt at solving previously unsolved problems, those who had had an intervening period of sleep completed more of the hard RATs relative to those who had had no sleep and those who completed all the tasks in the same session. The hard problems were characterised by having fewer stimulus words directly associated with the target word and, as a result, they required a broader spread of activation through the semantic network before the remotely associated target word was activated to facilitate the solution. The explanation for harder problems being facilitated by sleep is also relevant for false memory since associative activation is one of the key processes proposed as underlying the creation of false memories, as mentioned above.

Taken together, if increases in the inter-item connectivity in DRM lists and the number of words in each list result in closer associations, such increases offer easy access to the critical lure that is similar in meaning. On the other hand, if sleep is more beneficial with more complex tasks, then one should expect that sleep would lead to increases in false memories for lists with lower connectivity and those with fewer words. This is because activation of the critical lure would be more difficult to achieve. Consequently, more spreading activation would be required over semantic networks. To examine this idea further, two separate experiments were conducted to explore the connection between sleep, false memory and these two factors (i.e. inter-item relations and list length). In Experiment 5, the level of inter-item connectivity and backward associative strength (BAS) of the lists were manipulated. While in Experiment 6, the length of the lists was manipulated to examine how this will affect false memory after sleep. The main hypotheses for these experiments are:

1. Higher inter-item connectivity, higher BAS and longer lists will produce higher level of false recall.

2. Sleep compared to wake will result in higher levels of false memory.

3. Based on the findings of Sio et al. (2013), there will be a greater increase in

false memory for critical lures related to low inter-item connectivity and shorter lists after sleep.

5.2 Experiment 3: The Role of Inter-item connectivity and BAS on DRM False Recall After Sleep

5.2.1 Introduction

DRM study lists constructed to vary list BAS and inter-item connectivity were used to investigate changes in the information that becomes accessible following sleep. I expect that lists with a weaker semantic relationship (sparser semantic network) will give rise to more false memories following sleep than following an equivalent period of wakefulness. Since it was found that the effect of sleep is less prominent for recognition tests than recall tests (Newbury & Monaghan, 2019), it was decided to use a recall test in this experiment.

5.2.2 Method

5.2.2.1 Participants

A total of 232 native English-speaking participants from the UK were recruited online via Prolific to participate in the study. They were aged 18-65 years and received a payment of £3 for completing both phases of the study. The screening criteria were the same as in Experiment 2. Participants who slept for less than 6 hours in the sleep group (n=1), and those who reported sleep difficulties (n=53) or were taking medication that might affect their sleep (n=2) were excluded from the final analysis. Additionally, 108 participants were excluded because they failed to complete the second part of the experiment. The total number of participants who were eligible and successfully completed the two phases of the experiment was 68 participants (mean age 34 years and standard deviation 11.27; 26 males and 42 females), 34 participants in each group. A sample size calculation was conducted using G*Power (version 3). The power analysis showed that a minimum of 62 participants are needed to find a medium effect size of sleep (.06) with power of .80 and an alpha value of .05 (see Appendix R). However, to keep it consistent with previous experiments 68 participants were recruited.

5.2.2.2 Stimuli

Sixteen DRM lists of 8 words were presented to the participants. These consisted of 4 lists with high BAS/high connectivity, 4 with high BAS/low connectivity, 4 with low BAS/high connectivity, and 4 with low BAS/low connectivity (see Table 5.1). The word lists were taken from (Knott et al., 2012). The lists used in this experiment can be found in Appendix (B).

Table 5.1 Means of BAS and inter-item connectivity for each word list

	BAS	Connectivity
BAS: High/Connectivity: High		
Chair	0.27	2.7
Cold	0.37	2.6
King	0.30	2.6
Sweet	0.23	3.0
BAS: High/Connectivity: Low		
Slow	0.26	.8
Sleep	0.13	.9
Needle	0.25	.8
Smell	0.33	1.1
BAS: Low/Connectivity: High		
Thief	0.09	3.2
Army	0.13	2.1
Bread	0.11	2.2
Chemistry	0.11	2.8
BAS: Low/Connectivity: Low		
Cup	.11	.6
Pen	.09	.8
Smoke	.11	.8
Anger	.09	1.1

5.2.2.3 Materials

Materials were the same as Experiment 2.

5.2.2.4 Design

The experiment had three independent variables. The first variable ‘group’ had two levels: sleep and wake. The second independent variable ’BAS’ had two levels: high and low. The third independent variable ’connectivity’ had two levels: high and low. The dependent variable ‘type of item’ had two levels: accurate response to presented words / false response to lures.

The experiment comprised two phases. In the first phase, the study phase, participants were presented with lists of words and were asked to remember them. In the second phase, the retrieval phase, participants performed a free recall test in which they required to type as many words as they could remember from the study phase of the experiment.

5.2.2.5 Procedure

In order to collect data from sleep and wake groups, the experiment was distributed at two different times, at 9am and 9pm. Once participants signed up to the experiment, they were automatically transferred to the Qualtrics web page where the experiment was created. All information about the study was provided on the first page of the web-based experiment (see Appendix G). Participants were required to consent by checking one of two boxes: ‘I have read the information on this page and I consent to take part in this study’ or ‘I do not wish to take part in this study’ (Appendix P). Checking the latter box, exited the participant from the experiment. Checking the first box allowed the participant to move to the next page where they were required to answer some questions about their sleep before moving on to study the lists. If the participants' answers fell within the criteria of the study, they were allowed to proceed to the study phase where the word lists were presented. The word lists were presented orally in a digital female voice one at a time in a random order using Qualtrics with a delay of 10 seconds between lists and 2 seconds between words. There were 10 words in each list and they were presented in order of descending backward associative strength to the critical lure. Lists were divided into four blocks of four lists, the presentation order of word lists was random across participants. A phrase "NEW LIST" appeared on the screen before each new list. The sleep group studied the lists at 9pm and took the test at 9am the next morning, while the wake group studied the lists at 9am and completed the test at 9pm in the same evening. During the test phase, participants were required to write all the words they could remember from the lists they had heard previously. Before starting the test, they were informed that this part had no time limit. Debriefing information was provided at the end of this section (see Appendix M for a copy of the debrief sheet).

5.2.3 Statistical Analysis

The data were analysed using IBM SPSS (version 24). The data was analysed using a 2 × 2 × 2 mixed ANOVA with the between-participants independent variable of Condition (sleep vs. wake); and the within-participants independent variables of BAS (low vs. high) and Connectivity (low vs high). A series of 2 (group: sleep vs. wake) x2 (BAS: low vs high) x2 (Connectivity: low vs. high) mixed ANOVA's were conducted separately on the correct recall data then the false recall data. The number of related and unrelated intrusions was calculated for sleep/wake groups (see Table 5.2). Means and standard deviations are reported for correct and false recall in Table 5.3. Data is available on the Open Science Framework (https://osf.io/r5cd2/).

5.2.4 Results

Intrusions

Table 5.2 shows the intrusion data, with participants in both groups making more unrelated intrusions as opposed to related intrusions. These intrusions were not analysed further.

Table 5.2 Frequencies of Intrusions as a function of sleep/wake conditions.

	Sleep	**Wake**	**Total**
Related Intrusions	28	30	58
Unrelated Intrusions	90	78	188
Total Intrusions	118	108	246

Correct recall of studied words

The proportion of times that participants recalled a studied word is shown in Figure 5.1 as a function of connectivity, BAS and sleep/wake groups. The 2 × 2 × 2 ANOVA showed that there was no effect of group on correct recall F(1, 66) = .412, p > .05, η^2_p = .006. However, there was a significant effect of BAS F(1, 66) = .4.451, p = .039, η^2_p = .063, with correct recall greater for high BAS (M = .105, SE = .015), compared to low BAS (M = .083, SE = .012). There was also a significant main effect of connectivity $F(1, 66) = 7.641$, $p = .007$, $\eta^2_p = .104$, with correct recall greater for high connectivity ($M = .106$, $SE = .014$) compared to low connectivity ($M = .081$, $SE = .013$). No significant interactions were found between group and BAS $F(1, 66) = .190$, $p > .05$, $\eta^2_p = .003$, group and connectivity $F(1, 66) = 1.88$, $p > .05$, $\eta^2_p = .028$ or between BAS and connectivity $F(1, 66) = .080$, $p > .05$, $\eta^2_p = .001$.

Table 5.3 Means (standard deviations) for correct and false recall as a function of connectivity, BAS and sleep/wake groups.

	Sleep		Wake	
Connectivity	Low	High	Low	High
Correct recall				
BAS Low	.068 (.018)	.085 (.021)	.075 (.018)	.102 (.021)
BAS High	.090 (.024)	.099 (.023)	.092 (.024)	.139 (.023)
False recall				
BAS Low	.081 (.020)	.022 (.012)	.037 (.020)	.022 (.012)
BAS High	.169 (.025)	.147 (.030)	.081 (.025)	.140 (.030)

False recall of critical lures

The proportion of times that participants falsely recalled a critical lure is shown in Figure 5.1 as a function of connectivity, BAS and sleep/wake groups. The 2 × 2 × 2 ANOVA showed that there was a significant main effect of group on false recall $F(1, 66) = 5.942$, $p = .017$, $\eta^2_p = .083$, as participants in the sleep condition ($M = .105$, $SE = .010$) had more false recall than participants in the wake condition (M = .070, SE = .010). A significant main effect of BAS was also found $F(1, 66) = 36.049$, $p < .001$, η^2_p = .353, because more false recall was made in high BAS lists (M = .134, SE = .012) compared to low BAS lists (M = .040, SE = .008). However, there was no main effect of connectivity $F(1, 66)$ = .288, $p > .05$, η^2_p = .004, nor any significant interactions between group and BAS $F(1, 66)$ = .679, $p > .05$, η^2_p = .010 or between BAS and connectivity $F(1, 66)$ = 2.56, $p > .05$, η^2_p = .037 . However, the interaction between group and connectivity approaching but not reaching significance $F(1, 124) = 3.007$, p = .073, η^2_p =.024.

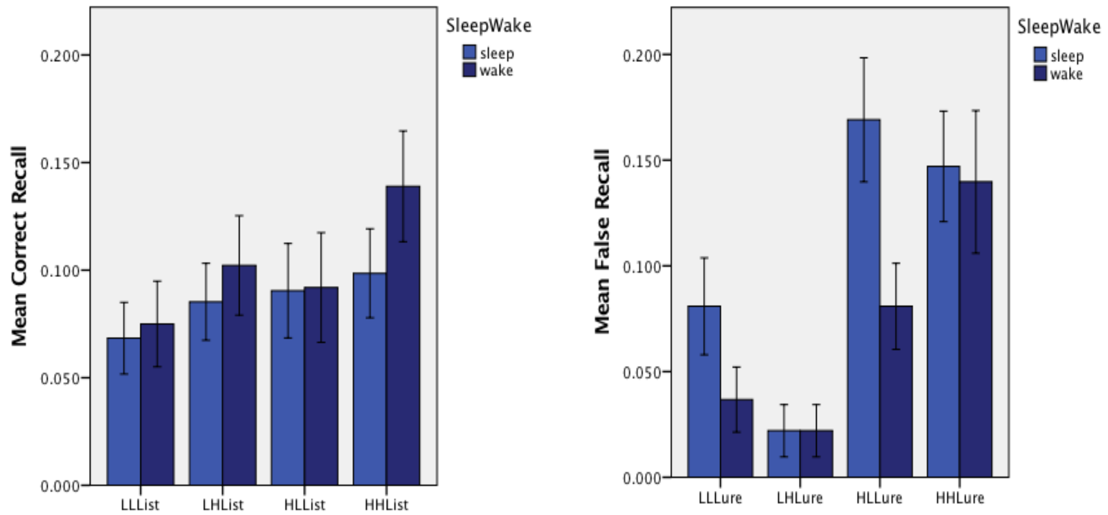


Figure 5.1 Correct recall (left) and false recall as a function of BAS, Connectivity and sleep/wake group (Error bars represent standard error).

Note: LL = Low BAS/Low Connectivity; LH = Low BAS/High Connectivity; HL = High BAS/Low Connectivity; HH = High BAS/High Connectivity.

5.2.5 Exploratory Analyses

As noted above, in false recall the interaction between group and connectivity fell just short of statistical significance (p = .073). Therefore, an exploratory analysis was conducted to examine this interaction further. Post-hoc tests were conducted on all interactions. The interaction finding between Sleep/wake group and BAS was followed up with two Independent t-tests. In order to do this, a *low BAS* mean correct recall score was computed by averaging across the low BAS/low Connectivity variable and the low BAS/high Connectivity variable. The *high BAS* mean correct recall score was computed by averaging across the high BAS/low Connectivity variable and the high BAS/high Connectivity variable. Similarly, the interaction between Sleep/wake group and Connectivity was followed up with two post-hoc Independent t-tests. In order to do this, a *low Connectivity* mean correct recall score was computed using the average of the low BAS/low Connectivity variable and the high BAS/low Connectivity variable. The *high Connectivity* mean correct recall score was computed by averaging across the low BAS/high Connectivity variable and the high BAS/high Connectivity variable. For false recall, the interaction between Sleep/wake group and BAS/Connectivity was followed up with two Independent t-tests. Mean false recall for *low BAS*, *high BAS low connectivity* and *high connectivity* was calculated in the same way as for correct recall.

Correct Recall

There was no significant difference in mean correct recall for low BAS between the sleep and wake group $t(66) = -.490$, $p = .626$. Similarly, no significant difference was observed between the sleep and wake groups for high BAS $t(66) = -.682$, $p = .498$ (see Figure 5.2).

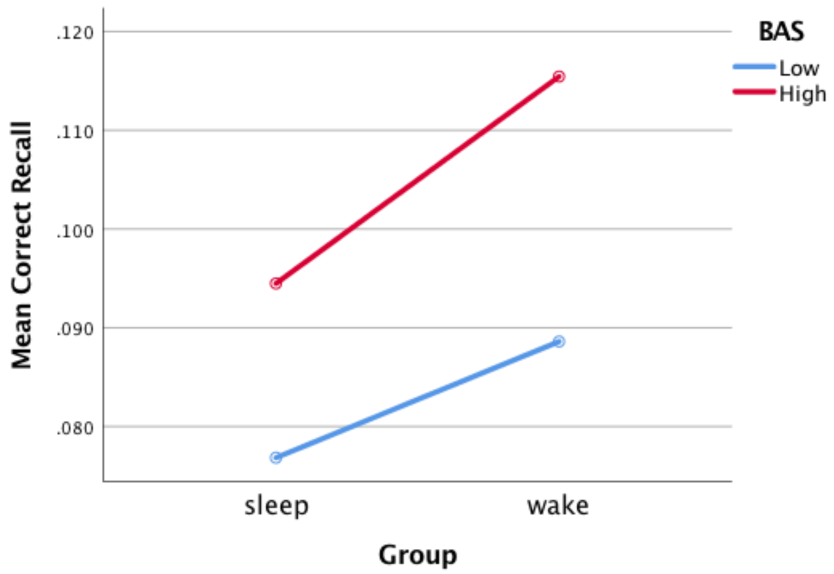


Figure 5.2 Correct recall as a function of BAS and sleep/wake group.

For low connectivity, no significant difference was found in mean correct recall between the sleep and wake groups t(66) = -.151, p = 880. Similarly, no significant difference was observed between the sleep and wake groups for high connectivity t(66) = -1.051, p = .297 (see Figure 5.3).

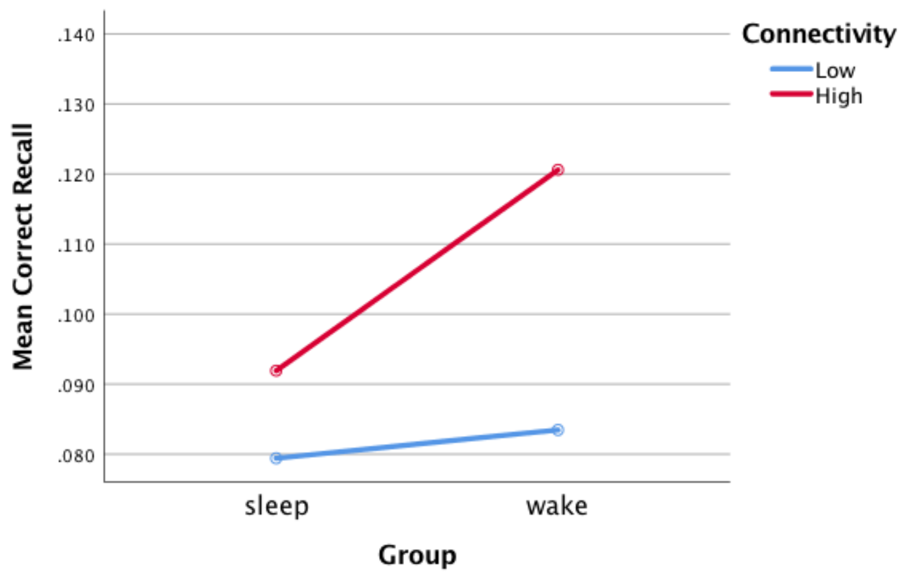


Figure 5.3 Correct recall as a function of Connectivity and sleep/wake group.

False Recall

The independent t-tests showed that there was no significant difference between the sleep and wake groups for mean low BAS false recall $t(66)$ = 1.310, p = .195. The comparison between sleep and wake groups for mean high BAS false recall also failed to reach significance $t(66)$ = 1.928, p = .058 (see Figure 5.4).

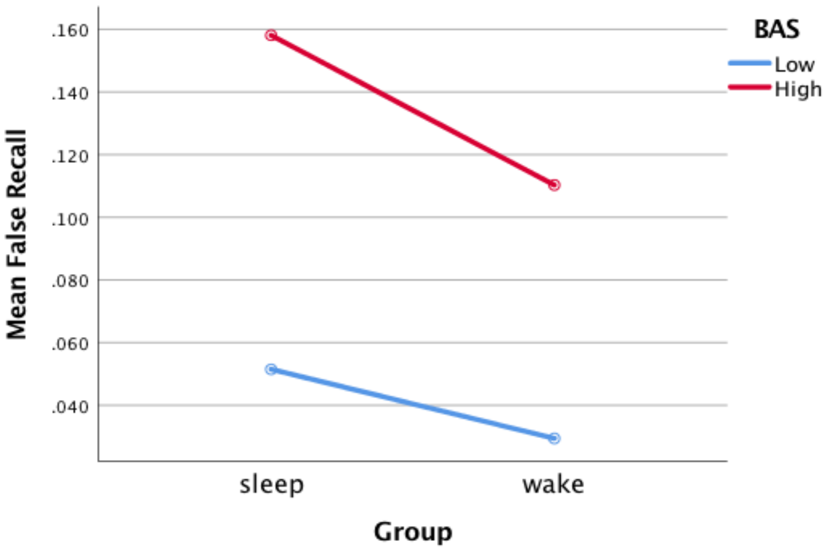


Figure 5.4 False recall as a function of BAS and sleep/wake group

A significant difference in mean false recall for low connectivity was found between sleep and wake groups. Participants in the sleep group (M = .125, SD = .102) had greater mean false recall for low connectivity compared to the wake group (M = .059, SD = .063, $t(66)$ = 3.213, p = .002). There was no significant difference in mean false recall for high connectivity between sleep and wake groups $t(66)$ = .153, p = .878 (see Figure 5.5).

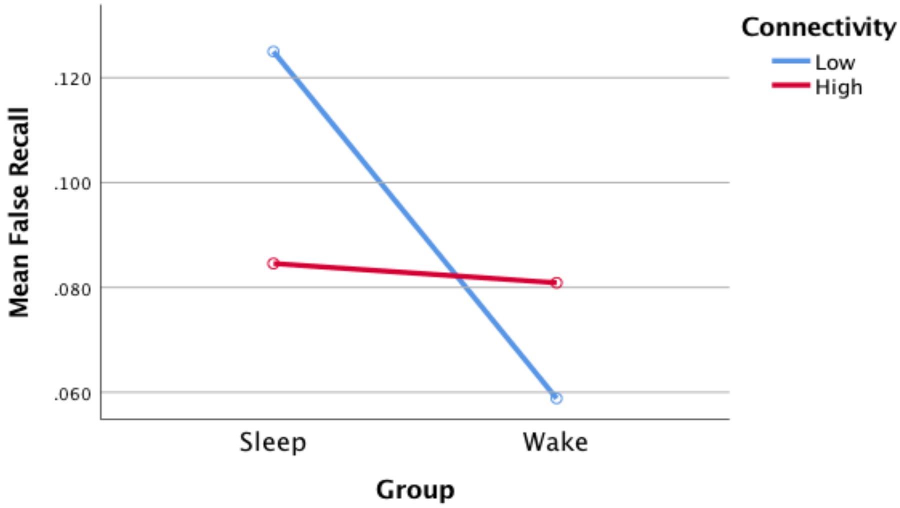


Figure 5.5 False recall as a function of Connectivity and sleep/wake group

5.2.6 Discussion

Using the DRM paradigm, this experiment aimed to explore the role of manipulated levels of BAS and inter-item connectivity on the effect of sleep on false recall. The principal findings from the experiment suggest that high inter-item connectivity aids recall as lists designed with high inter-item connectivity produced the highest level of correct recall. False recall typically increased following a period of sleep compared to an equivalent period of wakefulness. The hypothesis that high connectivity and high BAS would increase false recall in general, was partially supported with the role of high BAS. This is because results indicated that more false recall was made in high BAS but there was no main effect of connectivity.

The finding of the impact of sleep on false memory, which was the key focus of the studies, not only supported the hypothesis that sleep would increase false memory level, but was also consistent with other studies on DRM false memory (Darsaud et al., 2011; Diekelmann et al., 2010; Pardilla-Delgado & Payne, 2017; Payne et al., 2009; Shaw & Monaghan, 2017). Sleep is believed to provide a period free of distracting stimuli for the brain to consolidate certain types of memory, such as the episodic, semantic, or procedural (Rasch & Born, 2013; Stickgold, 2005). Neurological studies of this phenomenon have pointed to increased activity in the hippocampus during sleep as a strong indicator of memory mechanisms being active in this period (Buzsák, 1998). Klinzing, Niethard and Born (2019) have further suggested that, following sleep, such memories may become more 'gist-like' in nature. This could be interpreted as being consistent with FTT, which differentiates two different types of memory: verbatim (specific and contextual) and gist (general and semantic). According to FTT, gist traces are responsible for the formation of false memories, therefore this might serve as an explanation of how memory consolidation during sleep can lead to false recall.

In contrast to the effect that sleep has on false recall, sleep did not affect correct recall. This dissociation may seem contradictory with the idea that sleep significantly affects memory consolidation. However, several DRM studies have shown that sleep can impact false memory without affecting the ability to remember studied words (Diekelmann et al., 2010; Fenn et al., 2009; Lo et al., 2014). This result may suggest that sleep may not strengthen memory, it only provides a period of time in which the memory is protected from the effects of interference (Stepan, Dehnke, & Fenn, 2017).

The next finding, that lists with high connectivity yielded higher correct recall, appears to be consistent with the findings of some DRM studies, which suggest that lists with high inter-item connectivity are actually positively correlated with an increase in correct recall and decrease in false recall compared to those with lower inter-item connectivity (Deese, 1959; McEvoy et al., 1999). This result suggests that the denser the connections between the words in a list, the greater the recall of those words, as recalling one word from the list will cue another word that is semantically related.

The finding that false recall was greater when lists were designed with high BAS compared to low BAS is in line with those of Knott et al.'s (2012) from whom the word lists for the current study were obtained. The role of a higher BAS in increasing false recall has been well established in previous studies (Beato & Arndt, 2017; Hicks & Hancock, 2002; Knott et al., 2012). It has been suggested that BAS is a critical factor for false memory generation (Howe et al., 2009). AMT theory supports this, as BAS pertains to the strength of the association between words and the critical lure, meaning that a higher chance of false recall would be the result of association in the semantic networks (Coane, McBride, Raulerson III, & Jordan, 2007). FTT may also come into play here, as verbatim memory may be affected by higher BAS between the word list and critical lure, with this typically increasing following sleep and the transfer of memory to becoming more gist-like (Arndt, 2012). Moreover, the absence of an effect of connectivity on false recall may be because that BAS is an overwhelming factor in false recall, and any other role of connectivity largely gets lost where BAS is an influence. It has been suggested that the strong association between the critical lure and list words is the most powerful predictor of false memories formation (Smith, Gerkens, Pierce & Choi, 2002; Knott et al., 2012).

5.3 Experiment 4: The Role of List Length on DRM False Recall After Sleep

5.3.1 Introduction

As mentioned earlier, Newbury and Monaghan's (2019) meta-analysis aimed to investigate the role of sleep on the consolidation of correct and false memory in studies that used the DRM paradigm. They included multiple factors (i.e. list length, emotionality of the lists and modality of the list's presentation) to examine how they influenced the effect that sleep has on false memory. The findings suggested that shorter lists enhance sleep-based increases in false memory with DRM paradigm. The aim of this experiment was to explore the impact of list length on the size of the effect that sleep has on DRM false recall. Following the results of Newbury and Monaghan (2019) meta-analysis, the main hypotheses for this experiment are:

1. Shorter lists will give rise to more false memories after sleep than longer lists.
2. Sleep compared to wake will result in higher levels of false recall.

5.3.2 Method

5.3.2.1 Participants

A total of 435 native-English speaking participants from the UK aged between 18 to 65 years were recruited to participate in the study via Prolific. A total of £3 payment was paid to the participants who successfully completed the whole experiment. The same screening criteria as Experiment 2 were applied.

Participants in the sleep group who slept for less than 6 hours (n=7), who reported sleep problems (n=105) and those who were taking medication that might affect their sleep (n=7) were excluded from the final analysis. A further 188 participants were excluded because they failed to complete the second part of the experiment. The total number of participants who successfully completed the two phases of the experiment was 128 participants (mean age 36 years and standard deviation 13.07; 51 males and 77 females). There were two groups in this study the ‘sleep’ and the ‘wake’ groups. The sleep group (n=84) studied the list at 9pm and invited back to take the test at 9am the next morning, while the wake group (n=84) studied the list at 9am and invited back to complete the test at 9pm that evening. A sample size calculation was conducted using G*power (version 3). The power analysis showed that a minimum of 125 participants are needed to find a medium effect size of interaction (.06) with power of .80 and an alpha value of .05 (Appendix S).

5.3.2.2 Stimuli

The eight DRM lists used at the encoding were taken from (Stadler, Roediger,

& McDermott, 1999) norming study. These lists were the top lists that produced the highest level of false recall in their study (see Table 5.4). However, the list “sleep” was ranked as one of the top lists but as this experiment was studying the effect of sleep, the word ‘sleep’ was constantly repeated (i.e. in the study title and information sheet). Therefore, we used the list ‘needle’ instead, which was the next list in the order. The full lists used in this experiment can be found in Appendix (D).

Table 5.4 Lists that produced the highest percentage of false recall in Stadler et al’s (1999) study.

List	***%***	***List***	***%***	***List***	***%***
Window	65	**Smoke**	54	**City**	46
Sleep	61	**Rough**	53	**Cup**	45
Smell	60	**Needle**	52	**Cold**	44
Doctor	60	**Anger**	49	**Mountain**	42
Sweet	54	**Trash**	49	**Slow**	42
Chair	54	**Soft**	46	**River**	42

5.3.2.3 Materials

Materials were the same as Experiment 2.

5.3.2.4 Design

The experiment had two independent variables. The first variable ‘group’ had two levels: sleep and wake. The second independent variable ‘length’ had two levels: long and short. The dependent variable ‘type of item’ had two levels: accurate response to presented words and false response to lures. The experiment again comprised two phases. The study and retrieval phases were exactly as reported in Experiment 2.

5.3.2.5 Procedure

The procedure was the same as previous recall experiments reported in this thesis. However, in this experiment half of the participants in each group (sleep and wake) studied lists that consisted of 12 words “long lists condition”, while the other half studied lists that consist of 6 words “short lists condition”. The lists were presented orally by a digital female voice in a random order using Qualtrics. The words were presented with a delay of 10 seconds between lists and 2 seconds between words. The order of word presentation was from the strongest to weakest associative strength to the critical lure. The phrase “NEW LIST” appeared on the screen before each new list.

5.3.3 Statistical Analysis

The data was analysed using IBM SPSS (version 24). Two 2 (group: sleep / wake) ×2 (length: long / short) mixed ANOVAs were conducted on correct recall rate and false recall rate. Bonferroni post-hoc tests were used to follow up any main effects. The number of intrusions was calculated for sleep/wake groups (see Table 5.5). Means and standard deviations for the control data are reported in Table 5.6. Data is available on the Open Science Framework (https://osf.io/n5fvj/).

5.3.4 Results

Intrusions

Table 5.5 shows the intrusion data, with participants making more unrelated intrusions as opposed to related intrusions in general. These intrusions were not analysed further.

Table 5.5 Frequencies of Intrusions as a function of sleep/wake conditions

Intrusions	Sleep		Wake		Total	
	6	*12*	*6*	*12*	*6*	*12*
Related Intrusions	7	9	15	13	22	22
Unrelated Intrusions	71	52	40	82	111	134
Total Intrusions	78	61	55	95	133	156

Correct recall of studied words

The proportion of times that participants recalled a studied word is shown in Figure 5.6 as a function of word length and sleep/wake group. The 2 x 2 mixed ANOVA showed that there was no significant main effect of group on mean correct recall $F(1, 124) = 1.658$, $p > .05$, $\eta^2_p = .013$. However, there was a significant main effect of length $F(1, 124) = 8.203$, $p = .005$, $\eta^2_p = .062$, with correct recall greater for 6 words (*M* = .191, *SE* = .016), compared to 12 words (*M* = .128, *SE* = .016). The interaction between condition and length was not significant $F(1, 124) = .068$, $p > .05$, η^2_p =.001.

Table 5.6 Means (standard deviations) for correct and false recall as a function of condition and word length.

	Sleep		Wake	
Length	6	12	6	12
Correct recall	.174 (.022)	.116 (.022)	.208 (.022)	.139 (.022)
False recall	.194 (.025)	.264 (.025)	.084 (.025)	.242 (.025)

False recall of critical lures

The proportion of times that participants falsely recalled a critical lure is shown in Figure 5.6 a function of word length and sleep/wake group. The 2 x 2 mixed ANOVA showed that there was a significant main effect of group on mean false recall $F(1, 124) = 6.750$, $p = .011$, η^2_p = .052, because false recall was greater in the sleep (*M* = .229, *SE* = .018) compared to wake condition (*M* = .163, *SE* = .018). Moreover, there was also a significant main effect of length $F(1, 124) = 20.241$, $p < .001$, η^2_p = .140, with false recall greater for 12 words (*M* = .253, *SE* = .018), compared to 6 words (*M* = .139, *SE* = .018). The interaction between condition and length tended to approach significance $F(1, 124) = 3.007$, $p = .085$, η^2_p =.024.

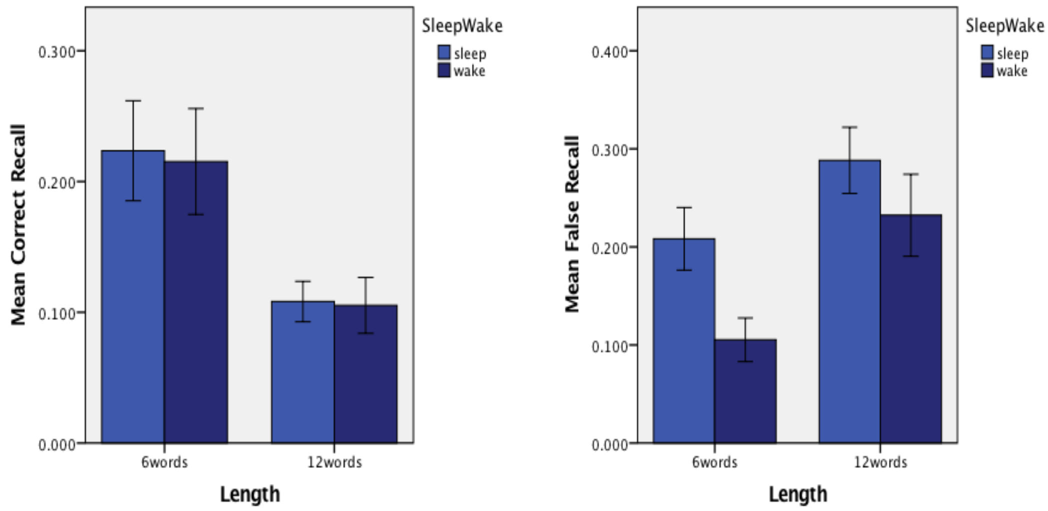


Figure 5.6 Correct recall (left) and false recall as a function of word length and sleep/wake group (Error bars represent standard error).

5.3.5 Exploratory Analyses

As noted above, the interaction between group and length for false recall nearly reached a significant level (*p* = .085). Therefore, the interaction between Sleep/wake group and length (short vs long) for both correct and false recall was followed up with two post-hoc Independent t-tests.

Correct Recall

There was no significant difference in mean correct recall for 6 words (short lists) between the sleep and wake groups *t*(62) = -.934, *p* = .354. Similarly, no significant difference was found between the sleep and wake groups for 12 words (long lists) *t*(62) = -.919, *p* = .362.

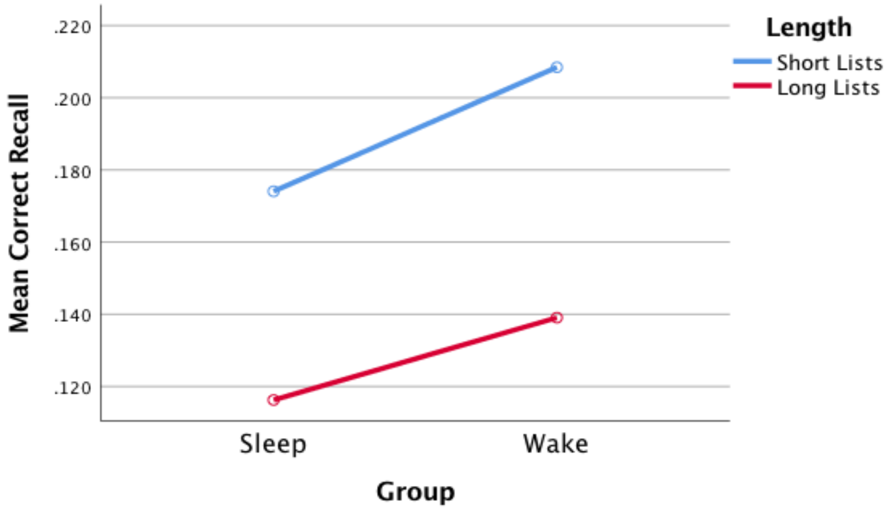


Figure 5.7 Correct recall as a function of length and sleep/wake group

False Recall

The independent t-tests showed that there was a significant difference between Sleep/wake group and length. Participants in the sleep group (*M* = .194, *SD* = .141) had significantly greater mean false recall compared to participants in the wake group (*M* = .084, *SD* = .089) for 6words (short list) $t(62) = 3.718$, $p < .001$. Finally, no significant difference was found between the sleep and wake groups for mean false recall for 12 words (long list) $t(62) = .532$, $p = .597$.

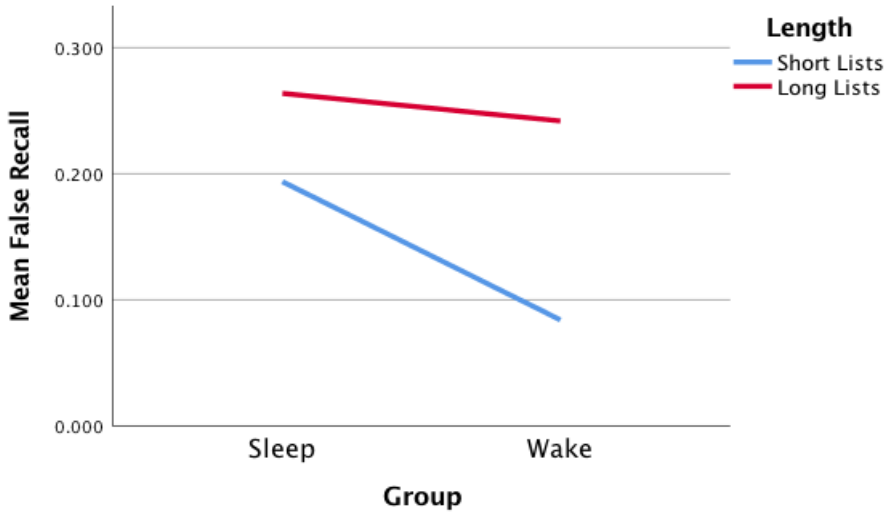


Figure 5.8 False recall as a function of length and sleep/wake group

5.3.6 Discussion

In this experiment, two different lengths of DRM lists were used to examine the pattern of false recall that occurs after sleep when the length of the lists are manipulated.

The findings from the current experiment provide evidence for the role of sleep in increasing false recall regardless of length of word list (although levels of correct recall were consistent across both groups). The current findings also support the view that the length of the word list has a significant effect on correct recall, as participants tended to correctly recall more studied words from shorter lists compared to long lists. Furthermore, the results of the current experiment indicated that long lists produce more false recall than shorter lists. These findings are unpacked in more detail below.

The finding of the effect of sleep on false memory supports the hypothesis of the experiment that sleep would increase false recall. It also replicates the findings from Experiment 3 and is consistent with the findings from other DRM studies (Darsaud et al., 2011; Diekelmann et al., 2010; Pardilla-Delgado & Payne, 2017; Payne et al., 2009; Shaw & Monaghan, 2017). One possible explanation might relate to AMT, which suggests that, during retrieval, 'semantic networks' are activated, resulting in many semantically related words being recalled while we attempt to retrieve the specific word needed. The monitoring aspect refers to our ability to identify the source of the activation (i.e. the memory itself). However, sometimes this can be disrupted, and the source of the activation can no longer be distinguished (i.e. the specific aspect of the memory), resulting in false recall. As mentioned earlier, sleep leads to greater spread of activation, therefore, it may increase the likelihood of such disruption occurring as per AMT (Klinzing et al., 2019; Underwood, 1965).

The next finding, that shorter lists yielded higher correct recall, is consistent with studies that have looked at the effect of list length. For instance, Robinson and Roediger (1997) found a consistent negative correlation between list length and increases in both correct recall and recognition. Sugrue et al. (2009) also found a similar pattern of results, consistent with the findings of the current experiment for shorter lists. It has been established that increasing list length, systematically decreases the probability of recalling studied items (Murdock Jr, 1961). This may be explained that in short lists -compared to longer lists- there are only small number of associated words that need to be recalled.

The next key finding of the current experiment addresses the impact of list length on false recall, namely demonstrating that longer lists lead to an increase in false memory. This is consistent with the findings of previous studies (Gallo & Roediger, 2003; Hutchison & Balota, 2005; Libby & Neisser, 2001; Marsh & Bower, 2004; Robinson & Roediger, 1997; Sugrue et al., 2009; Watson et al., 2003), which found that increased length results in increased false memory. These findings can be explained by the role of association because it seems that studying lists with a large number of words increases the likelihood of developing associations between words and, as a result, recalling more non-presented related words.

5.4 Summary of the Results

In this chapter, inter-item connectivity, BAS and list length were manipulated to examine whether they influence the size of sleep effects on the formation of false memories within the DRM paradigm. Specifically, in Experiment 3 the inter-item connectivity and BAS of the lists were manipulated. While in Experiment 4 different lists lengths were used (6 vs. 12). The findings indicated that false recall was higher following a period of sleep than following an equivalent period of wakefulness. This supports the findings of previous studies that investigated the effect of sleep on false memory (Darsaud et al., 2011; Diekelmann et al., 2010; Pardilla-Delgado & Payne, 2017; Payne et al., 2009; Shaw & Monaghan, 2017). Long lists and high BAS lists were found to be associated with a higher prevalence of false recall. This confirms previous findings that the length of the lists and the level of lists' associative strength have an impact on false memory generation.

Overall, these experiments supported the previous finding that sleep enhances false memories compared to an equivalent period of wakefulness (Darsaud et al., 2011; Diekelmann et al., 2010; Pardilla-Delgado & Payne, 2017; Payne et al., 2009; Shaw & Monaghan, 2017). Furthermore, when conducting a follow-up exploratory analysis on the connectivity and list length factors, the findings indicated that lists with low connectivity and fewer words led to enhanced false memory after sleep. No difference was found between sleep and wake groups for false recall for high connectivity lists or longer lists. AMT may serve again to explain these findings. As outlined above, strong association between the words can make the activation of the critical lure easier, while, for less closely related lists, more spreading activation would be required over semantic networks. Also, extracting the theme word from a long list is more achievable, as activation of the lure word is facilitated by the large number of related words in the list. Therefore, spreading activation may occur equally among both wake and sleep groups for high connectivity and long lists leading to near ceiling activation of the critical lure. Hence, the advantage of sleep-related spreading activation may not be easily detected for high connectivity and long lists as compared to low connectivity and shorter lists. These findings support the findings of previous research that harder problems are facilitated by sleep (Sio et al., 2013). The current experiments suggest that this might be due to sleep's ability to consolidate information that has been practiced and rehearsed during waking hours; which make it especially effective at dealing with more difficult tasks like studying low connectivity or shorter lists (Newbury & Monaghan, 2019; Sio et al., 2013; Stickgold & Walker, 2004).

Chapter 6: Experiment 5: Effect of Early and Late Sleep on False Memory Formation.

(Online Study)

6.1 Introduction

In the previous chapter the level of BAS and inter-item connectivity were manipulated within DRM lists to explore the link between sleep and false memory. The results indicated that sleep increases false memories compared to an equivalent period of wakefulness and that more false recall was generally made in high compared to low BAS lists. This finding raises a question regarding whether specific sleep stages are involved in false memory consolidation and, if so, which stage of sleep is engaged?

Consolidation of memories in the different memory systems has been suggested to differentially depend on specific sleep stages (Stickgold & Walker, 2007). Human sleep consists of cycles that last approximately 90 minutes. These cycles consist of REM sleep and 4 stages of non-REM sleep, with stages 3 and 4 being the deepest sleep known as SWS. Work conducted by Born and Gais (2003) studied the effects of late and early sleep on various types of memory. They investigated the degree of memory processing following sleep in the first half of the night, 'early sleep', compared to the memory processing following sleep in the second half of the night, 'late sleep'. Most SWS sleep occurs in the 'early sleep' period, the first part of the night, whereas REM sleep occurs most in the later part of the night, the 'late sleep' period. The time spent in SWS is 5 times longer during the early sleep period when compared to the late sleep period. Whereas the time spent in REM sleep is twice as long during the late sleep period than it is during the early sleep period Born & Gais, 2003).

In Plihal and Born's (1997) experiment, the recall of information connected to declarative memory and also to procedural memory was assessed. Recall of paired-associate lists (i.e. tea – ball) were used in order to assess declarative memory, whereas mirror-tracing skills was used in order to assess procedural memory. Participants were tested after either a period of early or late sleep. The results indicated that sleep, in general, increased recall when compared to states of wakefulness, but that recall depended on the specific phase of sleep and also on the type of memory. The recall of the paired-associate lists was better in the early sleep group, but conversely, the late sleep group showed better performance on the mirror-tracing skills (Plihal & Born, 1997). This suggests that declarative or episodic memory increased after a session of NREM early sleep, but procedural/implicit memory did not. Procedural/implicit memory increased after REM late sleep (Born & Gais, 2003; Stickgold, 2005). However, as SWS facilitates declarative memory consolidation, it is possibly linked to the consolidation of false memory.

An experiment conducted by Payne et al. (2009) used the DRM word lists to investigate the effect of sleep on false memory. In their second experiment the researchers sought to establish which stages of sleep affect the consolidation of false memory. Therefore, a sleep group was tested to consider the time spent in SWS using the PSG while sleeping in the lab (Payne et al., 2009). This study found that the percentage of time spent in SWS was negatively correlated with DRM correct recall. In other words, correct recall correlated with less SWS, not more. It has been suggested that within the DRM paradigm, correct memory tends to be inversely related to false memory (Roediger et al., 2001). As such, one would anticipate that longer SWS times may be correlated with false recall. SWS is considered to be important for sleep-dependent memory processing (Payne et al., 2009). Because SWS primarily occurs mostly in the early period of the night this perhaps points to some connection between this period and false memory. However, this suggests that although false memory within DRM found to be enhanced by sleep in general (McKeon et al., 2012; Payne et al., 2009), it may be hindered by a lower percentage of time spent in SWS.

As outlined in Chapter 1, the BAS and association density of the words in the lists play an important role in the generation of false memory (Howe, 2006). The role that gist extraction plays in the creation of false memory in research using the DRM paradigm is related to the strength and density of inter-item connectivity. The more densely the list words are related semantically, the higher the probability of correct memory and the lower the likelihood a participant will remember the critical lure (Knott et al., 2012). Additionally, the likelihood of false memory has been shown to vary with the strength of connections from the list words to the critical items. The stronger the connection between the list and lure, the greater the likelihood that the brain will create a false memory based on abstraction (McEvoy et al., 1999). While SWS has a beneficial effect on episodic memory, deprivation of REM sleep has been found to have a deleterious effect on learning, especially if the task is difficult (Rauchs, Desgranges, Foret, & Eustache, 2005). Stickgold, Scott, Rittenhouse and Hobson (1999) investigated the effects of semantic priming between strongly and weakly related pairs when participants were awakened from distinct sleep stages. They found that participants who were awakened from REM sleep produced more priming by weak-related word pairs (e.g. thief and wrong) than by strongly-related pairs (e.g., hot and cold), compared to those who are awake or awakened from non-REM sleep. There was no significant group effect for strong priming. This finding was explained in terms of spreading activation between concepts favoured by sleep: during REM sleep, the activation spreads widely from the presented stimuli to more remote associates, instead of activation being limited to only closely associated concepts (Sio et al., 2013).

Therefore, the current study set out to assess in detail any connection between early and late sleep periods and an increase or decrease in false recognition with manipulated levels of inter-item connectivity and BAS. As outlined above, sleep consists of SWS, during which new memories are reactivated for consolidation and which dominates in the first half of the night. During this sleep phase, a narrow spread of activation may occur (comparable to the awake state). REM, by contrast, facilitates spreading activation and promotes memory insight and dominates in the second half of the night. However, following the findings of Stickgold et al. (1999) , the main hypotheses for this experiment are:

1. Late sleepers would exhibit increased false memories for the less closely related (sparser) DRM lists relative to the early sleepers.
2. Both groups would show equivalent increases in false memory for the more closely related (denser) DRM lists.

6.2 Method:

6.2.1 Participants

A total of 679 native English-speaking participants aged 18-65 years from the UK signed up via Prolific to participate in the study. Participants were paid £5 for full participation. This study had two groups 'early sleep' (n=34) and 'late sleep' (n=34) groups. Both groups studied the word lists at 12 pm and were invited back to take the test at the same time next day. All participants in both groups were asked to sleep for at least 6 hours after the study phase. Participants who slept for less than 6 hours (n=2), were classified as medium sleepers (n=230), reported sleep problems (n=253) or were taking medication that may affect their sleep (n=40) were excluded from the final analysis. Additionally, outliers were removed from the final data set (n=1; remembered all the words), and 85 participants were excluded because they failed to complete the whole experiment. The total number of participants who successfully completed the experiment was 68 participants (Mean age= 33 years, Standard Deviation= 11.32) and there were 33 males, 35 females and 1 participant who preferred not to say. A sample size calculation was conducted using G*Power (version 3). The power analysis showed that a minimum of 62 participants are needed to find a medium effect size of sleep (.06) with power of .80 and an alpha value of .05 (see Appendix R). However, to be consistent with previous experiments 68 participants were recruited.

6.2.2 Stimuli

The lists used were the same lists used in Experiment 3.

6.2.3 Materials

As in the previous online experiments reported in this thesis, the participants were screened for self-reported sleep. Additionally, because only early and late sleepers were needed in this experiment, the participants were asked to complete the Morningness-Eveningness questionnaire (MEQ) in order to find their sleep type. A score ranging from 16 to 41 on the MEQ indicates an evening type, whereas scores ranging from 42 to 58 and from 59 to 86 indicate medium and morning types, respectively. The participants were categorised as either early sleepers or late sleepers according to their score on the MEQ.

6.2.4 Design

The experiment had two independent variables. The first variable ‘group’ had two levels: early sleep and late sleep. The second independent variable ’BAS’ had two levels: high and low. The third independent variable ’connectivity’ had two levels: high and low. The dependent variable ‘type of item’ had two levels: accurate response to presented words and false response to lures. The experiment again comprised two phases. The study and retrieval phases were exactly as reported in Experiment 2.

6.2.5 Procedure

The data was collected using Prolific and the experiment was created using Qualtrics. On the first page of the web-based experiment, the information sheet was presented (see Appendix I), and consent was obtained (see Appendix P). On the next page, participants were required to answer some questions about their sleep (i.e. how many hours they sleep at night). They were also asked to complete the MEQ in order to check their eligibility before they presented with the lists of words. Qualtrics was programmed to calculate the questionnaire scores. If the participants' score did not fit within the criteria for the study (i.e. they were a medium sleeper), they were automatically exited from the experiment. Participants who satisfied the criteria (i.e. they were morning or evening sleepers) were able to proceed and move to the study phase where they were presented with the word lists. The word lists were presented orally in a digital female voice one at a time using Qualtrics with a delay of 10 seconds between lists and 2 seconds between words. There were 10 words in each list and they were presented in order of descending backward associative strength to the critical lure. Lists were divided into four blocks of four lists, the presentation order of word lists was random across participants. A phrase "NEW LIST" appeared on the screen before each new list. After 24 hours participants were invited back -using their prolific ID- to complete the second phase of the study (the test phase). During the test phase, participants were presented with 64 words in a random order with 32 being studied words (three words were chosen randomly from each studied list), 16 being critical lures and 16 being unrelated distractors (see Appendix E for the words presented in the recognition test). All participants were asked to complete the recognition test (for recognition test procedure see Experiment 1). At the end of this phase, participants were provided with debriefing information (see Appendix N).

6.3 Statistical Analysis

The data was analysed using IBM SPSS (version 24). Means for correct recognition are reported in Table 6.1, and Table 6.2 reports the means and standard deviations for false recognition. Total correct recognition was computed as the sum of correct recognition for Remember, Know and Guess judgments. Total false recognition of the critical lures was calculated as the sum of false recognition for Remember, Know and Guess judgments. A series of 2 (group: early vs late) by 2 (BAS: low vs high) by 2 (connectivity: low vs high) mixed model ANOVA's were conducted separately on total correct recognition rates and separately for correct recognition rates for Remember and Know judgements. The same analyses were carried out for the false recognition rates. Groups (early and late) was entered into the analysis as a between-participants variable and BAS (high and low) and connectivity (high and low) were entered as within-participants variables. Additionally, independent-samples t-tests were run on the filler correct recognition trials to compare these scores between the early and late groups. Paired-samples t-tests with Bonferroni corrections followed up any significant interactions between BAS and connectivity. Data is available on the Open Science Framework (https://osf.io/gr6ch/).

6.4 Results

6.4.1 Proportion of recognition responses

Correct recognition of studied words

The proportion of times that participants recognised a studied word is shown in Figure 6.1 as a function of BAS and connectivity. The 2 x 2 x 2 ANOVA showed that there was no significant main effect of group on total correct recognition $F(1, 66) = .001$, $p > .05$, $\eta^2 p = .000$. Similarly, there was no significant main effect of BAS $F(1, 66) = .605$, $p > .05$, $\eta^2_p = .004$. However, the main effect of connectivity was significant $F(1, 66) = 19.098$, $p < .001$, $\eta^2_p = .224$, as correct recognition was greater for high connectivity (*M* = .626, *SE* = .024) compared to low connectivity (*M* = .526, *SE* = .023). This was qualified by a significant interaction between BAS and Connectivity $F(1, 66) = 4.270$, $p = .04$, $\eta^2_p = .061$ (see figure 5.1). Paired-sample t-tests revealed that there was a significant difference in correct recognition between low BAS/low connectivity (*M* = .49, *SD* = .23) and low BAS/high connectivity (*M* = .64, *SD* = .23, $t(67) = -5.1$, $p < .001$). There was no difference in correct recognition between high BAS/low connectivity and high BAS/high connectivity $t(67) = -1.432$, $p > .025$. The were no significant interactions between group and BAS $F(1, 66) = .634$, $p > .05$, $\eta^2_p = .010$ or between group and connectivity $F(1, 66) = .709$, $p > .05$, $\eta^2_p = .011$.

Table 6.1 Mean correct recognition for early and late groups as a function of BAS and connectivity.

	Early				Late			
BAS/Connectivity	LL	LH	HL	HH	LL	LH	HL	HH
Remember	0.25	0.31	0.21	0.25	0.24	0.33	0.22	0.26
Know	0.13	0.23	0.21	0.21	0.14	0.20	0.17	0.21
Guess	0.10	0.11	0.14	0.17	0.13	0.11	0.16	0.13
Total	0.47	0.65	0.56	0.63	0.52	0.64	0.55	0.59

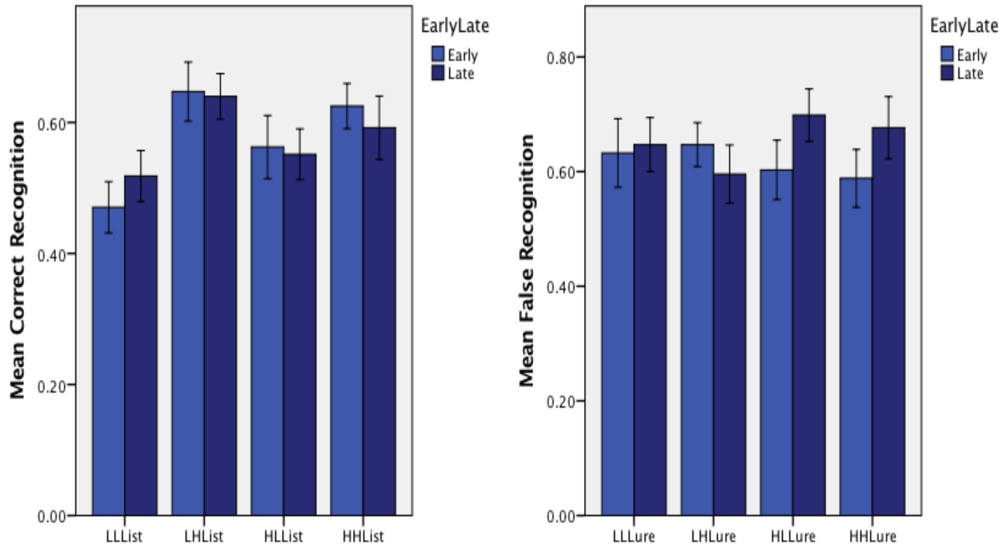


Figure 6.1 Correct recall (left) and false recall as a function of BAS, Connectivity and early/late groups (Error bars represent standard error).

Note: LL = Low BAS/Low Connectivity; LH = Low BAS/High Connectivity; HL = High BAS/Low Connectivity; HH = High BAS/High Connectivity.

Correct recognition for Remember judgments

Next, a separate 2 x 2 x 2 mixed ANOVA found no significant main effect of group on Remember responses $F(1, 66) = .044$, $p > .05$, $\eta^2_p = .001$. However, a significant main effect of BAS was found $F(1, 66) = 4.175$, $p = .05$, $\eta^2_p = .059$, as correct recognition was larger for low BAS (*M* = .28, *SE* = .02) compared to high BAS (*M* = .24, *SE* = .02). Similarly, there was a significant main effect of connectivity $F(1, 66) = 6.273$, $p = .015$, $\eta^2_p = .087$ because correct recognition was greater for high connectivity (*M* = .29, *SE* = .03) than low connectivity (*M* = .23, *SE* = .02). The were no significant interactions between group and BAS $F(1, 66) = .002$, $p > .05$, $\eta^2_p = .000$, group and connectivity $F(1, 66) = .146$, $p > .05$, $\eta^2_p = .002$ or between BAS and connectivity $F(1, 66) = .723$, $p > .05$, $\eta^2_p = .011$.

Correct recognition for Know judgments

A 2 x 2 x 2 mixed ANOVA found no significant main effect of group on Know responses $F(1, 66) = .365$, $p > .05$. Similarly, no main effect of BAS occurred $F(1, 66) = 2.63$, $p > .05$. However, there was a main effect of connectivity for Know correct recognition rates $F(1, 66) = 10.779$, $p < .01$, $\eta^2_p = .140$. Correct recognition was higher for high connectivity (*M* = .212, *SE* = .015) compared to low connectivity (*M* = .162, *SE* = .012) for Know responses. The were no significant interactions between group and BAS $F(1, 66) = .176$, $p > .05$, $\eta^2_p = .003$, group and connectivity $F(1, 66) = .032$, $p > .05$, $\eta^2_p = .000$ or between BAS and connectivity $F(1, 66) = 2.494$, $p > .05$, $\eta^2_p = .036$.

False recognition of critical lures

The proportion of times that participants falsely recognised a critical lure is shown in Figure 6.1 as a function of BAS and connectivity. The 2 x 2 x 2 ANOVA showed that there was no significant main effect of group on total false recognition $F(1, 66) = .556$, $p > .05$, $\eta^2_p = .008$. Similarly, there was no significant main effect of BAS $F(1, 66) = .121$, $p > .05$, $\eta^2_p = .002$, and no main effect of connectivity $F(1, 66) = .406$, $p > .05$, $\eta^2_p = .006$. Furthermore, there were no significant interactions between group and BAS $F(1, 66) = 3.03$, $p > .05$, $\eta^2_p = .044$, group and connectivity $F(1, 66) = .406$, $p > .05$, $\eta^2_p = .006$ or between BAS and connectivity $F(1, 66) = .406$, $p > .05$, $\eta^2_p = .006$.

Table 6.2 Mean false recognition for early and late groups as a function of BAS and connectivity.

	Early				Late			
BAS/Connectivity	LL	LH	HL	HH	LL	LH	HL	HH
Remember	0.19	0.29	0.24	0.23	0.21	0.31	0.32	0.22
Know	0.25	0.26	0.21	0.24	0.34	0.15	0.21	0.32
Guess	0.19	0.10	0.15	0.13	0.10	0.15	0.17	0.14
Total	0.63	0.65	0.60	0.59	0.65	0.60	0.70	0.68

Note: LL = Low BAS/Low Connectivity; LH = Low BAS/High Connectivity; HL = High BAS/Low Connectivity; HH = High BAS/High Connectivity.

False recognition for Remember judgments

A 2 x 2 x 2 ANOVA found no significant main effect of group on false recognition rates for Remember responses, $F(1, 66) = .267$, $p > .05$. Also, there was no main effect of BAS $F(1, 66) = .055$, $p > .05$, and no main effect of connectivity $F(1, 66) = .636$, $p > .05$ on Remember responses. A significant interaction found was between BAS and connectivity for Remember responses $F(1, 66) = .8972$, $p < .01$, $\eta^2_p = .120$, see Figure 5.2). Paired-samples t-tests found that false recognition was greater for low BAS/high connectivity ($M = .30$, $SD = .28$) compared to low BAS/low connectivity ($M = .20$, $SD = .25$, $t(67) = -2.65$, $p = .01$. However, there was no difference between high BAS/low connectivity and high BAS/high connectivity $t(67) = 1.654$, $p > . 025$.

False recognition for Know judgments

A 2 x 2 x 2 ANOVA showed no significant main effect of group on false recognition rates for Know judgements, $F(1, 66) = . 208$, $p > .05$. Also, there was no main effect of BAS $F(1, 66) = .086$, $p > .05$, and no main effect of connectivity $F(1, 66) = . 129$, $p > .05$ on Know responses. There was also a significant interaction between BAS and connectivity for Know responses $F(1, 66) = 10.308$, $p < .01$, $\eta^2_p = .135$, see figure 5.3). Paired-samples t-tests found that false recognition was greater for low BAS/low connectivity ($M = .29$, $SD = .25$) compared to low BAS/high connectivity ($M = .20$, $SD = .19$, $t(67) = 2.347$, $p = .02$. However, there was no difference between high BAS/low connectivity and high BAS/ high connectivity $t(67) = 1.669$, $p > .025$. Finally, a significant interaction was found between group, BAS and connectivity for Know responses $F(1, 66) = 7.688$, $p < .01$, $\eta^2_p = .104$, see figure 4). The interaction was followed up with 2-way ANOVAs with BAS and connectivity as factors separately for early and late groups. For the late group only, a significant interaction between BAS and connectivity was found $F(1, 33) = 17.439$, $p < .001$, $\eta^2_p = .346$. Paired-samples t-tests showed that false recognition for Know responses was greater for low BAS/low connectivity ($M = .34$, $SD = .26$) compared to low BAS/high connectivity ($M = .15$, $SD = .16$, $t(33)$ 3.199, $p = .003$.

Recognition of filler judgments

Four independent-samples t-tests were conducted on total correct recognition, Remember, and Know filler judgments for participants in the early vs late groups. No significant difference in total filler correct recognition was found between the early ($M = .31$, $SD = .19$) and late groups ($M = .26$, $SD = .23$, $t(66) = .883$, $p > .05$). Similarly, no difference was found for Remember judgments, early ($M = .06$, $SD = .09$), late ($M = .05$, $SD = .08$), $t(66) = .702$, $p > .05$), or for Know judgments, early ($M = .13$, $SD = .14$), late ($M = .09$, $SD = .11$, $t(66) = 1.264$, $p > .05$).

6.4.2 Signal Detection Analyses

Again, as for Experiment 1, signal detection analysis was conducted for correct old responses to studied words and false old responses to the critical lure in order to examine participant's response bias. For this, signal detection parameters d' (discrimination ability) and C (response bias) were computed using the BAS/connectivity list items/critical lures as hits and the filler items as false alarms (Snodgrass & Corwin,1988 correction). Signal detection parameters d' and C were analysed separately for correct and false recognition using the same repeated measure ANOVAs as before. Paired-samples t-tests with Bonferonni correction adjusting the alpha at 0.05 followed up any significant interactions between BAS and connectivity and two 2-way ANOVAs separately for early and late conditions, with the factors BAS and connectivity followed up the 3-way interaction in false responses for Know judgments. The results of d' and C are summarised in Table 6.3.

Table 6.3 Signal detection of discriminability (d') and Criterion Bias(C) for studies words and critical lures

	Early						Late					
	d'			C			d'			C		
	M	LB	UB	M	LB	UB	M	LB	UB	M	LB	UB
Studied words												
LL	.52	.20	.83	.34	.16	.53	.80	.48	1.12	.44	.26	.62
LH	.58	.16	1.00	.31	.15	.48	1.06	.64	1.49	.31	.14	.47
HL	.53	.20	.86	.34	.14	.54	.79	.46	1.12	.44	.24	.64
HH	.98	.64	1.32	.11	-.08	.31	.82	.49	1.17	.42	.23	.62
Critical Lures												
LL	.21	-.17	.60	.50	.31	.69	.62	.23	1.00	.53	.30	.72
LH	.61	.24	.98	.30	.13	.47	.48	.11	.85	.60	.43	.77
HL	.34	-.02	.70	.43	.24	.63	.52	.16	.88	.58	.38	.77
HH	.38	.00	.76	.41	.22	.60	.39	.01	.77	.65	.46	.83

Note: M = Mean; LB= Lower Bound; UB = Upper Bound for 95% confidence intervals.

Correct responses to studied words

For discrimination ability d' of participants' correct old responses to the studied words, there was no significant main effect of group on total correct recognition $F(1, 66) = 1.151$, $p > .05$, $\eta^2p = .017$. Similarly, there was no significant main effect of BAS $F(1, 66) = .229$, $p > .05$, $\eta^2_p = .003$. However, the main effect of connectivity was significant $F(1, 66) = 7.322$, $p = .009$, $\eta^2_p = .100$ with better discrimination ability towards high connectivity (M = .865) compared to low connectivity lists (*M* = .661). There were no significant interactions between group and BAS $F(1, 66) = .374$, $p > .05$, $\eta^2_p = .054$, group and connectivity $F(1, 66) = .530$, $p > .05$, $\eta^2_p = .008$, BAS and connectivity $F(1, 66) = .202$, $p > .05$, $\eta^2_p = .003$ or between group, BAS and connectivity $F(1, 66) = 2.946$, $p > .05$, $\eta^2_p = .043$.

For response bias C, there was no main effect of group on correct responses $F(1, 66) = 1.268$, $p > .05$, $\eta^2_p = .019$. Similarly, there was no significant main effect of BAS $F(1, 66) = .242$, $p > .05$, $\eta^2_p = .004$. However, the main effect of connectivity was significant $F(1, 66) = 7.384$, $p = .008$, $\eta^2_p = .100$ showing a relatively liberal bias in the high connectivity (*M* = .288) compared to low connectivity lists (*M* = .390). The were no significant interactions between group and BAS $F(1, 66) = 3.80$, $p > .05$, $\eta^2_p = .054$, group and connectivity $F(1, 66) = .538$, $p > .05$, $\eta^2_p = .008$, BAS and connectivity $F(1, 66) = .219$, $p > .05$, $\eta^2_p = .003$ or between group, BAS and connectivity $F(1, 66) = 2.940$, $p > .05$, $\eta^2_p = .043$.

Correct Remember responses

For discrimination ability d' of participants' Remember responses to the studied words, there was no main effect of group on correct responses $F(1, 66) = .554$, $p > .05$, $\eta^2_p = .008$. However, the main effect of BAS was significant $F(1, 66) = 4.117$, $p < .05$, $\eta^2_p = .059$ with better discrimination ability in low BAS (M = .942) compared to high BAS lists (*M* = .797). Similarly, the main effect of connectivity was significant $F(1, 66) = 5.764$, $p < .05$, $\eta^2_p = .080$ with better discrimination ability in high connectivity (M = .949) compared to low connectivity lists (*M* = .790).). There were no significant interactions between group and BAS $F(1, 66) = .003$, $p > .05$, $\eta^2_p = .00$, group and connectivity $F(1, 66) = .082$, $p > .05$, $\eta^2_p = .001$, BAS and connectivity $F(1, 66) = .529$, $p > .05$, $\eta^2_p = .008$ or between group, BAS and connectivity $F(1, 66) = .146$, $p > .05$, $\eta^2_p = .002$.

For response bias C, there was no main effect of group on Remember responses $F(1, 66) = .037$, $p > .05$, $\eta^2_p = .001$. There was a significant main effect of BAS $F(1, 66) = 4.147$, $p < .05$, $\eta^2_p = .059$, showing a relatively liberal bias in the low BAS (*M* = 1.129) compared high BAS lists (*M* = 1.202). In addition, there was a significant main effect of connectivity $F(1, 66) = 5.717$, $p < .05$, $\eta^2_p = .080$, showing a relatively liberal bias in the high connectivity (*M* = 1.126) compared to low connectivity lists (*M* = 1.206). The were no significant interactions between group and BAS $F(1, 66) = .003$, $p > .05$, $\eta^2_p = .00$, group and connectivity $F(1, 66) = .092$, $p > .05$, $\eta^2_p = .001$, BAS and connectivity $F(1, 66) = .537$, $p > .05$, $\eta^2_p = .008$ or between group, BAS and connectivity $F(1, 66) = .149$, $p > .05$, $\eta^2_p = .002$.

Correct Know responses

For discrimination ability d' of participants' Know responses to the studied words, there was there was no main effect of group $F(1, 66) = .942$, $p > .05$, $\eta^2_p = .014$. Also, there was no significant main effect of BAS $F(1, 66) = 2.466$, $p > .05$, $\eta^2_p = .036$. There was a significant main effect of connectivity $F(1, 66) = 11.068$, $p = .001$, $\eta^2_p = .144$ with better discrimination ability in the high connectivity lists ($M = .501$) compared to low connectivity ($M = .334$). There were no significant interactions between group and BAS $F(1, 66) = .255$, $p > .05$, $\eta^2_p = .004$, group and connectivity $F(1, 66) = .037$, $p > .05$, $\eta^2_p = .001$, BAS and connectivity $F(1, 66) = 2.500$, $p > .05$, $\eta^2_p = .036$ or between group, BAS and connectivity $F(1, 66) = .875$, $p > .05$, $\eta^2_p = .013$.

For response bias C, there was no main effect of group on Know responses $F(1, 66) = 1.426$, $p > .05$, $\eta^2_p = .021$. Also, there was no significant main effect of BAS $F(1, 66) = 2.430$, $p > .05$, $\eta^2_p = .036$. However, there was a significant main effect of connectivity $F(1, 66) = 11.155$, $p = .001$, $\eta^2_p = .145$, showing a relatively liberal bias in the high connectivity lists ($M = 1.110$) compared to low connectivity ($M = 1.193$). The were no significant interactions between group and BAS $F(1, 66) = .260$, $p > .05$, $\eta^2_p = .004$, group and connectivity $F(1, 66) = .039$, $p > .05$, $\eta^2_p = .001$, BAS and connectivity $F(1, 66) = 2.504$, $p > .05$, $\eta^2_p = .037$ or between group, BAS and connectivity $F(1, 66) = .851$, $p > .05$, $\eta^2_p = .013$.

False responses to critical lures

For discrimination ability d' of participants' false old responses to the critical lure words, there was no main effect of group on false responses $F(1, 66) = .307$, $p > .05$, $\eta^2_p = .005$. Similarly, there was no significant main effect of BAS $F(1, 66) = .576$, $p > .05$, $\eta^2_p = .009$, and no main effect of connectivity $F(1, 66) = .197$, $p > .05$, $\eta^2_p = .003$. Furthermore, there were no significant interactions between group and BAS $F(1, 66) = .820$, $p > .05$, $\eta^2_p = .001$, group and connectivity $F(1, 66) = 3.613$, $p > .05$, $\eta^2_p = .052$, BAS and connectivity $F(1, 66) = .558$, $p > .05$, $\eta^2_p = .014$ or between group, BAS and connectivity $F(1, 66) = .969$, $p > .05$, $\eta^2_p = .014$.

For response bias C, there was no main effect of group on false responses $F(1, 66) = 3.004$, $p > .05$, $\eta^2_p = .004$. Similarly, there was no significant main effect of BAS $F(1, 66) = .572$, $p > .05$, $\eta^2_p = .008$, and no main effect of connectivity $F(1, 66) = .201$, $p > .05$, $\eta^2_p = .003$. Also, there were no significant interactions between group and BAS $F(1, 66) = .55$, $p > .05$, $\eta^2_p = .001$, group and connectivity $F(1, 66) = 3.602$, $p > .052$, $\eta^2_p = .001$, BAS and connectivity $F(1, 66) = .953$, $p > .05$, $\eta^2_p = .014$ or between group, BAS and connectivity $F(1, 66) = .947$, $p > .05$, $\eta^2_p = .014$.

False Remember responses

For discrimination ability d', there was no main effect of group $F(1, 66) = .104$, $p > .05$, $\eta^2_p = .002$, BAS $F(1, 66) = .283$, $p > .05$, $\eta^2_p = .004$ or connectivity $F(1, 66) = .283$, $p > .05$, $\eta^2_p = .004$. A significant interaction was found between BAS and connectivity for Remember responses $F(1, 66) = 8.776$, $p = .004$, η^2_p = .117. Paired-samples t-tests showed a better discrimination ability in low BAS/high connectivity (*M* = 1.00, *SD* = .73) compared to low BAS/low connectivity (*M* = .78, *SD* = .58, $t(67) = -2.408$, $p = .019$. However, there was no difference between high BAS/low connectivity and high BAS/high connectivity $t(67) = 1.708$, p = . 092. There were no significant interactions between group and BAS $F(1, 66) = .692$, $p > .05$, η^2_p = .010, group and connectivity $F(1, 66) = 1.425$, $p > .05$, η^2_p = .004 or between group, BAS and connectivity $F(1, 66) = .266$, $p > .05$, η^2_p = .004 .

For response bias C, again, there was no main effect of group $F(1, 66) = .325$, $p > .05$, $\eta^2_p = .005$, BAS $F(1, 66) = .287$, $p > .05$, $\eta^2_p = .004$ or connectivity $F(1, 66) = .286$, $p > .05$, $\eta^2_p = .004$. A significant interaction was found between BAS and connectivity for Remember responses $F(1, 66) = 8.804$, $p = .004$, η^2_p = .118. Paired-samples t-tests showed a more liberal bias in low BAS/high connectivity (*M* = 1.10, *SD* = .42) compared to low BAS/low connectivity (*M* = 1.21, *SD* = .41, $t(67) = 2.414$, $p = .019$. However, there was no difference between high BAS/low connectivity and high BAS/high connectivity $t(67) = -1.713$, p = . 091. The were no significant interactions between group and BAS $F(1, 66) = .700$, $p > .05$, η^2_p = .010, group and connectivity $F(1, 66) = 1.447$, $p > .05$, η^2_p = .021 or between group, BAS and connectivity $F(1, 66) = .268$, $p > .05$, η^2_p = .004 .

False Know responses

The discrimination ability d’ of participants’ Know responses to the critical lure, there was no significant main effect of group $F(1, 66) = 2.157$, $p > .05$, $\eta^2_p = .032$. Also, there was no main effect of BAS $F(1, 66) = .717$, $p > .05$, $\eta^2_p = .011$, and no main effect of connectivity $F(1, 66) = .191$, $p > .05$, $\eta^2_p = .003$ on Know responses. However, there was a significant interaction between BAS and connectivity $F(1, 66) = 7.011$, $p = .010$, $\eta^2_p = .096$. Paired-samples t-tests found a better discrimination ability for low BAS/low connectivity (*M* = .79, *SD* = .81) compared to low BAS/high connectivity (*M* = .58, *SD* = .66, $t(67) = 2.140$, $p = .036$. However, there was no difference between high BAS/low connectivity and high BAS/ high connectivity $t(67) = -1.439$, $p = .155$. Finally, a significant interaction was found between group, BAS and connectivity $F(1, 66) = 4.787$, $p = .032$, $\eta^2_p = .068$. The interaction was followed up with 2-way ANOVAs with BAS and connectivity as factors separately for early and late groups. For the late group only, a significant interaction between BAS and connectivity was found $F(1, 33) = 11.231$, $p = .002$, $\eta^2_p = .254$. Paired-samples t-tests showed a better discrimination ability for low BAS/low connectivity (*M* = .94, *SD* = .80) compared to low BAS/high connectivity (*M* = .52, *SD* = .59, $t(33)$ 2.970, $p = .006$. There was no difference between high BAS/low connectivity and high BAS/ high connectivity $t(33) = -1.423$, $p = .164$. There were no significant interactions between group and BAS $F(1, 66) = 2.590$, $p > .05$, $\eta^2_p = .038$ or between group and connectivity $F(1, 66) = 1.142$, $p > .05$, $\eta^2_p = .017$.

For response bias C, there was no main effect of group on Know responses $F(1, 66) = .578$, $p > .05$, $\eta^2_p = .009$. Also, there was no main effect of BAS $F(1, 66) = .716$, $p > .05$, $\eta^2_p = .011$, and no main effect of connectivity $F(1, 66) =$.

178, $p > .05$, $\eta^2_p = .003$ on Know responses. However, there was a significant interaction between BAS and connectivity $F(1, 66) = 6.958$, $p = .010$, $\eta^2_p = .095$. Paired-samples t-tests found a more liberal bias for low BAS/low connectivity (*M* = .97, *SD* = .41) compared to low BAS/high connectivity (*M* = 1.07, *SD* = .40, $t(67) = -2.126$, $p = .037$. However, there was no difference between high BAS/low connectivity and high BAS/ high connectivity $t(67) = 1.448$, $p = .152$. Finally, a significant interaction was found between group, BAS and connectivity $F(1, 66) = 4.888$, $p = .031$, $\eta^2_p = .069$. The interaction was followed up with 2-way ANOVAs with BAS and connectivity as factors separately for early and late groups. For the late group only, a significant interaction between BAS and connectivity was found $F(1, 33) = 11.297$, $p = .002$, $\eta^2_p = .255$. Paired-samples t-tests showed a more liberal bias for low BAS/low connectivity (*M* = .97, *SD* = .83) compared to low BAS/high connectivity (*M* = 1.18, *SD* = .34, $t(33)$ -2.975, $p = .005$. There was no difference between high BAS/low connectivity and high BAS/ high connectivity $t(33) = 1.446$, $p = .158$. However, there were no significant interactions between group and BAS $F(1, 66) = 2.607$, $p > .05$, $\eta^2_p = .038$ or between group and connectivity $F(1, 66) = 1.138$, $p > .05$, $\eta^2_p = .017$.

6.5 Discussion

The aim of the current experiment was to explore the role of early and late sleep on false memory with manipulated levels of BAS and inter-item connectivity of the DRM word lists. The findings of the experiment showed that regardless of BAS, a high level of inter-item connectivity aids recognition. The greatest correct recognition was found for lists with high inter-item connectivity. No difference was found between early and late sleepers in terms of either correct or false recognition.

The results for correct recognition showed that inter-item connectivity was the most important factor. Correct recognition was found to be greater when lists were designed with high inter-item connectivity compared to lists with low inter-item connectivity. This was supported by the findings of signal detection analysis that criterion bias was more liberal for items in the high connectivity compared to low connectivity lists. This result is in accordance with those of McEvoy et al. (1999), who found that participants were more likely to correctly recognise list words from strongly interconnected lists than those from the weakly interconnected lists. A similar result was also found by Knott et al. (2012), suggesting that the denser the interconnections, the greater the recognition of list words. This result can be explained in terms of spreading activation that the high connection between the words in the list results in strong activation to those words during study (Nelson, McKinney, Gee, & Janczura, 1998). Such activation makes these words more available for processing and the amount of activation accruing to these words is determined by the level of connections among the words themselves (McEvoy et al., 1999). Moreover, signal detection analysis showed that discrimination ability (more correct memories to studied items and fewer false alarms to unrelated fillers) was higher for high inter-item connectivity compared to low inter-item connectivity. This suggests that high inter-item connectivity may permit better encoding of list words, making the presented words more distinguishable from other words.

The lack of a significant difference between early sleep (rich in SWS) and late sleep (rich in REM) on both correct and false memory is a finding which at first appears counterintuitive. However, previous research on episodic memory has provided mixed results regarding the role of SWS on episodic memory. Barrett and Ekstrand (1972) investigated word lists consolidation in three different groups. The interval corresponds to waking duration for the first group and early sleep and late sleep for the second and third group respectively. Performance improved after sleeping instead of after wakefulness and there was greater improvement when retention intervals were filled with SWS instead of with REM sleep. Similar results were also found in previous research demonstrating that SWS improves declarative memory (Alger, Lau, & Fishbein, 2012; Plihal & Born, 1997; for review see Rasch & Born, 2013; Yaroush, Sullivan, & Ekstrand, 1971). However, other research has produced conflicting results. For instance, the ability to remember words has been found to be significantly reduced subsequent to being deprived of REM sleep (Empson & Clarke, 1970). Similarly, REM sleep has been found to benefit the recognition of schema-conformant melodies (Durrant, Cairney, McDermott, & Lewis, 2015). Therefore, this may indicate the beneficial effect of sleep in general, instead of during a distinct sleep stage. This idea was supported by the findings of Rauchs et al. (2004) who investigated the consolidation of episodic memories by using a task accounting for three aspects (factual, spatial, and temporal) of episodic memory. The recognition task was used for all of the responses. The results reveal that spatial component of episodic memory (i.e. recognising the item locations) benefits from late REM sleep but SWS seems to be of greater benefit to the temporal dimension of episodic memory. Therefore, the results suggest that real episodic memory consolidation (i.e. that which includes every dimension) requires both REM and SWS sleep because these stages likely deal with distinct elements of memory (Rauchs et al., 2004). Moreover, although it is well known that the first half of the night is rich in SWS, it does still contain a considerable amount of REM sleep. This is because normal sleepers usually cycle throughout two NREM/REM periods during the early sleep hours (Tucker et al., 2006). Therefore, this may suggest that it is not the NREM sleep alone, but the combination of SWS and REM sleep that benefit the memory following sleep.

As previously outlined, FTT claims that the gist traces are responsible for the formation of false memories (Reyna & Brainerd, 1995). Meanwhile, the AMT proposes that study of words associated with a lure leads to activation of the lure representation in semantic memory. This activation increases the probability that errors will be made in association with that lure in subsequent memory tests (Roediger et al., 2001). However, it has been suggested that relational memory formation and memory integration, rule and pattern extraction, and generalisation occur mainly in the process of SWS hippocampal-neocortical transfer (Landmann et al., 2014). On the other hand, it has been demonstrated that REM sleep facilitates spreading activation and the associative processes mainly occur during REM sleep (Rauchs et al., 2005). Hence, false memory can be formed as a result of gist extraction during SWS or from spreading activation during REM sleep. It is also possible that a combination of SWS and REM sleep affect the false memory formation as well.

This is because the reactivation during SWS is believed to aid episodic memory transformation and synaptic consolidation processes which happen during subsequent REM sleep assist with stabilising the newly transformed representation (Rasch & Born, 2013).

Overall, the results of this experiment support the findings of the empirical research regarding the role of inter-item connectivity on correct memory. The results also indicate that various neural activities patterns over the sleep period in general, instead of during a distinct sleep stage, are likely to have a function in false and correct memory consolidation. This also supports the sequential hypothesis mentioned in Chapter 1 which proposed that memory is benefitted by sleep through both REM and SWS sleep in cyclic succession.

Chapter 7: Summary of Results, General Discussion and Conclusion

7.1 Introduction

This thesis has reported the findings from five experimental studies using the DRM paradigm which investigated some possible factors that might influence the effect of sleep on false memory. The studies also explored whether specific sleep stages are involved in false memory consolidation. In addition, total sleep time (TST) scores obtained from sleep diary and actigraphy watch were compared to examine the extent to which there is congruence between these two tools. In this final chapter, the findings will be summarised and then the interpretation and discussion of the findings will be addressed prior to drawing final conclusions.

7.2 Summary of the Results

Previous research has suggested multiple factors that could affect false memory formation within the DRM paradigm. This thesis looked at some of these potential factors to examine how they might influence the size of the effect of sleep on memory activation for broad semantic associations. In Chapters 3 and 4 an investigation into the effects of sleep and emotional versus neutral DRM memories was conducted. Previous research which investigated the selective enhancement of emotional memories, only used negative and neutral stimuli (e.g. McKeon et al., 2012; Nishida et al., 2009; Payne et al., 2008). Thus, this chapter aimed to explore the role of sleep on false memories for negative, neutral *and* positive DRM word lists using both recall and recognition tasks. The results demonstrated that valence has a significant impact on true recognition. Specifically, the highest level of correct recognition was observed for negative words, followed by positive and neutral words. With respect to false recognition, a significant main effect of valence was observed, with a significantly higher level of false recognition found for negative lures, compared to neutral lures. However, no significant effect of sleep on either correct or false recognition was found. When the recall test was used, the results showed that valence had a significant impact on correct recall, with a higher recall for neutral compared to positive lists. When false recall was assessed, no significant main effects of valence were observed. As with the recognition test, no significant differences between the sleep and wake groups were observed in terms of correct and false recall, indicating that valence did not have a significant impact on the effect of sleep on DRM false memory. Additionally, the findings from the AM and PM control groups showed that the time of day had no significant effect on correct or false recall.

Chapter 5 aimed to explore the connection between sleep, false memory formation, BAS, inter-item relations within DRM lists and list length. This was achieved by conducting two experiments; in the first, inter-item connectivity and BAS of the lists were manipulated; and in the second, the length of the lists were manipulated. The results indicated that the greatest correct recall was associated with high inter-item connectivity and both correct and false recall were greater for lists with high BAS. In addition, false recall was higher among the group which had a period of sleep, compared to the group assigned to a comparative period of wakefulness. Moreover, the results demonstrated that more studied words were recalled correctly when shorter DRM lists were used while false recall was greater for longer lists. Additionally, sleep increased false recall, irrespective of the list length, although correct recall levels were unchanged between sleep and wakefulness conditions. Interestingly, none of the manipulated factors interacted with sleep, however, the interaction p-values tended to approach significance. When conducting a follow-up exploratory analysis, findings revealed that lists designed with low connectivity and fewer words enhanced false memory after sleep. These results may suggest that inter-item connectivity and list length mediated the effect of sleep on false memory.

Chapter 6 extended the findings of Chapter 5 to investigate the specific sleep architecture involved in DRM false memory with manipulated BAS and inter-item connectivity. The findings indicated that correct recognition was greater for high connectivity compared to low connectivity. No difference was found between the early and late sleepers in terms of either correct or false recognition. This suggests that both SWS and REM sleep play a role in the creation of false memory.

Additionally, in Chapter 3, the actigraphy and sleep diary data from Experiment 1 was used to investigate the compatibility between these two tools in measuring TST on a day-by-day basis. There were significant differences in the overall TST scores between the actigraphy watch and sleep diary. Interestingly, when the daily scores of TST between the actigraphy watch and sleep diary were compared, the differences between them were not consistently different. A significant difference was found between the two measurements on days three and five, with sleep diary overestimated TST than actigraphy. Whereas, on days one, two and four there was congruence in the TST scores for both actigraphy and the sleep diary.

7.3 Interpretation and Discussion of the Results

7.3.1 The Role of Sleep and Emotional Stimuli in DRM False Memory

Previous research suggests that memories with an emotional tone are more likely to be remembered than neutral memories (Adelman & Estes, 2013; Walker & van Der Helm, 2009; Wiesner et al., 2015), with evidence indicating that sleep, in particular, selectively enhances emotional memory (Rasch & Born, 2013; Walker & van Der Helm, 2009). However, this enhancement of emotional memory during sleep may not be limited to correct memories; it could also lead to enhanced emotional false memories. This is because the activation of broad semantic networks spreads during sleep giving an opportunity for false memories to be activated (Newbury & Monaghan, 2019) and integration of memories with other events or associations mainly occurs during sleep. Within the DRM paradigm, the emotionality of the word lists has previously been found to be a factor that could increase the probability of false memory (Brainerd et al., 2008; Dehon et al., 2010; Howe et al., 2010; Sharkawy et al., 2008). However, investigations into the role of valence in DRM false memories are yet to examine the role of sleep and its impact on negative, positive and neutral DRM lists.

Therefore, Experiments 1 and 2 reported in this thesis investigated the influence of negative, positive and neutral lists on sleep-dependent false memory performance. In Experiment 1, a recognition test was used as a measure of false memory. A recognition test was chosen rather than a recall test because a greater influence of emotionality has been found to exist during recognition tests relative to recall tests (Howe et al., 2010; Sharkawy et al., 2008). The results from this experiment revealed a specific influence of valence on correct memory recognition, in particular, the highest level of correct recognition was found in negative valence lists followed by the positive and then neutral lists. This suggests that selective enhancement occurs for emotional negative and positive information compared to neutral information. These findings are consistent with previous research which has found that emotionality of words increases memory accuracy. For example, Adelman and Estes (2013) found that negative and positive words were remembered significantly more than neutral words. Moreover, the high level of correct memory for negative stimuli over positive and neutral stimuli was also reported in previous studies (Brainerd et al., 2008; Dewhurst & Parry, 2000; Kensinger & Corkin, 2004; Ochsner et al., 2004). This finding suggests that, compared to positive or neutral information; negative information makes the memory more distinctive. It has been claimed that people are more likely to report detailed memories for events that have negative emotion even after many years (Brewer, 2006); this is because the negative qualities of the information prevent memories from distorting (Bookbinder & Brainerd, 2016). Additionally, the current findings indicated that emotional valence significantly impacted on false recognition, with emotional lures falsely recognised more than neutral lures. This is in line with the findings from previous studies which also found that participants tended to recognise emotional lures at a higher rate than neutral lures (Brainerd et al., 2008; Howe et al., 2010; Sharkawy et al., 2008). The current findings support the suggestion that gist representations of emotional lists are stronger than non-emotional lists (Goodman et al., 2011). However, no significant difference was found between negative and positive false recognition. This is not in line with the research by Brainerd et al. (2008), which found that, in comparison to words with neutral valence, negative valence results in increases in false recognition, whereas positive valence reduces false memory effects. In contrast, the current results suggest that there could be an overall influence of emotionality on increasing associations, which leads to general increases in false memories for emotional words with positive and negative valence, in comparison to neutral words.

Another objective of Experiment 1 was to investigate the effects of sleep on emotional and non-emotional memory processing. The results showed that there were no overall differences between sleep and wake groups in terms of either correct or false memory for any of the list types. This finding is in contrast with the only study that has examined the role of sleep and emotional stimuli on DRM false memory (McKeon et al., 2012). However, previous research has suggested that recognition memory does not seem to be as affected by post-learning sleep than cued or free recall (Drosopoulos et al., 2005; Hu et al., 2006). Therefore, it is possible that the effects of post-learning sleep on generating false memories could be uncovered with methods that test with greater sensitivity. However, the current findings are consistent with the findings from the study by Diekelmann et al. (2008), in which no difference was observed between the sleep and wake group in terms of either correct or false recognition.

Since the effect of sleep might be less prominent for recognition tests than recall tests, Experiment 2 replicated the previous experiment but with a recall test instead of a recognition test in an attempt to test the above suggestion and to recreate the effect of sleep that McKeon et al. (2012) found. The findings from Experiment 2 revealed that valence affected correct recall, as demonstrated by the fact that correct recall for the neutral lists was significantly higher than for the negative or positive lists. This is consistent with the findings from previous studies, which also found that neutral lists were preferentially recalled more effectively than emotional lists (Howe et al., 2010; McKeon et al., 2012). It has been argued that emotional information yields more gist than neutral information; hence, correct recall for neutral lists is better than for emotional lists (Howe et al., 2010). This is because false memory is based on gist trace while correct memory is based on verbatim traces. However, no significant effects of valence were found on false recall. While this finding contradicts certain previous findings (Dehon et al., 2010), it is consistent with others (Sharkawy et al., 2008). Interestingly, the difference found in the effect of valence on false memory between the test of recognition and test of recall supports the previous findings where the effect of emotion on false memory was only observed in the recognition test (Howe et al., 2010; Sharkawy et al., 2008).

As with Experiment 1 and false recognition, the results from Experiment 2 showed no overall effect of sleep on false recall. This is inconsistent with previous findings in which sleep was found to increase false recall (i.e. McKeon et al., 2012; Payne et al., 2009). However, Experiment 2A showed no difference between AM and PM groups in terms of either correct and false recall; which is consistent with the findings of prior studies (Fenn et al., 2009; McKeon et al., 2012; Payne et al., 2009). These findings suggest that the absence of sleep effect on false memory found in Experiment 1 and 2 are not due to circadian rhythms effects. They also provided evidence that online collection is valid with such an experimental design.

Furthermore, in a study by Diekelmann et al. (2010), the enhancement of false recall after a period of sleep was only observed in participants with low general memory performance. Hence, Diekelmann et al.'s strategy of grouping participants according to their memory performance was followed on the Experiment 2 data, in order to ascertain if the same results would be found. The results from the high/low memory performance analysis were the same as the initial findings for the sleep and wake groups: no effect of sleep was found on false recall. The finding that sleep had no effect on false recall seems counterintuitive to the current understanding that the sleep effect is greater in tests of recall than in recognition tests (Newbury &

Monaghan, 2019) as well as being inconsistent with the previous findings (Diekelmann et al., 2010; McKeon et al., 2012; Payne et al., 2009). A number of potential explanations for this result can be suggested. Firstly, it may be that a specific stage of sleep results in the effect on emotional false memory rather than sleep in general. Previous studies have found that REM sleep strengthens emotional memories (Groch et al., 2013; Nishida et al., 2009; Wagner et al., 2001). This role of REM sleep on emotional memory enhancement may also lead to enhanced emotional false memory as REM sleep has been found to facilitate spreading activation and promote memory insight (Carr & Nielsen, 2015; Sio et al., 2013). Secondly, it may be that emotional stimuli hinder the effect of sleep on false memory. When McKeon et al. (2012) studied the effect of sleep and the emotional content of DRM lists on the production of false memories, they found that false memories for both negative and neutral lures were increased after a period of sleep. Their results suggested that emotional valence does not impact the formation of false memories after sleep. However, this could help to explain the lack of an effect of sleep on false memory in the current experiment. Specifically, McKeon et al. (2012) tested their participants using five negative lists and five neutral lists, while in the current experiment two types of emotional lists: positive and negative, with four lists for each were used. Consequently, using more emotional lists may have a cumulative effect on the effect that sleep has on false memory. Hence, if there is no effect of emotional stimuli on false memory after sleep when using a small number of emotional lists, then the absence of the effect of sleep on false memory found in the current experiment could be due to the inclusion of a greater proportion of emotional lists in the experimental setup. Furthermore, Lehmann, Seifritz and Rasch (2016) found no group differences in a study where participants first gave ratings to neutral words for arousal. The words were paired with images which were either positive, negative, or neutral. Then, after sleeping or waking, the words were again rated for arousal which indicated that words linked to either positive or negative images were now rated as more arousing that those linked with a neutral image. Yet, the differences did not change depending on whether individuals were in a sleeping or waking group, indicating that sleep did not increase how much the emotion of each image was 'transferred' to the neutral words linked to it.

Another possible explanation is that it may be that sleep deprivation instead of sleep interacts with the effect of emotion on false memories. It is widely recognised that disturbances in sleep may yield profound deficits in the executive functions, attention span, vigilance and working memory of an individual, particularly when sleep deprivation occurs over an extended period of time (Kerkhof & Van Dongen, 2011). Accordingly, it may be expected that periods of sleep deprivation may lead to an increased risk of false memories due to poor retrieval of memories. A reduction in source and reality monitoring have been associated with sleep deprivation in the context of memory retrieval, both of which have been identified as contributing to the generation of false memories (Kerkhof & Van Dongen, 2011). Interestingly, sleep deprivation is known to have a large influence on the ability to process emotional information (Kahn, Sheppes, & Sadeh, 2013). Sleep loss appears to decrease the ability to recognise emotions from a facial expression (Babson & Feldner, 2015; Minkel, Htaik, Banks, & Dinges, 2011) . Dahl and Lewin (2002) have suggested that sleep loss leads to a deficit in the capability to downregulate emotions through its effect on the prefrontal cortex, an area known to have specific sensitivity to the effect of sleep loss (for a review, see Franzen, Siegle, & Buysse, 2008; Ochsner et al., 2004). Franzen et al. (2008) published research based on a study in which participants viewed emotionally provoking photos with and without being deprived of a night's sleep. The group differences were detected in the pupillary response. They stated that pupil dilation occurred as a response to negative photos for the sleep-deprived group relative to the control group, which suggests that reductions in sleep are linked to increased emotional activation and a response to unpleasant stimuli. Thus, without sufficient sleep, there is greater reactivity to emotional stimuli, although the precise effects vary between the research. Some research has indicated that being deprived of sleep primarily has an effect on a response to a negative stimulus and a lack of sleep featured in a more negative interpretation of neutral stimuli. Some studies have discovered that sleep deprivation increases the response to positive stimuli. Daniela et al. (2010) found that sleep deprivation causes participants to rate a neutral image more negatively without having an effect on how a negative image is rated. Zohar, Tzischinsky, Epstein, and Lavie (2005) utilised experiential sampling to collect data on emotion and sleep among medical professionals who had sleep fluctuations due to their work shifts. They report that less sleep is linked to decreases in positive emotions after a desirable event and increases in negative emotions after an undesirable event. Gujar, Yoo, Hu and Walker (2011) found that sleep loss leads to an increase in emotional reactivity for both negative and positive information. They suggested that sleep deprivation, in comparison to normal sleep, decreases functional connectivity between the brain's frontal regions. This lack of functional connectivity leads to an increase in emotional reactivity. Therefore, it could be that emotional stimuli interact with sleep deprivation to produce false memory, but they don't have much influence on the size of the effect of sleep on false memory, which would explain the current findings of the absence of the effect of sleep on false memory when emotional lists were used. One could argue that this is inconsistent with the findings of Howe et al. (2010, Exp. 3) whereby false recognition for the negative lures increased over one week. However, their study did not explicitly measure sleep and did not screen out participants with sleep problems. Thus, it is possible that sleep deprivation may have had an impact on their findings, or that the effects they found emerged over a longer time period than 24 hours.

7.3.2 The Roles of Backward Associative Strength, Inter-item Connectivity and List Length on DRM False Memory

It has been suggested that the density of inter-item connections plays a considerable role in the production of false memory within the DRM paradigm (Deese, 1959; McEvoy et al., 1999). In Experiment 3, the level of inter-item connectivity and BAS of the lists were manipulated, and in Experiment 4, the length of the lists were manipulated, lists of 6 and 12 words were used.

The results from Experiment 3 suggested that both inter-item connectivity and BAS affect recall, as the greatest correct recall was associated with high inter-item connectivity and high BAS lists compared to low connectivity and low BAS. The finding that high connectivity yielded higher correct recall is consistent with the findings of prior studies (Deese, 1959; Knott et al., 2012; McEvoy et al., 1999; Roediger et al., 2001). This suggests that lists that have stronger connectivity between the words increase the opportunity to activate similar presented words. This is due to the fact that activation is spread amongst the associates within semantic memory. Moreover, the higher the levels of inter-item association, the more likely it is that the production of a list item will cue recall of other list items compared to the non-presented lure. The role of high BAS on correct recall was also found in previous research (McEvoy et al., 1999). It has been suggested that list words share more connections when they are more connected to their lure words and higher connections between the list words increase the probability of correctly recalling them (McEvoy et al., 1999). This was demonstrated in the current findings where high inter-item connectivity was found to increase correct recall.

However, BAS has been identified as a key factor in false memory formation within the DRM paradigm (Roediger et al., 2001). This was supported by the current findings as high BAS lists were found to produce more recall of a critical lure than low BAS. Also, a substantial body of research supports this finding and confirms that BAS plays a considerable role in the production of false memory within the DRM paradigm (Arndt, 2012; Beato & Arndt, 2017; Deese, 1959; Hicks & Hancock, 2002; Hicks & Marsh, 1999; Howe et al., 2009; Knott et al., 2012; McEvoy et al., 1999; Roediger et al., 2001). BAS refers to the strength of the association between the list words and their associated critical lure, meaning that there is a high chance of the critical lure being recalled in high BAS lists as a result of the strong association in the semantic networks as per AMT. In contrast to the effect that inter-item connectivity has on correct recall, connectivity did not affect false recall. This is inconsistent with the findings of prior studies (Knott et al., 2012; McEvoy et al., 1999). However, it has been suggested that the strong association between the critical lure and list words is the primary predictor of false memories (Smith et al., 2002; Knott et al., 2012). This may explain the absence of an effect of connectivity on false recall. BAS may be an overwhelming factor in false recall, meaning that any other role of connectivity may get lost where BAS is manipulated.

In Experiment 4, DRM list length was manipulated to explore its impact on the effect of sleep on correct and false recall. The results confirmed the findings by Robinson and Roediger (1997) and Sugrue et al. (2009) that short lists yielded higher correct recall compared to longer lists. This suggests that decreasing the number of words in the lists increases the probability that studied words will be recalled as there are only a small number of words to be recalled. However, the results for false recall support the prediction that studying lists with a large number of words increases false recall. Long lists have been found previously to increase false memory (Gallo & Roediger, 2003; Hutchison & Balota, 2005; Libby & Neisser, 2001; Marsh & Bower, 2004; Robinson & Roediger III, 1997; Sugrue et al., 2009; Watson et al., 2003). It can be concluded from this that the greater the number of words in a list, and therefore the more associations there are with the critical lure, the greater the likelihood that false memory for an associated lure will occur.

Interestingly, the results from the experiments reported in this chapter provide evidence for the role of sleep in increasing false recall, with both experiments revealing higher false recall in the sleep groups relative to the wake groups. This effect of sleep on false memory has also been identified in other DRM studies (Darsaud et al., 2011; Diekelmann et al., 2010; Pardilla-Delgado & Payne, 2017; Payne et al., 2009; Shaw & Monaghan, 2017). According to FTT, gist traces are responsible for the formation of false memories, and sleep is believed to enhance associative processes that permit more gist extraction (McKeon et al., 2012). As a result, more false memory would be generated after sleep. Moreover, if critical lures are activated due to their similarity to presented words, then the increase in false recall after sleep is expected as sleep promotes a greater spread of activation of associates within semantic memory (Sio et al., 2013).

In Experiment 3 and 4, the interaction p-values for false recall approached significance. Therefore, follow-up exploratory analyses were conducted to further examine the interaction between sleep and BAS, connectivity and list length. The findings showed that low connectivity and shorter lists enhanced false memory formation after sleep compared to an equivalent period of wakefulness. This supports the finding of previous studies that sleep is more beneficial with more complex tasks (Sio et al., 2013). In DRM lists, decreasing the level of inter-item connectivity and the number of words in a list results in less strong associations, making the activation of the critical lure more difficult to achieve. Hence, one possible explanation for the current findings is that the activation of broad semantic networks spreads during sleep may give an opportunity for false memories to be activated in lists with lower connectivity and fewer words. The current findings suggest that the density of inter-connections and list length might be moderator variables which affect the role of sleep on DRM false memory.

7.3.3 Sleep Stages, Inter-item Connectivity and BAS

Experiment 5 built on the finding from Experiment 3 that revealed that sleep increases false recall. Also, the finding that inter-item connectivity mediated the effect of sleep on false memory as low connectivity lists found to enhance false recall after sleep. It should be noted that the later finding emerged from an exploratory analysis.

In this experiment, the same materials were used as Experiment 3 to investigate whether a specific sleep stage was responsible for this increase. This investigation was carried out using the night-half comparison procedure which allows for dissociation of the effects of SWS and REM sleep. Most SWS sleep occurs in the 'early sleep' period, the first part of the night, whereas REM sleep occurs most in the later part of the night, the 'late sleep' period (Plihal & Born, 1997). The results were consistent with the findings from Experiment 5 and from previous studies regarding the role of high inter-item connectivity in correct recognition (McEvoy et al., 1999). Correct recall was greater for high connectivity lists compared to low connectivity, however, there was no effect of connectivity on false recognition. It has been argued that strong relationships between list items may allow for better encoding of studied words, meaning that those words become more distinguishable from the lure words (Knott et al., 2012). Regarding BAS, although high BAS was found to increase both correct and false recall in Experiment 3, in this experiment no such effect was found on correct and false recognition. It may be that BAS is somehow affected by the differences between the test of recall and recognition after sleep. However, no difference was found, in this study, between the early and late sleepers in terms of either correct or false recognition. This suggests that both stages play a role in memory consolidation. Declarative memory was found to be enhanced by SWS sleep in some studies (Plihal & Born, 1997, 1999) and by REM sleep in other studies (Durrant et al., 2015; Empson & Clarke, 1970; Plihal & Born, 1999; Géraldine Rauchs et al., 2004). In terms of false memory, while SWS was found to enhance memory integration and gist extraction (Landmann et al., 2014), REM sleep on other hand appears to facilitates spreading activation, associative processes and the promotion of memory insight (Geraldine Rauchs et al., 2005). Therefore, both SWS and REM sleep can affect false memory formation within DRM as false memory can be either formed as a result of gist extraction during SWS or from spreading activation during REM sleep. Furthermore, the results can be further explained in the light of the sequential hypothesis that proposes that during sleep, memories are processed in two sequential phases occurring through SWS and REM sleep respectively. First, SWS weakens memories encoded during wakefulness to distinguish them from irrelevant information. Then, the retained memories are integrated with pre-existing memories during REM sleep. Hence, both SWS and REM may contribute to false memory generation during the weakening process which occurs during SWS sleep and/or through the memory integration during REM sleep.

7.3.4 Measuring Total Sleep Time with Actigraphy Watch and Sleep Diary

Analysing the data from the sleep measurement part of experiment 1 enabled exploration of the extent to which there is congruence in TST scores between actigraphy and sleep diaries on a day-by-day basis in healthy young adults for five consecutive days. The findings showed that there was a significant difference in the total TST scores across the five nights between the actigraphy watch and sleep diaries. Participants generally reported a longer TST in the sleep diary relative to that recorded by the actigraphy watch. Previous research supports this finding, with results indicating that sleep diaries tend to overestimate the TST compared to actigraphy (Campanini et al., 2017; Lichstein et al., 2006; Maes et al., 2014; Maich et al., 2018; Short et al., 2012). Daily sleep diaries rely on self-reporting that is subject to inaccuracies caused by human judgment and interpretation. Also, individuals completing sleep diaries are likely to report retrospective data which increases the chance of error (Wolfson et al., 2003). However, the differences between actigraphy scores and diary scores were not consistently different in the study reported in this thesis. The results indicated that there was no difference in TST scores generated by both tools on the first, second and fourth days. Meanwhile, a significant difference was observed on the third and fifth days. The findings suggested that there was not a consistent level of agreement or differentiation between the TST scores from actigraphy and sleep diaries, but rather variation on a day-by-day basis. They also suggested that when using a sleep diary to measure the TST, the level of accuracy may fluctuate across a long recording span. However, the current finding also suggests that if the TST is to be measured for one night – which is the case in online experiments reported in this thesis – a sleep diary can be a reliable tool to use, since the highest agreement found between sleep diary and actigraphy watch was on day one. This suggestion is supported by a more recent study comparing the sleep diary and actigraphy TST scores for one night in 21 young participants; the study found no difference between the TST scores reported by the two methods (Williams et al., 2020).

The experiments reported in this thesis looked beyond the conflicting findings available in the literature. They aimed to investigate the reasons for the mixed results found on whether sleep increases, decreases or has no effect on false memory with DRM. Based on the author's knowledge, these experiments are the first that have examined potential factors that influence the effect of sleep on false memory within DRM paradigm. Table 7.1 provides a summary of the key findings of the DRM experiments reported in this thesis.

Table 7.1 key findings of the DRM experiments

Experiment number	Aim of the Experiment	Key Findings	Results are consistent with:
1 and 2	To investigate whether emotional word lists have a significant impact on false memories after sleep.	No difference between sleep and wake groups in terms of either correct or false memory, this was seen with both recognition and recall tests.	Diekelmann et al. 2008.
		Emotionality of the lists does not play a role in the influence of sleep on false memory.	McKeon et al. 2012.
		Correct recognition was higher for negative lists than positive and neutral lists.	Brainerd et al. 2008.
		False recognition was higher for emotional lures than neutral lures.	Brainerd et al. 2008. Howe et al. 2010. Sharkawy et al. 2008.
		Correct recall was higher for neutral lists compared to positive lists	Howe et al. 2010. McKeon et al. 2012.
3	To investigate whether manipulation of the associative strength from the words in the list to the critical lure and the connectivity between words in the list can affect the role of sleep in the creation of false memories	Sleep was found to increase false recall when these factors were manipulated.	Darsaud et al. 2011. Diekelmann et al. 2010. Pardilla-Delgado & Payne, 2017. Payne et al., 2009. Shaw & Monaghan, 2017
		Correct recall was greater for high connectivity.	Deese, 1959. McEvoy et al. 1999. Knott et al. 2012.
		High BAS was associated with higher false recall.	Knott et al. 2012.
4	To investigate how the manipulation of list length affects the role of sleep in the creation of false memories	Sleep was found to increase false recall when the length of lists was manipulated.	Darsaud et al. 2011. Diekelmann et al., 2010. Pardilla-Delgado & Payne, 2017. Payne et al. 2009. Shaw & Monaghan, 2017.
		Correct recall was greater for shorter lists. False recall was greater for longer lists.	Robinson and Roediger, 1997. Sugrue et al. 2009.

5	To investigate the role of SWS and REM on false memory formation with manipulated levels of inter-item connectivity and BAS	No specific effect of sleep stages on correct and false recognition.	No relevant research available
		Correct recognition was greater for high connectivity lists.	Deese, 1959. McEvoy et al. 1999. Knott et al. 2012.

The findings from these experiments allow for the exploration of the underlying factors of how sleep interacts with other variables to effect DRM false memories in a way that has not been previously explored. These experiments provide an important platform on which further work can expand. Newbury and Monaghan (2018) conducted a meta-analysis to investigate the role of sleep on the consolidation of correct and false memory in studies that used the DRM paradigm. They included six possible factors to examine how they influenced the effect that sleep has on false memory. These factors were: (1) the different test used (i.e. recall or recognition); (2) list length; (3) number of presented lists; (4) the modality of the list's presentation; (5) the emotionality of the used word lists; and (6) time of sleep (i.e. overnight or nap). The authors determined that only two factors, recall tests and lists with fewer words, were found to enhance sleep-based increases of false memory within DRM. The current findings are in line with the meta-analysis, as no effect of list emotionality on false memory after sleep has been found, whereas shorter lists found to enhance false memory after sleep. However, it should be stated that these are the only factors shared between the study and the meta-analysis. Moreover, as outlined through the discussion, the results are consistent with the theories of false memory described in Chapter 1, suggesting that AMT and FTT can provide an explanation for the current findings. For instance, the finding that false recognition was higher in emotional lures than neutral lures could be due to the fact that gist representations of emotional lists are stronger than non-emotional lists or that emotional lists have tighter associative connections, leading to a stronger activation of their associate critical lure. These explanations may also apply to the absence of the effect of sleep on false memory with emotional lists where the strong gist representations and spread of activation occur for both sleep and wake groups. AMT theory supports the finding that high BAS was associated with higher false recall as BAS pertains to the strength of the association between words and the critical lure, meaning that a higher chance of false recall would be the result of association in the semantic networks. FTT also supports this finding; it has been suggested that BAS is an index of gist (Knott et al., 2012). Gist extraction will be easier with high BAS lists because the associative strength from the list words to the critical lure is strong, thereby making it more accessible to extract the theme word from semantically related words. The finding that false recall was greater for longer lists could be also explained in the light of both FTT and AMT as studying lists with a large number of related words increases the likelihood of extracting the general meaning as well as activating their associated words, resulting in producing more false memory, as per FTT and AMT.

7.4 Limitations

One potential limitation of Experiment 2 was that the data were collected online; participants in the sleep group were required to adhere to a minimum amount of sleep – no less than 6 hours – while participants in the wake group were required not to have naps during the day. Adherence to these criteria were only verified through self-reporting methods. It is possible that participants lapsed from this; therefore, the absence of the effect of sleep on false recall of emotional lists may have resulted from inaccurate reporting by participants about adherence to the minimum sleep duration. However, previous research found that conducting perceptual and cognitive experiments online did not reduce data quality (Germine et al., 2012). Furthermore, Zwaan et al., (2018) replicated nine preregistered cognitive psychology experiments online, one of those experiments was the Roediger and McDermott (1995) DRM experiment. They found that neither the online environment not the use of nonnaïve participants undermined the results. Zwaan et al. stated that "the jury is still out on whether such secondary findings (i.e. emotionality of the stimuli) are as robust as the more basic findings we have presented here" (p. 1971). Hence, it may be that while online data collection is an excellent option in many cases, some factors such as emotion may be too subtle to be revealed using this method. However, the results from Experiment 1 reported in this thesis was conducted in a lab and confirm the findings from experiment 2 of no effect of sleep on emotional false memory. Experiments 3 and 4 left emotion behind and were also conducted online and an effect of sleep on false recall was found, adding further weight to the idea that online data collection has not undermined these findings.

Another possible limitation is that the mood status of the participants was not measured in Experiments 1 and 2. Mood-congruence has been found to have an effect on false memory for emotional lists in previous DRM studies (Zhang et al., 2017). This may have influenced the findings. However, there was no difference between the sleep and wake groups for correct emotional lists, suggesting that mood is unlikely to have had an effect on the overall false memory results.

In Experiment 1 and 2, the number of female participants was more than the double that of the male participants. Out of 68 participants, there were 52 females in Experiment 1 and 48 females in Experiment 2. Previous DRM studies that looked at false memory and gender differences found no gender differences in the level of false memory (Bauste & Ferraro, 2004; Howe et al., 2010; Kreiner, Price, Gross & Appleby, 2004). Dewhurst, Anderson and Knott (2012) also found no difference between male and female in the false recall of neutral critical lures, although they did find that females falsely recalled more negative lures than males. However, the current findings revealed that there was no significant effect of valence on false recall. Thus, if the gender balance of the sample was an issue, then I might have expected to see an increase in false recall for negative lures. Since this did not happen, gender balance is unlikely to have had an effect on the current results.

Another possible limitation is related to the size of the effect used in the experiments reported in this thesis (i.e. medium). A medium effect size was chosen for pragmatic reasons. However, on consulting previous DRM studies that found an effect of sleep on false memory, I found that a medium effect size was used in some studies (e.g. Payne et al., 2009) and a small effect size was used in other studies (e.g. Fenn et al., 2009). However, a more recent study by

Monaghan et al. (2017) used a large effect size. Therefore, it seems that there is some variability in the effect sizes across studies. Effect size specifies the size of the difference between the study groups and is a true measure of the significance of the difference. Therefore, variation in the effect size might explain variability in the literature and the current research. Larger samples, of sufficient size to detect small effects, should be recruited in future research in order to draw clearer conclusions about the effect of sleep on false memory formation.

A final limitation concerns the generalizability of the results. Due to the fact that participants with sleep problems were excluded from the experiments as well as participants who were classified as ‘medium sleepers’ in Experiment 5, the findings from the experiments reported in this thesis are necessarily constrained. This is important in order to be able to tease out the processes involved in false memory creation, but when considering possible wider applications for the findings as discussed in 7.6 below caution should be taken when generalising beyond a very specific population.

7.5 Further Directions

To fully understand the way in which emotional false memories within the DRM paradigm may be generated, future research could extend the experiments reported in Chapter 3 by comparing the effects of wakefulness, sleep, and sleep deprivation state on participants’ performance for emotional DRM lists. Indeed, it can be reasonably hypothesised that although sleep may lead to a risk of false memories being generated, as a result of consolidation and retrieval processes, sleep deprivation may have a more profound and damaging effect on memory (Darsaud et al., 2011). Sleep has been found to have no impact on emotional false memory in this thesis; sleep deprivation, on the other hand, has been found in previous studies to greatly influence the ability to process emotional information (Kahn et al., 2013). Retrieval of emotional memories may be adversely affected by sleep deprivation, as profound effects have been observed on a number of emotional processes leading to an increased error rate (Minkel et al., 2011; Zohar et al., 2005). Furthermore, sleep deprivation may provide additional insight into how memories are affected by sleep and forced wakefulness events.

Investigating the effects of sleep stages on emotional memory is another avenue for research that needs be considered in future. To gain a complete understanding of the effect of sleep on processing emotional memories, further research could investigate how various sleep stages influence false, correct, and emotional memories. Experiments 1 and 2 indicated that sleep in general does not have an effect on emotional false memory. Yet, it is unclear whether sleep stages are involved in emotional false memory generation. Sleep and memory theories show that both SWS and REM sleep are important for consolidating memories; repeated activation of memories in the hippocampus occurs during SWS, and the stabilisation of these memories happens during REM sleep. REM sleep has also been correlated with emotional memory, yet the specific effects of particular sleep stages on consolidating both positive and negative false memory remain unclear. The use of EEG or PSG would allow a better investigation of this question.

A potential future direction would be to investigate the different effects of recall versus recognition on false memory with a manipulated level of BAS and inter-item connectivity. It has been demonstrated that the recognition test has less vulnerability to post-sleep false memories, possibly because of increased source monitoring abilities (Fenn et al., 2009; Johnson et al., 1993). However, the recall test is more likely to elicit false memories within DRM after sleep. Indeed, Experiment 3 found an effect of sleep on false recall with manipulated levels of list density and BAS. However, the impact of inter-item relationships and BAS on correct and false memories has been previously identified using recognition testing (Knott et al., 2012; McEvoy et al., 1999) Using the same type of stimuli, future research could investigate the potential differences between recall and recognition tests on false memory after sleep with manipulated levels of list density and BAS. This would help us to understand whether the effect of sleep found with manipulated inter-item connectivity and BAS lists is limited to recall, or whether there is a similar finding when recognition is assessed.

Finally, as Zwaan et al., (2018) suggest, the jury is out regarding how sensitive online data collection is to manipulations even of robust findings such as false memory in the DRM paradigm. Future research could usefully explore methodological aspects of data collection further which in turn could facilitate larger sample sizes.

7.6 Implications

In addition to extending our understanding of sleep and memory processes, the findings reported in this thesis have implications for applied settings such as the courtroom and classroom.

In court, the default assumption, which is also supported by the literature, is that people are more susceptible to correctly remember negative information (Brainerd et al., 2008; Dewhurst & Parry, 2000; Kensinger & Corkin, 2004; Ochsner et al., 2004). The current findings support this, whereby high level of correct recognition and recall were found for negative lists. However, the current findings revealed that false recognition was higher for emotional lures than neutral lures, whereas valence had no impact on levels of false recall. This may be related to the debate around ‘leading questions’, the idea that presenting individuals with logically associated elements that weren’t present in a memory test may lead to false recognition (Blagrove & Akehurst, 2000). Therefore, caution is needed in the way that witnesses are questioned. Furthermore, the finding that, with emotional stimuli, there was no difference between the sleep and wake group in terms of either correct or false memory has a strong legal implication. It suggests that staying awake does not impact the accuracy -or inaccuracy- of the emotional information.

The current findings also have implications for classroom learning, which is critical to understand if and how sleep contributes to educational outcomes. Although many studies have emphasised that sleep after learning strengthens declarative memory (Ellenbogen, Hulbert, Stickgold, Dinges & Thompson-Schill, 2006; Backhaus, Hoeckesfeld, Born, Hohagen & Junghanns, 2008), the current findings suggest that sleep after learning does not affect memory accuracy. Across all the experiments reported in this thesis, no specific effect of sleep -compared to an equivalent period of wakefulness- on correct memory was found and this finding was also observed in previous studies (Diekelmann et al., 2008; Fenn et al., 2009; Diekelmann et al., 2010; Lo et al., 2014). Furthermore, the findings that correct recall was greater for shorter lists and false recall was higher for longer lists may also provide insight into school learning strategies as it may indicate the importance of reconsidering the amount of information taught to students in a class for each subject. Hence, the smaller the amount of information presented to students, the more likely to be correctly remembered over time. Therefore, reducing the duration of the classes may be an effective way to enhance educational outcomes.

A growing body of literature has linked chronotype and sleep to school performance, late sleepers have been found to obtain lower school grades than early sleepers (Borisenkov, Perminova, & Kosova, 2010; Roeser, Schlarb, & Kübler, 2013; Rahafar, Maghsudloo, Farhangnia, Vollmer, & Randler, 2016). The effect of chronotype on school performance found in previous research could be influenced by other factors, such as sleep duration. Díaz-Morales and Escribano (2015) noticed that the significant correlation between chronotype and school achievements disappeared when sleep duration was added to the analysis. They stated that it is difficult to disentangle the effects of both chronotype and sleep duration on school performance. Chronotype is tightly linked to total sleep time, and late sleepers have less amount of sleep in comparison to early sleepers (Zerbini & Merrow, 2017). The findings of this thesis revealed that no difference between early and late sleepers in terms of either correct or false memory. It should be noted that early and late sleep groups experienced similar amounts of sleep. Therefore, schools could delay their starting time in order for late sleepers to have sufficient time to sleep. Another solution to enhance students' performance would be to consider the optimal time of day for teaching and testing according to students' time-of-day preferences in learning.

Finally, the current findings have strong implications for future false memory research. The work reported determined some of the factors that influence the size of the effect of sleep on false memory. This will help researchers interested in studying false memories and sleep by knowing which factors to include in future studies.

7.7 Conclusion

This thesis has conducted a series of experiments designed to expand our understanding of the factors that influence sleep and false memory formation within the DRM paradigm. The findings have been inconsistent across these experiments, some of the experiments found a main effect of sleep on false memory while no effect of sleep was found in other experiments. This inconsistency is also found in the literature. However, the thesis reports the first experiment that explored the effect of sleep and three types of emotional DRM lists. This thesis also reports the first experiments to investigate how sleep interacts with BAS, inter-item connectivity and list length to affect false memory. It also reports the first experiment to explore the connection between early and late sleep periods and an increase or decrease in false memory. The findings from this thesis add considerable weight to the existing literature on DRM false memories by suggesting that emotionality of the lists doesn't play a role in the influence of sleep on false memory. In contrast, inter-item connectivity and list length do interfere with the effect of sleep on false memory as the follow-up exploratory analysis on these factors revealed that low inter-item connectivity and shorter lists enhanced false memory after sleep. The findings from this thesis also suggest that the observed effect of sleep on false memory is not due to a specific sleep stage. Furthermore, this thesis has extended the literature on the validation of sleep diaries for measuring total sleep time by suggesting that there is not a consistent level of differentiation between the scores generated from actigraphy and sleep diaries, but instead variation on a day-by-day basis.

The novel findings reported here have demonstrated a variety of ways in which sleep interacts with other variables to create false memories in the DRM paradigm.

References

Acebo, C., Sadeh, A., Seifer, R., Tzischinsky, O., Wolfson, A. R., Hafer, A., & Carskadon, M. A. (1999). Estimating sleep patterns with activity monitoring in children and adolescents: how many nights are necessary for reliable measures? *Sleep, 22*(1), 95-103.

Adelman, J. S., & Estes, Z. (2013). Emotion and memory: A recognition advantage for positive and negative words independent of arousal. *Cognition, 129*(3), 530-535.

Akhtar, S., Howe, M. L., & Hoepstine, K. (2020). Can false memories prime problem solutions for healthy older adults and those with Alzheimer's disease? *The Journals of Gerontology: Series B, 75*(4), 743-752.

Alger, S. E., Lau, H., & Fishbein, W. (2012). Slow wave sleep during a daytime nap is necessary for protection from subsequent interference and long-term retention. *Neurobiology of learning and memory, 98*(2), 188-196.

Ancoli-Israel, S., Cole, R., Alessi, C., Chambers, M., Moorcroft, W., & Pollak, C. P. (2003). The role of actigraphy in the study of sleep and circadian rhythms. *Sleep, 26*(3), 342-392.

Arndt, J. (2012). The influence of forward and backward associative strength on false recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(3), 747.

Arndt, J. (2015). The influence of forward and backward associative strength on false memories for encoding context. *Memory, 23*(7), 1093-1111.

Arora, T., Broglia, E., Pushpakumar, D., Lodhi, T., & Taheri, S. (2013). An investigation into the strength of the association and agreement levels between subjective and objective sleep duration in adolescents. *PloS one, 8*(8).

Babin, L., Lee, S., Halko, S., Boudreau, A., & George, C. (1997). Determining sleep-wake activity using actiwatch. *Sleep Res, 26*, 640.

Babson, K., & Feldner, M. (2015). *Sleep and affect: assessment, theory, and clinical implications*: Academic Press.

Baddeley, A. D. (2017). The concept of working memory: A view of its current state and probable future development. In *Exploring working memory* (pp. 99-106): Routledge.

Baehr, E. K., Revelle, W., & Eastman, C. I. (2000). Individual differences in the phase and amplitude of the human circadian temperature rhythm: with an emphasis on morningness–eveningness. *Journal of sleep research, 9*(2), 117-127.

Backhaus, J., Hoeckesfeld, R., Born, J., Hohagen, F., & Junghanns, K. (2008). Immediate as well as delayed post learning sleep but not wakefulness enhances declarative memory consolidation in children. *Neurobiology of learning and memory, 89*(1), 76-80

Baekeland, F., & Hoy, P. (1971). Reported vs recorded sleep characteristics. *Archives of General Psychiatry, 24*(6), 548-551.

Bailey, S. L., & Heitkemper, M. M. (2001). Circadian rhythmicity of cortisol and body temperature: morningness-eveningness effects. *Chronobiology international, 18*(2), 249-261.

Baran, B., Pace-Schott, E. F., Ericson, C., & Spencer, R. M. (2012). Processing of emotional reactivity and emotional memory over sleep. *Journal of Neuroscience, 32*(3), 1035-1042.

Baranski, M., & Was, C. A. (2017). Mindfulness meditation may not increase false-memory and may instead protect from false-memory susceptibility. Mindfulness, 8(6), 1569-1579.

Barrett, T. R., & Ekstrand, B. R. (1972). Effect of sleep on memory: III. Controlling for time-of-day effects. *Journal of experimental psychology, 96*(2), 321.

Bauer, L. M., Olheiser, E. L., Altarriba, J., & Landi, N. (2009). Word type effects in false recall: Concrete, abstract, and emotion word critical lures. *The American journal of psychology*, 469-481.

Bauste, G., & Ferraro, F. R. (2004). Gender differences in false memory production. *Current Psychology*, *23*(3), 238-244.

Beato, M. S., & Arndt, J. (2017). The role of backward associative strength in false recognition of DRM lists with multiple critical words. *Psicothema, 29*(3), 358-363.

Blagrove, M., & Akehurst, L. (2000). Effects of sleep loss on confidence–accuracy relationships for reasoning and eyewitness memory. *Journal of Experimental Psychology: Applied, 6(*1), 59.

Bookbinder, & Brainerd, C. (2017). Emotionally negative pictures enhance gist memory. *Emotion, 17*(1), 102.

Bookbinder, H., & Brainerd, C. J. (2016). Emotion and false memory: The context–content paradox. *Psychological Bulletin, 142*(12), 1315.

Borisenkov, M. F., Perminova, E. V., & Kosova, A. L. (2010). Chronotype, sleep length, and school achievement of 11-to 23-year-old students in northern European Russia. Chronobiology international, 27(6), 1259-1270.

Born, J., & Gais, S. (2003). Roles of early and late nocturnal sleep for the consolidation of human memories. *Sleep and brain plasticity*, 65-85.

Born, J., Rasch, B., & Gais, S. (2006). Sleep to remember. *The Neuroscientist, 12*(5), 410-424.

Born, J., & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological research, 76*(2), 192-203.

Brainerd, Holliday, R., Reyna, V., Yang, Y., & Toglia, M. (2010). Developmental reversals in false memory: Effects of emotional valence and arousal. *Journal of Experimental Child Psychology, 107*(2), 137-154.

Brainerd, & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science, 11*(5), 164-169.

Brainerd, Stein, L. M., Silveira, R., Rohenkohl, G., & Reyna, V. F. (2008). How does negative emotion cause false memories? *Psychological Science, 19*(9), 919-925.

Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review, 10*(1), 3-47.

Brainerd, C. J., & Reyna, V. F. (2005). The science of false memory: *Oxford University Press.38,* 361–422.

Bravi, R., Cohen, E. J., Martinelli, A., Gottard, A., & Minciacchi, D. (2017). When non-dominant is better than dominant: Kinesiotape modulates asymmetries in timed performance during a synchronization-continuation task. *Frontiers in integrative neuroscience, 11*, 21.

Brena, R., DeYoung, P., Orr, J., Schmickl, C., Bosompra, N.-O., Gilbertson, D., . . . Owens, R. (2020). Actigraphy vs Polysomnography for Assessing

Sleep in People Living with HIV. In *B69. SRN: diagnosis and monitoring of sleep and sleep disorders* (pp. A4145-A4145): American Thoracic Society.

Brewer, W. F. (2006). Brown and Kulik—theory Circumstances (news reception context) The core phenomenon described by Brown and Kulik (1977) is that certain events give rise to memories that show little forgetting. These flashbulb memories include both the central event and the circumstances in which one. *Affect and accuracy in recall: Studies of'flashbulb'memories, 4*, 274.

Buchanan, T. W., Etzel, J. A., Adolphs, R., & Tranel, D. (2006). The influence of autonomic arousal and semantic relatedness on memory for emotional words. *International journal of psychophysiology, 61*(1), 26-33.

Buchner, A., Rothermund, K., Wentura, D., & Mehl, B. (2004). Valence of distractor words increases the effects of irrelevant speech on serial recall. *Memory & Cognition, 32*(5), 722-731.

Burke, J. F., Long, N. M., Zaghloul, K. A., Sharan, A. D., Sperling, M. R., & Kahana, M. J. (2014). Human intracranial high-frequency activity maps episodic memory formation in space and time. *Neuroimage, 85*, 834-843.

Buysse, D. J., Ancoli-Israel, S., Edinger, J. D., Lichstein, K. L., & Morin, C. M. (2006). Recommendations for a standard research assessment of insomnia. *Sleep, 29*(9), 1155-1173.

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 res, 28*(2), 193-213.

BuzsÁk, G. (1998). Memory consolidation during sleep: a neurophysiological perspective. *Journal of sleep research, 7*(S1), 17-23.

Cairney, S. A., Durrant, S. J., Jackson, R., & Lewis, P. A. (2014). Sleep spindles provide indirect support to the consolidation of emotional encoding contexts. *Neuropsychologia, 63*, 285-292.

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*(6), 1565-1575.

Calado, B., Otgaar, H., & Muris, P. (2018). Are children better witnesses than adolescents? Developmental trends in different false memory paradigms. *Journal of Child Custody, 15*(4), 330-348.

Calvillo, D. P., & Parong, J. A. (2016). The misinformation effect is unrelated to the DRM effect with and without a DRM warning. *Memory, 24*(3), 324-333.

Campanini, M. Z., Lopez-Garcia, E., Rodríguez-Artalejo, F., González, A. D., Andrade, S. M., & Mesas, A. E. (2017). Agreement between sleep diary and actigraphy in a highly educated Brazilian population. *Sleep medicine, 35*, 27-34.

Capek, S., & Guenther, R. K. (2009). Caffeine's effects on true and false memory. *Psychological reports, 104*(3), 787-795.

Carney, C. E., Buysse, D. J., Ancoli-Israel, S., Edinger, J. D., Krystal, A. D., Lichstein, K. L., & Morin, C. M. (2012). The consensus sleep diary: standardizing prospective sleep self-monitoring. *Sleep, 35*(2), 287-302.

Carr, M., & Nielsen, T. (2015). Morning REM sleep naps facilitate broad access to emotional semantic networks. *Sleep, 38*(3), 433-443.

Carrier, A. M., Humphreys, R. W., & Petersen, P. M. (1997). Detectable water-treatment polymers and methods for monitoring the concentration thereof. In: Google Patents. *Biotechnology Advances*, *15*(1), 301.

Carskadon, M. A., Dement, W. C., Mitler, M., Guilleminault, C., Zarcone, V. P., & Spiegel, R. (1976). Self-reports versus sleep laboratory findings in 122 drug-free subjects with complaints of chronic insomnia. *Am J Psychiatry, 133*(12), 1382-1388.

Cerin, E., Cain, K. L., Oyeyemi, A. L., Owen, N., Conway, T. L., Cochrane, T., . . . Toftager, M. (2016). Correlates of agreement between accelerometry and self-reported physical activity. *Medicine and science in sports and exercise, 48*(6), 1075.

Coane, J. H., McBride, D. M., Raulerson III, B. A., & Jordan, J. S. (2007). False memory in a short-term memory task. *Experimental Psychology, 54*(1), 62-70.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review, 82*(6), 407.

Corson, Y., & Verrier, N. (2007). Emotions and false memories: Valence or arousal? *Psychological Science, 18*(3), 208-211.

Currie, S. R., Malhotra, S., & Clark, S. (2004). Agreement among subjective, objective, and collateral measures of insomnia in postwithdrawal recovering alcoholics. *Behavioral Sleep Medicine, 2*(3), 148-161.

Dahl, R. E., & Lewin, D. S. (2002). Pathways to adolescent health sleep regulation and behavior. *Journal of adolescent health, 31*(6), 175-184.

Daniela, T., Alessandro, C., Giuseppe, C., Fabio, M., Cristina, M., & Michele, F. (2010). Lack of sleep affects the evaluation of emotional stimuli. *Brain research bulletin, 82*(1-2), 104-108.

Darsaud, A., Dehon, H., Lahl, O., Sterpenich, V., Boly, M., Dang-Vu, T., . . . Peters, F. (2011). Does sleep promote false memories? *Journal of Cognitive Neuroscience, 23*(1), 26-40.

Davis, M. H., Di Betta, A. M., Macdonald, M. J., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of cognitive neuroscience, 21*(4), 803-820.

de Souza, L., Benedito-Silva, A. A., Pires, M. L. N., Poyares, D., Tufik, S., & Calil, H. M. (2003). Further validation of actigraphy for sleep studies. *Sleep, 26*(1), 81-85.

Deese, J. (1959). Influence of inter-item associative strength upon immediate free recall. *Psychological Reports, 5*(3), 305-312.

Dehon, H., Larøi, F., & Van der Linden, M. (2010). Affective valence influences participant's susceptibility to false memories and illusory recollection. *Emotion, 10*(5), 627.

Dementienko, V. V., Maev, R. G., Bugaev, A. S., Indursky, P., Markelov, V. V., Shakhnarovich, V. M., & Strumban, E. E. (2018). Method and apparatus for improving slow wave sleep stage. In: Google Patents.

Dewhurst, S. A., Anderson, R. J., & Knott, L. M. (2012). A gender difference in the false recall of negative words: Women DRM more than men. *Cognition & emotion*, *26*(1), 65-74.

Dewhurst, S. A., & Parry, L. A. (2000). Emotionality, distinctiveness, and recollective experience. *European Journal of Cognitive Psychology, 12*(4), 541-551.

Díaz-Morales, J. F., & Escribano, C. (2013). Predicting school achievement: The role of inductive reasoning, sleep length and morningness–eveningness. *Personality and Individual Differences, 55*(2), 106-111.

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience, 11*(2), 114-126.

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. *Behavioural brain research, 208*(2), 425-429.

Diekelmann, S., Landolt, H.-P., Lahl, O., Born, J., & Wagner, U. (2008). Sleep loss produces false memories. *PloS one, 3*(10).

Diekelmann, S., Wilhelm, I., & Born, J. (2009). The whats and whens of sleep-dependent memory consolidation. *Sleep medicine reviews, 13*(5), 309-321.

Domes, G., Heinrichs, M., Rimmele, U., Reichwald, U., & Hautzinger, M. (2004). Acute stress impairs recognition for positive words—association with stress-induced cortisol secretion. *Stress, 7*(3), 173-181.

Driller, M. W., O'Donnell, S., & Tavares, F. (2017). What wrist should you wear your actigraphy device on? Analysis of dominant vs. non-dominant wrist actigraphy for measuring sleep in healthy adults. *Sleep Science, 10*(3), 132.

Drosopoulos, S., Wagner, U., & Born, J. (2005). Sleep enhances explicit recollection in recognition memory. *Learning & Memory, 12*(1), 44-51.

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science, 18*(1), 35-39.

Durrant, S. J., Cairney, S. A., McDermott, C., & Lewis, P. A. (2015). Schema-conformant memories are preferentially consolidated during REM sleep. *Neurobiology of Learning and Memory, 122*, 41-50.

Edinger, J. D., & Fins, A. I. (1995). The distribution and clinical significance of sleep time misperceptions among insomniacs. *Sleep, 18*(4), 232-239.

Elbaz, M., Yauy, K., Metlaine, A., Martoni, M., & Leger, D. (2012). Validation of a new actigraph motion watch versus polysomnography on 70 healthy and suspected sleep-disordered subjects. *J Sleep Res, 21*(Suppl 1), 218.

Ellenbogen, J. M., Hulbert, J. C., Stickgold, R., Dinges, D. F., & Thompson-Schill, S. L. (2006). Interfering with theories of sleep and memory: sleep, declarative memory, and associative interference. *Current Biology, 16*(13), 1290-1294.

Empson, J., & Clarke, P. (1970). Rapid eye movements and remembering. *Nature, 227*(5255), 287-288.

Evans, B., Mason, K., & Buroker, O. (2018). Understanding the Effects of Meditation on False Memories Using the DRM Paradigm.

Fafrowicz, M., Bohaterewicz, B., Ceglarek, A., Cichocka, M., Lewandowska, K., Sikora-Wachowicz, B., . . . Marek, T. (2019). Beyond the low frequency fluctuations: morning and evening differences in human brain. *Frontiers in Human Neuroscience, 13*, 288.

Falck, R. S., Best, J. R., Davis, J. C., Eng, J. J., Middleton, L. E., Hall, P. A., & Liu-Ambrose, T. (2019). Sleep and cognitive function in chronic stroke: a comparative cross-sectional study. *Sleep, 42*(5), 040.

Falck, R. S., Best, J. R., Davis, J. C., & Liu-Ambrose, T. (2018). The independent associations of physical activity and sleep with cognitive function in older adults. *Journal of Alzheimer's Disease, 63*(4), 1469-1484.

Fenn, K. M., Gallo, D. A., Margoliash, D., Roediger, H. L., & Nusbaum, H. C. (2009). Reduced false memory after sleep. *Learning & Memory, 16*(9), 509-513.

Ficca, G., Lombardo, P., Rossi, L., & Salzarulo, P. (2000). Morning recall of verbal material depends on prior sleep organization. *Behavioural brain research, 112*(1-2), 159-163.

Ficca, G., & Salzarulo, P. (2004). What in sleep is for memory. *Sleep medicine, 5*(3), 225-230.

Fogel, S. M., Smith, C. T., & Cote, K. A. (2007). Dissociable learning-dependent changes in REM and non-REM sleep in declarative and procedural memory systems. *Behavioural brain research, 180*(1), 48-61.

Fonseca, P., Weysen, T., Goelema, M. S., Møst, E. I., Radha, M., Lunsingh Scheurleer, C., . . . Aarts, R. M. (2017). Validation of photoplethysmography-based sleep staging compared with polysomnography in healthy middle-aged adults. *Sleep, 40*(7), 097.

Franzen, P. L., Siegle, G. J., & Buysse, D. J. (2008). Relationships between affect, vigilance, and sleepiness following sleep deprivation. *Journal of sleep research, 17*(1), 34-41.

Gais, S., & Born, J. (2004). Declarative memory consolidation: mechanisms acting during human sleep. *Learning & Memory, 11*(6), 679-685.

Gais, S., Lucas, B., & Born, J. (2006). Sleep after learning aids memory recall. *Learning & memory, 13*(3), 259-262.

Gais, S., Plihal, W., Wagner, U., & Born, J. (2000). Early sleep triggers memory for early visual discrimination skills. *Nature neuroscience, 3*(12), 1335-1339.

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition, 38*(7), 833-848.

Gallo, D. A., Foster, K. T., & Johnson, E. L. (2009). Elevated false recollection of emotional pictures in young and older adults. *Psychology and aging, 24*(4), 981.

Gallo, D. A., Foster, K. T., Wong, J. T., & Bennett, D. A. (2010). False recollection of emotional pictures in Alzheimer's disease. *Neuropsychologia, 48*(12), 3614-3618.

Gallo, D. A., & Roediger, H. L. (2003). The effects of associations and aging on illusory recollection. *Memory & Cognition, 31*(7), 1036-1044.

Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic bulletin & review*, *19*(5), 847-857.

Giannotti, F., Cortesi, F., Sebastiani, T., & Ottaviano, S. (2002). Circadian preference, sleep and daytime behaviour in adolescence. *Journal of sleep research, 11*(3), 191-199.

Girschik, J., Fritschi, L., Heyworth, J., & Waters, F. (2012). Validation of self-reported sleep against actigraphy. *Journal of epidemiology, 22*(5), 462-468.

Giuditta, A., Ambrosini, M. V., Montagnese, P., Mandile, P., Cotugno, M., Zucconi, G. G., & Vescia, S. (1995). The sequential hypothesis of the function of sleep. *Behavioural brain research, 69*(1-2), 157-166.

Gomes, C. F., Brainerd, C. J., & Stein, L. M. (2013). Effects of emotional valence and arousal on recollective and nonrecollective recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 663.

Gonzalez, R., Tamminga, C., Tohen, M., & Suppes, T. (2013). Comparison of objective and subjective assessments of sleep time in subjects with bipolar disorder. *Journal of affective disorders, 149*(1-3), 363-366.

Goodman, G. S., Ogle, C. M., Block, S. D., Harris, L. S., Larson, R. P., Augusti, E.-M., . . . Urquiza, A. (2011). False memory for trauma-related Deese–Roediger–McDermott lists in adolescents and adults with histories of child sexual abuse. *Development and Psychopathology, 23*(2), 423-438.

Grandner, M. A., Kripke, D. F., Yoon, I.-Y., & Youngstedt, S. D. (2006). Criterion validity of the Pittsburgh Sleep Quality Index: Investigation in a non-clinical sample. *Sleep and biological rhythms, 4*(2), 129-136.

Green, D. M., & Swets, J. A. (1966). Signal detection theory and psychophysics (Vol. 1). New York: Wiley.

Groch, S., Wilhelm, I., Diekelmann, S., & Born, J. (2013). The role of REM sleep in the processing of emotional memories: evidence from behavior and event-related potentials. *Neurobiology of learning and memory, 99*, 1-9.

Groch, S., Zinke, K., Wilhelm, I., & Born, J. (2015). Dissociating the contributions of slow-wave sleep and rapid eye movement sleep to emotional item and source memory. *Neurobiology of Learning and Memory, 122*, 122-130.

Gujar, N., Yoo, S.-S., Hu, P., & Walker, M. P. (2011). Sleep deprivation amplifies reactivity of brain reward networks, biasing the appraisal of positive emotional experiences. *Journal of Neuroscience, 31*(12), 4466-4474.

Halász, P., Bódizs, R., Parrino, L., & Terzano, M. (2014). Two features of sleep slow waves: homeostatic and reactive aspects–from long term to instant sleep homeostasis. *Sleep medicine, 15*(10), 1184-1195.

Hansen, C. H., & Hansen, R. D. (1994). Automatic emotion: Attention and facial efference. In: Niedenthal PM, Kitayama S, editors. *The Heart's Eye: Emotional Influences in Perception and Attention,* 217–43.

Hardt, O., Nader, K., & Nadel, L. (2013). Decay happens: the role of active forgetting in memory. *Trends in cognitive sciences, 17*(3), 111-120.

Hege, A. C., & Dodson, C. S. (2004). Why distinctive information reduces false memories: evidence for both impoverished relational-encoding and distinctiveness heuristic accounts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(4), 787.

Hellenthal, M. V., Knott, L. M., Howe, M. L., Wilkinson, S., & Shah, D. (2019) The effects of arousal and attention on emotional false memory formation. *Journal of Memory and Language*, *107*, 54–68.

Hicks, J. L., & Hancock, T. W. (2002). Backward associative strength determines source attributions given to false memories. *Psychonomic Bulletin & Review, 9*(4), 807-815.

Hicks, J. L., & Marsh, R. L. (1999). Attempts to reduce the incidence of false recall with source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(5), 1195.

Hong, H., Zhang, L., Zhao, H., Chu, H., Gu, C., Brown, M., . . . Li, C. (2019). Microwave sensing and sleep: Noncontact sleep-monitoring technology with microwave biomedical radar. *IEEE Microwave Magazine, 20*(8), 18-29.

Horne, J. A., & Östberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International journal of chronobiology*. *4,* 97–110.

Howe, M. L. (2006). Developmentally invariant dissociations in children's true and false memories: Not all relatedness is created equal. *Child Development, 77*(4), 1112-1123.

Howe, M. L., & Akhtar, S. (2020). Priming older adults and people with mild to moderate Alzheimer's disease problem-solving with false memories. *Cortex, 125*, 318-331.

Howe, M. L., Candel, I., Otgaar, H., Malone, C., & Wimmer, M. C. (2010). Valence and the development of immediate and long-term false memory illusions. *Memory, 18*(1), 58-75.

Howe, M. L., Wimmer, M. C., Gagnon, N., & Plumpton, S. (2009). An associative-activation theory of children's and adults' memory illusions. *Journal of Memory and Language, 60*(2), 229-251.

Hu, P., Stylos-Allan, M., & Walker, M. P. (2006). Sleep facilitates consolidation of emotional declarative memory. *Psychological science, 17*(10), 891-898.

Huber, R., Ghilardi, M. F., Massimini, M., & Tononi, G. (2004). Local sleep and learning. *Nature, 430*(6995), 78-81.

Hutchison, K. A., & Balota, D. A. (2005). Decoupling semantic and associative information in false memories: Explorations with semantically ambiguous and unambiguous critical lures. *Journal of Memory and Language, 52*(1), 1-28.

Huth, A. G., De Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature, 532*(7600), 453-458.

Ibáñez, V., Silva, J., & Cauli, O. (2018). A survey on sleep assessment methods. *PeerJ, 6*, e4849.

Inaba, M., Nomura, M., & Ohira, H. (2005). Neural evidence of effects of emotional valence on word recognition. *International journal of psychophysiology, 57*(3), 165-173.

Inostroza, M., & Born, J. (2013). Sleep for preserving and transforming episodic memory. *Annual review of neuroscience, 36*, 79-102.

Israel, L., & Schacter, D. L. (1997). Pictorial encoding reduces false recognition of semantic associates. *Psychonomic Bulletin & Review, 4*(4), 577-581.

Izawa, C., & Ohta, N. (2014). *Human Learning and Memory: Advances in Theory and Applications: The 4th Tsukuba International Conference on Memory*: Psychology Press.

Janiszewska, A. (2014). From Liberal to Conservative: Effects on False Memories and Response Bias by Emotion and EMDR in DRM Paradigm.

Jenkins, J. G., & Dallenbach, K. M. (1924). Obliviscence during sleep and waking. *The American Journal of Psychology, 35*(4), 605-612.

Jinmei, L., Rong, H., Xu, Z., Yi, X., & Jiong, Z. (2014). STOP-Bang questionnaire is superior to Epworth sleepiness scales, Berlin questionnaire, and STOP questionnaire in screening obstructive sleep apnea hypopnea syndrome patients. *Chinese medical journal, 127*(17), 3065-3070.

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

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological bulletin, 114*(1), 3.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological review, 88*(1), 67.

Kahn, M., Sheppes, G., & Sadeh, A. (2013). Sleep and emotions: bidirectional links and underlying mechanisms. *International Journal of Psychophysiology, 89*(2), 218-228.

Kanady, J. C., Drummond, S. P., & Mednick, S. C. (2011). Actigraphic assessment of a polysomnographic-recorded nap: A validation study. *Journal of sleep research, 20*(1pt2), 214-222.

Kawada, T. (2015). Validation study on sleep parameters by actigraphy for normal subjects. *Sleep and Breathing, 19*(1), 11-11.

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences, 101*(9), 3310-3315.

Kerkhof, G. A., & Van Dongen, H. (2011). *Human sleep and cognition: Basic research*: Elsevier.

Khanna, M. M., & Cortese, M. J. (2009). Children and adults are differentially affected by presentation modality in the DRM paradigm. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 23*(6), 859-877.

Kilman, L., Zekveld, A. A., Hällgren, M., & Rönnberg, J. (2015). Episodic long-term memory by native and non-native stories masked by speech. *Frontiers In Psychology*, *38*(7), 902-922.

King, B. R., Saucier, P., Albouy, G., Fogel, S. M., Rumpf, J.-J., Klann, J., . . . Karni, A. (2017). Cerebral activation during initial motor learning forecasts subsequent sleep-facilitated memory consolidation in older adults. *Cerebral cortex, 27*(2), 1588-1601.

Klinzing, J. G., Niethard, N., & Born, J. (2019). Mechanisms of systems memory consolidation during sleep. *Nature neuroscience, 22*(10), 1598-1610.

Knott, L. M., Dewhurst, S. A., & Howe, M. L. (2012). What factors underlie associative and categorical memory illusions? The roles of backward associative strength and interitem connectivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(1), 229.

Knott, L. M., & Thorley, C. (2014). Mood-congruent false memories persist over time. *Cognition & Emotion, 28*(5), 903-912.

Knutson, K. L., Rathouz, P. J., Yan, L. L., Liu, K., & Lauderdale, D. S. (2006). Stability of the Pittsburgh Sleep Quality Index and the Epworth Sleepiness Questionnaires over 1 year in early middle-aged adults: the CARDIA study. *Sleep, 29*(11), 1503-1506.

Knutson, K. L., Rathouz, P. J., Yan, L. L., Liu, K., & Lauderdale, D. S. (2007). Intra-individual daily and yearly variability in actigraphically recorded sleep measures: the CARDIA study. *Sleep, 30*(6), 793-796.

Kosmadopoulos, A., Sargent, C., Darwent, D., Zhou, X., & Roach, G. D. (2014). Alternatives to polysomnography (PSG): a validation of wrist actigraphy and a partial-PSG system. *Behavior research methods, 46*(4), 1032-1041.

Kreiner, D. S., Price, R. Z., Gross, A. M., & Appleby, K. L. (2004). False recall does not increase when words are presented in a gender-congruent voice. *Journal of Articles in Support of the Null Hypothesis*, *3*(1), 1-11.

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental Psychology: General, 143*(3), 1065.

Kushida, C. A., Chang, A., Gadkary, C., Guilleminault, C., Carrillo, O., & Dement, W. C. (2001). Comparison of actigraphic, polysomnographic, and subjective assessment of sleep parameters in sleep-disordered patients. *Sleep medicine, 2*(5), 389-396.

Kölling, S., Endler, S., Ferrauti, A., Meyer, T., & Kellmann, M. (2016). Comparing subjective with objective sleep parameters via multisensory actigraphy in German physical education students. *Behavioral Sleep Medicine, 14*(4), 389-405.

Laberge, L., Carrier, J., Lesperance, P., Lambert, C., Vitaro, F., Tremblay, R., & Montplaisir, J. (2000). Sleep and circadian phase characteristics of adolescent and young adult males in a naturalistic summertime condition. *Chronobiology international, 17*(4), 489-501.

Landmann, N., Kuhn, M., Piosczyk, H., Feige, B., Baglioni, C., Spiegelhalder, K., . . . Nissen, C. (2014). The reorganisation of memory during sleep. *Sleep medicine reviews, 18*(6), 531-541.

Landry, G. J., Best, J. R., & Liu-Ambrose, T. (2015). Measuring sleep quality in older adults: a comparison using subjective and objective methods. *Frontiers in aging neuroscience, 7*, 166.

Lauderdale, D. S., Knutson, K. L., Yan, L. L., Liu, K., & Rathouz, P. J. (2008). Sleep duration: how well do self-reports reflect objective measures? The CARDIA Sleep Study. *Epidemiology (Cambridge, Mass.), 19*(6), 838.

Lawrence, G., & Muza, R. (2018). Assessing the sleeping habits of patients in a sleep disorder centre: a review of sleep diary accuracy. *Journal of thoracic disease, 10*(Suppl 1), S177.

Leduc, C., Jones, B., Robineau, J., Piscione, J., & Lacome, M. (2019). Sleep quality and quantity of international rugby sevens players during pre-season. *The Journal of Strength & Conditioning Research, 33*(7), 1878-1886.

Lehmann, M., Seifritz, E., & Rasch, B. (2016). Sleep benefits emotional and neutral associative memories equally. *Somnologie, 20*(1), 47-53.

Lewis, P. A., Cairney, S., Manning, L., & Critchley, H. D. (2011). The impact of overnight consolidation upon memory for emotional and neutral encoding contexts. *Neuropsychologia, 49*(9), 2619-2629.

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in cognitive sciences, 15*(8), 343-351.

Libby, L. K., & Neisser, U. (2001). Structure and strategy in the associative false memory paradigm. *Memory, 9*(3), 145-163.

Lichstein, K. L., Stone, K. C., Donaldson, J., Nau, S. D., Soeffing, J. P., Murray, D., . . . Aguillard, R. N. (2006). Actigraphy validation with insomnia. *Sleep, 29*(2), 232-239.

Light, L. L. (2015). Long-Term Memory. *The Encyclopedia of Adulthood and Aging*, 1-6.

Lo, J. C., Sim, S. K., & Chee, M. W. (2014). Sleep reduces false memory in healthy older adults. *Sleep, 37*(4), 665-671.

Lockley, S. W., Skene, D. J., & Arendt, J. (1999). Comparison between subjective and actigraphic measurement of sleep and sleep rhythms. *Journal of sleep research, 8*(3), 175-183.

Lyon, D. R. (1977). Individual differences in immediate serial recall: A matter of mnemonics? *Cognitive Psychology, 9*(4), 403-411.

Maes, J., Verbraecken, J., Willemen, M., De Volder, I., Van Gastel, A., Michiels, N., . . . Haex, B. (2014). Sleep misperception, EEG characteristics and autonomic nervous system activity in primary insomnia: a retrospective study on polysomnographic data. *International Journal of Psychophysiology, 91*(3), 163-171.

Maich, K. H., Lachowski, A. M., & Carney, C. E. (2018). Psychometric properties of the consensus sleep diary in those with insomnia disorder. *Behavioral sleep medicine, 16*(2), 117-134.

Mallinson, D. C., Kamenetsky, M. E., Hagen, E. W., & Peppard, P. E. (2019). Subjective sleep measurement: comparing sleep diary to questionnaire. *Nature and science of sleep, 11*, 197.

Mantua, J., Gravel, N., & Spencer, R. (2016). Reliability of sleep measures from four personal health monitoring devices compared to research-based actigraphy and polysomnography. *Sensors, 16*(5), 646.

Marino, M., Li, Y., Rueschman, M. N., Winkelman, J. W., Ellenbogen, J., Solet, J. M., . . . Buxton, O. M. (2013). Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep, 36*(11), 1747-1755.

Marsh, E., & Bower, G. (2004). The role of rehearsal and generation in false memory creation. *Memory, 12*(6), 748-761.

Marshall, L., & Born, J. (2007). The contribution of sleep to hippocampus-dependent memory consolidation. *Trends in cognitive sciences, 11*(10), 442-450.

McCrae, C. S., Rowe, M. A., Tierney, C. G., Dautovich, N. D., DeFinis, A. L., & McNamara, J. P. (2005). Sleep complaints, subjective and objective sleep patterns, health, psychological adjustment, and daytime functioning in community-dwelling older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 60*(4), P182-P189.

McEvoy, C. L., Nelson, D. L., & Komatsu, T. (1999). What is the connection between true and false memories? The differential roles of interitem associations in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(5), 1177.

McGaugh, J. L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annu. Rev. Neurosci., 27*, 1-28.

McKeon, S., Pace-Schott, E. F., & Spencer, R. M. (2012). Interaction of sleep and emotional content on the production of false memories. *PloS one, 7*(11).

Meddis, R. (2018). The evolution and function of sleep. In *Brain, Behaviour and Evolution* (pp. 99-125): Routledge.

Mednick, S., Nakayama, K., & Stickgold, R. (2003). Sleep-dependent learning: a nap is as good as a night. *Nature neuroscience, 6*(7), 697-698.

Metzger, R. L., Warren, A. R., Shelton, J. T., Price, J., Reed, A. W., & Williams, D. (2008). Do children" DRM" like adults? False memory production in children. *Developmental Psychology, 44*(1), 169.

Middelkoop, H. A., Van Dam, E. M., SMILDE-VAN DEN DOEL, D. A., & VAN DIJK, G. (1997). 45-hour continuous quintuple-site actimetry: relations between trunk and limb movements and effects of circadian sleep-wake rhythmicity. *Psychophysiology, 34*(2), 199-203.

Miller, C. B., Kyle, S. D., Melehan, K. L., & Bartlett, D. J. (2015). Methodology for the assessment of sleep. In *Sleep and Affect* (pp. 65-90): Elsevier.

Minkel, J., Htaik, O., Banks, S., & Dinges, D. (2011). Emotional expressiveness in sleep-deprived healthy adults. *Behavioral sleep medicine, 9*(1), 5-14.

Mollayeva, T., Thurairajah, P., Burton, K., Mollayeva, S., Shapiro, C. M., & Colantonio, A. (2016). The Pittsburgh sleep quality index as a screening tool for sleep dysfunction in clinical and non-clinical samples: A systematic review and meta-analysis. *Sleep medicine reviews, 25*, 52-73.

Monaghan, P., Shaw, J. J., Ashworth-Lord, A., & Newbury, C. R. (2017). Hemispheric processing of memory is affected by sleep. *Brain and language, 167*, 36-43.

Moscovitch, M., Cabeza, R., Winocur, G., & Nadel, L. (2016). Episodic memory and beyond: the hippocampus and neocortex in transformation. *Annual review of psychology, 67*, 105-134.

Murdock Jr, B. B. (1961). The retention of individual items. *Journal of experimental psychology, 62*(6), 618.

Natale, V., Grandi, C. A., Fabbri, M., Tonetti, L., Martoni, M., & Esposito, M. J. (2006). Additional validity evidence for the reduced version of the Morningness-Eveningness Questionnaire (MEQr). *Sleep and Hypnosis, 8*(2), 47.

Natale, V., Martoni, M., Esposito, M. J., Fabbri, M., & Tonetti, L. (2007). Circadian motor asymmetries before and after prolonged wakefulness in humans. *Neuroscience letters, 423*(3), 216-218.

Nelson, D. L., McKinney, V. M., Gee, N. R., & Janczura, G. A. (1998). Interpreting the influence of implicitly activated memories on recall and recognition. *Psychological review, 105*(2), 299.

Newbury, C. R., & Monaghan, P. (2019). When does sleep affect veridical and false memory consolidation? A meta-analysis. *Psychonomic bulletin & review, 26*(2), 387-400.

Nishida, M., Pearsall, J., Buckner, R. L., & Walker, M. P. (2009). REM sleep, prefrontal theta, and the consolidation of human emotional memory. *Cerebral cortex, 19*(5), 1158-1166.

Njamnshi, A. K., Etet, P. F. S., Perrig, S., Acho, A., Funsah, J. Y., Mumba, D., . . . Bentivoglio, M. (2012). Actigraphy in human African trypanosomiasis as a tool for objective clinical evaluation and monitoring: a pilot study. *PLoS neglected tropical diseases, 6*(2).

Norman, D. A., & Shallice, T. (1986). Attention to action. In *Consciousness and self-regulation* (pp. 1-18): Springer.

Norris, D. (2017). Short-term memory and long-term memory are still different. *Psychological bulletin, 143*(9), 992.

Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D., & Gross, J. J. (2004). For better or for worse: neural systems supporting the cognitive down-and up-regulation of negative emotion. *Neuroimage, 23*(2), 483-499.

Okado, Y., & Stark, C. (2003). Neural processing associated with true and false memory retrieval. *Cognitive, Affective, & Behavioral Neuroscience, 3*(4), 323-334.

Oliver, M. C., Bays, R. B., & Zabrucky, K. M. (2016). False memories and the DRM paradigm: Effects of imagery, list, and test type. *The Journal of General Psychology, 143*(1), 33-48.

Otgaar, H., Moldoveanu, G., Wang, J., & Howe, M. L. (2017). Exploring the consequences of nonbelieved memories in the DRM paradigm. *Memory, 25*(7), 922-933.

Palmer, J. E., & Dodson, C. S. (2009). Investigating the mechanisms fuelling reduced false recall of emotional material. *Cognition and Emotion, 23*(2), 238-259.

Pardilla-Delgado, E., Alger, S. E., Cunningham, T. J., Kinealy, B., & Payne, J. D. (2016). Effects of post-encoding stress on performance in the DRM false memory paradigm. *Learning & memory, 23*(1), 46-50.

Pardilla-Delgado, E., & Payne, J. D. (2017). The impact of sleep on true and false memory across long delays. *Neurobiology of learning and memory, 137*, 123-133.

Pataka, A., Daskalopoulou, E., Kalamaras, G., Passa, K. F., & Argyropoulou, P. (2014). Evaluation of five different questionnaires for assessing sleep apnea syndrome in a sleep clinic. *Sleep medicine, 15*(7), 776-781.

Payne, J., Chambers, A. M., & Kensinger, E. A. (2012). Sleep promotes lasting changes in selective memory for emotional scenes. *Frontiers in integrative neuroscience, 6*, 108.

Payne, J. D., & Kensinger, E. A. (2010). Sleep's role in the consolidation of emotional episodic memories. *Current Directions in Psychological Science, 19*(5), 290-295.

Payne, J. D., Kensinger, E. A., Wamsley, E. J., Spreng, R. N., Alger, S. E., Gibler, K., . . . Stickgold, R. (2015). Napping and the selective consolidation of negative aspects of scenes. *Emotion, 15*(2), 176.

Payne, J. D., Schacter, D. L., Propper, R. E., Huang, L.-W., Wamsley, E. J., Tucker, M. A., . . . Stickgold, R. (2009). The role of sleep in false memory formation. *Neurobiology of learning and memory, 92*(3), 327-334.

Payne, J. D., Stickgold, R., Swanberg, K., & Kensinger, E. A. (2008). Sleep preferentially enhances memory for emotional components of scenes. *Psychological Science, 19*(8), 781-788.

Peigneux, P., Laureys, S., Delbeuck, X., & Maquet, P. (2001). Sleeping brain, learning brain. The role of sleep for memory systems. *Neuroreport, 12*(18), A111-A124.

Peltz, J. S., Rogge, R. D., & Sturge-Apple, M. L. (2018). Transactions within the family: Coparenting mediates associations between parents' relationship satisfaction and the parent–child relationship. *Journal of Family Psychology, 32*(5), 553.

Pesta, B. J., Murphy, M. D., & Sanders, R. E. (2001). Are emotionally charged lures immune to false memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(2), 328.

Pixton, T. S. (2011). Happy to see me, aren't you, Sally? Signal detection analysis of emotion detection in briefly presented male and female faces. *Scandinavian journal of psychology*, *52*(4), 361-368.

Plihal, W., & Born, J. (1997). Effects of early and late nocturnal sleep on declarative and procedural memory. *Journal of cognitive neuroscience, 9*(4), 534-547.

Plihal, W., & Born, J. (1999). Effects of early and late nocturnal sleep on priming and spatial memory. *Psychophysiology, 36*(5), 571-582.

Popa, D., Duvarci, S., Popescu, A. T., Léna, C., & Paré, D. (2010). Coherent amygdalocortical theta promotes fear memory consolidation during paradoxical sleep. *Proceedings of the National Academy of Sciences, 107*(14), 6516-6519.

Porter, S., Spencer, L., & Birt, A. R. (2003). Blinded by emotion? Effect of the emotionality of a scene on susceptibility to false memories. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 35*(3), 165.

Prehn-Kristensen, A., Munz, M., Molzow, I., Wilhelm, I., Wiesner, C. D., & Baving, L. (2013). Sleep promotes consolidation of emotional memory in healthy children but not in children with attention-deficit hyperactivity disorder. *PLoS One, 8*(5), e65098.

Radvansky, G. A. (2017). Event segmentation as a working memory process. *Journal of Applied Research in Memory and Cognition*, *6*(2), 121-123.

Rahafar, A., Maghsudloo, M., Farhangnia, S., Vollmer, C., & Randler, C. (2016). The role of chronotype, gender, test anxiety, and conscientiousness in academic achievement of high school students. *Chronobiology International, 33*(1), 1-9.

Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological reviews, 93*(2), 681-766.

Rasch, B., Büchel, C., Gais, S., & Born, J. (2007). Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science, 315*(5817), 1426-1429.

Rauchs, G., Bertran, F., Guillery-Girard, B., Desgranges, B., Kerrouche, N., Denise, P., . . . Eustache, F. (2004). Consolidation of strictly episodic memories mainly requires rapid eye movement sleep. *Sleep, 27*(3), 395-401.

Rauchs, G., Desgranges, B., Foret, J., & Eustache, F. (2005). The relationships between memory systems and sleep stages. *Journal of sleep research, 14*(2), 123-140.

Reyna, V. F. (2008). A theory of medical decision making and health: fuzzy trace theory. *Medical decision making, 28*(6), 850-865.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and individual Differences, 7*(1), 1-75.

Robinson, K. J., & Roediger III, H. L. (1997). Associative processes in false recall and false recognition. *Psychological Science, 8*(3), 231-237.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of experimental psychology: Learning, Memory, and Cognition, 21*(4), 803.

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic bulletin & review, 8*(3), 385-407.

Roediger III, H. L., Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories. The nature of remembering: Essays in honor of Robert G. Cr*owder* (pp. 95–115). Washington, DC: American Psychological Association.

Roeser, K., Schlarb, A. A., & Kübler, A. (2013). The Chronotype-Academic Performance Model (CAM): Daytime sleepiness and learning motivation link chronotype and school performance in adolescents. *Personality and Individual Differences, 54*(7), 836-840.

Rogers, A. E., Caruso, C. C., & Aldrich, M. S. (1993). Reliability of sleep diaries for assessment of sleep/wake patterns. *Nursing research*.

Rosenstreich, E. (2016). Mindfulness and false-memories: The impact of mindfulness practice on the DRM paradigm. The Journal of psychology, 150(1), 58-71.

Roveda, E., Bruno, E., Galasso, L., Mulè, A., Castelli, L., Villarini, A., . . . Pasanisi, P. (2019). Rest-activity circadian rhythm in breast cancer survivors at 5 years after the primary diagnosis. *Chronobiology international, 36*(8), 1156-1165.

Ruci, L., Tomes, J. L., & Zelenski, J. M. (2009). Mood-congruent false memories in the DRM paradigm. *Cognition and Emotion, 23*(6), 1153-1165.

Rupp, T. L., & Balkin, T. J. (2011). Comparison of Motionlogger Watch and Actiwatch actigraphs to polysomnography for sleep/wake estimation in healthy young adults. *Behavior research methods, 43*(4), 1152-1160.

Sadeh, A., & Acebo, C. (2002). The role of actigraphy in sleep medicine. *Sleep medicine reviews, 6*(2), 113-124.

Sadeh, A., Dan, O., & Bar-Haim, Y. (2011). Online assessment of sustained attention following sleep restriction. *Sleep medicine, 12*(3), 257-261.

Sadeh, A., Hauri, P. J., Kripke, D. F., & Lavie, P. (1995). The role of actigraphy in the evaluation of sleep disorders. *Sleep, 18*(4), 288-302.

Sadeh, A., Raviv, A., & Gruber, R. (2000). Sleep patterns and sleep disruptions in school-age children. *Developmental psychology, 36*(3), 291.

Sadeh, A., Sharkey, M., & Carskadon, M. A. (1994). Activity-based sleep-wake identification: an empirical test of methodological issues. *Sleep, 17*(3), 201-207.

Salvato, G., Scarpa, P., Francione, S., Mai, R., Tassi, L., Scarano, E., . . . Bottini, G. (2016). Declarative long-term memory and the mesial temporal lobe: Insights from a 5-year postsurgery follow-up study on refractory temporal lobe epilepsy. *Epilepsy & Behavior, 64*, 102-109.

Schacter, D. L., Kagan, J., & Leichtman, M. D. (1995). True and false memories in children and adults: A cognitive neuroscience perspective. *Psychology, Public Policy, and Law, 1*(2), 411.

Sergi, I., Senese, V., Pisani, M., & Nigro, G. (2014). Assessing activation of true and false memory traces: A study using the DRM Paradigm. *Psychologica Belgica, 54*(1).

Shah, D., & Knott, L. M. (2018). The role of attention at retrieval on the false recognition of negative emotional DRM lists. *Memory, 26*(2), 269-276.

Shahid, A., Wilkinson, K., Marcu, S., & Shapiro, C. M. (2012). *STOP, THAT and one hundred other sleep scales*: Springer Science & Business Media.

Sharkawy, J. E., Groth, K., Vetter, C., Beraldi, A., & Fast, K. (2008). False memories of emotional and neutral words. *Behavioural Neurology, 19*(1-2), 7-11.

Shaw, J. J., & Monaghan, P. (2017). Lateralised sleep spindles relate to false memory generation. *Neuropsychologia, 107*, 60-67.

Sherman, S. M., & Grange, J. A. (2020) "Exploring the impact of mindfulness on false-memory susceptibility." *Psychological Science* 31, no. 8: 968-977.

Short, M. A., Arora, T., Gradisar, M., Taheri, S., & Carskadon, M. A. (2017). How many sleep diary entries are needed to reliably estimate adolescent sleep? *Sleep, 40*(3), 006.

Short, M. A., Gradisar, M., Lack, L. C., Wright, H., & Carskadon, M. A. (2012). The discrepancy between actigraphic and sleep diary measures of sleep in adolescents. *Sleep medicine, 13*(4), 378-384.

Silva, G. E., Goodwin, J. L., Vana, K. D., & Quan, S. F. (2016). Obstructive sleep apnea and quality of life: comparison of the SAQLI, FOSQ, and SF-36 questionnaires. *Southwest journal of pulmonary & critical care, 13*(3), 137.

Sio, U. N., Monaghan, P., & Ormerod, T. (2013). Sleep on it, but only if it is difficult: effects of sleep on problem solving. *Memory & Cognition, 41*(2), 159-166.

Smeets, T., Jelicic, M., & Merckelbach, H. (2006a). Stress-induced cortisol responses, sex differences, and false recollections in a DRM paradigm. *Biological psychology, 72*(2), 164-172.

Smeets, T., Jelicic, M., & Merckelbach, H. (2006b). The effect of acute stress on memory depends on word valence. *International journal of psychophysiology, 62*(1), 30-37.

Smith, R. E., Hunt, R. R., & Gallagher, M. P. (2008). The effect of study modality on false recognition. *Memory & cognition, 36*(8), 1439-1449.

Smith, S. M., Gerkens, D. R., Pierce, B. H., & Choi, H. (2002). The roles of associative responses at study and semantically guided recollection at test in false memory: The Kirkpatrick and Deese hypotheses. *Journal of Memory and Language, 47*(3), 436-447.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of experimental psychology: General*, *117*(1), 34.

Sopp, M. R., Michael, T., Weeß, H.-G., & Mecklinger, A. (2017). Remembering specific features of emotional events across time: The role of REM sleep and prefrontal theta oscillations. *Cognitive, Affective, & Behavioral Neuroscience, 17*(6), 1186-1209.

Spaniol, J., Voss, A., & Grady, C. L. (2008). Aging and emotional memory: Cognitive mechanisms underlying the positivity effect. *Psychology and aging, 23*(4), 859.

Spira, A. P., Beaudreau, S. A., Stone, K. L., Kezirian, E. J., Lui, L.-Y., Redline, S., . . . Stewart, A. (2012). Reliability and validity of the Pittsburgh Sleep Quality Index and the Epworth Sleepiness Scale in older men. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences, 67*(4), 433-439.

Squire, L. R. (1998). Memory systems. *Comptes Rendus de l'Académie des Sciences-Series III-Sciences de la Vie, 321*(2-3), 153-156.

Squire, L. R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. *Annual review of psychology, 44*(1), 453-495.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences, 93*(24), 13515-13522.

Stadler, M. A., Roediger, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition, 27*(3), 494-500.

Stepan, M. E., Dehnke, T. M., & Fenn, K. M. (2017). Sleep and eyewitness memory: Fewer false identifications after sleep when the target is absent from the lineup. *PloS one, 12*(9).

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature, 437*(7063), 1272-1278.

Stickgold, R., Scott, L., Rittenhouse, C., & Hobson, J. A. (1999). Sleep-induced changes in associative memory. *Journal of cognitive neuroscience, 11*(2), 182-193.

Stickgold, R., & Walker, M. (2004). To sleep, perchance to gain creative insight? *Trends in cognitive sciences, 8*(5), 191-192.

Stickgold, R., & Walker, M. P. (2007). Sleep-dependent memory consolidation and reconsolidation. *Sleep medicine, 8*(4), 331-343.

Stickgold, R., Whidbee, D., Schirmer, B., Patel, V., & Hobson, J. A. (2000). Visual discrimination task improvement: A multi-step process occurring during sleep. *Journal of cognitive neuroscience, 12*(2), 246-254.

Sugrue, K., Strange, D., & Hayne, H. (2009). False memories in the DRM paradigm: Age-related differences in lure activation and source monitoring. *Experimental Psychology, 56*(5), 354-360.

Swannell, E. R., & Dewhurst, S. A. (2012). Phonological false memories in children and adults: Evidence for a developmental reversal. *Journal of Memory and Language, 66*(2), 376-383.

Tanum, T. K., & Ringen, H. (2017). *Attentional function following partial sleep deprivation: does extraversion make a difference?* NTNU, Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychological Bulletin, 110*(1), 67.

Tempesta, D., Socci, V., De Gennaro, L., & Ferrara, M. (2019). The Role of Sleep in Emotional Processing. In *Sleep, Memory and Synaptic Plasticity* (pp. 125-170): Springer.

Thun, E., Bjorvatn, B., Osland, T., Steen, V. M., Sivertsen, B., Johansen, T., . . . Pallesen, S. (2012). An actigraphic validation study of seven morningness-eveningness inventories. *European Psychologist*.

Thurman, S. M., Wasylyshyn, N., Roy, H., Lieberman, G., Garcia, J. O., Asturias, A., . . . Grafton, S. T. (2018). Individual differences in compliance and agreement for sleep logs and wrist actigraphy: A longitudinal study of naturalistic sleep in healthy adults. *PLoS One, 13*(1), e0191883.

Tonetti, L., Mingozzi, R., & Natale, V. (2016). Comparison between paper and electronic sleep diary. *Biological Rhythm Research, 47*(5), 743-753.

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.

Tucker, M. A., Hirota, Y., Wamsley, E. J., Lau, H., Chaklader, A., & Fishbein, W. (2006). A daytime nap containing solely non-REM sleep enhances declarative but not procedural memory. *Neurobiology of learning and memory, 86*(2), 241-247.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie canadienne, 26*(1), 1.

Tworoger, S. S., Davis, S., Vitiello, M. V., Lentz, M. J., & McTiernan, A. (2005). Factors associated with objective (actigraphic) and subjective sleep quality in young adult women. *Journal of psychosomatic research, 59*(1), 11-19.

Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of experimental psychology, 70*(1), 122.

Unsworth, N., & Brewer, G. A. (2010). Individual differences in false recall: A latent variable analysis. *Journal of Memory and Language, 62*(1), 19-34.

Vallières, A., & Morin, C. M. (2003). Actigraphy in the assessment of insomnia. Sleep, *26*(7), 902-906.

Van De Water, A. T., Holmes, A., & Hurley, D. A. (2011). Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography–a systematic review. *Journal of sleep research, 20*(1pt2), 183-200.

Van Dongen, E. V., Thielen, J.-W., Takashima, A., Barth, M., & Fernández, G. (2012). Sleep supports selective retention of associative memories based on relevance for future utilization. *PloS one, 7*(8).

Vignatelli, L., Plazzi, G., Barbato, A., Ferini-Strambi, L., Manni, R., Pompei, F., & D'Alessandro, R. (2003). Italian version of the Epworth sleepiness scale: external validity. *Neurological Sciences, 23*(6), 295-300.

Violani, C., Testa, P., & Casagrande, M. (1998). Actigraphic motor asymmetries during sleep. *Sleep, 21*(5), 472-476.

Wagner, U., Gais, S., & Born, J. (2001). Emotional memory formation is enhanced across sleep intervals with high amounts of rapid eye movement sleep. *Learning & Memory, 8*(2), 112-119.

Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature, 427*(6972), 352-355.

Wagner, U., Hallschmid, M., Verleger, R., & Born, J. (2003). Signs of REM sleep dependent enhancement of implicit face memory: a repetition priming study. *Biological psychology, 62*(3), 197-210.

Wagner, U., Kashyap, N., Diekelmann, S., & Born, J. (2007). The impact of post-learning sleep vs. wakefulness on recognition memory for faces with different facial expressions. *Neurobiology of learning and memory, 87*(4), 679-687.

Walker, M. P., Brakefield, T., Morgan, A., Hobson, J. A., & Stickgold, R. (2002). Practice with sleep makes perfect: sleep-dependent motor skill learning. *Neuron, 35*(1), 205-211.

Walker, M. P., & Stickgold, R. (2004). Sleep-dependent learning and memory consolidation. *Neuron, 44*(1), 121-133.

Walker, M. P., & Stickgold, R. (2006). Sleep, memory, and plasticity. *Annu. Rev. Psychol., 57*, 139-166.

Walker, M. P., & Stickgold, R. (2010). Overnight alchemy: sleep-dependent memory evolution. *Nature Reviews Neuroscience, 11*(3), 218-218.

Walker, M. P., & van Der Helm, E. (2009). Overnight therapy? The role of sleep in emotional brain processing. *Psychological bulletin, 135*(5), 731.

Wang, J. Y., Weber, F. D., Zinke, K., Inostroza, M., & Born, J. (2018). More effective consolidation of episodic long-term memory in children than adults—Unrelated to sleep. *Child development, 89*(5), 1720-1734.

Ward, T. M., Lentz, M., Kieckhefer, G. M., & Landis, C. A. (2012). Polysomnography and actigraphy concordance in juvenile idiopathic arthritis, asthma and healthy children. *Journal of sleep research, 21*(1), 113-121.

Warrington, E. K. (2014). The double dissociation of short-and long-term memory. *Human Memory and Amnesia (PLE: Memory)*, 61.

Watson, J. M., Balota, D. A., & Roediger III, H. L. (2003). Creating false memories with hybrid lists of semantic and phonological associates: Over-additive false memories produced by converging associative networks. *Journal of Memory and Language, 49*(1), 95-118.

Watson, J. M., Mcdermott, K. B., & Balota, D. A. (2004). Attempting to avoid false memories in the Deese/Roediger—McDermott paradigm: Assessing the combined influence of practice and warnings in young and old adults. *Memory & cognition, 32*(1), 135-141.

Weiss, A. R., Johnson, N. L., Berger, N. A., & Redline, S. (2010). Validity of activity-based devices to estimate sleep. *Journal of Clinical Sleep Medicine, 6*(4), 336-342.

Werner, H., Molinari, L., Guyer, C., & Jenni, O. G. (2008). Agreement rates between actigraphy, diary, and questionnaire for children's sleep patterns. *Archives of pediatrics & adolescent medicine, 162*(4), 350-358.

Wiesner, C. D., Pulst, J., Krause, F., Elsner, M., Baving, L., Pedersen, A., . . . Göder, R. (2015). The effect of selective REM-sleep deprivation on the consolidation and affective evaluation of emotional memories. *Neurobiology of learning and memory, 122*, 131-141.

Williams, J. M., Taylor, D. J., Slavish, D. C., Gardner, C. E., Zimmerman, M. R., Patel, K., . . . Estevez, R. (2020). Validity of actigraphy in young adults with insomnia. *Behavioral sleep medicine, 18*(1), 91-106.

Wilson, B. M., Mickes, L., Stolarz-Fantino, S., Evrard, M., & Fantino, E. (2015). Increased false-memory susceptibility after mindfulness meditation. Psychological Science, 26(10), 1567-1573.

Wilson, J. K., Baran, B., Pace-Schott, E. F., Ivry, R. B., & Spencer, R. M. (2012). Sleep modulates word-pair learning but not motor sequence learning in healthy older adults. *Neurobiology of aging, 33*(5), 991-1000.

Winters, B. D., Saksida, L. M., & Bussey, T. J. (2008). Object recognition memory: neurobiological mechanisms of encoding, consolidation and retrieval. *Neuroscience & Biobehavioral Reviews, 32*(5), 1055-1070.

Wohlgemuth, W. K., Edinger, J. D., Fins, A. I., & Sullivan, R. J. (1999). How many nights are enough? The short-term stability of sleep parameters in elderly insomniacs and normal sleepers. *Psychophysiology, 36*(2), 233-244.

Wolfson, A. R., Carskadon, M. A., Acebo, C., Seifer, R., Fallone, G., Labyak, S. E., & Martin, J. L. (2003). Evidence for the validity of a sleep habits survey for adolescents. *Sleep, 26*(2), 213-216.

Yap, Y., Rice-Lacy, R., Bei, B., & Wiley, J. (2018). 0178 Bidirectional Relations between Stress and Sleep: An Intensive Daily Study. *Sleep, 41*(suppl_1), A70-A70.

Yaroush, R., Sullivan, M. J., & Ekstrand, B. R. (1971). Effect of sleep on memory: II. Differential effect of the first and second half of the night. *Journal of experimental psychology, 88*(3), 361.

Yeh, L.-H., & Lu, A. Y.-C. (2019). Association of trait mindfulness with the false recognition of emotional words in bilinguals. *International Journal of Research, 8*(2), 67-80.

Zerbini, G., & Merrow, M. (2017). Time to learn: How chronotype impacts education. *PsyCh journal, 6*(4), 263-276.

Zhang, W., Gross, J., & Hayne, H. (2017). The effect of mood on false memory for emotional DRM word lists. *Cognition and Emotion, 31*(3), 526-537.

Zinkhan, M., & Kantelhardt, J. W. (2016). Sleep assessment in large cohort studies with high-resolution accelerometers. *Sleep medicine clinics, 11*(4), 469-488.

Zohar, D., Tzischinsky, O., Epstein, R., & Lavie, P. (2005). The effects of sleep loss on medical residents' emotional reactions to work events: a cognitive-energy model. *Sleep, 28*(1), 47-54.

Zwaan, R. A., Pecher, D., Paolacci, G., Bouwmeester, S., Verkoeijen, P., Dijkstra, K., & Zeelenberg, R. (2018). Participant nonnaiveté and the reproducibility of cognitive psychology. Psychonomic Bulletin & Review, 25(5), 1968-1972

Appendices

Appendix A

Emotional DRM word lists used at the encoding phase in Experiments 1, 2 and 2A taken from Zhang et al. (2017).

Negative word lists

Critical lure: anger

List word	BAS	Arousal	Valence
rage	0.617	4.517	6.367
mad	0.600	4.417	6.350
enrage	0.483	4.450	6.117
fury	0.350	4.550	5.867
wrath	0.233	4.383	5.717
hatred	0.233	5.033	7.050
hate	0.183	4.683	7.050
annoy	0.117	4.833	6.083
fight	0.100	4.333	5.967
fear	0.017	4.483	6.517
Mean	0.293	4.568	6.308

Critical lure: sick

List word	BAS	Arousal	Valence
ill	0.683	5.100	6.617
illness	0.617	5.117	6.933
flu	0.450	5.167	6.600
nausea	0.417	5.083	6.583
cough	0.350	5.333	6.167
virus	0.300	4.300	6.567
fever	0.267	4.783	6.383
vomit	0.250	4.650	6.917
cold	0.017	5.150	5.800
hurt	0.017	5.033	6.700
Mean	0.337	4.972	6.527

Critical lure: lie

List word	BAS	Arousal	Valence
fib	0.667	5.167	6.117
deceive	0.317	4.933	6.467
deception	0.317	5.000	6.400
cheat	0.283	4.783	6.750
untruthful	0.283	5.150	6.717
dishonest	0.267	5.050	6.883
mislead	0.183	4.917	6.567
FALSE	0.133	5.117	5.800
betray	0.067	4.617	7.100
trick	0.050	4.483	4.717
Mean	0.257	4.922	6.352

Critical lure: sad

List word	BAS	Arousal	Valence
unhappy	0.550	5.567	6.717
sorrow	0.433	5.683	7.033
depression	0.400	5.367	7.500
cry	0.400	5.400	6.550
upset	0.367	5.333	6.917
tears	0.293	5.345	6.741
down	0.117	5.433	5.433
suicide	0.100	4.667	7.900
low	0.033	5.450	5.900
distress	0.033	4.300	6.767
Mean	0.273	5.254	6.746

Positive word lists

Critical lure: beautiful

List word	BAS	Arousal	Valence
stunning	0.417	4.217	3.467
gorgeous	0.350	4.200	3.283
pretty	0.217	4.183	3.267
attractive	0.167	4.217	3.333
elegant	0.133	4.767	3.533
breathtaking	0.117	3.583	3.050
lovely	0.100	4.750	3.367
exquisite	0.100	4.867	3.567
striking	0.050	4.300	3.950
appealing	0.033	4.533	3.633
Mean	0.168	4.362	3.445

Critical lure: nice

List word	BAS	Arousal	Valence
pleasant	0.433	5.400	3.817
polite	0.317	5.133	4.000
compliment	0.317	4.583	3.533
considerate	0.167	5.467	3.750
genial	0.083	5.167	4.783
sweet	0.067	4.500	4.033
thoughtful	0.050	5.383	4.117
friend	0.033	4.300	3.250
amiable	0.033	5.383	4.483
smile	0.000	4.533	3.233
Mean	0.150	4.985	3.900

Critical lure: love

List word	BAS	Arousal	Valence
affection	0.533	4.600	3.550
adore	0.483	4.683	3.417
passion	0.383	4.000	3.683
kiss	0.350	3.767	3.267
heart	0.233	4.667	3.917
care	0.217	5.317	3.733
devoted	0.183	4.600	3.800
friendship	0.150	4.283	3.433
embrace	0.133	4.867	3.500
like	0.100	4.983	4.100
Mean	0.277	4.577	3.640

Critical lure: happy

List word	BAS	Arousal	Valence
glad	0.517	4.917	3.833
pleased	0.483	4.900	3.517
joyful	0.433	4.433	3.433
content	0.350	5.633	3.750
elated	0.300	4.333	3.900
satisfied	0.250	5.050	3.550
laugh	0.217	3.550	3.133
enjoyment	0.217	4.317	3.633
ecstatic	0.200	3.850	3.317
enjoyable	0.183	4.317	3.433
Mean	0.315	4.530	3.550

Neutral word lists

Critical lure: foot

List word	BAS	Arousal	Valence
toe	0.567	5.433	5.133
shoe	0.367	5.267	4.533
ankle	0.250	5.250	5.117
sock	0.200	5.417	4.817
hand	0.183	5.150	4.417
sandals	0.167	5.117	4.300
boot	0.117	5.317	4.783
kick	0.100	4.583	4.867
knee	0.050	5.383	5.083
walk	0.033	5.200	4.233
Mean	0.203	5.212	4.728

Critical lure: car

List word	BAS	Arousal	Valence
sedan	0.650	5.367	4.833
automobile	0.567	5.150	4.600
vehicle	0.517	5.083	4.600
garage	0.450	5.167	4.633
drive	0.433	4.900	4.417
van	0.317	5.383	4.767
jeep	0.283	4.650	4.383
highway	0.183	5.117	4.700
truck	0.083	5.350	5.017
bus	0.050	5.583	4.717
Mean	0.353	5.175	4.667

Critical lure: mountain

List word	BAS	Arousal	Valence
peak	0.483	4.567	4.267
climber	0.350	4.817	4.367
summit	0.317	4.417	3.917
climb	0.267	4.467	4.200
hill	0.167	5.033	4.750
steep	0.100	4.683	5.017
ski	0.100	4.217	3.917
top	0.083	5.133	4.617
valley	0.067	5.350	4.300
range	0.050	5.033	4.583
Mean	0.198	4.772	4.393

Critical lure: music

List word	BAS	Arousal	Valence
band	0.500	4.500	4.183
instrument	0.433	4.567	3.967
concert	0.383	4.067	3.583
piano	0.367	5.250	3.983
jazz	0.367	4.867	3.817
melody	0.350	5.333	4.050
rhythm	0.317	4.517	3.750
sound	0.267	4.767	4.183
symphony	0.267	4.883	3.850
radio	0.250	4.833	4.100
Mean	0.350	4.758	3.947

Appendix B

Words lists used at the recognition task in Experiment 1.

Lure words	**Studied words**	**Filler words**
Anger Sick Lie Sad Beautiful Nice Love Happy Foot Car Mountain Music	Mad Wrath Fight Ill Fever Hurt Deception Dishonest Betray Unhappy Upset Suicide Gorgeous Elegant Striking Pleasant Thoughtful Smile Passion Care Embrace Glad Elated Enjoyment Shoe Hand Knee Sedan Jeep Bus Summit Steep Valley Band Jazz Sound	Waste Rubbish Ugly Slob Pain Fall Robber Jail Sun Girls Feather Dog Child Mother Yellow Frog Vegetable Basket Seat Stool Door House Write Crayon

Appendix C

DRM word lists used in Experiments 3 and 5, taken from Knott et al., (2012).

BAS: High/Connectivity: High				BAS: High/Connectivity: Low			
Chair	**Cold**	**King**	**Sweet**	**Slow**	**Sleep**	**Needle**	**Smell**
table seat stool desk couch sit bench cushion sofa wood	hot chill warm winter ice snow heat freeze frost cool	queen crown dictator emperor throne monarch rule royal leader England	sour candy sugar bitter tangy tart dessert chocolate cake taste	fast lethargic snail turtle quick sluggish swift slug stall delay	bed nap rest snooze blanket yawn dream relax lazy quiet	thread pin sewing sharp prick thimble haystack injection knitting pine	nose sniff aroma nostril scent fragrance perfume salts rose stench
BAS: Low/Connectivity: High				BAS: Low/Connectivity: Low			
Thief	**army**	**Bread**	**Chemistry**	**Cup**	**Pen**	**Smoke**	**Anger**
steal robber burglar cop bad rob jail crime theft fraud	navy soldier military march captain war uniform pilot officer marine	butter toast eat sandwich jam milk biscuit jelly cheese margarine	beaker element lab physics molecule electron biology experiment science chemical	mug saucer glass tea coaster handle straw jug soup lid	pencil fountain highlight scribble crayon tip marker ruler cap letter	cigarette tobacco blaze ashes fire habit nicotine flames stain billows	mad fear Hate temper wrath mean fury fight happy

Critical lures are in **bold**

Appendix D

DRM word lists used in Experiments 4, taken from Stadler et al. (1999).

Windows	**Smell**	**Doctor**	**Sweet**
door glass pane shade ledge sill house open curtain frame view breeze	nose breathe sniff aroma hear see nostril whiff scent reek stench fragrance	nurse sick lawyer medicine health hospital dentist physician ill patient office stethoscope	sour candy sugar bitter good taste tooth nice honey soda chocolate heart
Chair	**Smoke**	**Rough**	**Needle**
table sit legs seat couch desk recliner sofa wood cushion swivel stool	cigarette puff blaze billows pollution ashes cigar chimney fire tobacco stink pipe	smooth bumpy road tough sandpaper jagged ready coarse uneven riders rugged sand	thread pin eye sewing sharp point prick thimble haystack thorn hurt injection

Critical lures are in **bold**

Appendix E

Words lists used at the recognition task in Experiment 5.

Lure words	Studied words	Filler words
Anger	Emotion	Sedan
Cup	Fury	Garage
Pen	Handle	Truck
Smoke	Lid	Bus
Army	Marker	Shoe
Bread	Highlight	Ankle
Chemistry	Billows	Sock
Thief	Cigarette	Sandals
Needle	Captain	Clime
Sleep	Navy	Steep
Slow	Jam	Top
Smell	Margarine	Valley
Chair	Crime	Piano
Cold	Steal	Rhythm
King	Lab	Sound
Sweet	Molecule	Radio
	Nap	
	Relax	
	Quick	
	Swift	
	Rose	
	Sniff	
	Sharp	
	Kitting	
	Bench	
	Wood	
	Heat	
	Snow	
	Throne	
	England	
	Chocolate	
	Sugar	

Appendix F

Information sheet for Experiment 1

Study Title: Investigation into the effect of sleep on memory

Aims of the Research

The aim of this research is to understand whether sleep increases or decreases memory performance.

Invitation

You are being invited to consider taking part in a research study that is looking at the effect sleep has on memory. This project is being undertaken by Zainab Alyobi, a PhD student in the School of Psychology at Keele University.

Before you decide whether or not you wish to take part, it is important for you to understand why this research is being done and what it will involves. Please take time to read this information carefully, ask me if there is anything that is unclear or if you would like more information.

Why have I been invited?

You are expected to complete 'The Pittsburg Sleep Quality Index' and The Epworth Sleepiness Scale. The results will determine whether or not you are an appropriate candidate for this study. All successful participants must be a student or other member of Keele University, be between the ages of 18-65, and you must also be a native English speaker.

Participants with skin conditions that will be irritated by the continuous wearing of a watch should not take part in the study.

Do I have to take part?

Taking part in this study is completely voluntary and if you do decide to take part you will be given this information sheet to keep and be asked to sign a form to give consent for your participation. If you decide to take part, you are free to withdraw at any time until the end of your participation without giving your reason. If you do choose to withdraw then your data will not be included in the study and will be destroyed immediately.

What will happen if I take part?

If you decide to take part in this study after reading the information sheet, you will be asked to sign a consent form, after that you will be invited to a session where everything regarding your participation and the things you need to do will be explained to you. You will be asked to fill in a sleep diary and wear an actigraphy watch five days before and throughout the experiment.

The actigraphy will be placed on your non-dominant wrist and should not be transferred or reattached on the dominant wrist during the assessment period. You will also be asked to abstain from alcohol and caffeine two days before and until the experiment is over. On experiment day you will be asked to come along at 9 pm and return at 9 am next morning*. After the recognition test you will be asked to hand your sleep diary and actigraphy watch and in return you will be handed a debrief sheet.

* Deleted according to the group participants are allocated to (i.e. sleep/wake)

What are the benefits and risks in this study?

There are no anticipated benefits or risks to participants in this study. However, you will be contributing to our understanding of the impact of sleep on memory.

If you are interested in this topic yourself, you can request a copy of the report about the findings. If you wish to request the written report my contact details can be found below.

What will happen to the results of the research study?

Your data will be kept confidential and no one will be able to identify them other than myself and my supervisor. The data you have provided will be stored on a password protected computer. The findings will be used in my PhD report and will also be presented at academic conferences and published in academic journals. You will not be identified in any research presentation or publication as the study is completed anonymously.

On completion of the project the data collected will be posted online in a professional repository such as https://github.com/. Researchers in psychology are being encouraged to make their data "Open", so that other researchers can check the accuracy of published analytical findings independently, ensuring our research is reproducible and verifiable.

The data posted online will consist of a database containing non-sensitive, non-identifiable, information. For example, the database may contain only participant numbers, tags for experimental conditions (independent variables), and raw response times and accuracy scores. It is a requirement of the American Psychological Association—a governing body who sets standards across a wide range of psychological journals—that authors allow their data to be accessed by independent researchers. For more information, see https://osf.io/ and https://osf.io/vmrgu/.

Who is funding and organising the research?

This research is funded by the Saudi Arabia Cultural Bureau and organised by Keele University.

What if there is a problem?

If you have any problems or concerns do not hesitate to contact me or my supervisor. Our contact details can be found below. If you remain unhappy about the research and/or wish to raise a complaint about any aspect during the course of the study, please write to Nicola Leighton who is the University's contact for complaints regarding research at the address can be found below.

Researcher: Zainab Alyobi
Email: z.m.a.alyobi@keele.ac.uk

Supervisor: Dr. Sue Sherman
Email: s.m.sherman@keele.ac.uk

Nicola Leighton
Research Governance Officer
Directorate of Engagement and Partnerships IC2 Building
Keele University
ST5 5NH
E-mail: n.leighton@ keele.ac.uk
Tel: 01782 733306

Appendix G

Information sheet for Experiment 2 ,3, and 4

Study Title: Investigation into the effect of sleep on memory

Aims of the Research

The aim of this research is to understand whether sleep increases or decreases memory performance.

Invitation

This project is being undertaken by Zainab Alyobi, a PhD student in the School of Psychology at Keele University. Before you decide whether or not you wish to take part, it is important for you to understand why this research is being done and what it will involves. Please take time to read this information carefully, contact the researcher if there is anything that is unclear or if you would like more information.

Why have I been invited?

You are being invited to consider taking part in a research study that is looking at the effect of sleep on memory. You have been invited to take part in this experiment because you fit the age criteria (between 18-65) and are a native English speaker. English language is necessary because I will be using English words and it's important that they are clearly understood. Please be aware that if you have history of psychiatric or sleep disorders, drug abuse, and current use of antidepressant or hypnotic medications you should not take part in this study.

Do I have to take part?

Taking part in this study is completely voluntary and if you do decide to take part you will be rewarded with £3 for full participation.

Please be aware that you are free to withdraw at any time prior to or during the study. However, as the data are anonymous you cannot withdraw once you have completed the study. If you do choose to withdraw then your data will not be included in the study and it will be destroyed immediately.

What will happen if I take part?

Stage 1- Suitability check and study phase:

This study is completed entirely online on Qualtrics (a survey tool). Firstly, you will be asked to answer some questions about your sleep. If your answers don't fit the criteria for the study, you will not be able to proceed but you will be given 50p payment. If you meet the eligibility criteria, then you will be able to continue to complete the study. In the next step you will be asked to hear lists of words on Qualtrics for a later memory test.

Stage 2- Test phase:

This phase should be completed 12 hours later. You will be asked to type as many words as you can remember from the lists you heard previously. After you complete the test a debrief sheet will appear on the screen for you to read.

What will happen to the results of the research study?

The data you have provided will be anonymous; the findings will be used in my PhD report and will also be presented at academic conferences and published in academic journals. You will not be identified in any research presentation or publication as the study is completed anonymously.

On completion of the project the data collected will be posted online in a professional repository such as https://github.com/. Researchers in psychology are being encouraged to make their data "Open", so that other researchers can check the accuracy of published analytical findings independently, ensuring our research is reproducible and verifiable.

The data posted online will consist of a database containing non-sensitive, non-identifiable, information. For example, the database may contain only participant numbers, tags for experimental conditions (independent variables), and raw response times and accuracy scores. It is a requirement of the American Psychological Association—a governing body who sets standards across a wide range of psychological journals—that authors allow their data to be accessed by independent researchers. For more information, see https://osf.io/ and https://osf.io/vmrgu/.

Who is funding and organising the research?

This research is funded by the Saudi Arabia Cultural Bureau and organised by Keele University.

What if there is a problem?

If you have any problems or concerns do not hesitate to contact me or my supervisor. Our contact details can be found below. If you remain unhappy about the research and/or wish to raise a complaint about any aspect during the course of the study please write to Research Integrity Team, the address can be found below.

Researcher: Zainab Alyobi
Email: z.m.a.alyobi@keele.ac.uk

Supervisor: Dr. Sue Sherman
Email: s.m.sherman@keele.ac.uk

Research Integrity Team
Directorate of Research, Innovation and Engagement
IC2 Building
Keele University
ST5 5NE
Tel: 01782 733371
Email: research.governance@keele.ac.uk

Appendix H

Information sheet for Experiment 2A

Study Title: Investigation into the effect of sleep on memory

Aims of the Research

The aim of this research is to understand whether sleep increases or decreases memory performance.

Invitation

This project is being undertaken by Zainab Alyobi, a PhD student in the School of Psychology at Keele University. Before you decide whether or not you wish to take part, it is important for you to understand why this research is being done and what it will involves. Please take time to read this information carefully, contact the researcher if there is anything that is unclear or if you would like more information.

Why have I been invited?

You are being invited to consider taking part in a research study that is looking at the effect of sleep on memory. You have been invited to take part in this experiment because you fit the age criteria (between 18-65) and are a native English speaker. English language is necessary because I will be using English words and it's important that they are clearly understood. Please be aware that if you have history of psychiatric or sleep disorders, drug abuse, and current use of antidepressant or hypnotic medications you should not take part in this study.

Do I have to take part?

Taking part in this study is completely voluntary and if you do decide to take part you will be rewarded with £2 for full participation.

Please be aware that you are free to withdraw at any time prior to or during the study. However, as the data are anonymous you cannot withdraw once you have completed the study. If you do choose to withdraw then your data will not be included in the study and it will be destroyed immediately.

What will happen if I take part?

Stage 1- Suitability check and study phase:

This study is completed entirely online on Qualtrics (a survey tool). Firstly, you will be asked to answer some questions about your sleep. If your answers don't fit the criteria for the study, you will not be able to proceed but you will be given 50p payment. If you meet the eligibility criteria, then you will be able to continue to complete the study. In the next step you will be asked to hear lists of words on Qualtrics for a later memory test.

Stage 2- Test phase:

This phase should be completed 20 minutes later. You will be asked to type as many words as you can remember from the lists you heard previously. After you complete the test a debrief sheet will appear on the screen for you to read.

What will happen to the results of the research study?

The data you have provided will be anonymous; the findings will be used in my PhD report and will also be presented at academic conferences and published in academic journals. You will not be identified in any research presentation or publication as the study is completed anonymously.

On completion of the project the data collected will be posted online in a professional repository such as https://github.com/. Researchers in psychology are being encouraged to make their data "Open", so that other researchers can check the accuracy of published analytical findings independently, ensuring our research is reproducible and verifiable.

The data posted online will consist of a database containing non-sensitive, non-identifiable, information. For example, the database may contain only participant numbers, tags for experimental conditions (independent variables), and raw response times and accuracy scores. It is a requirement of the American Psychological Association—a governing body who sets standards across a wide range of psychological journals—that authors allow their data to be accessed by independent researchers. For more information, see https://osf.io/ and https://osf.io/vmrgu/.

Who is funding and organising the research?

This research is funded by the Saudi Arabia Cultural Bureau and organised by Keele University.

What if there is a problem?

If you have any problems or concerns do not hesitate to contact me or my supervisor. Our contact details can be found below. If you remain unhappy about the research and/or wish to raise a complaint about any aspect during the course of the study please write to Research Integrity Team, the address can be found below.

Researcher: Zainab Alyobi
Email: z.m.a.alyobi@keele.ac.uk

Supervisor: Dr. Sue Sherman
Email: s.m.sherman@keele.ac.uk

Research Integrity Team
Directorate of Research, Innovation and Engagement
IC2 Building
Keele University
ST5 5NE
Tel: 01782 733371
Email: research.governance@keele.ac.uk

Appendix I

Information sheet for Experiment 5

Study Title: Investigation into the effect of early and late sleep on memory

Aims of the Research

The aim of this research is to understand the impact of Early vs Late sleep on memory.

Invitation

This project is being undertaken by Zainab Alyobi, a PhD student in the School of Psychology at Keele University.

Before you decide whether or not you wish to take part, it is important for you to understand why this research is being done and what it will involves. Please take time to read this information carefully, contact the researcher if there is anything that is unclear or if you would like more information.

Why have I been invited?

You are being invited to consider taking part in a research study that is looking at the effect of early and late sleep on memory. You have been invited to take part in this experiment because you fit the age criteria (between 18-65) and are a native English speaker. English language is necessary because I will be using English words and it's important that they are clearly understood. Please be aware that if you have history of psychiatric or sleep disorders, drug abuse, and current use of antidepressant or hypnotic medications you should not take part in this study.

What are the benefits and risks in this study?

Taking part in this study is completely voluntary and if you do decide to take part you will be rewarded with £5 for full participation.

Please be aware that you are free to withdraw at any time prior to or during the study. However, as the data are anonymous you cannot withdraw once you have completed the study. If you do choose to withdraw then your data will not be included in the study and it will be destroyed immediately.

There are no anticipated risks to participants in this study. However you will be contributing to our understanding of the impact of early vs late sleep on memory. If you are interested in this topic yourself, you can request a copy of the report about the findings. If you wish to request the written report my contact details can be found below.

What will happen if I take part?

Stage 1- Suitability check and study phase:

This study is completed entirely online on Qualtrics (a survey tool). Firstly, you will be asked to complete an initial questionnaire, 'The Morningness-Eveningness questionnaire', and answer some questions about your sleep. If your answers don't fit the criteria for the study, you will not be able to proceed but you will be given £2 payment. If you meet the eligibility criteria, then you will be able to continue to complete the rest of the study. In the next phase you will be asked to study lists of words on Qualtrics for a later memory test.

Stage 2- Test phase:

This phase should be completed 24 hours later. You will be presented with words and asked to identify whether you have heard the words in the previous study phase or not and whether you had a detailed memory of them. After you complete the test a debrief sheet will appear on the screen for you to read.

What will happen to the results of the research study?

The data you have provided will be anonymous; the findings will be used in my PhD report and will also be presented at academic conferences and published in academic journals. You will not be identified in any research presentation or publication as the study is completed anonymously.

On completion of the project the data collected will be posted online in a professional repository such as https://github.com/. Researchers in psychology are being encouraged to make their data "Open", so that other researchers can check the accuracy of published analytical findings independently, ensuring our research is reproducible and verifiable.

The data posted online will consist of a database containing non-sensitive, non-identifiable, information. For example, the database may contain only participant numbers, tags for experimental conditions (independent variables), and raw response times and accuracy scores. It is a requirement of the American Psychological Association—a governing body who sets standards across a wide range of psychological journals—that authors allow their data to be accessed by independent researchers. For more information, see https://osf.io/ and https://osf.io/vmrgu/.

Who is funding and organising the research?

This research is funded by the Saudi Arabia Cultural Bureau and organised by Keele University.

What if there is a problem?

If you have any problems or concerns do not hesitate to contact me or my supervisor. Our contact details can be found below. If you remain unhappy about the research and/or wish to raise a complaint about any aspect during the course of the study please write to Research Integrity Team, the address can be found below.

Researcher: Zainab Alyobi
Email: z.m.a.alyobi@keele.ac.uk

Supervisor: Dr. Sue Sherman
Email: s.m.sherman@keele.ac.uk

Research Integrity Team
Directorate of Research, Innovation and Engagement
IC2 Building
Keele University
ST5 5NE
Tel: 01782 733371
Email: research.governance@keele.ac.uk

Appendix J

Debrief sheet for Experiment 1

Keele University

DE-BRIEF: The effect of sleep on false memory

Thank you for taking part in this study. This study was an investigation into whether memory, specifically false memory is increased or decreased by sleep. Here is a good explanation of how false memory works: “Memory is not a literal record of the world, but what is retrieved from memory can substantially differ from what was originally encoded, and under certain circumstances people even claim to remember events that in fact never happened. Such false memories are typically semantically associated with actually encoded events, and subjects are highly confident about the correctness of these memories” (Diekelmann, Born & Wagner, 2010, p. 1)

False memories can be created in the laboratory using the DRM paradigm (Deese 1959, Roediger & McDermott 1995). Using the DRM paradigm participants are presented with word lists for example ‘bed, ‘wake’, ‘dream’, ‘snooze’. When they are later tested on their memory for the words, they often have a false memory for the non-presented word ‘sleep’. Past research has shown that this method is reliable when studying the formation of false memory in a controlled laboratory setting. Previous research has reported mixed findings on the effect of sleep on false memory, some of them indicated that sleep increases false memory (Darsaud et al., 2011) whereas some of them found that sleep decreases false memory (Fenn, Gallo, Margoliash, Roediger & Nusbaum, 2009). Diekelmann et al. (2008) actually found that sleep had no effect in creating false memory. This study aims to clarify these discrepancies.

You have been asked to wear an actigraphy watch and fill in a sleep diary, which will be compared to each other and will report the duration of your sleep. You have also been asked to abstain from caffeine and alcohol for two days prior to the experiment. This is to prevent it from having an effect on your sleep and data.

During the experiment you were presented with words from three types of word lists, the types consisted of negative, positive and neutral word lists. An example of the negative word lists is mad, fear, hate etc. which were related to the non-presented word 'anger'. You were then asked to return for testing 12 hours later. In the next session you were presented with a list of words which included some of the words that you studied in the first session (e.g., mad, fear) together with related words that were not presented to you during the first stage of the experiment (e.g., anger) and some completely new words (e.g., robber). There were two groups in this experiment, one group came at 9 am and the other group came at 9 pm, all to be tested 12 hours later. This is in order to see how sleep or wakefulness differently impact on different types of correct and false memories (negative, positive and neutral).

If you would like further information about this study, do contact myself (Zainab Alyobi, z.m.a.alyobi@keele.ac.uk) or my supervisor (Dr. Sue Sherman, s.m.sherman@keele.ac.uk).

Thanks again for your participation.

References

Darsaud, A., Dehon, H., Lahl, O., Sterpenich, V., Boly, M., Dang-Vu, T., ... & Schabus, M. (2011). Does sleep promote false memories?. *Journal of Cognitive Neuroscience*, *23*(1), 26-40.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17-21

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. *Behavioural Brain Research*, *208*(2), 425-429.

Diekelmann, S., Landolt, H. P., Lahl, O., Born, J., & Wagner, U. (2008). Sleep loss produces false memories. *PLoS One*, *3*(10), e3512.

Fenn, K. M., Gallo, D. A., Margoliash, D., Roediger, H. L., & Nusbaum, H. C. (2009). Reduced false memory after sleep. *Learning & Memory*, *16*(9), 509-513.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of experimental psychology: Learning, Memory, and Cognition*, *21*(4), 803-823

Appendix K

Debrief sheet for Experiment 2

Keele University

DE-BRIEF: The effect of sleep on false memory

Thank you for taking part in this study. This study was an investigation into whether memory, specifically false memory is increased or decreased by sleep. Here is a good explanation of how false memory works: “Memory is not a literal record of the world, but what is retrieved from memory can substantially differ from what was originally encoded, and under certain circumstances people even claim to remember events that in fact never happened. Such false memories are typically semantically associated with actually encoded events, and subjects are highly confident about the correctness of these memories” (Diekelmann, Born & Wagner, 2010, p. 1)

False memories can be created in the laboratory using the DRM paradigm (Deese 1959, Roediger & McDermott 1995). Using the DRM paradigm participants are presented with word lists for example ‘bed, ‘wake’, ‘dream’, ‘snooze’. When they are later tested on their memory for the words, they often have a false memory for the non-presented word ‘sleep’. Past research has shown that this method is reliable when studying the formation of false memory in a controlled laboratory setting. Previous research has reported mixed findings on the effect of sleep on false memory, some of them indicated that sleep increases false memory (Darsaud et al., 2011) whereas some of them found that sleep decreases false memory (Fenn, Gallo, Margoliash, Roediger

& Nusbaum, 2009). Diekelmann et al. (2008) actually found that sleep had no effect in creating false memory. This study aims to clarify these discrepancies.

During the experiment you were presented with words from three types of word lists, the types consisted of negative, positive and neutral word lists. An example of the negative word lists is mad, fear, hate etc. which were related to the non-presented word 'anger'. You were then asked to complete the second part of the study 12 hours later. In the second part you were asked to write as many words as you can remember from the lists you heard previously. This is in order to see if you can recall any of the non- presented words. There were two groups in this experiment, one group completed the first part at 9 am and the other group complete it at 9 pm, all to be tested 12 hours later. This is in order to see how sleep or wakefulness differently impact on different types of correct and false memories (negative, positive and neutral).

If you would like further information about this study, do contact myself (Zainab Alyobi, z.m.a.alyobi@keele.ac.uk) or my supervisor (Dr. Sue Sherman, s.m.sherman@keele.ac.uk).

Thanks again for your participation.

References

Darsaud, A., Dehon, H., Lahl, O., Sterpenich, V., Boly, M., Dang-Vu, T., ... & Schabus, M. (2011). Does sleep promote false memories? *Journal of Cognitive Neuroscience*, *23*(1), 26-40.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17-21

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. *Behavioural Brain Research*, *208*(2), 425-429.

Diekelmann, S., Landolt, H. P., Lahl, O., Born, J., & Wagner, U. (2008). Sleep loss produces false memories. *PLoS One*, *3*(10), e3512.

Fenn, K. M., Gallo, D. A., Margoliash, D., Roediger, H. L., & Nusbaum, H. C. (2009). Reduced false memory after sleep. *Learning & Memory*, *16*(9), 509-513.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of experimental psychology: Learning, Memory, and Cognition*, *21*(4), 803-823

Appendix L

Debrief sheet for Experiment 2A

Keele University

DE-BRIEF: The effect of sleep on false memory

Thank you for taking part in this study. This study was an investigation into whether memory, specifically false memory is increased or decreased by sleep. Here is a good explanation of how false memory works: “Memory is not a literal record of the world, but what is retrieved from memory can substantially differ from what was originally encoded, and under certain circumstances people even claim to remember events that in fact never happened. Such false memories are typically semantically associated with actually encoded events, and subjects are highly confident about the correctness of these memories” (Diekelmann, Born & Wagner, 2010, p. 1)

False memories can be created in the laboratory using the DRM paradigm (Deese 1959, Roediger & McDermott 1995). Using the DRM paradigm participants are presented with word lists for example ‘bed, ‘wake’, ‘dream’, ‘snooze’. When they are later tested on their memory for the words, they often have a false memory for the non-presented word ‘sleep’. Past research has shown that this method is reliable when studying the formation of false memory in a controlled laboratory setting. Previous research has reported mixed findings on the effect of sleep on false memory, some of them indicated that sleep increases false memory (Darsaud et al., 2011) whereas some of them found that sleep decreases false memory (Fenn, Gallo, Margoliash, Roediger

& Nusbaum, 2009). Diekelmann et al. (2008) actually found that sleep had no effect in creating false memory. This study aims to clarify these discrepancies.

During the experiment you were presented with words from three types of word lists, the types consisted of negative, positive and neutral word lists. An example of the negative word lists are mad, fear, hate etc. which were related to the non-presented word 'anger'. You were then asked to complete the second part of the study 20 min later. In the second part you were asked to write as many words as you can remember from the lists you heard previously. This is in order to see if you can recall any of the non- presented words.

There were two groups in this study, one group completed the first part at 9 am and the other group complete it at 9 pm, all to be tested 20 min later. This is in order to see how time of the day differently impact on different types of correct and false memories (negative, positive and neutral).

If you would like further information about this study, do contact myself (Zainab Alyobi, z.m.a.alyobi@keele.ac.uk) or my supervisor (Dr. Sue Sherman, s.m.sherman@keele.ac.uk).

Thank you for participating. When you click on "Submit Responses" this will submit your data to the study and you will not be able to withdraw your data.

References:

Darsaud, A., Dehon, H., Lahl, O., Sterpenich, V., Boly, M., Dang-Vu, T., ... & Schabus, M. (2011). Does sleep promote false memories?. *Journal of Cognitive Neuroscience*, *23*(1), 26-40.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17-21

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. *Behavioural Brain Research*, *208*(2), 425-429.

Diekelmann, S., Landolt, H. P., Lahl, O., Born, J., & Wagner, U. (2008). Sleep loss produces false memories. *PLoS One*, *3*(10), e3512.

Fenn, K. M., Gallo, D. A., Margoliash, D., Roediger, H. L., & Nusbaum, H. C. (2009). Reduced false memory after sleep. *Learning & Memory*, *16*(9), 509-513.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of experimental psychology: Learning, Memory, and Cognition*, *21*(4), 803-823

Appendix M

Debrief sheet for Experiment 3 and 4

Keele University

DE-BRIEF: The effect of sleep on false memory

Thank you for taking part in this study. This study was an investigation into whether memory, specifically false memory is increased or decreased by sleep. Here is a good explanation of how false memory works: “Memory is not a literal record of the world, but what is retrieved from memory can substantially differ from what was originally encoded, and under certain circumstances people even claim to remember events that in fact never happened. Such false memories are typically semantically associated with actually encoded events, and subjects are highly confident about the correctness of these memories” (Diekelmann, Born & Wagner, 2010, p. 1)

False memories can be created in the laboratory using the DRM paradigm (Deese 1959, Roediger & McDermott 1995). Using the DRM paradigm participants are presented with word lists for example ‘bed, ‘wake’, ‘dream’, ‘snooze’. When they are later tested on their memory for the words, they often have a false memory for the non-presented word ‘sleep’. Past research has shown that this method is reliable when studying the formation of false memory in a controlled laboratory setting. Previous research has reported mixed findings on the effect of sleep on false memory, some of them indicated that sleep increases false memory (Darsaud et al., 2011) whereas some of them found that sleep decreases false memory (Fenn, Gallo, Margoliash, Roediger & Nusbaum, 2009). Diekelmann et al. (2008) actually found that sleep had no effect in creating false memory.

During the experiment you were presented with DRM word lists, you were then asked to complete the second part of the study 12 hours later. In the second part you were asked to write as many words as you can remember from the lists you heard previously. This is in order to see if you can recall any of the non- presented lure words. There were two groups in this experiment, one group completed the first part at 9 am and the other group completed it at 9 pm, all to be tested 12 hours later. This is in order to see how sleep and wakefulness differently impact correct and false memories

If you would like further information about this study, do contact myself (Zainab Alyobi, z.m.a.alyobi@keele.ac.uk) or my supervisor (Dr. Sue Sherman, s.m.sherman@keele.ac.uk).

Thanks again for your participation.

References

Darsaud, A., Dehon, H., Lahl, O., Sterpenich, V., Boly, M., Dang-Vu, T., ... & Schabus, M. (2011). Does sleep promote false memories?. *Journal of Cognitive Neuroscience*, *23*(1), 26-40.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17-21

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. *Behavioural Brain Research*, *208*(2), 425-429.

Diekelmann, S., Landolt, H. P., Lahl, O., Born, J., & Wagner, U. (2008). Sleep loss produces false memories. *PLoS One*, *3*(10), e3512.

Fenn, K. M., Gallo, D. A., Margoliash, D., Roediger, H. L., & Nusbaum, H. C. (2009). Reduced false memory after sleep. *Learning & Memory*, *16*(9), 509-513.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of experimental psychology: Learning, Memory, and Cognition*, *21*(4), 803-823

Appendix N

Debrief sheet for Experiment 5

DE-BRIEF: Investigation into the effect of early and late sleep on false memory

Thank you for taking part in this study. This study was an investigation into Thank you for taking part in this study. This study was an investigation into the effect of early and late sleep on false memory. Here is a good explanation of how false memory works: “Memory is not a literal record of the world, but what is retrieved from memory can substantially differ from what was originally encoded, and under certain circumstances people even claim to remember events that in fact never happened. Such false memories are typically semantically associated with actually encoded events, and subjects are highly confident about the correctness of these memories” (Diekelmann, Born & Wagner, 2010, p. 1).

False memories can be created in the laboratory using the DRM paradigm (Deese 1959, Roediger & McDermott 1995). Using the DRM paradigm participants are presented with word lists for example ‘bed, ‘wake’, ‘dream’, ‘snooze’. When they are later tested on their memory for the words, they often have a false memory for the non-presented word ‘sleep’. Past research has shown that this method is reliable when studying the formation of false memory in a controlled laboratory setting. Human sleep consists of cycles that last approximately 90 minutes. These cycles consist of REM (Rapid Eye

Movement) sleep and 4 stages of non-REM sleep, with stages 3 and 4 being the deepest sleep known as slow-wave sleep (SWS). For the first half of nocturnal sleep, longer periods of SWS are experienced, whilst for the second half, REM dominates. Pervious research has suggested that SWS impairs declarative memory consolidation (of facts and events) and is therefore possibly linked to the consolidation of false memory. As SWS occurs mostly in the early period of sleep this perhaps points to some connection between this period and false memory. Hence, the current research aims to assess in detail any connection between early and late sleep periods and an increase or decrease in false memory.

You were divided into two groups (morning and evening) according to your results from the Morningness-Eveningness questionnaire. During the experiment you were presented with word lists. You were then asked to complete the second part 24 hours later. In the testing part you were presented with list of words which included some of the words that you studied in the first session (e.g., bed, wake) together with related words that were not presented to you during the first stage of the experiment (e.g., sleep) and some completely new words (e.g., chair). This is in order to see if you can recognise any of the non-presented words as an "old" word.

When you click on "Submit Responses" this will submit your data to the study and you will not be able to withdraw your data.

If you would like further information about this study, do contact myself (Zainab Alyobi, z.m.a.alyobi@keele.ac.uk) or my supervisor (Dr. Sue Sherman, s.m.sherman@keele.ac.uk).

Thanks again for your participation.

References

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17-21

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. *Behavioural Brain Research*, *208*(2), 425-429.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of experimental psychology: Learning, Memory, and Cognition*, *21*(4), 803-823

Appendix O

Consent form for Experiment 1

CONSENT FORM

Title of Project: Investigation into the effect of sleep on memory
Name and contact details of Principal Investigator: Zainab Alyobi - Dorothy Hodgkin Building 1.96
T: +44 (0)1782 734247 Email: z.m.a.alyobi@keele.ac.uk

	Please initial box if you agree with the statement
I confirm that I have read and understood the information sheet for the above study	☐
I understand that my participation is voluntary and that I am free to withdraw at any time until I have completed the study.	☐
I agree to take part in this study.	☐
I agree to allow the dataset collected to be used for future research projects	☐

Name of participant ____________ Date ____________ Signature ____________

Researcher ____________ Date ____________ Signature ____________

Appendix P

Consent form for Experiment 2,2A,3,4 and 5

Online consent information

Before taking part in this study, please read the consent information below and click on one of the boxes at the bottom of the page.

This study involves a web-based experiment designed to understand the impact of sleep on memory. Your participation in this study is voluntary. You have the right to withdraw at any point during the stage of the study, for any reason, and without any prejudice. However, as the data are anonymous you cannot withdraw once you have completed the study.

By clicking the first button below, you acknowledge that: your participation in the study is voluntary, you are 18 years old or over, you agree to allow the dataset collected to be used for future research projects, you are a native English speaker, you have no history of sleep disorders or current use of antidepressant or hypnotic medications and that you are aware of the exclusion criteria mentioned on the information sheet (which you read in the previous page). This means it is possible that you will not be asked to complete the rest of the study (and so you will only receive 50p payment for this stage). If your answers fit the criteria for the study, you will automatically move to the first part of the main experiment.

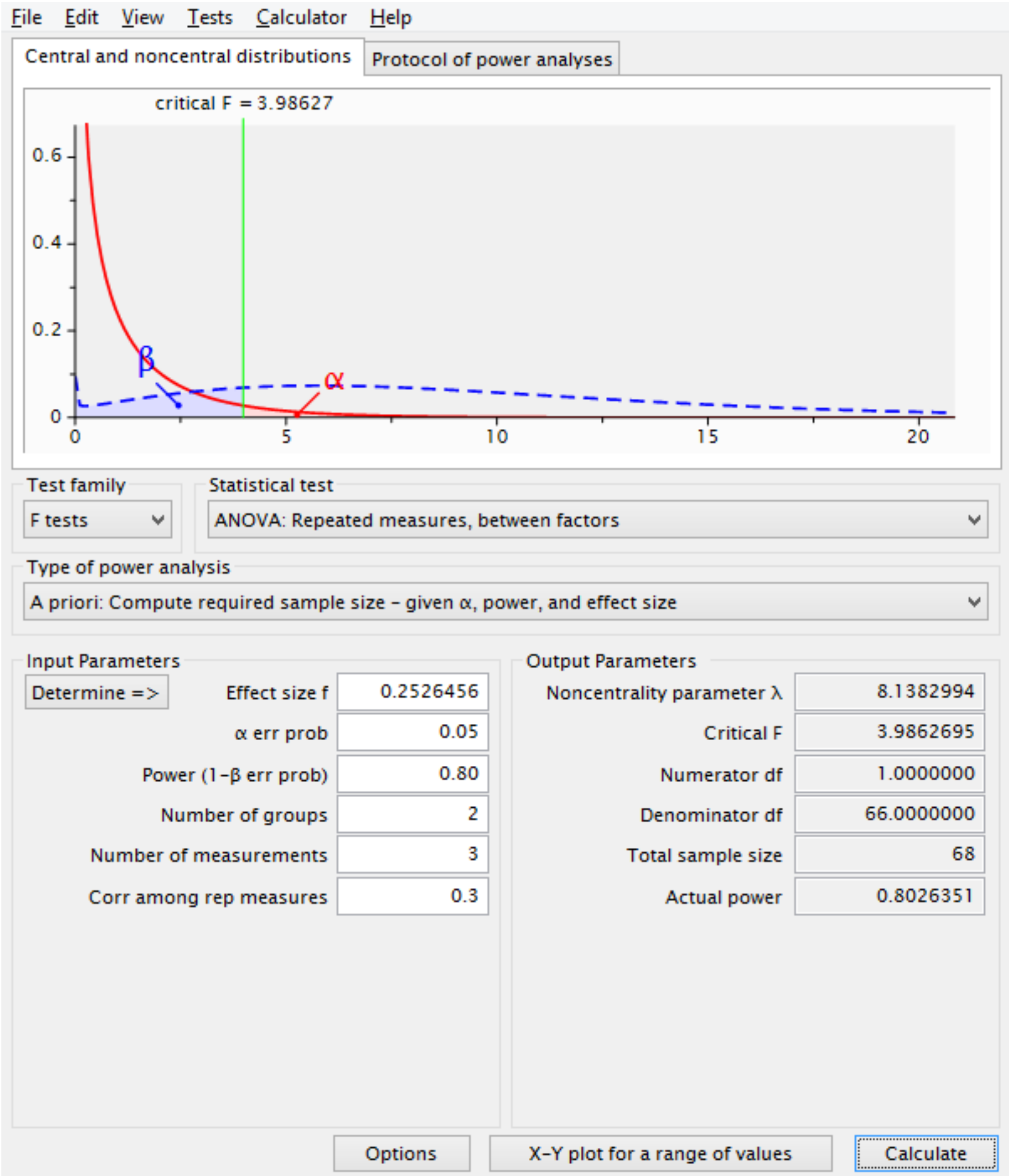
1- I have read the information on this page and I consent to take part in this study ☐

2- I do not wish to take part in this study ☐

(Ticking this box will exit the participant from the study)

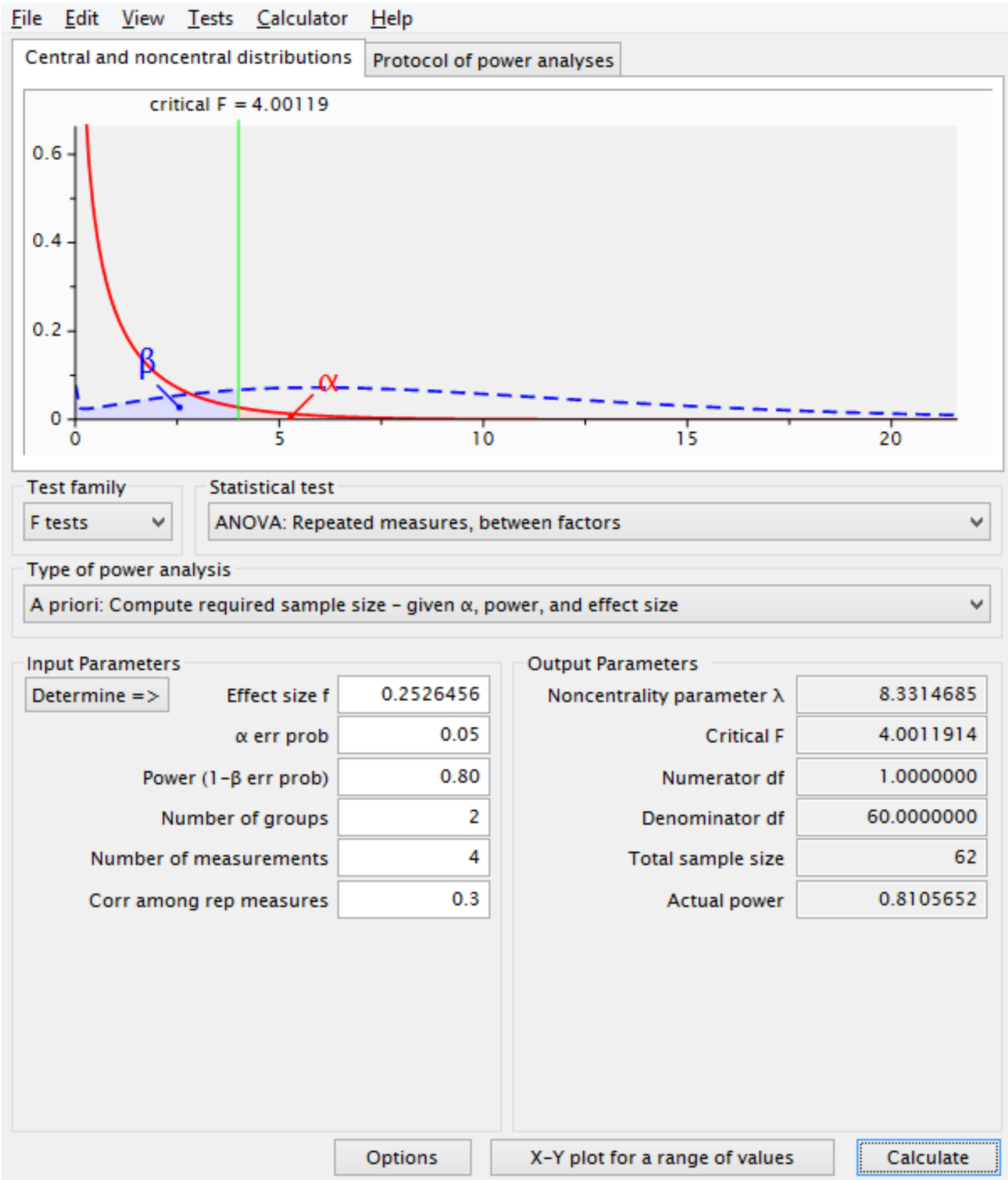
Appendix Q

Power Analysis for Experiments 1,2 and 2A



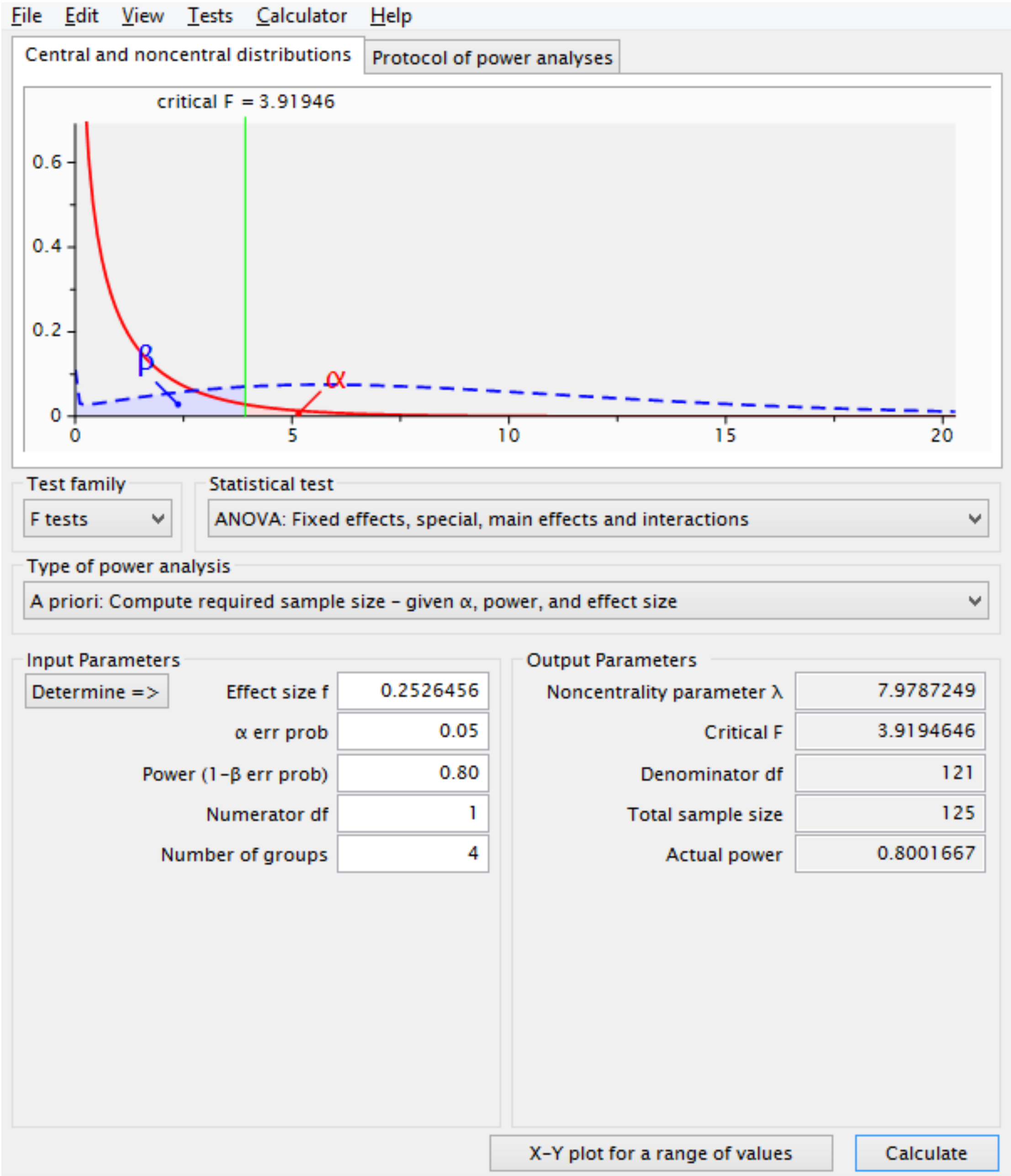
Appendix R

Power Analysis for Experiments 3 and 5



Appendix S

Power Analysis for Experiments 4



Appendix T

A Comparison between different brands of Actigraphy watches (Roy, 2020).

ACTIGRAPHY

Company		ActiGraph	Advanced Brain Monitoring Inc	Ambulatory Monitoring Inc	CamNtech Inc	Condor Instruments	Philips Respironics	SOMNOmedics America Inc
Actigraph		 CentrePoint Insight Watch	 Night Shift	 Micro Motionlogger Watch	 Motionwatch-8	 ActTrust	 Actiwatch Spectrum PRO	 SOMNOwatch plus
Website		www.actigraphcorp.com	www.advancedbrainmonitoring.com	www.ambulatory-monitoring.com	www.camntech.com	www.condorinst.com.br/en/acttrust	www.actigraphy.com	www.somnomedics.com
Dimensions (cm)		4.83 x 3.43 x 1.04	5.6 x 3.8 x 1.3	3.6 x 3.6 x 1.2	3.6 x 2.8 x 0.9 (without band)	4.7 x 3.1 x 1.2	4.8 x 3.7 x 1.5	4.5 x 4.5 x 1.6
Weight, In Use (grams)		35	44	30	9 (without band; different band options have different weights)	38	31	30
Features	Time/date display	X		X			X	
	Event marker button	X		X	X	X	X	X
	Sleep efficiency calculation	X	X	X	X	X	unknown	X
	Sleep latency calculation	X		X	X	X	unknown	X
	Temperature			X		X	unknown	
	Other	total sleep time (TST), wake after sleep onset (WASO), daytime activity (energy expenditure, steps taken, activity intensity, sedentary time), raw acceleration data	WASO, positional snoring (dB), sleep position, sleep time, awakenings/hr, vibro-tactile feedback, daily/monthly/365-day reporting	ambient light	circadian rhythm, 40 sleep parameters	light with RGB spectrum		PLMS scoring; programmable start and stop periods for several recording periods
Battery Options		rechargeable lithium polymer	rechargeable lithium polymer 200mAh	disposable 2430	CR2032 user replaceable	rechargeable, Li-Ion, 70mAh	CLB 2032 lithium ion rechargeable (factory replaced)	lithium ion-accu, rechargeable
Battery Life (days, during regular use)		30	3	30	90 (light sensor on), 120 (without light sensor)	90	50	26+
Battery Replacement Cost		N/A	N/A	$3/each	$1.50	N/A	N/A	N/A
Rechargeable Battery, Time from Depleted to Fully Charged (hours)		2-3	4	N/A	N/A	3	2	1.75
Memory Size (MB)		512	N/A	2	4	8	32	64
Recording Time at 1-minute Sample Interval Under Regular Usage (days)		30 (raw data)	6	30+	90	90	50	45
Logging Interval Options (seconds)		raw data sampled at 32 Hz - 256 Hz	Intervals of 30	1, 2, 5, 10, 30, 60	1, 2, 5, 10, 15, 30, 60	Any from 1 to 30,000	15, 30, 60	Any from 1 to 256
Light Sensor Wavelength Range (nm)		N/A	N/A	400 to 700 (520 peak)	wide spectrum visible	400 - 700	400 to 700	350 to 970
30-minute Moisture Immersion Protection (m)		water resistant, IP57 I meter, 30 min	not water resistant	100	waterproof to 3 bar; acceptable for swimming	1	waterproof at 1 m per IP27 IEC 60529	IP54-rated (water-resistant) (on demand, housing can be modified to be waterproof)
Communications Interface		USB, Bluetooth LE 5	USB	USB-IR	micro USB	USB with dock connector	USB 2.0 or higher; 150 mA	USB
Warranty (years)		1	2 (limited)	1	2	2	2	2
Platform Compatibility		Windows, iOS, Android	Windows, Mac	Windows	Windows	Windows, Mac	Windows	Windows
Peer-reviewed Sleep Validation Study		www.actigraphcorp.com/category/research-database/sleep	Levendowski DJ, Seagraves S, Popovic D, Westbrook PR. Assessment of a neck-based treatment and monitoring device for positional obstructive sleep apnea. *J Clin Sleep Med.* 2014;10(8):863-71.	Cole R, et al. Automatic sleep/wake identification from wrist activity. *Sleep.* 1992;15(5):461-9.	www.camntech.com/products/motionwatch/mw8-biblio	Rodrigues J, Eckeli AL. Validação de um actígrafo nacional. Universidade de São Paulo, Ribeirão Preto. 2018. Available at http://bdpi.usp.br/item/002889048.	Kushida et al. Refer to www.actigraphy.com/webliography for full details.	Dick R, et al. AASM standards of practice compliant validation of actigraphic sleep analysis from SOMNOwatch™ versus polysomnographic sleep diagnostics shows high conformity also among subjects with sleep disordered breathing. *Physiol Meas.* 2010;31(12):1623-33.
Additional Information		Real-time data uploads to cloud-based CentrePoint software platform via home data hub and mobile application. Slim and compact design, interchangeable wrist bands.	Worn around the neck, Night Shift monitors sleep and wake time. Records up to 365 nights' hours of use, sleep efficiency, % snoring >50 dB, neck position, and awakenings/hr. Delivers vibro-tactile feedback to discourage supine sleep.	Event marker with LED feedback, visual status indicator, multimode data collection, off-wrist detection. Five peer-reviewed articles validate this actigraph.	Contact 830-755-8036.	Computes several parameters as sleep and wake time, sleep efficiency, sleep latency, number and duration of awakening. Includes chronobiology functions as acrophase, MESOR, M10, L5, F5, IV, IS. Automated report with information required for clinical and research use.	None provided.	1 external channel possible (EEG, ECG, PLM, 2nd actigraphy-sensor, respiratory), linking of several SOMNOwatch plus recordings thanks to high synchronization rate.

Sleep Review strives for accuracy in all data but cannot be held responsible for claims made by manufacturers. All actigraphs may not be included. E-mail sroy[at]medqor.com to be considered for the next update.

30 • sleepreviewmag.com November/December 2018 November/December 2018

Roy, S. (2020). Actigraphy Side-by-Side Comparison Guide (November 2018). Retrieved 11 July 2020, from https://www.sleepreviewmag.com/sleep-diagnostics/screeners/actigraphy/actigraphy-guide/

Appendix U

Pittsburgh Sleep Quality Index (PSQI)

Sleep Quality Assessment (PSQI)

What is PSQI, and what is it measuring?

The Pittsburgh Sleep Quality Index (PSQI) is an effective instrument used to measure the quality and patterns of sleep in adults. It differentiates "poor" from "good" sleep quality by measuring seven areas (components): subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medications, and daytime dysfunction over the last month.

INSTRUCTIONS:

The following questions relate to your usual sleep habits during the past month only. Your answers should indicate the most accurate reply for the majority of days and nights in the past month. Please answer all questions.

During the past month,

1. When have you usually gone to bed? ______________________
2. How long (in minutes) has it taken you to fall asleep each night? ______________________
3. What time have you usually gotten up in the morning? ______________________
4. A. How many hours of actual sleep did you get at night? ______________________
   B. How many hours were you in bed? ______________________

5. During the past month, how often have you had trouble sleeping because you	Not during the past month (0)	Less than once a week (1)	Once or twice a week (2)	Three or more times a week (3)
A. Cannot get to sleep within 30 minutes				
B. Wake up in the middle of the night or early morning				
C. Have to get up to use the bathroom				
D. Cannot breathe comfortably				
E. Cough or snore loudly				
F. Feel too cold				
G. Feel too hot				
H. Have bad dreams				
I. Have pain				
J. Other reason (s), please describe, including how often you have had trouble sleeping because of this reason (s):				
6. During the past month, how often have you taken medicine (prescribed or "over the counter") to help you sleep?				
7. During the past month, how often have you had trouble staying awake while driving, eating meals, or engaging in social activity?				
8. During the past month, how much of a problem has it been for you to keep up enthusiasm to get things done?				
9. During the past month, how would you rate your sleep quality overall?	Very good (0)	Fairly good (1)	Fairly bad (2)	Very bad (3)

Appendix V

The Epworth Sleepiness Scale (EES)

The Epworth Sleepiness Scale

The Epworth Sleepiness Scale is widely used in the field of sleep medicine as a subjective measure of a patient's sleepiness. The test is a list of eight situations in which you rate your tendency to become sleepy on a scale of 0, no chance of dozing, to 3, high chance of dozing. When you finish the test, add up the values of your responses. Your total score is based on a scale of 0 to 24. The scale estimates whether you are experiencing excessive sleepiness that possibly requires medical attention.

How Sleepy Are You?

How likely are you to doze off or fall asleep in the following situations? You should rate your chances of dozing off, not just feeling tired. Even if you have not done some of these things recently try to determine how they would have affected you. For each situation, decide whether or not you would have:

- No chance of dozing =0
- Slight chance of dozing =1
- Moderate chance of dozing =2
- High chance of dozing =3

Write down the number corresponding to your choice in the right hand column. Total your score below.

Situation	Chance of Dozing
Sitting and reading	•
Watching TV	•
Sitting inactive in a public place (e.g., a theater or a meeting)	•
As a passenger in a car for an hour without a break	•
Lying down to rest in the afternoon when circumstances permit	•
Sitting and talking to someone	•
Sitting quietly after a lunch without alcohol	•
In a car, while stopped for a few minutes in traffic	•

Total Score = ____________________

Appendix W

Morningness-Eveningness Questionnaire (MEQ)

MORNINGNESS-EVENINGNESS QUESTIONNAIRE

For each question, please select the answer that best describes you by circling the point value that best indicates how you have felt in recent weeks.

1. *Approximately* what time would you get up if you were entirely free to plan your day?

    [5] 5:00 AM–6:30 AM *(05:00–06:30 h)*

    [4] 6:30 AM–7:45 AM *(06:30–07:45 h)*

    [3] 7:45 AM–9:45 AM *(07:45–09:45 h)*

    [2] 9:45 AM–11:00 AM *(09:45–11:00 h)*

    [1] 11:00 AM–12 noon *(11:00–12:00 h)*

2. *Approximately* what time would you go to bed if you were entirely free to plan your evening?

    [5] 8:00 PM–9:00 PM *(20:00–21:00 h)*

    [4] 9:00 PM–10:15 PM *(21:00–22:15 h)*

    [3] 10:15 PM–12:30 AM *(22:15–00:30 h)*

    [2] 12:30 AM–1:45 AM *(00:30–01:45 h)*

    [1] 1:45 AM–3:00 AM *(01:45–03:00 h)*

3. If you usually have to get up at a specific time in the morning, how much do you depend on an alarm clock?

    [4] Not at all

    [3] Slightly

    [2] Somewhat

    [1] Very much

4. How easy do you find it to get up in the morning (when you are not awakened unexpectedly)?
   [1] Very difficult
   [2] Somewhat difficult
   [3] Fairly easy
   [4] Very easy

5. How alert do you feel during the first half hour after you wake up in the morning?
   [1] Not at all alert
   [2] Slightly alert
   [3] Fairly alert
   [4] Very alert

6. How hungry do you feel during the first half hour after you wake up?
   [1] Not at all hungry
   [2] Slightly hungry
   [3] Fairly hungry
   [4] Very hungry

7. During the first half hour after you wake up in the morning, how do you feel?
   [1] Very tired
   [2] Fairly tired
   [3] Fairly refreshed
   [4] Very refreshed

8. If you had no commitments the next day, what time would you go to bed compared to your usual bedtime?
   [4] Seldom or never later
   [3] Less that 1 hour later
   [2] 1-2 hours later
   [1] More than 2 hours later

9. You have decided to do physical exercise. A friend suggests that you do this for one hour twice a week, and the best time for him is between 7-8 AM *(07-08 h).* Bearing in mind nothing but your own internal “clock,” how do you think you would perform?

[4] Would be in good form

[3] Would be in reasonable form

[2] Would find it difficult

[1] Would find it very difficult

10. At *approximately* what time in the evening do you feel tired, and, as a result, in need of sleep?

[5] 8:00 PM–9:00 PM *(20:00–21:00 h)*

[4] 9:00 PM–10:15 PM *(21:00–22:15 h)* [

[3] 10:15 PM–12:45 AM *(22:15–00:45 h)*

[2] 12:45 AM–2:00 AM *(00:45–02:00 h)*

[1] 2:00 AM–3:00 AM *(02:00–03:00 h)*

11. You want to be at your peak performance for a test that you know is going to be mentally exhausting and will last two hours. You are entirely free to plan your day. Considering only your “internal clock,” which one of the four testing times would you choose?

[6] 8 AM–10 AM *(08–10 h)*

[4] 11 AM–1 PM *(11–13 h)*

[2] 3 PM–5 PM *(15–17 h)*

[0] 7 PM–9 PM *(19–21 h)*

12. If you got into bed at 11 PM *(23 h),* how tired would you be?

[0] Not at all tired

[2] A little tired

[3] Fairly tired

[5] Very tired

13. For some reason you have gone to bed several hours later than usual, but there is no need to get up at any particular time the next morning. Which one of the following are you most likely to do?

[4] Will wake up at usual time, but will not fall back asleep

[3] Will wake up at usual time and will doze thereafter

[2] Will wake up at usual time, but will fall asleep again

[1] Will not wake up until later than usual

14. One night you have to remain awake between 4-6 AM *(04-06 h)* in order to carry out a night watch. You have no time commitments the next day. Which one of the alternatives would suit you best?

[1] Would not go to bed until the watch is over

[2] Would take a nap before and sleep after

[3] Would take a good sleep before and nap after

[4] Would sleep only before the watch

15. You have two hours of hard physical work. You are entirely free to plan your day. Considering only your internal "clock," which of the following times would you choose?

[4] 8 AM–10 AM *(08–10 h)*

[3] 11 AM–1 PM *(11–13 h)*

[2] 3 PM–5 PM *(15–17 h)*

[1] 7 PM–9 PM *(19–21 h)*

16. You have decided to do physical exercise. A friend suggests that you do this for one hour twice a week. The best time for her is between 10-11 PM *(22-23 h).* Bearing in mind only your internal "clock," how well do you think you would perform?

[1] Would be in good form

[2] Would be in reasonable form

[3] Would find it difficult

[4] Would find it very difficult

17. Suppose you can choose your own work hours. Assume that you work a five-hour day (including breaks), your job is interesting, and you are paid based on your performance. At *approximately* what time would you choose to begin?

[5] 5 hours starting between 4–8 AM *(05–08 h)*

[4] 5 hours starting between 8–9 AM *(08–09 h)*

[3] 5 hours starting between 9 AM–2 PM *(09–14 h*

[2] 5 hours starting between 2–5 PM *(14–17 h)*

[1] 5 hours starting between 5 PM–4 AM *(17–04 h)*

18. At *approximately* what time of day do you usually feel your best?

[5] 5–8 AM *(05–08 h)*

[4] 8–10 AM *(08–10 h)*

[3] 10 AM–5 PM *(10–17 h)*

[2] 5–10 PM *(17–22 h)*

[1]10 PM–5 AM *(22–05 h)*

19. One hears about "morning types" and "evening types." Which one of these types do you consider yourself to be?

[6] Definitely a morning type

[4] Rather more a morning type than an evening type

[2] Rather more an evening type than a morning type

[1] Definitely an evening type

_____ Total points for all 19 questions