PRE-LEARNING SLEEP EFFECTS ON DIRECTED FORGETTING

BY

VICTORIA A. J. KAVANAGH

A Thesis submitted to the School of Graduate Studies

In partial fulfillment of the

Requirements for the Master of Science in Experimental Psychology

Department of Psychology

Memorial University of Newfoundland

2019

Abstract

While it has long been known that many areas of cognition are negatively affected by poor sleep or an absence of sleep, the effect of sleep problems on intentional forgetting has yet to be established. A directed forgetting (DF) paradigm, in which participants are tested on items they were asked to both remember and forget, was used to compare the remembering and forgetting of participants with poor sleep quality and the presence of insomnia symptoms to those with good sleep quality. This study implemented the use of a point system in which participants were told that they would receive various points to incentivize performance in place of remember and forget instructions in a DF task with the goal of computing DF costs and benefits. The relations among memory, sleep, working memory capacity, and other demographic factors were also examined. A DF effect was found when comparing the positive point value to the negative and neutral point values, costs were found for participants without the presence of insomnia symptoms, and working memory capacity was found to only be related to remembering in the DF task. These results suggest that the current DF manipulation is useful in examining benefits but not costs in DF, the presence of insomnia symptoms does not affect performance on a DF task, and that working memory capacity does not differ between those with and without the presence of insomnia symptoms.

Acknowledgments

First and foremost, I would like to thank my supervisor, Dr. Kathleen Hourihan, without whom this would not have been possible. She is an extraordinary supervisor who was not only there whenever needed but also always seemed to find the gentlest way to criticize when warranted. Additionally, I would like to thank my supervisory committee members, Dr. Aimée Surprenant and Dr. Sheila Garland, who provided me with fantastic feedback, and my parents, who remained supportive and tried to understand my research despite not having any idea of what I was talking about. Lastly, I would like to thank NSERC, which provided the necessary funding.

Table of Contents

List of Tables

List of Figures

Lists of Appendices

Pre-learning Sleep Effects on Directed Forgetting

Forgetting is often seen as a flaw. When we cannot store information, process it, or retrieve it when wanted, we consider this to be a failure of our memory (e.g., Schacter, 2001). What is usually overlooked, however, is how beneficial forgetting can be; many everyday occurrences rely on us forgetting information in favour of other information. For example, if we park in a different parking spot every day, then we need to forget where we parked yesterday so that we are not wandering aimlessly in an attempt to find our car (Bjork, 1989). Without the ability to forget some things, we would be in a constant state of having to sort out one memory from the next.

While it may seem like forgetting is easy, it is actually a complicated, multifaceted process. Some forgetting is incidental and happens without effort (or intention), but other forms of forgetting are intentional and require effortful inhibition of memories (e.g., Wylie, Foxe, & Taylor, 2008). This type of forgetting can occur on one's own accord in an attempt to forget what is not wanted or it can be asked of by others, like in memory studies involving *directed forgetting* (DF). DF is a phenomenon where participants in an experimental setting reliably recall or recognize fewer items they are asked to forget compared to items they are asked to remember, even when asked or paid at test to ignore previous instructions to forget (MacLeod, 1998).

Over the years since DF was first used in experimental settings (e.g., Brown, 1954), the limits of the phenomenon have been examined by including numerous additional manipulations and individual difference measures. Two such individual difference measures include sleep and working memory (WM). Sleep has been incorporated into DF designs in order to examine how sleep can limit the effects of forgetting and facilitate the remembering of all information (Abel & Bäuml, 2013), while WM has been examined in order to determine whether the amount of information someone can hold in their limited WM capacity affects forgetting as a whole (Aslan, Zellner, & Bäuml, 2010). Limits to such studies to date, however, include that the role of sleep before learning has on forgetting has not been examined and that previous results involving DF and WM are not conclusive.

This thesis will examine the role that subjective sleep quality and insomnia symptoms play in the forgetting of information when problems with sleep occur prior to learning; it will also examine the role that WM plays in this. To begin, an overview of the previous literature on the areas of DF and the mechanisms behind them, sleep and cognition, and WM will be presented. Next, a connection between DF and WM will be made. This will then be followed by a discussion on sleep, what cognitive factors are impacted by problems with sleep, and what some of the causes of problems with sleep are. These discussions will conclude with a description of how previous studies have examined the role of sleep in DF experiments. Next, the study design and procedure will be explained and will be followed by the results of the study. The thesis will conclude with a discussion of the results, their implications for the previous literature, and the limitations of the study.

Directed Forgetting

One of the first experiments to use the Directed Forgetting (DF) paradigm was conducted by Brown (1954; MacLeod, 1998). In his experiment, Brown showed participants a series of paired items which included one number and one arrow. Either preceding or following the presentation of the items, participants were asked to remember one or both items; it was found that being told to remember only the numbers increased the recall of those numbers. Since this experiment, the DF paradigm has grown in both its methodology and its use in experimental settings. One development since this time is the classification of two different ways to implement being told what to remember and what to forget. These classifications are called the *list method* and the *item method* (MacLeod, 1998); the two methods differ in both methodology and how a DF effect is defined. Both methods will be described below.

List Method. In the list method, participants are shown two lists of items (MacLeod, 1998). For half of the participants, they are given an instruction, (or *cue)*, to forget the first list (L1) either as a direct instruction to forget the list or as an indirect instruction to forget under the guise of a computer glitch or mistake in presenting the list (e.g., Abel & Bäuml, 2013). This instruction to forget is called an *F* cue. The other half of the participants are asked to remember L1 for an upcoming test; this instruction to remember is called an *R* cue. Regardless of the L1 instruction, all participants are told to remember the second list (L2). At test, participants are informed that, regardless of earlier instructions involving the F cue, the test will involve all items studied and that they should try to respond to all items despite previous instructions. The list method is a between-subjects design with half of the participants in the experimental condition (L1-F cue) and half of the participants in the control condition (L1-R cue). For this method, whether DF is considered to have occurred is measured in terms of *costs* and *benefits* (MacLeod, 1998). Costs are associated with L1: If the recall of L1 items is lower for the experimental group than the control group, then DF costs are considered to have occurred. Benefits are associated with L2: If the recall of L2 items is higher for the experimental group than the control group, then DF benefits are considered to have occurred.

A recent explanation for why the DF effect occurs with the list method, specifically when the list method is paired with a recall test, is context change. As explained by Sahakyan and colleagues (Sahakyan, Delaney, Foster, & Abushanab, 2013, Sahakyan & Kelley, 2002), the lists that are studied during the list method are associated with specific study contexts. When an F cue is given, participants abandon the context that L1 was studied in, trying to forget that list, in favour of a new context to remember the new list (L2). At test, the context resembles that of L2 and not of the previously abandoned L1 as there has been no reason for them to change contexts again. This then leads to a lower recall of L1 (costs) compared to the control group, since recall is poor when there is little overlap in context between study and test (Godden & Baddeley, 1975). For those who were given an R cue, both study contexts are the same as the context at test, as the participants have no reason to abandon the study context between lists. Their recall of L1 is higher, but, since the participant has more information to remember, they end up with a lower recall of L2 compared to those who were given an F cue (benefits).

Item Method. More relevant to the current study is the item-method DF. In this method, as explained by MacLeod (1998), participants are shown a series of items (individual items can be a single item or a pair of items) and items are cued, immediately following the presentation of the item itself; half of the items are followed by an F cue and the other half of the items are followed by an R cue. Similar to the list method, this study phase is followed by a test where participants are instructed to respond to all item regardless of whether they were previously told to forget them. The item method is a within-subject design with all participants being given both cues throughout the experiment. Unlike the list method, which measures the DF effect in terms of costs and benefits, the item method typically measures the DF effect in terms of a difference magnitude. With this method, the DF effect is measured by the difference between the recall or recognition of R cued items and F cued items. If participants correctly respond to significantly more R items than F items, then a DF effect is considered to have occurred.

The most accepted explanation for why DF effects occur with the item method is selective rehearsal (Bjork, 1970). The process of selective rehearsal is heavily reliant on working memory (WM), which deals with the immediate processing and encoding, maintenance and manipulation, decision making, and response outputs based on the information (Sweller, 2011; Whitney & Rosen, 2013): Since WM has a limited capacity (Miller, 1956), it is necessary to discard irrelevant information (Baddeley, 1986; Ecker, Oberauer, & Lewandowsky, 2014). It was originally thought that the irrelevant information was just left to decay on its own (Baddeley, 2000) but now it is believed that the irrelevant information is purposefully removed from WM (Berman, Jonides, & Lewis, 2009; Ecker, Lewandowsky, & Oberauer, in press; Ecker, Lewandowsky, Oberauer, & Chee, 2010).

Considering the role that WM plays in selective rehearsal, the theory of selective rehearsal states that, when the item method is used, all items are in WM when they are first shown. After a cue is given, however, rehearsal of the item will either continue for the R cued items or stop for the F cued items (Woodward, Bjork, & Jongeward, 1973); essentially, an F cue renders the item irrelevant information that is to be discarded in favour of the R cued items. When first examined, researchers proposed that ‘non-important’ items were passively dropped from the participants data set, not to be thought of again (Elmes 1969; Fawcett & Taylor, 2008). More recently, however, some researchers now believe that an important part of selective rehearsal is the deliberate inhibition of the F cued items (Zacks, Radvansky, & Hasher, 1996). Instead of passively ignoring any item associated with an F cue in favour of the R cue, participants try to actively forget, or inhibit (though the use of this term is debated amongst researchers; for discussion, see Carr, 2007), the specified items. For example, among their wide array of experimental manipulations, Fawcett and Taylor (2008) examined how long it took participants to detect probes after being presented with a cue and found that it took longer to detect the probes after an F cue was given than after an R cue, showing that participants were more preoccupied cognitively when processing the F cue. Other support for this includes Fawcett and Taylor (2012) who had participants respond to probe words following a cue in an item method DF paradigm, and Taylor and Hamm (2016) who had participants locate words on a screen following a cue; both found that all tasks took longer following F cued items than R cued items.

Further support for the recent active rehearsal explanations behind the DF effect can also come from the use of developing technology, such as functional magnetic resonance imaging (fMRI) and electroencephalography (EEG). One important finding by Hauswald, Schulz, Iordanov, and Kissler (2011; Yang, Lei, & Anderson, 2016) has shown that, using EEG, R and F cues result in differential brain activity. For example, activity in the frontal and the parietal cortices increase more when a participant is given an F cue than when they are given an R cue, showing that not only do the different cues elicit different activity in the brain but also that there is an active process (inhibition) happening in the brain in response to the F cue. These differences are also supported by Reber et al. (2002) who used fMRI to show that R cues result in greater activity in the anterior ventral portion of the left inferior pre-frontal cortex (PFC). The PFC is most activated by tasks requiring encoding. Reber et al.'s results therefore provide further support for the idea that the different cues elicit different activity in the brain, particularly in the regions likely related to the encoding of information.

Another important finding supporting those of experimental studies compared neuroimaging of the PFC during a period of intentional forgetting to a period of incidental forgetting (Rizio & Dennis, 2013; Wylie et al., 2008). It was found that, while the PFC was not activated during incidental forgetting, there was activation during the period of intentional forgetting. This supports the idea that intentional forgetting is active and effortful and is more than just a passive decaying of information. Additional findings that support the idea of active inhibition occurring during DF include Paz-Caballero, Menor, and Jiménez (2004), who found higher levels of activity in the PFC in those who show higher levels of forgetting compared to those who show lower levels of forgetting, and Cheng and colleagues (Cheng, Liu, Lee, Hung, and Tzeng, 2011; Liu, Chen, & Cheng, 2017), who found higher levels of event-related brain potential (ERP) during the forgetting process compared to periods when the participant was not trying to remember or forget. Finally, fMRI imaging examining the PFC has also shown activity in this area to correspond to WM, further supporting the connection between WM and DF (Macmillan, 2000).

Expanding beyond the normal experiments used to test the mechanisms behind item method DF are newer alternatives to the R and F cues. One such example comes from Golding, Roper, and Hauselt (1996), who used 0% (F-cue), 50% (baseline), or 100% (R-cue) probabilities of words being tested and found results comparable to DF. That is, words with the 0% cue were remembered significantly worse than those with the 100% cue, with 50% cue words in between. Another example, and one focus of the current experiment, is Friedman and Castel's (2011) use of value directed remembering. Instead of being given a direct R or F cue, each item is given a point value and the participants are instructed that scores at test will fluctuate depending on the recall or recognition of the items associated with each point value. For example, in Experiment 3, Friedman and Castel (2011) identified items worth one of four point values: +10, +5, -5, or -10. Participants were told that if they recalled an item worth +10 points at test their score would go up by 10 points but, if they recalled an item worth -10 points, their score would go down by 10 points. In this instance, points that will increase their score should act as an R cue and points that decrease their score should act as an F cue. At test, however, participants were encouraged to recall any item they studied and were told that the scores did not matter. Friedman and Castel found that, while recall was highest for words worth +10 points, there was no difference between the other point values. As explained by Castel, Benjamin, Craik, and Watkins (2005), people choose to encode items associated with higher points, but put little value on lower point values. So, participants may have treated the +10 as an R cue, as it would maximize their score at test, but did not treat the +5 as an R cue or the negative values as an F cue, as they would not maximize their score at test. This would result in the participants ignoring the other point values rather than trying to actively remember the +5 point value or actively trying to suppress the negative point values, leading to recall being equal for all three.

Working Memory Capacity and Directed Forgetting. Working Memory Capacity (WMC) is determined by measuring how much information a person can hold and work with in WM at one time (Anderson, 1990). WMC tasks take a wide range of forms including having a participant count backwards while responding to cues on a screen (Aslan et al., 2010; Lee & Lee, 2011), recalling a specified item in a series of items (Noreen & De Fockert, 2017), and alternating between solving mathematics problems and recalling letters (Unsworth, Heitz, Schrock, & Engle, 2005). With increasing support for the role of WM in DF (e.g., Macmillan, 2000), especially in item method DF, some experimenters have chosen to use WMC tasks to further observe this role, and, overall, the results of the relation between the two is mixed.

While some support has been found for the relation between WM and DF using both the item and the list method (e.g. Aslan et al., 2010), some research actually shows that this relation is not as strong as it may seem (e.g., Lee, 2012). Even among the studies that have found a relation between WM and DF in the past, the explanations for why there is a relation have varied. For example, both Aslan et al. (2010) and Soriano and Bajo (2007; Experiments 2 and 3) presented participants with a WMC task and a list method DF task and found that participants with higher WMC showed greater levels of forgetting, meaning that participants recalled fewer items that were F cued. The explanation for both of these results was that inhibition, which is the purposeful discarding or suppressing of irrelevant information (Baddeley, 1986; Ecker et al., 2014), played a significant role. Essentially, it is thought that people with high WMC have greater control over their ability to inhibit information (Brewin & Beaton, 2002). For both Aslan et al. and Soriano and Bajo, this meant that the participants were better able to inhibit the F cued list which caused them to have lower recall at test because this list was less readily available. The participants with low WMC, however, were less able to inhibit information. Consequently, when they were presented with an F cue, they were less able to actively inhibit that list, and the list was more readily available at test.

In contrast to Aslan et al., (2010) and Soriano and Bajo (2007), Delaney and Sahakyan (2007) proposed that context is actually the reason for the connection between WMC and DF. Using a list method DF task, Delaney and Sahakyan found that participants who had a higher WMC showed greater costs to L1 than those who had a lower WMC by recalling significantly fewer items from that list. They did not, however, find any differences in L2 related to WMC. Instead of proposing that this was due to differences in inhibitory control, Delaney and Sahakyan proposed that the participants with high WMC were more context dependent and suffered to a greater extent when the context changed between L1 and L2 compared to those with lower WMC.

Marevic, Arnold, and Rummel (2018) also found results indicating that the relation between WMC and DF is unlikely to be due to inhibition. Using an item method DF task, Marevic et al. presented participants with a WMC task and found a correlation between WMC and DF. In contrast to Delaney and Sahakyan's list method study, Marevic et al. found that participants with higher WMC actually recalled significantly more R cued items than the participants with lower WMC and that there was no difference between the two groups on the recall of F cued items. They concluded that, while higher WMC led to greater memory, it did not affect the inhibition of items at all. It should be noted, however, that Marevic et al. used word pairs and cued recall in their study as opposed to single words and free recall, which are more typically used.

To further complicate the picture, some studies have found either no relation between WMC and DF or have found opposing results to the studies mentioned above. For example, Lee (2012) increased the cognitive load of participants which equated to them having a low WMC due to them being less able to work with information in WM. Then, using an item method DF task, they showed that these participants recalled fewer F cued items, resulting in greater DF compared to those who did not have their cognitive load increased. This finding was similar to that of Lee and Lee (2011) who also used an item method DF task and manipulated cognitive load. Unlike the studies mentioned previously which would suggest a positive correlation between WMC and high levels of either forgetting or remembering, the Lee and colleagues' results would suggest a negative correlation between WMC and high levels of forgetting. Overall, there is yet to be a conclusive answer to whether there is a relation between WMC and DF, especially for the item method, and, when a relationship is found what is the reason behind it.

Sleep, Sleep Quality, and its Effect on Cognition

Sleep is an important aspect of everyday life, without which normal functioning would be affected in numerous ways. Without sufficient sleep or good quality sleep, people can suffer physically (e.g., they can have increased blood pressure; Tockikubo, Ikeda, Miyajima, & Ishii, 1996; and a heightened activation of the sympathetic nervous system; Kato et al., 2000), psychologically (e.g., they can be diagnosed with a *Diagnostic and Statistical Manual of Mental Disorders* (DSM-V; American Psychiatric Association, 2013) disorder such as insomnia, which is characterized by difficulties in getting to sleep, staying asleep, or waking up early) or cognitively (e.g., they can have decreased alertness and trouble remembering new information; Diekelmann & Born, 2010; Killgore & Weber, 2014).

Although it may seem like sleep is something that just effortlessly occurs every night, it is actually a complicated process that gets more complicated when considering all of the different ways insufficient sleep or poor sleep quality can be defined. For example, one definition of sleep quality includes at least six different sub-definitions (Buysse, Reynolds, Monk, Berman, & Kupfer, 1989; Moon, Phelan, Lauver, Bratzke, 2015). In this definition, sleep quality can be determined by the general quality of the sleep, the duration (how long the sleep is), the sleep latency (how long it takes to get to sleep), how efficient the sleep is (how long sleep is relative to the amount of time spent in bed), whether sleep is disrupted throughout the night, and whether medication is required to aid in getting to sleep. In comparison, the Patient-Reported Outcomes Measurement Information System- Short Form (PROMIS-SF; Yu et. al, 2012), which is a scale that measures sleep disturbance and impairment due to problems with sleep, measures disturbance in terms of sleep quality itself, individual aspects of the six sub-definitions, and other things not listed by Buysse and colleagues. For example, the sleep disturbance scale asks about sleep quality, falling asleep, staying asleep, and getting enough sleep, all independent of one another, among other things. Given how varied these definitions can be both within and between who is defining it, it is easy to see just how complicated it is to define “problems with sleep.”

What Does Sleep Affect? The effects that sleep has on cognition are most relevant to the current study. Without proper sleep, individuals suffer in a wide range of cognitive areas compared to those who experience sufficient sleep and good quality sleep (Maquet, 2001; Pilcher & Huffcut, 1996; Stickgold, 2005). Deterioration has been seen both in studies in which participants undergo total sleep deprivation, which is a prolonged continuous period of wakefulness (e.g., staying awake all night; Babkoff, Mikulincer, Caspy, Kempinski, & Sing, 1988), and in studies in which participants undergo repeated partial sleep deprivation, which is repeated periods of not getting enough sleep (Belenky et al., 2003). Both total and repeated partial sleep deprivation have been found to have similar effects on cognition (Frenda & Fenn, 2016). Some of the areas that suffer include alertness and attention, which is severely lowered after 24 hours of total sleep deprivation (Durmer & Dinges, 2005; Joo, Yoon, Koo, Kim, & Hong, 2012), control over language and verbal memory, which is poor when sleep quality is poor (Lo, Groeger, Cheng, Dijk, & Chee, 2016; Tsapanou et al., 2017), control over inhibition, which decreases during sleep deprivation (Killgore & Weber, 2014; Lim & Dinges, 2008), and response time, which increases when tired (Alhola & Polo-Kantola, 2007). General executive functioning is also found to suffer after short sleep duration (Tsapanou et al., 2017) and longer sleep latency (Nebes, Buysse, Halligan, Houck, & Monk, 2009).

Although study results are not conclusive, WM is also thought to be affected by poor sleep (Ballesio, Cerolini, Ferlazzo, Cellini, & Lombardo, 2018; Lim & Dinges, 2010). One finding in support of this is that areas of the brain that are impacted by sleep loss are also important to WM, specifically the PFC which shows decreased levels of activation following sleep deprivation (Chee & Choo, 2004; Harrison & Horne, 2000a; Horne, 1993; Thomas et al., 2009, Whitney & Rosen, 2013; Wu et al., 1991). Petrov, Lichstein, and Baldwin (2014) also found a trend towards a correlation between occurrences of sleep disorders and lower WMC. Despite these findings, other studies have not found any correlation between WM or WMC and poor sleep (Drummond et al., 2012; Feld, Weis, & Born, 2016). One explanation for why it is hard to specify whether WM is affected by poor sleep is that studies tend to use ‘working memory’ as a general term even though WM is made up of a variety of components and can be measured in several different ways (Whitney & Rosen, 2013). As previously stated, WM involves processing, maintenance, decision making, and response outputs. When examining ‘working memory,’ however, most studies will use a measure that measures one component listed but state that they are looking at WM as a whole. It is possible, then, that results regarding WM and poor sleep are inconsistent due to the inconsistencies in measuring and labelling WM. Additionally, the same might also be true for the use of terms ‘sleep’ and ‘poor sleep,’ which are also complex and made up of a variety of components, leading to further inconsistences in the results due to the inconsistencies in the labelling and measuring of the terms.

Poor sleep also affects the encoding and storing of information into long term memory. Essentially, good overall sleep is necessary for the learning of new information and subsequent remembering of that information (Killgore & Weber, 2014). If sleep is insufficient in some way (whether it be quality or quantity) before learning information (also called the ‘pre-learning’ period), then the brain is less equipped to acquire new information (Walker & Stickgold, 2006). Support for this comes from the finding that both forms of sleep deprivation lead to a decrease in activity in the posterior hippocampus, which affects subsequent remembering (Alberca-Raina, Canteru, & Atienza, 2015; Drummond et al., 2000; Van Der Werf et al., 2009; Yoo, Hu, Gujar, & Jolesz, 2007). This is also supported by Harrison and Horne (2000b) who found that participants showed poor source recognition if they were sleep deprived before learning during the experiment. If sleep is insufficient in some way after learning (also called the ‘post-learning’ period), however, then the consolidation of that information is hindered (Diekelmann & Born, 2010); this consolidation is usually found to be affected by the overall amount of sleep obtained (Lahl, Wispel, Willigens, & Pietrowsky, 2008). After learning information, any interference can degrade how well that information is stored, but, if learning is followed by an interference-free period, then that information is stored more closely to its original form (Della Sala, Cowan, Beschin, & Perini, 2005); sleep provides this interference-free period, and, without proper sleep overall, this period is compromised.

What Affects Sleep? While it is important to understand the effect that poor sleep has on individuals, it is also important to understand what specifically causes poor sleep, regardless of how it is defined. One cause is thought to be the use of stimulants such as caffeine. Caffeine, which can be consumed via coffee, energy drinks, and even chocolate (Fredholm, Battig, Holmen, Nehlig, & Zvartau, 1999), is usually used by individuals to help them stay awake and has also been found to negatively impact overall sleep quality (Amlander & Fuller, 2005; Hersher & Chervin, 2014; Kerpershoek, Antypa, & Van den Berg, 2018). Excessive use of caffeine can then lead to issues with sleep latency, duration, efficiency, and even time spent in rapid eye movement (REM) sleep (Clark & Landolt, 2017; Duffy, Rimmer, & Czeisler, 2001). In addition to caffeine, sleep is also affected by the use of other substances such as alcohol (Amlander & Fuller, 2005; Hershner & Chervin, 2014) and drugs (Clegg-Kraynok, McBean, & Montgomery-Downs, 2011).

Sleep quality is also impacted by gender and sex. In terms of gender, the social roles that accompany female-identified individuals have been associated with poor sleep quality (de Souza, de Oliveira, de Sousa, & de Azevedo, 2018). As explained by de Sousa et al., female-identified individuals are expected to work outside of the home but also be the primary caregiver at home; this leads these individuals to be more tired throughout the day and also not get as much sleep at the same time. In terms of sex, females are found to have both a higher prevalence of insomnia disorder and anxiety disorders, which is also associated with sleep disturbance (Goldstein-Piekarski et al., 2018; Kessler et al., 1994; Neckelmann, Mykletun, & Dahl, 2007; Zhang & Wing, 2006). The consequences of poor sleep are greater for females than males because homeostatic pressure increases more quickly after sleep deprivation in females, leading to them being affected more by sleep deprivation than men (Tonetti, Fabbri, & Natale, 2008).

Sleep and Directed Forgetting

Unintentional forgetting can be caused by a number of different things, including interference with accessing information at retrieval (Brown, Neath, & Chater, 2007) and interference during the storing and consolidation of information (Wixted, 2010), which, as stated previously, can be prevented by proper sleep (Atherton et al., 2016; Diekelmann & Born, 2010). Intentional forgetting is commonly studied using DF, and researchers have examined how sleeping can influence DF effects in the list method. Generally, the effects of sleeping following learning have been examined in order to test whether the consolidation aspect of sleep can prevent the forgetting of information. For example, Abel and Bäuml (2013) had four groups of participants study at either 9:00 am or 9:00 pm and then test twelve hours later at either 9:00 pm or 9:00 am, respectively. Half of the participants experienced twelve hours of wakefulness in between study and test (am-pm) while the other half experienced up to twelve hours of sleep (pm-am), allowing for them to benefit from interference-free time to consolidate the studied information. These two groups were divided further with half of the participants receiving an F cue after L1 and half receiving an R cue. While the participants who stayed awake for the full twelve hours showed significant DF costs (with the group who were given the F cue recalling significantly fewer L1 items than those who were given the R cue), the participants who slept did not show any differences in memory regardless of what cue they received. This showed that the period of sleep protected the L1 items from being forgotten. This result was replicated by Saletin, Goldstein, and Walker (2011) who had participants study and test at the same times but allowed one group to nap for up to 100 minutes during the retention interval and had the other group stay awake for the duration. Reduced intentional forgetting when sleep follows learning has also been observed in a Think/No-Think experiment (Fischer, Diekelmann, & Born, 2011), and an experiment using paired association (Wilhelm et al., 2011).

Limitations and Theoretical Motivations

A limitation to the previous studies that have examined the effect of sleep on forgetting using DF is that they do not examine the whole picture. First, these studies have examined post-learning sleep (the effect that sleep has on forgetting after learning has already occurred; e.g., Abel & Bäuml, 2013; Rauchs et al., 2011; Saletin et al., 2011); no study has examined the effect of pre-learning sleep on forgetting (how prior sleep quality affects learning and subsequent forgetting). While it is important to know that post-learning sleep can protect against forgetting (Abel & Bäuml, 2013), it is also important to know whether prior self-reported sleep quality will affect forgetting, especially for those who can benefit from forgetting (e.g., those who experience a trauma or jurors who are instructed to forget testimony in court). Second, these studies have manipulated sleep between participants (e.g. Hupbach, 2017; Wilhelm et al., 2011). Participants are randomly assigned to whether they will experience a period of wakefulness or a period of sleep; none of the past studies have tested participants based on natural sleep quality as opposed to manipulated sleep quality. It is important to examine participants without manipulating sleep quality because this manipulation is unrepresentative of how learning normally occurs (Hershner & Chervin, 2014). Finally, the majority of the DF experiments that deal with sleep use the list method as opposed to the item method, so the effect of sleep on factors specifically related to the item method (e.g., selective rehearsal; Bjork, 1970) is unknown.

The Present Study

The current study examined how naturally occurring self-reported sleep quality influences item method DF by focusing on the issue of pre-learning sleep effects, as opposed to post-learning sleep effects, in order to examine whether non-manipulated prior sleep quality (in the form of sleep disturbance and impairment due to problems with sleep, and the presence of insomnia symptoms) affects subsequent forgetting. To clearly examine this, however, this study used a point system variant of the item method DF paradigm (Friedman & Castel, 2011), modifying it by adding in a value of 0 points, to allow for a baseline condition in an item method DF experiment (see also Foster & Sahakyan, 2012, who used all positive point values as a baseline condition). By having three point values (a negative point value, a positive point value, and a neutral point value), costs can be measured by comparing the negative point value to the neutral point value and benefits can be measured by comparing the positive point value to the neutral point value. Sleep measures (SCI, Espie et al., 2014; PROMIS, Yu et al., 2012) were used to determine how pre-learning sleep influences DF costs and benefits. The DF paradigm was reported first below in order to lay the framework for how the sleep measures were incorporated into the analysis. In addition, the relation between self-reported sleep and WMC and the relation between WMC and DF was examined. Lastly, though not pertinent to the aims of this study, other information was also examined in order to control for their potential effects on the main measures used. This information included gender, sex, age, three types of medication use, and psychological distress.

Hypotheses and Predictions. For the DF task, it was expected that a DF effect would be observed. Unlike Friedman and Castel (2011), who found there only to be an effect of the highest point value, it was expected that both costs and benefits would be found. Specifically, it was thought that recognition for the highest point value (the positive point value) would be higher than the other two point values (with the difference between this value and the neutral point value creating a benefit for the positive point value). It was also thought, however, that recognition for the neutral point value would be higher than recognition for the negative point value creating a cost to that point value. This result was expected because it was thought that the neutral point value would act as a cue for the participant to neither actively rehearse nor actively inhibit the items associated with that point value. The items associated with the positive point value, however, would be actively rehearsed and encoded, and recognition would therefore be higher for these items. Similarly, the items worth the negative point value would be actively inhibited and the recognition for the items would suffer compared to the neutral items.

For the data reflecting the sleep quality and the presence of insomnia symptoms (which were collected using self-report questionnaires), it was expected that each would affect the DF results in one or two ways. Specifically, it was believed that, for those with the worst sleep disturbance and impairment in daytime functioning due to problems with sleep or with the presence of insomnia symptoms, there would be lowered benefits and/or reduced costs in comparison to those who do not exhibit problems with sleep based on our measures. First, since problems with sleep are believed to affect the ability to encode information (Walker & Stickgold, 2006), participants who exhibit problems on either of the used sleep questionnaires would recognize fewer of the items associated with the positive point values and have lowered, if any, benefits for these items. This could be caused by the participants' reduced ability to encode these items, leading to ineffective rehearsal of the items. Second, since problems with sleep impair the ability to inhibit information (Killgore & Weber, 2014; Lim & Dinges, 2008), and selective rehearsal relies on the ability to inhibit information (Fawcett & Taylor, 2008; Zacks, Radvansky, & Hasher, 1996), those who have problems with their sleep would not have any costs to the items worth negative points. If participants cannot actively inhibit the encoding of these items, then there would be higher recognition of these items at test and the costs for these items should be reduced.

For the role of WMC in the current study, there were a number of possible relations to be found. First, to the extent WMC played a role in selective rehearsal, there would be a positive correlation found between WMC and DF costs and benefits (Aslan et al., 2010; Macmillan, 2000), regardless of sleep quality or presence of insomnia. Second, to the extent that WM was affected by sleep, there would be a negative correlation found between WMC and the sleep measures (i.e., higher scores related to greater problems with sleep being associated with reduced WMC). Third, the relation between WMC and either DF or sleep could be unique making it not present in one of the first two ways described above. Last, WMC might not have been found to be correlated with either DF or sleep confirming studies that have found similar results (e.g., Drummond et al., 2012; Lee, 2012).

Lastly, in regard to the other information examined, it was expected that the measures would be related to the sleep measures as indicated by the discussion above. That is, it was expected that participant performance on the sleep measures would be related to gender, sex, and stimulant use to stay awake, as well as medication use to go to sleep and psychological distress.

It was not expected, however, that any of these other measures would be related to the DF measure, the WMC measure, or to how either DF or WMC relate to the sleep measures.

Method

Participants

Participants were undergraduate students recruited from Memorial University of Newfoundland (where ethics approval was obtained from the Interdisciplinary Committee on Ethics in Human Research). Ninety-six participants were tested in total but only 94 were used in the analyses. Two participants had to be excluded due computer errors. Of the 94 participants used, 71 identified as female while 23 identified as male. All participants identified with the sex to which they were born except for two, who declined to disclose; these participants were excluded from analyses involving sex. The average age of the participants was 20.67 years old (*SD* = 3.46). For relevant analyses, participants were split into two groups based on how they scored on the SCI. In line with Espie et al. (2014), participants scoring 16 or less were considered as having insomnia symptoms (31 participants) and those scoring higher than 16 were considered to not have insomnia symptoms (63 participants). Factors such as race, nationality, and first language were not recorded for the purposes of this experiment. While most participants were recruited through the university's Psychology Research Experience Pool (PREP), where they received one course credit (equivalent to one added percentage on total course mark, in most cases) for their participation, several participants were recruited through a paid recruitment process and were paid $10.00 for their time.

Justification of Sample Size. Total sample size was determined based on the between-subjects factor of whether participants have insomnia (yes or no), as opposed to the within-subject DF factors, due to the former requiring more participants for the proper power analyses.

Since no previous study involving sleep and DF has examined naturally occurring problems with sleep, we based our minimum sample size on the average number of participants per group in several similar studies that have manipulated sleep (Abel & Bäuml, 2011; Cairney, Durrant, Musgrove, & Lewis, 2011; Hupbach, 2017; Rauchs et al., 2011; Saletin, Goldstein, & Walker, 2011; Wilhelm et al., 2011). The average number of participants per group was calculated to be 29. The sample size was confirmed using a G*Power (Erdfelder, Faul, & Buchner, 1996) A-priori, Analysis of Variance (ANOVA) analysis. Using an effect size *f* of 0.4, calculated from the main effects of the previous studies, and setting power to be .95, the minimum total sample size was calculated to be 58 (29 per group). Actual sample size exceeded this minimum and led to unequal sample sizes due to the experimenters having no control over whether the participants entering the lab had naturally occurring problems with sleep or not; testing continued until each group had at least 29 participants.

Tasks and Materials

Cognitive Tasks. For the DF task, 180 English words (see Appendix A) were obtained from the MRC Psycholinguistic Database (Coltheart, 1981). The words ranged in length from four to seven letters (M = 4.70, SD = 0.77). The frequency of the words ranged from 50 to 80 occurrences per million words in the English language (M = 62.46, SD = 8.49; Kučera & Francis, 1967). Finally, the words were equated based on familiarity, imageability, concreteness, and meaningfulness. Words were vetted and were excluded from use if they were proper nouns or were too similar to other words used either semantically or phonetically. E-Prime software version 3.0 (Psychology Software Tools, Pittsburgh, PA) was used to present stimuli and record responses in the DF task.

The DF task used a within-subject design. Participants were told that each word was worth either +10 points, 0 points, or -10 points for an upcoming test with the intent that participants would try to remember the words valued at +10 points, forget the words worth -10 points, and that the words worth 0 points would act as a control. One measure taken during this experiment was the recognition hit rate of words at each point value. These hit rates were used to compute DF costs and DF benefits. Costs were calculated by subtracting the hit rate for words worth -10 points from the hit rate for words worth 0 points. Benefits were calculated by subtracting the hit rates for words worth 0 points from the hit rates for words worth 10 points.

The WMC task in this experiment was an automated working memory capacity task called the Automated Operation Span Task (AOSPAN). This task was programmed using the E-Prime Software by Unsworth et al. (2005) and was successfully validated by Unsworth et al. against other WMC measures such as the paper based versions of the OSPAN task, the Raven Progressive Matrices, and the RSPAN task. This experiment used the partial WMC score given by the task as recommended by Unsworth et al.

Four questionnaires were also used in the experiment, as described below.

Demographics questionnaire. This questionnaire, found in Appendix B, was constructed by the experimenters. This questionnaire consisted of six questions that asked, in order, the participant’s age, gender, sex, whether they were taking medications that could affect cognitive functioning, whether they were taking any substances (over-the-counter or prescribed) to help them sleep, and whether they were taking any substances (e.g., caffeine) to help them stay awake. The first three questions had open-ended responses while the last three were *yes/no* responses. All responses were coded numerically so that they could be used in correlational analyses. For gender and sex, female was coded as 1 while male was coded as 2; no response was coded as 0. For the medication use, taking the medication was coded as 1 while not taking the medication was coded as 2.

Kessler Psychological Distress Scale (K10; Kessler et al., 2003). This scale, found in Appendix C, screens participants for non-specific psychological distress and is made up of ten questions that assess how often participants feel things such as nervous, hopeless, tired, depressed, and worthless. It was used by the experimenters to observe the possibility of severe psychological distress, and its affects, among participants. Responses to each question are on a scale of 1-5; responses range from *None of the time* to *All of the time*. Higher scores are indicative of a greater chance of the presence of a mental illness. Scores can be used continuously or can be categorized into one of four categories: Likely to be well (a score of 10-19), likely to have a mild disorder (20-24), likely to have a moderate disorder (25-29), and likely to have a severe disorder (30-50). This study used the scores continuously. Reliability and validity for this scale was assessed against other validated measures of psychological stress by Kessler et al.; overall, they found that the K10 was a statistically significant predictor of psychological distress.

Sleep Condition Indicator (SCI; Espie et. al, 2014). The SCI, found in Appendix D, is a subjective measure of insomnia that was developed based on the DSM-V (American Psychiatric Association, 2013) criteria for insomnia. While it is a subjective measure, the SCI has been found to have moderate correlations with objective measures of insomnia such as actigraphy and the Consensus Sleep Diary (Wong et al., 2017). Wong et al. also found the SCI to be able to discriminate between participants having insomnia or not as effectively as structured clinical interviews used in the diagnosis of insomnia. The SCI contains eight questions: four questions ask about sleep quality, such as how long it takes to fall asleep and how one would rate their sleep quality, three ask about how sleep affects mood, energy, concentration, and productivity, and one asks about how long one has had a problem with their sleep. Responses are scored on a scale of 0-4 and vary depending on the question asked: some questions are associated with time frames (e.g., 0 means *greater than 60 minutes* and 4 means *between 0 and 15 minutes*), while some are associated with quality ratings (e.g., 0 means *Very much* and 4 means *Not at all*). A total score of 16 or less is indicative of possible insomnia (and can be used as such as a categorical variable), though scores can also be used continuously with lower scores indicating a greater presence of insomnia symptoms (both were used in this study). Reliability and validity were assessed using five randomized controlled trials by Espie et al (2014).

PROMIS-SF (Yu et. al, 2012). The PROMIS-SF, found in Appendix D, uses subjective measures of sleep disturbance (SD) and sleep-related impairment (SRI). It contains two item banks, one for each of what it measures. Both the SD and the SRI item banks contain eight questions. The SD asks about things such as whether the person's sleep within the last seven days was refreshing or restless, whether the person had trouble sleeping or got enough sleep, and what their sleep quality was. Responses are rated on a scale of 1-5 (or 5-1 on reverse scored questions); responses range from either *Not at all* to *Very much*, *Never* to *Always*, or *Very poor* to *Very good* depending on the format of the question. For scoring, all points are added up and transformed into a T-Score; higher scores are indicative of greater levels of sleep disturbance. The SRI asks about whether sleep causes problems such as irritability, problems during the day, daytime sleepiness, and not being alert upon waking. Responses are also rated on a scale of 1-5 (or 5-1 on reverse scored questions), with responses ranging from *Not at all* to *Very much*. Again, for scoring, points are added up and transformed into a T-Score; higher scores are indicative of greater levels of sleep related impairment. T-scores for both the SD and the SRI can be used continuously or can fall into one of four categories: Normal levels of disturbance/impairment (scores of 20-55), mild levels of disturbance/impairment (55-60), moderate levels of disturbance/impairment (60-70), and severe levels of disturbance/impairment (70-80). This study used the scores continuously. Reliability and validity of these short form item banks were successfully assessed against the long form of each by Yu et al (2012).

Procedure

Participants first completed an informed consent form upon entering the lab. The experimenter then provided a brief verbal overview of the experiment and showed the participant to the computer to begin the first task. All participants completed the tasks in the same order: 1) AOSPAN; 2) DF; and, 3) Questionnaires (demographics, K10, SCI, and PROMIS-SF).

The first part of the experiment was the Automated Operation Span (AOSPAN) task which began with three practice sessions. The first practice session consisted of three trials of studying two letters, one by one, and then selecting which letters were presented in the correct presentation order. The second practice session consisted of 15 trials of looking at simple math equations, followed by verifying whether the equation was true or false. The final practice session combined the two previous sessions. In three separate trials, participants were presented with single letters to remember immediately after verifying an equation; each trial consisted of two equations and two letters, followed by selecting which letters were presented in the correct order. After all three practice sessions had concluded, participants were presented with on-screen instructions for the main task which was an extended version of the third practice trial. In total, there were 15 trials with the number of letters ranging from three to seven in length, with there being three trials of each length.

The AOSPAN task was then followed by the DF task. Participants were first presented with instructions on the screen; these instructions stated that the participant would be shown a series of words, one at a time, to study for an upcoming task and that these words would be followed by a point value. If the participant remembered these words at test, their score would either increase, decrease, or stay the same based on the point value assigned to that specific word. They were then asked to press the spacebar when they were ready to begin the study session. All 90 words in this session were presented in the centre of the screen in black, 22-point *Courier New* font for two seconds. The presentation of each word was followed by a point value in the centre of the screen for two seconds: +10, which was shown in green, 0, which was shown in black, or -10, which was shown in red. The 90 words were selected at random out of the 180-word pool by the program. The point value for each word was also assigned at random, with the only condition being that there were 30 words per point value. After the presentation of the words concluded, participants were given one minute to work on a connect-the-dot distraction task, using paper and pencil, to momentarily prevent rehearsal of the study words between study and test. This task was then followed by a recognition test in which participants were presented with all 180 words (i.e., 90 studied and 90 new), one at a time. Participants were instructed to press the “z” key if they were shown a word that had not been studied previously and to press the “m” key if they were shown a word that had been studied previously. They were also instructed to disregard the earlier information that there was a score for the test and to respond to words that were associated with 0 points and -10 points as being previously studied as they would not be penalized for it.

After participants were finished the DF task they were handed paper copies of the demographics questionnaire, the K-10, and the SCI and PROMIS-SF questionnaires, and were asked to either circle their response or check a box next to their response where applicable. Completion of these questionnaires was followed by a quick debriefing period and additional verbal consent for use of data where deception was used during the DF task.

Analytic Strategy

First, we analyzed the data related to DF using Analysis of Variances (ANOVAs). We began by looking at the DF performance for point value, and then for benefits and costs. Next, we split the participants into two groups (whether or not they had the presence of insomnia symptoms) based on the SCI measure and compared how the groups performed on the DF task, again by using ANOVAs. Third, we looked at the relation between the DF measures and the other measures using correlations, starting with WMC and concluding with the demographic questionnaire. Fourth, we analyzed the sleep measure data using correlations. We looked at the relation between the SCI and the PROMIS measures and then the relation between these measures and the other measures. Fifth, we analyzed the WMC data in relation to the data collected from the demographic questionnaires using correlations. Finally, we analyzed the relations among the data collected from the demographic questionnaires using correlations. For the main analyses of the results, an alpha level of .05 was used.

Results

Directed Forgetting

Figure 1 shows the recognition hit rate for the different point values. Hit rates were analyzed using a one-way repeated measures ANOVA, with the point value (10 vs. 0 vs. -10) as the within-subject factor. The main effect of point value was significant, $F(2, 186) = 110.207$, $MSE = 0.013$, $p < .001$, $\eta^2_p = .542$. Follow up paired sample t-tests conducted on point value found that all three values differed significantly from one another. Point value +10 differed from both point value -10 (*t* (93) = 10.791, *p* < .001, Cohen's *d* = 1.113), and point value 0, (*t*(93) = 12.968, *p* < .001, Cohen's *d* = 1.338), with hit rates being higher for the +10 point value. Point value -10 also differed from point value 0, (*t*(93) = 2.487, *p* = .015, Cohen's *d* = 0.257), with hit rates being higher for the -10 point value. The mean false alarm rate was .275 (*SE* = .016).

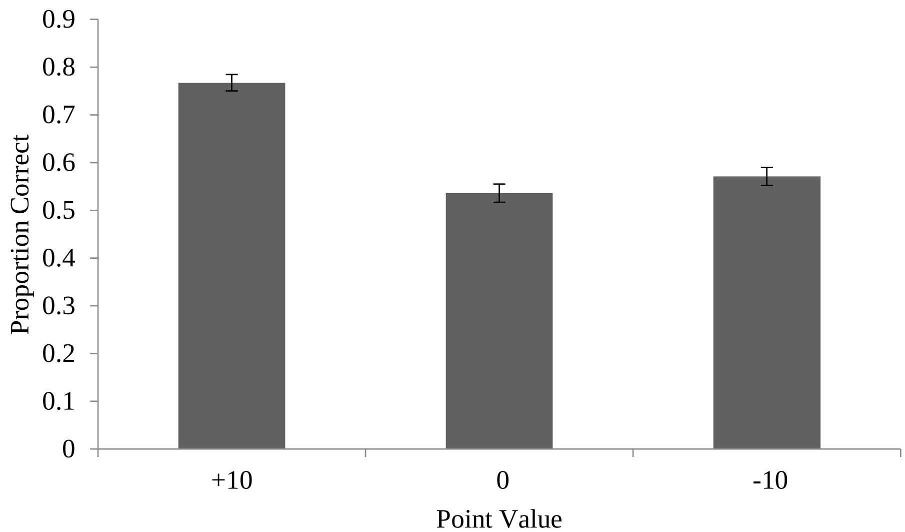

Figure 1. Mean proportion correct for the three point values. Error bars represent standard error of the mean.

Figure 2 shows the computed DF benefits and costs. A paired sample t-test showed that, benefits differed significantly from costs, *t* (93) = 10.066, *p* < .001, Cohen's *d* = 1.000. One sample t-tests on benefits and costs showed that both were significantly different from zero; for benefits, *t* (93) = 12.968, *p* < .001, Cohen's *d* = 1.337, and for costs, *t* (93) = -2.487, *p* = .015, Cohen's *d* = -0.256.

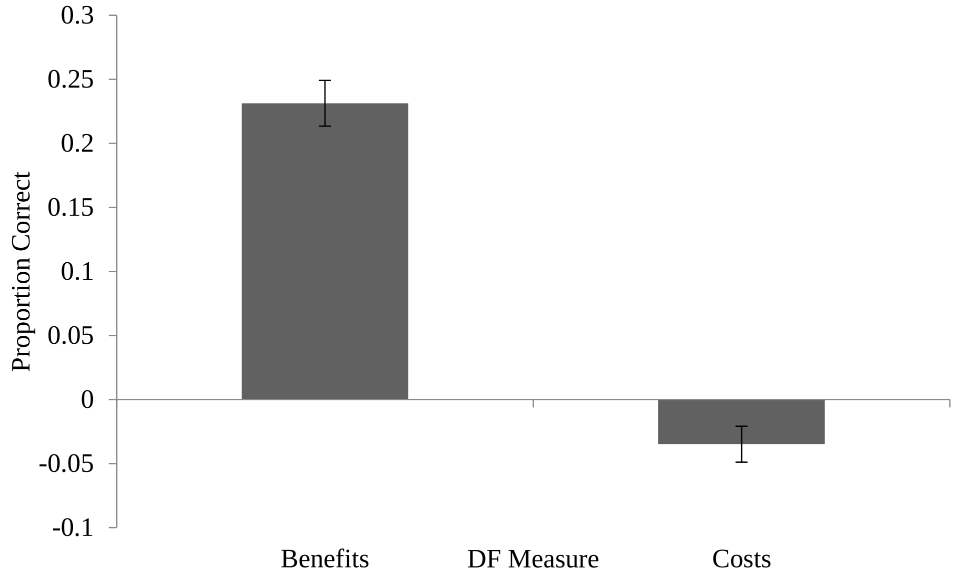

Figure 2. Mean proportion correct for benefits and costs. Error bars represent standard error of the mean.

Table 1 shows the relations between each point value as well as the benefits and costs, and other questionnaire measures (discussed below). Of note, Table 1 shows that the +10 point value was significantly positively correlated with both the 0 and -10 point values, and that the -10 point value was significantly positively correlated with the 0 point value. It also showed that benefits were significantly negatively correlated with costs, with higher differences between the +10 and 0 point values related to lower differences between the 0 and -10 point values. Based on these correlations, other significant correlations among the point values and costs and benefits were expected. No other correlations for the DF measures were significant.

Table 1

Correlations Among the DF Measures, WMC, and the Other Measures

	+10	-10	0	Benefit	Cost	WMC	SCI	K10	Gen.	Sex	Age	MCI	MS	MW
+10	1.00													
-10	.492**	1.00												
0	.520**	.726**	1.00											
Benefit	.406**	-.303**	-.507**	1.00										
Cost	.047	-.355**	.386**	-.368**	1.00									
WMC	.467**	.055	.065	.379*	.014	1.00								
SCI	-.145	-.076	-.196	.071	-.164	.055	1.00							
K10	.191	.123	.229*	.-.061	.147	.076	-.517**	1.00						
Gen.	-.032	.055	-.080	.055	-.182	.141	.075	.019	1.00					
Sex	-.011	.089	-.032	.023	-.163	.121	.056	.028	.951 **	1.00				
Age	-.106	-.046	-.117	.023	-.097	-.196	-.057	-.226*	.040	.013	1.00			
MCI	-.032	.170	.055	-.090	-.153	-.194	.170	-.172	.067	.049	.043	1.00		
MS	.081	-.025	-.003	.081	.029	.077	.388**	-.123	.047	.032	-.215*	.257*	1.00	
MW	.081	.084	.035	.040	-.065	.044	.078	-.203*	.137	.176	.034	.140	.017	1.00

Note: * denotes $p < .05$; ** denotes $p < .01$; Gen. = Gender, MCI = Medications Impacting Cognitive Functioning, MS = Medications Inducing Sleep, MW = Medications Used to Stay Awake

Table 2

Descriptive Statistics for Select Continuous Measures

	Mean	Median	Mode	Standard Deviation	Minimum	Maximum	Skew	Standard Error Skew	Kurtosis	Standard Error Kurtosis
SCI	19.10	19.50	25.00	6.81	5.00	32.00	-0.210	0.249	-0.800	0.493
Dist.	51.50	51.65	47.90	6.95	28.90	67.50	-0.308	0.249	0.484	0.493
Imp.	58.10	58.20	51.60	8.07	30.00	75.00	-0.472	0.249	0.505	0.493
WMC	56.10	58.50	67.00	13.50	14.00	75.00	-1.08	0.249	0.640	0.493
K10	23.30	22.00	16.00	7.62	11.00	42.00	0.478	0.249	-0.614	0.493
Gen	1.24	1.00	1.00	0.43	1.00	2.00	1.207	0.249	-0.555	0.493
Sex	1.22	1.00	1.00	0.47	0.00	2.00	0.683	0.249	0.032	0.493
Age	20.70	20.00	20.00	3.46	17.00	41.00	3.18	0.249	14.70	0.493
MCI	1.93	2.00	2.00	0.26	1.00	2.00	-3.295	0.249	9.046	0.493
MS	1.94	2.00	2.00	0.25	1.00	2.00	-3.627	0.249	11.395	0.493
MW	1.53	2.00	2.00	0.50	1.00	2.00	-0.130	0.249	-2.027	0.493

Note: Gen. = Gender, MCI = Medications Impacting Cognitive Functioning, MS = Medications Inducing Sleep, MW = Medications Used to Stay Awake; t-scores were used for the Disturbance and Impairment measures; see coding for demographic measures above

Directed Forgetting and Sleep Measures. Table 1 also presents the relations among the SCI and the DF measures; none of these correlations were significant. Before examining how sleep, as defined by our sleep measures, affected performance on the DF task, participants were split into two groups based on the SCI (used as a categorical variable) as described above; see Table 2 for descriptive statistics. Participants' memory for the study words depending on point value based on this split is shown in Figure 3 below. Hit rates were analyzed using a 2 (insomnia: yes vs. no) x 3 (point value: +10 vs. 0 vs. -10) mixed factors ANOVA, with point value again as the within-subject factor and insomnia as a between-subjects factor. The main effect of point value was again significant, $F(2, 184) = 95.270$, $MSE = 0.013$, $p < .001$, $\eta^2_p = .509$. However, neither the main effect of insomnia, $F(1, 92) = 0.300$, $MSE = 0.069$, $p = .585$, $\eta^2_p = .003$, nor the interaction, $F(2, 186) = 1.686$, $MSE = 0.013$, $p = .188$, $\eta^2_p = .018$, were significant.

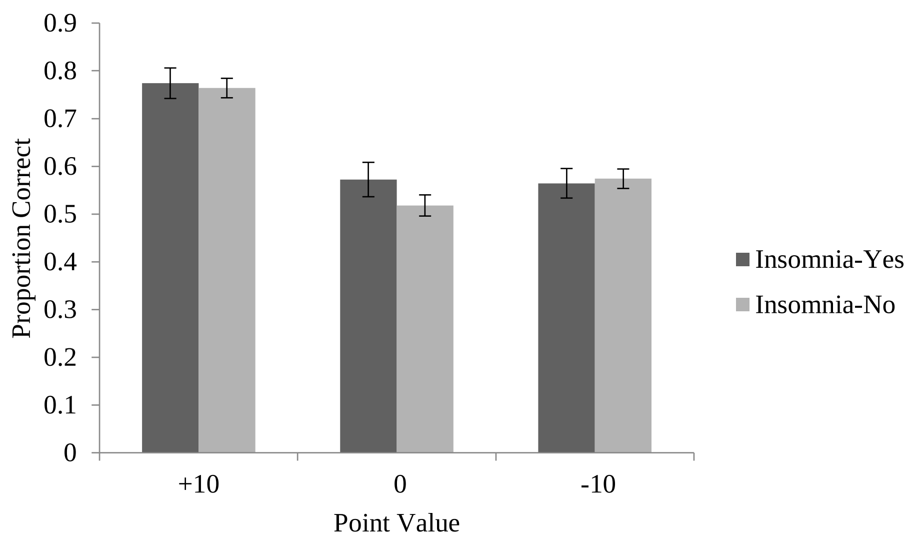

Figure 3. Mean proportion correct for the three point values split between participants with insomnia symptoms and those without. Error bars represent standard error of the mean.

Figure 4 shows benefits and costs split between participants with insomnia symptoms and those without (again using the SCI as a categorical variable). Means were analyzed using a 2 (insomnia: yes vs. no) x 2 (DF Measure: benefits vs. costs) mixed factors ANOVA. DF Measure was the within-subject factor and insomnia was the between-subjects factor. Again, there was a main effect of DF Measure, $F(1, 92) = 79.979$, $MSE = 0.032$, $p < .001$, $\eta^2_p = .465$, but no main effect of insomnia, $F(1, 92) = 0.004$, $MSE = 0.016$, $p = .615$. There was also a trend towards a significant interaction, $F(1, 92) = 3.798$, $MSE = 0.032$, $p = .054$, $\eta^2_p = .040$. Follow up independent samples t-tests on the interaction showed no differences between those with and without insomnia symptoms for benefits ($t(92) = -1.167$, $p = .246$, Cohen's $d = -0.256$), but did show significant differences between those with and without insomnia symptoms for costs ($t(92) = 2.179$, $p = .032$, Cohen's $d = 0.478$); those without the presence of insomnia symptoms showed significantly smaller costs (indeed, a negative cost) relative to those with the presence of insomnia symptoms.

Lastly, differences in false alarms for participants with and without the presence of insomnia symptoms were analyzed using an Independent Samples t-test. This analysis showed that false alarms for participants without the presence of insomnia symptoms ($M = .284$) did not differ from false alarms for participants with insomnia symptoms ($M = .257$); $t(92) = .787$, $p = .433$, Cohen's $d = 0.173$.

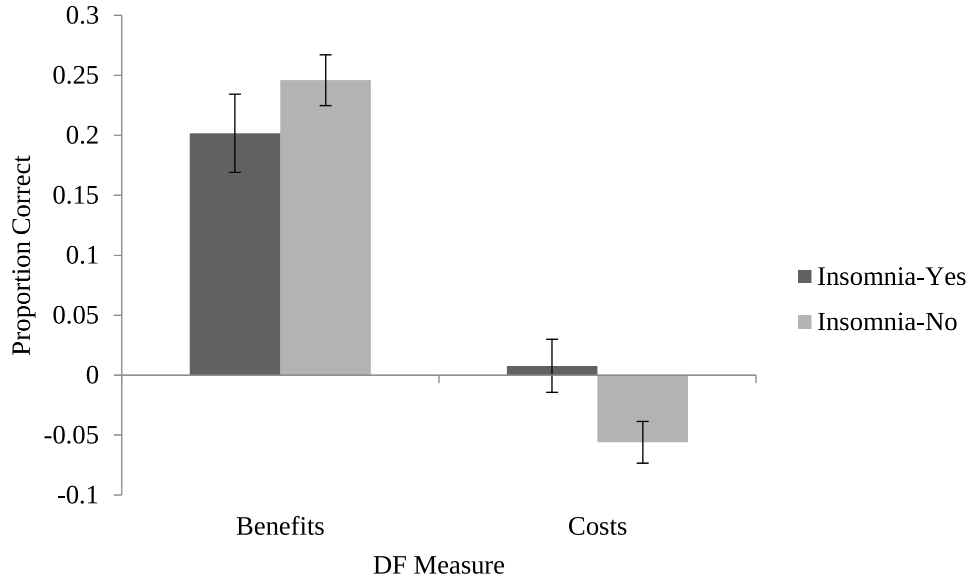

Figure 4. Mean proportion correct for benefits and costs split between participants with insomnia symptoms and those without. Error bars represent standard error of the mean.

Directed Forgetting and WMC. Table 1 also presents the correlations between the WMC measure and the DF measures. The descriptive statistics for the WMC measure can be found in Table 2. Table 1 shows that WMC was significantly positively correlated with the +10 point value, with a higher WMC related to higher hit rates for that point value, and benefits, with higher WMC related to higher differences between the +10 point value and the 0 point value. WMC was not, however correlated with the -10 point value, the 0 point value, or costs.

Directed Forgetting, WMC, and Sleep. Tables 3 and 4 present the correlations between the WMC measure and the DF measures for participants with and without insomnia symptoms, respectively. The SCI was again used as a categorical variable. The separate groups show the same overall pattern of correlations as in the entire sample: the WMC measure was significantly positively correlated with the +10 point value and benefits, but not, however, significantly correlated with the -10 point value, the 0 point value, or costs.

Table 3

Correlations Among WMC and the DF Measures For Participants With Insomnia Symptoms

	WMC	+10	-10	0	Benefit	Cost
WMC	1.00					
+10	.644**	1.00				
-10	.074	.600**	1.00			
0	.161	.546**	.792**	1.00		
Benefit	.453*	.377*	-.287	-.570**	1.00	
Cost	.158	.050	-.109	.521**	-.527**	1.00

Note: * denotes $p < .05$; ** denotes $p < .001$

Table 4

Correlations Among WMC and the DF Measures For Participants Without Insomnia Symptoms

	WMC	+10	-10	0	Benefit	Cost
WMC	1.00					
+10	.382**	1.00				
-10	.045	.443**	1.00			
0	.034	.506**	.715**	1.00		
Benefit	.332**	.433**	-.320*	-.558**	1.00	
Cost	-.018	.038	-.457**	.296*	-.273*	1.00

Notes: * denotes $p < .05$; ** denotes $p < .001$

Directed Forgetting and the Other Measures. The correlations between the point values and the other measures taken during this study can also be found in Table 1; these other measures include: K10 score, gender, sex, age, medications impacting cognitive functioning, medications used to induce sleep, and medications used to stay awake. As mentioned, the descriptive statistics for the other measures can be found in Table 2. The only significant correlation was between the 0 point value and the K10 score, with higher levels of psychological distress related to higher hit rates for the 0 point value. No other correlations between the point values and questionnaires were significant. (The remaining correlations between the questionnaires will be discussed below.)

Relations between the benefits and costs and the other measures, presented in Table 1, are similar to the relations with the +10 and -10 point values. This table shows that there were no correlations among the other measures and either benefits or costs.

Sleep Measures

Table 5 presents the correlations between the SCI measure (as a continuous measure) and both components of the PROMIS measure. This table shows that the SCI measure is significantly negatively correlated with the disturbance component of the PROMIS, and the impairment component of the PROMIS, with greater presence of insomnia symptoms related to worse levels of sleep disturbance and overall impairment due to problems with sleep. The disturbance component of the PROMIS was significantly positively correlated with the impairment component of the PROMIS, with worse levels of sleep disturbance related to overall impairment due to problems with sleep.

Sleep Measures and WMC. Table 5 also presents the correlations between WMC and the three sleep measures (with the SCI again being used as a continuous variable). This table shows that WMC is not correlated with any of the sleep measures.

Sleep Measures and the Other Measures. The correlations between the three sleep measures and the other measures taken during the study can also be found in Table 5. This table shows that the SCI measure was significantly negatively correlated with the K10 score, with greater presence of insomnia symptoms related to higher levels of psychological distress, and significantly positively correlated with the medications used to induce sleep, with greater presence of insomnia symptoms related to a greater use of medications to induce sleep. The SCI measure, however, was not correlated with any of the other measures.

Lastly, Table 5 also shows that both of the components of the PROMIS were significantly positively correlated with the K10 score, with worse levels of sleep disturbance and overall impairment due to problems with sleep related to greater levels of psychological distress, and significantly negatively correlated with medications used to induce sleep, with worse levels of sleep disturbance and overall impairment due to problems with sleep related to greater use of medications to induce sleep. These components were not, however, correlated with any of the other measures.

Table 5

Correlations Among the Sleep Measures, WMC, and the Other Measures

	SCI	Dist	Imp.	WMC	K10	Gen.	Sex	Age	MCI	MS	MW
SCI	1.00										
Dist.	-.763**	1.00									
Imp.	-.668**	.656**	1.00								
WMC	.055	-.100	-.078	1.00							
K10	-.517**	.515**	.447**	.076	1.00						
Gen.	.075	-.091	-.128	.141	.019	1.00					
Sex	.056	-.074	-.105	.121	.028	.951**	1.00				
Age	-.057	-.007	-.098	-.196	-.226*	.040	.013	1.00			
MCI	.170	-.008	.047	-.194	-.172	.067	.049	.043	1.00		
MS	.388**	-.233*	-.241*	.077	-.123	.047	.032	-.215*	.257*	1.00	
MW	.078	-.059	-.170	.044	-.203*	.137	.176	.034	.140	.017	1.00

Note: * denotes $p < .05$; ** denotes $p < .01$; Dist = Disturbance Component of PROMIS, Gen. = Gender, Imp = Impairment Component of PROMIS, MCI = Medications Impacting Cognitive Functioning, MS = Medications Inducing Sleep, MW = Medications Used to Stay Awake.

Working Memory Capacity and the Other Measures

The correlations between the WMC measure and the other measures can be found in Tables 1 and 4. These tables show that WMC was not correlated with any these measures.

The Other Measures

The correlations between the other measures taken during this study can also be found in Tables 1 and 4. These tables show that the K10 score was significantly negatively correlated with age, with younger participants showing higher levels of psychological distress, and negatively correlated with medications used to stay awake, higher levels of psychological distress related to greater use of medications used to stay awake. Gender was statistically positively correlated with sex. Age was significantly negatively correlated with medications used to induce sleep, with increased age related to greater use of medications to help induce sleep. Medications impacting cognitive functioning were significantly positively correlated with medications used to induce sleep, with use of medications impacting cognitive functioning related to use of medications used to help induce sleep. No other correlations were significant.

Discussion

There were several goals of this study. First, it was believed that the use of a control condition in an item method DF experiment would allow the analysis of both costs and benefits in this type of experiment. Second, this study sought to observe the effects of natural occurring sleep problems on forgetting as opposed to manipulating sleep deprivation. Last, the relation of WMC to both sleep and DF was further explored. It was hypothesized that both costs and benefits would be observed in the DF data, that participants with worse levels of problems with sleep would have worse levels of remembering overall and would be unable to inhibit information, and that WMC would be related to problems with sleep and, subsequently, DF as well.

Overall, it was found that participants only showed benefits in relation to the DF data, with memory being better for the words associated with the +10 point value than the 0 or -10 point values. It was also found that problems with sleep had no impact on remembering with overall recognition being similar between groups. Finally, it was found that WMC was not related to problems with sleep but was related to the remembering aspect of DF. These findings, in relation to the other measures, and their implications will be discussed in more detail below.

Directed Forgetting Costs and Benefits

Of what was found in the DF analysis, some of the results fall in line with what was hypothesized, and some do not. First, as hypothesized and mentioned previously, benefits were found with participants recognizing more of the words associated with the +10 point value than the 0 point value. This result is similar to that of Friedman and Castel (2011), who also found memory to be highest for the highest point value (+10) used in their study. Unexpected, however, was the presence of negative costs with participants recognizing more of the words associated with the -10 point value than the 0 point value. This result was neither the desired outcome in this study, nor does it match up with what was found by Friedman and Castel. As stated previously, Friedman and Castel found that, while memory was highest for the +10 point value, memory for all other point values was similar to one another. This study, on the other hand, found a difference among all three point values.

Among the possible reasons for why this study found negative costs when using a point value of 0 is that a point value of -10 carried with it some kind of meaning, leading to it being encoded, but a point value of 0 did not. It is possible that participants chose to encode the words associated with the +10 point value as it would help their score the most at test but also to encode the words associated with the -10 point value in the study phase so that they knew to not respond to those words at test in order to also better their score. Since the 0 point value would neither benefit nor harm them at test, little encoding was done for the words associated with that point value. With Friedman and Castel`s procedure (2011), however, it would not have been an effective strategy to attempt to encode the words that would harm their score as there were two negative point values; to encode those along with the positive ones would have meant that they would have had to attempt to remember all of the words (as there were only positive and negative point values), along with each word's associated point value. It would have been easier for the participants in the current study to attempt to encode two-thirds of the study words as opposed to the participants in Friedman and Castel's study to attempt to encode all of the study words. This explanation, however, cannot explain why the participants in Friedman and Castel's study also chose not to encode the lower positive point value (+5) and only focused on the highest positive point value.

Another possible explanation for the negative costs found in this study lies in the fact that there was a large sample size for the within-subject analysis conducted on the difference between the -10 and 0 point values but also a small numerical difference and effect size found. When comparing this small effect size to the larger effect sizes found for the other DF results, it is likely, then, that the statistical difference that was found between the -10 and 0 point sizes was due to the large sample size that was used in a within-subject analysis and that the significance that was found should be taken with caution.

Regardless of what explanation is behind the negative costs found in this study, it does not appear that the use of a point system in item method DF, using the method laid out above, is an effective way to be able to analyze both costs and benefits similar to that of list method DF. However, it is important to point out other attempts at a control condition using the item method show that there may be hope for the point system. Namely, several studies have shown costs and benefits using a between-subjects item method design with one group being given the traditional remember or forget cues following each individual item and the control group getting only the remember cues following each individual item (e.g., Sahakyan & Foster, 2009; Taylor, Quinlan, & Vullings, 2018, Experiment 2; see also Fawcett, Taylor, & Nadel, 2013).

Performance for their forget-half groups was compared to their remember-all groups with costs occurring if memory was lower for the forget items in the forget-half group compared to memory in the remember-all group. Benefits occurred if memory was higher for the remember items in the forget-half group compared to memory in the remember-all group. In contrast to the current study, Taylor et al. (2018) found costs but no benefits. Participants in their forget-half group remembered fewer forget items compared to the remember-all group but performed similarly when responding to their remember items. Sahakyan and Foster (2009), on the other hand, found both costs and benefits when the participants just read the action phrases but found neither when participants acted out the phrases. These results show that it possible to observe both costs and benefits, or at least costs, when applying a control variable to an item method task, so, given the right manipulation, the same results might be able to be replicated using a point system design.

Natural Occurring Problems with Sleep and Forgetting

In order to explore how forgetting in this study was impacted by naturally occurring problems with sleep, SCI scores were used to compare DF costs and benefits between those with and without insomnia symptoms. Similar to the complete set of DF data, some results from the group comparisons were unexpected and some were expected. First, unexpectedly, these results showed that the participants with the presence of insomnia symptoms did not have lower levels of remembering overall. Those with the presence of insomnia symptoms were just as able to remember the studied words as were the participants without the presence of insomnia symptoms. Despite this being unexpected, however, there is past research that would support this conclusion (Schmidt, Richter, Gendolla, & Van Der Linden, 2010). Schmidt et al. found in their study that young poor sleepers were able to focus all of their cognitive resources to the memory task at hand for short periods of time. So, while these young adults scored with at least threshold levels of insomnia on the insomnia measure given, they were able to perform similarly to young adults with no sleep problems when the resources were present to do so. It is possible that the participants in this study were able to do the same thing: They were able to divert their cognitive resources to the memory task at hand and were able to perform similarly to their counterparts with no presence of insomnia symptoms.

A second, and expected, result was that the participants with the presence of insomnia symptoms were unable to inhibit the words associated with the -10 point value compared to the 0 point value, as memory for both was similar as can be seen from the cost analysis. Unexpected, however, was that the participants without the presence of insomnia symptoms not only did not inhibit memory for the words associated with the -10 point value, but also showed significant negative costs with memory in fact being better for the words associated with the -10 point value than the 0 point value. This result falls in line with the results from the full sample analysis and is likely the driving factor behind that result.

One explanation for why the participants with insomnia symptoms showed no costs while the participants without symptoms showed negative costs could be just an inability for those participants to inhibit information. However, a different explanation follows the logic behind the explanation presented previously for the difference between the -10 and 0 point values and the logic laid out by Schmidt et al. (2010). If the participants were attempting to encode the words associated with the -10 point value in order to successfully reject them at test, this would again explain why the participants without insomnia symptoms showed negative costs, recognizing more of the words associated with the -10 point value once instructed to disregard the point values at test. If then also taking into account that the participants with the presence of insomnia symptoms were diverting their cognitive resources towards remembering what they could, resulting in the similar levels of memory overall, it would make sense that they had no resources left to put towards remembering information in order to avoid it later, resulting in no negative costs.

The Role of WMC

The last hypothesis made was in regard to the effect of problems with sleep on WMC and how that related to DF. While it was thought that having the presence of insomnia symptoms would lead to a decrease in WMC, there was unexpectedly no relation between the sleep measures used and WMC found at all. It was expected that a relation would be found because previous research would suggest there to be one (e.g., Ballesio et al., 2018; Lim & Dinges, 2010), but, as pointed out by Whitney and Rosen (2013), and discussed above, the past literature tends to use WM as a global term without fully defining which part of WM is being measured. It is possible that WMC is not generally found to be affected by poor sleep and that it should not have been expected to in this study either.

Additionally, because there was no significant relation between WMC and sleep, this relation could therefore not affect DF. What was found, however, was that WMC did have an effect on DF independent of sleep. High WMC was related to higher recognition of the words associated with the +10 point value and the calculated benefits. This finding is important because it is in line with the findings of Marevic et al. (2018), who, as mentioned previously, also found WMC to only be related to the ‘remember’ items in their item method design. It should be noted that the importance of this finding is further highlighted by the fact that this is one of few studies to actually examine the role of WMC using an item method design and it replicated the previous findings. In terms of what the relation between WMC and DF would suggest to the overall literature, it is plausible that WMC is related to the active rehearsal of information and not the attempted inhibition or suppression of that information. This would explain why WMC was correlated with recognition of the +10 point value words and not the -10 point value words.

Additional Analyses

Though not central to the original hypotheses, the other analyses conducted showed mostly expected results. To summarize, it was found that the sleep measures were all correlated with one another. This showed that the presence of insomnia symptoms, as measured by the SCI, was related to both sleep disturbance and daily impairment due to problems with sleep, as measured by the PROMIS. This was expected as what defines insomnia overlaps with issues related to sleep disturbance (American Psychiatric Association, 2013) and it is easy to see how dealing with not being able to get to sleep or stay asleep can lead to daily impairment. It is believed that this is the first study to compare the SCI to the sleep bank portion of the PROMIS. The sleep measures were also expectedly related to psychological distress and the use of medications to go to sleep. Unexpected, however, was that the sleep measures were not related to gender, sex, or the use of medications to stay awake, despite previous research (see above for discussion).

Additionally, it was found that the WMC measure was not related to any use of medication, psychological distress, or gender or sex, which was also expected. It was also found that the 0 point value was related to psychological distress, though the reason for this is unknown and not taken to be meaningful. Age was found to be related to psychological distress and use of medications to sleep, which is not taken to be meaningful given the mean of the ages of the participants and the cluster of the ages in a small age range. Psychological distress was also related to the use of medications to stay awake. Lastly, the use of medications to go to sleep was related to the use of medications that impaired cognition.

Limitations

Several limitations to the current study should be noted. First, in regard to the DF data, the analyses were conducted with the assumption that the participants treated the +10 point value and the -10 point value as direct remember and forget instructions, respectively. While these point values were treated analytically as direct cues, as that was their intended purpose, it is possible that the participant did not treat them as such. If the participants did not treat the point values as intended, then this could be a reason behind why the manipulation involving the -10 point value did not work. Further work on these point values is required to address this issue. Second, since there was a small number of participants with the presence insomnia symptoms compared to those without, analyses based on this group could have been limited in power. This limit could be supported by the small effect size exhibited.

A third limit is that the sleep measures used asked participants questions based on wide time frames (e.g., seven weeks, one month, or no time limit at all). It is possible that the average responses given by participants were less reflective of how they were actually affected by their sleep in the moment. Similar to this point, a last limit to the current study is the use of self-reported sleep measures themselves. Indeed, self-reported measures can be problematic with regards to issues such as the reporter having faulty memory or under/over-reporting. It is possible that some of the participants in this study overestimated their sleep problems and rated themselves higher on the SCI than was true to their actual sleep problems. Similarly, some of the participants could have underestimated their sleep problems and rated themselves lower on the SCI than was true to their actual sleep problems. If objective measures, such as actigraphy (which monitors physiological functions during sleep and can be used as a watch; Vallieres & Morin, 2003), were used as opposed to self-report measures, the distribution of participants to the presence of insomnia symptoms/no presence of symptoms groups could have been more refined and the differences between the two groups in relation to the performance on the DF task could have been more apparent. It should be noted, however, that both measures used were successfully validated against other objective sleep measures, as mentioned above (Espie et al., 2014; Yu et al., 2012), so it is possible that the distribution of the participants to the two groups would not have changed substantially had objective measures been used.

Future Directions

Based on the limitations listed above and other observations, there are several things that should be explored further given this study. First, since it is possible that the participants did not treat the point values as remember/forget cues, the use of a 0 point value in order to compute costs and benefits in an item method DF experiment should be replicated, but by either questioning the participant on what their strategies were given the cues or by using more salient instructions involving the other point values (e.g., explicitly instructing participants to try to forget items associated with negative point values). If it is assured that the participants are treating the other point values as remember/forget cues then it is possible that memory might be found to be worse for the negative point value compared to the neutral point value, as intended. Second, the effect of naturally occurring sleep on forgetting should be observed once more using more refined measures of sleep. Differences in the relation between sleep and forgetting might be found if measures used in the diagnosing of sleep problems are used. This suggestion is limited, however, by the access of those sleep measures. Third, given the general use of the term 'WM' in the sleep literature, the exact effect of sleep on each part of WM should be separated out for easier clarification in the future as to what should be expected when testing the effect of sleep on WMC.

Last, given Schmidt et al.'s (2014) finding that poor sleepers tend to divert resources in order to cope with the effects of poor sleep on memory, another direction that should be considered is examining the effect of the capacity of available resources in poor sleepers in relation to the current study. This would be done in order to see if the null finding of the effect of sleep on overall memory is caused by the participants with the presence of insomnia symptoms actually using all of their resources in order to match the performance of their counterparts. This could be explored further using the tasks presented in Fawcett and Taylor (2008; 2011), which examined the attentional aftereffects of remember and forget instructions. Though their studies used these tasks to show that forgetting is effortful, this could be applied to demonstrating reduced resources in poor sleepers. The more resources one has, the more distractible they are at a given task (Lavie, Hirst, De Fockert, & Viding, 2004), so, if the participants with the presence of insomnia symptoms have less resources, they would be less distracted by the probes used in the task, leading to faster reaction times on the task. Those without the presence of insomnia symptoms, on the other hand, should have slower reaction times on the task as they are more distracted. If this were the case, it would support the resource diversion explanation for the current null finding.

Final Conclusions

Regardless of which hypotheses were supported and which were not, the current study presented some interesting findings that were both unexpected and novel that should be explored more in future research. Among these findings, it was shown that the current manipulation of point values in a DF experiment could produce benefits using the item method but not costs. This finding suggests that, with further investigation, the use of a control variable in the within-subject experiment might be plausible. It was also shown that the participants with naturally occurring insomnia symptoms were just as able to remember information as were the participants without insomnia symptoms and neither group showed evidence of inhibiting information. With further support for this finding, this could mean that people who exhibit insomnia symptoms do not find their sleep problems to be as problematic in relation to memory and inhibition as one would expect. Finally, it was shown that WMC was related to remembering in the DF task and not the sleep measures used. This would suggest that the extent of the amount of information one can deal with in WM is not affected by the presence of insomnia symptoms, nor does it have an affect on one’s ability to forget information.

References

Abel, M., & Bäuml, K. T. (2013). Sleep can eliminate list-method directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 946-952. doi: 10.1037/a0030529

Alberca-Reina, Cantero, & Atienza. (2015). Impact of sleep loss before learning on cortical dynamics during memory retrieval. *NeuroImage, 123*, 51-62. doi: 10.1016/j.neuroimage.2015.08.033

Alhola, P., & Polo-Kantola, P. (2007). Sleep deprivation: Impact on cognitive performance. *Neuropsychiatric Disease and Treatment*, *3*, 553-567.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Amlander, C. J., Fuller, P. M. (Eds.). (2005). *Basics of sleep guide* (2nd ed.). Westchester, IL: Sleep Research Society.

Anderson, J. R. (1990). *Cognitive psychology and its implications*. New York: Freeman.

Aslan, A., Zellner, M., & Bäuml, K. (2010). Working memory capacity predicts listwise directed forgetting in adults and children. *Memory (Hove, England), 18*, 442-50. doi: 10.1080/09658211003742698

Atherton, K. E., Nobre, A. C., Lazar, A. S., Wulff, K., Whittaker, R. G., Dhawan, V., . . . Butler, C. R. (2016). Slow wave sleep and accelerated forgetting. *Cortex, 84*, 80-89. doi: 10.1016/j.cortex.2016.08.013

Babkoff, H., Mikulincer, M., Caspy, T., Kempinski, D., Sing, H. (1988). The topology of performance curves during 72 hours of sleep loss: a memory and search task. *Q J Exp Psychol A, 40*, 737–756.

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, 417–423.

Ballesio, A., Cerolini, S., Ferlazzo, F., Cellini, N., & Lombardo, C. (2018). The effects of one night of partial sleep deprivation on executive functions in individuals reporting chronic insomnia and good sleepers. *Journal of Behavior Therapy and Experimental Psychiatry*, *60*, 42-45. doi: 10.1016/j.jbtep.2018.02.002

Belenky, G., Wesensten, N., Thorne, D., Thomas, M., Sing, H., Redmond, D., . . . Balkin, T. (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: A sleep dose response study. *Journal of Sleep Research*, *12*, 1-12. doi: 10.1046/j.1365-2869.2003.00337.x

Berman, M. G., Jonides, J., & Lewis, R. L. (2009). In search of decay in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 317–333

Bjork, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning & Verbal Behavior*, *9*, 255-268. doi:10.1016/S0022-5371(70)80059-7

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In: Roediger HL III, Craik FIM, editors. *Varieties of Memory and Consciousness: Essays in Honour of Endel Tulving.Erlbaum*; Hillsdale, NJ: pp. 309–330

Brewin, C. R., & Beaton, A. (2002). Thought suppression, intelligence, and working memory capacity. *Behaviour Research and Therapy, 40*, 923–930. doi:10.1016/S0005-7967(01)00127-9

Brown, J. (1954). The nature of set-to-learn and of intra-material interference in immediate memory. *Quarterly Journal of Experimental Psychology*, *6*, 141-148. doi: 10.1080/17470215408416659

Brown, G. D., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114*, 539-576. doi: 10.1037/0033-295X.114.3.539

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*, 193-213. doi: 10.1016/0165-1781(89)90047-4

Cairney, S., Durrant, A., Musgrove, S., & Lewis, J. (2011). Sleep and environmental context: Interactive effects for memory. *Experimental Brain Research, 214*, 83-92. doi: 10.1007/s00221-011-2808-7

Carr, T. H. (2007). Is it time to inhibit inhibition? Lessons from a decade of research on the place of inhibitory processes in cognition. In *Inhibition in cognition.* (pp. 307–316). Washington, DC: American Psychological Association. https://doi.org/10.1037/11587-015

Castel, A. D., Benjamin, A. S., Craik, F. I. M., & Watkins, M. J. (2002). The effects of aging on selectivity and control in short-term recall. *Memory & Cognition, 30*, 1078–1085. doi:10.3758/BF03194325

Chee, M., & Choo, W. (2004). Functional imaging of working memory after 24 hr of total sleep deprivation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 24*, 4560-7.

Chee, M. W., & Chuah, Y. M. L. (2007). Functional neuroimaging and behavioural correlates of capacity decline in visual short-term memory following sleep deprivation. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 9487-9492.

Cheng, S., Liu, I., Lee, J., Hung, D., & Tzeng, O. (2011). Intentional forgetting might be more effortful than remembering: An ERP study of item-method directed forgetting. *Biological Psychology, 89*, 283- 292. doi: 10.1016/j.biopsycho.2011.10.019

Clark, I., & Landolt, H. P. (2017). Coffee, caffeine, and sleep: a systematic review of epidemiological studies and randomized controlled trials. *Sleep Medicine Reviews, 31*, 70–78.

Clegg-Kraynok, M. M., McBean, A. L., & Montgomery-Downs, H. E. (2011). Sleep quality and characteristics of college students who use prescription psychostimulants non-medically. *Sleep Med., 12*, 598–602.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, *33*A, 497-505. doi:10.1080/14640748108400805

de Souza, J. C., de Oliveira, M. L. C., de Sousa, I. C., & de Azevedo, C. V. M. (2018). Gender differences in sleep habits and quality and daytime sleepiness in elementary and high school teachers. *Chronobiology International*, *35*, 486–498. doi: 10.1080/07420528.2017.1415921

Delaney, P., & Sahakyan, F. (2007). Unexpected costs of high working memory capacity following directed forgetting and contextual change manipulations. *Memory & Cognition, 35*, 1074-1082. doi: 10.3758/BF03193479

Della Sala, S., Cowan, N., Beschin, N., & Perini, M. (2005). Just lying there, remembering: Improving recall of prose in amnesic patients with mild cognitive impairment by minimising interference. *Memory (Hove, England), 13*, 435-40.

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience, 11*, 114-26. doi:10.1038/nrn2762

Drummond, S., Anderson, D., Straus, L., Vogel, E., Perez, V., & Goel, N. (2012). The effects of two types of sleep deprivation on visual working memory capacity and filtering efficiency (sleep deprivation effects on visual working memory). *PLoS ONE, 7*, 35653. doi: 10.1371/journal.pone.0035653

Drummond, S., Brown, G., Gillin, J., Stricker, J., Wong, E., & Buxton, R. (2000). Altered brain response to verbal learning following sleep deprivation. *Nature, 403*, 655-7.

Duffy, J. F., Rimmer, D. W., & Czeisler, C. A. (2001). Association of intrinsic circadian period with morningness-eveningness, usual wake time, and circadian phase. *Behavioral Neuroscience, 115*, 895–899.

Durmer, J. S., & Dinges, D. F. (2005). Neurocognitive consequences of sleep deprivation. *Seminars in Neurology, 25,* 117–129. doi: 10.1055/s-2005-867080

Ecker, U. K. H., Lewandowsky, S., & Oberauer, K. (in press). Removal of information from working memory: A specific updating process. *Journal of Memory and Language*.

Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 170–189.

Ecker, U., Oberauer, K., & Lewandowsky, S. (2014). Working memory updating involves item-specific removal. *Journal of Memory and Language, 74*, 1-15. doi: 10.1016/j.jml.2014.03.006

Elmes, D. (1969). Cueing to forget in short-term memory. *Journal of Experimental Psychology, 80*, 561-562.

Erdfelder, E., Faul, F. & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*, *28*, 1-11. doi: 10.3758/BF03203630

Espie, C., Kyle, S., Hames, P., Gardani, M., Fleming, L., & Cape, J. (2014). The sleep condition indicator: A clinical screening tool to evaluate insomnia disorder. *BMJ Open, 4*. doi: 10.1136/bmjopen-2013-004183

Fawcett, J. M., & Taylor, T. L. (2008). Forgetting is effortful: Evidence from reaction time probes in an item-method directed forgetting task. *Memory & Cognition*, *3*6, 1168–1181

Fawcett, J., & Taylor, T. (2012). The control of working memory resources in intentional forgetting: Evidence from incidental probe word recognition. *Acta Psychologica, 139*, 84-90. doi: 10.1016/j.actpsy.2011.10.001

Fawcett, J., Taylor, T., & Nadel, L. (2013). Event-method directed forgetting: Forgetting a video segment is more effortful than remembering it. *Acta Psychologica*, *144*, 332-343.

Feld, G., Weis, P., & Born, J. (2016). The limited capacity of sleep-dependent memory consolidation. *Frontiers In Psychology, 7*, 1368. doi: 10.3389/fpsyg.2016.01368

Fischer, S., Diekelmann, S., & Born, J. (2011). Sleep's role in the processing of unwanted memories. *Journal of Sleep Research*, *20*, 267-74. doi: 10.1111/j.1365-2869.2010.00881.x

Foster, N. L., & Sahakyan, L. (2012). Metacognition influences item-method directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1309–1324. doi: 10.1037/a0027868

Fredholm, B. B., Battig, K., Holmen, J., Nehlig, A., & Zvartau, E. E. (1999). Actions of caffeine in the brain with special reference to factors that contribute to its widespread use. *Pharmacological Reviews*, *51*, 83–133.

Friedman, M., & Castel, C. (2011). Are we aware of our ability to forget? Metacognitive predictions of directed forgetting. *Memory & Cognition*, *39*, 1448-1456. doi: 10.3758/s13421-011-0115-y

Frenda, S., & Fenn, K. (2016). Sleep less, think worse: The effect of sleep deprivation on working memory. *Journal of Applied Research in Memory and Cognition.*, *5*, 463-469. doi: 10.1016/j.jarmac.2016.10.001

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

Godden, D., & Baddeley, A. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology, 66*, 325. doi: 10.1111/j.2044-8295.1975.tb01468.x

Golding, J., Roper, K. L., & Hauselt, J. (1996). To forget or not to forget: The effect of probability of test on directed forgetting. *The Quarterly Journal of Experimental Psychology Section A*, *49*, 326-340.

Goldstein-Piekarski, A. N., Greer, S. M., Saletin, J. M., Harvey, A. G., Williams, L. M., & Walker, M. P. (2018). Sex, sleep deprivation, and the anxious brain. *J Cogn Neurosci.*, *30*, 565–578. doi:10.1162/jocn_a_01225

Harrison, Y., & Horne, J. (2000a). Sleep loss and temporal memory. *The Quarterly Journal of Experimental Psychology. A. Human Experimental Psychology, 53*, 271-9. doi: 10.1080/027249800390772

Harrison, Y., & Horne, J. (2000b). The impact of sleep deprivation on decision making: A review. *Journal of Experimental Psychology. Applied, 6,* 236-49.

Hershner, S., & Chervin, R. (2014). Causes and consequences of sleepiness among college students. *Nature and Science of Sleep, 6,* 73-84. doi: 10.2147/NSS.S62907

Horne, J. A. (1993). Human sleep, sleep loss and behaviour. Implications for the prefrontal cortex and psychiatric disorder. *The British Journal of Psychiatry, 162,* 413-419.

Hupbach, A. (2017). Long-term effects of directed forgetting. *Memory (Hove, England),* 1-9. doi: 10.1080/09658211.2017.1358748

Joo, E., Yoon, C., Koo, D., Kim, D., & Hong, S. (2012). Adverse effects of 24 hours of sleep deprivation on cognition and stress hormones. *Journal of Clinical Neurology (Seoul, Korea), 8,* 146-50. doi: 10.3988/jcn.2012.8.2.146

Kato, M. G., Phillips, B. A., Sigurdsson, G. K., Narkiewicz, K., Pesek, C., & Somers, V. (2000). Effects of sleep deprivation on neural circulatory control. *Hypertension: Journal of the American Heart Association, 35,* 1173-1175.

Kerpershoek, M. L., Antypa, N., Van den Berg, J. F. (2018). Evening use of caffeine moderates the relationship between caffeine consumption and subjective sleep quality in students. *J Sleep Res., 27*. doi: 10.1111/jsr.12670

Kessler, R., Barker, P., Colpe, L., Epstein, J., Gfroerer, J., Hiripi, E., . . . Zaslavsky, A. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry, 60*, 184-9.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H. U., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Archives of General Psychiatry, 51*, 8–19.

Killgore, W., & Weber, M. (2014). Sleep deprivation and cognitive performance. In Bianchi, M. (Ed.), *Sleep deprivation and disease* (pp. 209-229). New York, NY: Springer.

Kučera, H., & Francis, W. N. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.

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

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*, 339-354.

Lee, Y. (2012). Cognitive load hypothesis of item-method directed forgetting. *The Quarterly Journal Of Experimental Psychology*, *65*, 1110-1122. doi: 10.1080/17470218.2011.644303

Lee, Y., & Lee, H. (2011). Divided attention facilitates intentional forgetting: Evidence from item-method directed forgetting. *Consciousness And Cognition: An International Journal*, *20*, 618-626. doi: 10.1016/j.concog.2010.09.008

Lim, J., & Dinges, D. (2008). Sleep deprivation and vigilant attention. *Annals of the New York Academy of Sciences*, *1129*, 305-322. doi: 10.1196/annals.1417.002

Lim, J., & Dinges, D. F. (2010). A meta-analysis of the impact of short-term sleep deprivation on cognitive variables. *Psychological Bulletin*, *136*, 375-389.

Liu, T., Chen, N., & Cheng, S. (2017). Selective rehearsal is affected by the emotionality of the encoding context in item-method directed forgetting: An event-related potential study. *Biological Psychology, 123*, 15-24. doi: 10.1016/j.biopsycho.2016.11.012

Lo, J. C., Groeger, J. A., Cheng, G. H., Dijk, D. J., & Chee, M. W. L. (2016). Self-reported sleep duration and cognitive performance in older adults: A systematic review and meta-analysis. *Sleep Medicine, 17*, 87-98. doi: 10.1016/j.sleep.2015.08.021

MacLeod, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches*. Mahwah: Lawrence Erlbaum. pp. 157.

Macmillan, M. (2000). *An odd kind of fame: Stories of Phineas Gage.* Cambridge, MA: MIT Press.

Maquet, P. (2001). The role of sleep in learning and memory. *Science (Washington), 294*, 1048-1052.

Marevic, I., Arnold, N., & Rummel, J. (2018). Item-method directed forgetting and working memory capacity: A hierarchical multinomial modeling approach. *Quarterly Journal of Experimental Psychology, 71*, 1070-1080. doi: 10.1080/17470218.2017.1310270

Medeiros, A. L. D., Mendes, D. B., Lima, P. F., Araujo, J. F. (2001). The relationships between sleep-wake cycle and academic performance in medical students. *Biol Rhythm Res., 32*, 263–270.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Moon, C., Phelan, C. H., Lauver, D. R., & Bratzke, L. C. (2015). Is sleep quality related to cognition in individuals with heart failure? *Heart & Lung - The Journal of Acute and Critical Care*, *44*, 212-218. doi: 10.1016/j.hrtlng.2015.02.005

Nebes, R. D., Buysse, D. J., Halligan, E. M., Houck, P. R., Monk, T. H. (2009). Self-reported sleep quality predicts poor cognitive performance in healthy older adults. *J Gerontol B Psychol Sci Soc Sci.,64*, 180-187. doi: 10.1093/geronb/gbn037

Neckelmann, D., Mykletun, A., & Dahl, A. A. (2007). Chronic insomnia as a risk factor for developing anxiety and depression. *Sleep, 30*, 873–880.

Noreen, S., & De Fockert, J. (2017). The role of cognitive load in intentional forgetting using the think/no-think task. *Experimental Psychology, 64*, 14-26. doi: 10.1027/1618-3169/a000347

Petrov, M. E., Lichstein, K. L., & Baldwin, C. M. (2014). Prevalence of sleep disorders by sex and ethnicity among older adolescents and emerging adults: Relations to daytime functioning, working memory and mental health. Journal of Adolescence, 37, 587-597. doi: 10.1016/j.adolescence.2014.04.007

Pilcher, J., & Huffcutt, A. (1996). Effects of sleep deprivation on performance: A meta-analysis. *Sleep, 19*, 318-326.

Psychology Software Tools, Inc. [E-Prime 3.0]. (2016). Retrieved from https://www.pstnet.com.

Rauchs, G., Feyers, D., Landeau, B., Bastin, C., Luxen, A., Maquet, P., & Collette, F. (2011). Sleep contributes to the strengthening of some memories over others, depending on hippocampal activity at learning. *The Journal Of Neuroscience*, *31*, 2563-2568. doi:10.1523/JNEUROSCI.3972-10.2011

Reber, P. J., Siwiec, R. M., Gitelman, D. R., Parrish, T. B., Mesulam, M. M., Paller, P. (2002). Neural correlates of successful encoding identified using functional magnetic resonance imaging. *J Neurosci 22*, 9541–9548.

Rizio, A. A., Dennis, N. A., (2013). The neural correlates of cognitive control: successful remembering and intentional forgetting. *J. Cogn. Neurosci. 25*, 297 – 312.

Sahakyan, L., Delaney, P., Foster, N., & Abushanab, B. (2013). List-method directed forgetting in cognitive and clinical research: A theoretical and methodological Review. *Psychology of Learning and Motivation, 59*, 131-189. doi: 10.1016/B978-0-12-407187-2.00004-6

Sahakyan, L., & Foster, N. L. (2009). Intentional forgetting of actions: Comparison of list-method and item-method directed forgetting. *Journal of Memory and Language, 61*, 134-152. doi: 10.1016/j.jml.2009.02.006

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 1064–1072. doi: 10.1037/0278-7393.28.6.1064

Saletin, J., Goldstein, A., & Walker, M. (2011). The role of sleep in directed forgetting and remembering of human memories. *Cerebral Cortex, 21*, 2534-2541. doi: 10.1093/cercor/bhr034

Schacter, D. (2001). *The seven sins of memory : How the mind forgets and remembers*. Boston: Houghton Mifflin

Schmidt, R., Richter, M., Gendolla, G., & Van Der Linden, M. (2010). Young poor sleepers mobilize extra effort in an easy memory task: Evidence from cardiovascular measures. *Journal of Sleep Research, 19*, 487-495. doi: 10.1111/j.1365-2869.2010.00834.x

Soriano, M. F., & Bajo, M. T. (2007). Working memory resources and interference in directed forgetting. *Psicologica: International Journal of Methodology and Experimental Psychology, 28*, 63-85.

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature, 437*, 1272-8. doi: 10.1038/nature04286

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York, NY: Springer.

doi: 10.1007/978-1-4419-8126-4

Taylor, T., & Hamm, L. (2016). Selection for encoding: No evidence of greater attentional capture following forget than remember instructions. *Attention, Perception, & Psychophysics, 78,* 168-186. doi: 10.3758/s13414-015-0984-4

Taylor, T., Quinlan, L., & Vullings, C. (2018). Decomposing item-method directed forgetting of emotional pictures: Equivalent costs and no benefits. *Memory & Cognition, 46*, 132-147. doi: 10.3758/s13421-017-0751-y

Thomas, M., Sing, H., Belenky, G., Holcomb, H., Mayberg, H., Dannals, R., …. Redmond, D. (2000). Neural basis of alertness and cognitive performance impairments during sleepiness. I. Effects of 24 h of sleep deprivation on waking human regional brain activity. *Journal of Sleep Research, 9*, 335–352. doi:10.1046/j.1365-2869.2000.00225.x

Tonetti, L., Fabbri, M., & Natale, V. 2008. Sex difference in sleep-time preference and sleep need: A cross-sectional survey among Italian Pre-Adolescents, adolescents, and adults. *Chronobiol Int, 25,* 745–759.

Tsapanou, A., Gu, Y., O'Shea, D. M., Yannakoulia, M., Kosmidis, M. H., Dardiotis, E., Hadjigeorgiou, G., Sakka, P… & Scarmeas, N. (2017). Dataset on the associations between sleep quality/duration and cognitive performance in cognitively healthy older adults. *Data in Brief, 14,* 720-723. doi: 10.1016/j.dib.2017.08.028

Unsworth, N., Heitz, R., Schrock, J., & Engle, R. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*, 498-505. doi: 10.3758/BF03192720

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

Van Der Werf, Y. D., Altena, E., Schoonheim, M. M., Sanz-Arigita, E. J., Vis, J. C., De Rijke, W., & Van Someren, E. W. (2009). Sleep benefits subsequent hippocampal functioning. *Nature Neuroscience*, *12*, 122-123. doi:10.1038/nn.2253

Walker, M., & Stickgold, R. (2006). Sleep, memory, and plasticity. *Annual Review of Psychology*, *57,* 139-66.

Whitney, P., & Rosen, P. J. (2013). Sleep deprivation and performance: The role of working memory. In T. Packiam Alloway & R. G. Alloway (Eds.), *Working memory: The connected intelligence.* (pp. 175–186). New York, NY: Psychology Press.

Wilhelm, I., Diekelmann, S., Molzow, I., Ayoub, A., Mölle, M., & Born, J. (2011). Sleep selectively enhances memory expected to be of future relevance. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 31*, 1563-9. doi: 10.1523/JNEUROSCI.3575-10.2011

Wixted, J. T. (2010). The role of retroactive interference and consolidation in everyday forgetting. In S. Della Sala (Ed.), *Forgetting*. New York: Psychology Press.

Wong, M. L., Lau, K. N. T., Espie, C. A., Luik, A. I., Kyle, S. D., & Lau, E. Y. Y. (2017). Psychometric properties of the Sleep Condition Indicator and Insomnia Severity Index in the evaluation of insomnia disorder. *Sleep Medicine, 33*, 76-81. doi: 10.1016/j.sleep.2016.05.019

Woodward, A. E., Bjork, R. A., & Jongeward, R. H. (1973). Recall and recognition as a function of primary rehearsal. *Journal of Verbal Learning & Verbal Behavior, 12*, 608–617.

Wu, J. C., Gillin, J. C., Buchsbaum, M. S., Hershey, T., Hazlett, E., Sicotte, N., & Bunney, W. E. J. (1991). The effect of sleep deprivation on cerebral glucose metabolic rate in normal humans assessed with positron emission tomography. *Sleep, 14,* 155-162.

Wylie, G. R., Foxe, J. J., & Taylor, T. L. (2008). Forgetting as an active process: An fMRI investigation of item-method directed forgetting. Cerebral Cortex, 18, 670 –682.

Yang, T., Lei, X., & Anderson, M. (2016). Decreased inhibitory control of negative information in directed forgetting. *International Journal of Psychophysiology, 100*, 44-51. doi: 10.1016/j.ijpsycho.2015.09.007

Yoo, S., Hu, P. T., Gujar, N., Jolesz, F. A., & Walker, M. P. (2007). A deficit in the ability to form new human memories without sleep. *Nature Neuroscience*, *10*, 385-392.

Yu, L., Buysse, D., Germain, A., Moul, D., Stover, A., Dodds, N., . . . Pilkonis, P. (2012). Development of short forms from the PROMIS sleep disturbance and sleep-related impairment item banks. *Behavioral Sleep Medicine*, *10*, 6-24. doi: 10.1080/15402002.2012.636266

Zacks, R. T., Radvansky, G. A., & Hasher, L. (1996). Studies of directed forgetting in older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 143–156. doi:10.1037/0278-7393.22.1.14

Zhang, B., & Wing, Y. K. (2006). Sex differences in insomnia: A meta-analysis. *Sleep*, *29*, 85–93.

Appendix A: Words

Word	Length	Frequency
agree	5	51
alive	5	57
allow	5	72
angle	5	51
anode	5	77
apart	5	57
appeal	6	62
apply	5	56
aside	5	67
avoid	5	58
baby	4	62
band	4	53
beach	5	61
beat	4	68
birth	5	66
bitter	6	53
block	5	66
boat	4	72
bottle	6	76
breath	6	53
brief	5	73
broke	5	72
busy	4	58
camp	4	75
career	6	67
cell	4	65
chain	5	50
chair	5	66
choose	6	50
circle	6	60
clean	5	70
coast	5	61
coffee	6	78
cool	4	62
cross	5	55
crowd	5	53
daughter	8	72
dear	4	54
depth	5	53
desk	4	65

double	6	56
dozen	5	52
draw	4	56
dream	5	64
dress	5	67
drop	4	59
drove	5	62
dust	4	70
edge	4	78
enter	5	78
evil	4	72
fair	4	77
fast	4	78
fellow	6	63
fill	4	50
flat	4	67
flesh	5	52
flow	4	67
foot	4	70
fort	4	55
forth	5	71
frame	5	74
fund	4	62
goal	4	60
gold	4	52
grass	5	53
grew	4	64
grey	4	80
gross	5	66
guess	5	56
hence	5	58
hero	4	52
hill	4	72
hole	4	58
huge	4	54
hung	4	65
ideal	5	61
inner	5	55
seat	4	54
join	4	65
jury	4	67
kill	4	63
lady	4	80

laid	4	77
lake	4	54
leader	6	74
league	6	69
listen	6	51
loose	5	53
metal	5	61
mine	4	59
minor	5	58
model	5	77
moon	4	60
motor	5	56
narrow	6	63
nice	4	75
nose	4	60
offer	5	80
onto	4	60
page	4	66
pair	4	50
pale	4	58
palm	6	56
phase	5	72
pick	4	55
player	6	51
please	6	62
proud	5	50
prove	5	58
pull	4	51
pure	4	56
quick	5	68
rain	4	70
raise	5	52
rear	4	51
rich	4	74
rifle	5	63
risk	4	54
rock	4	75
roof	4	59
rule	4	73
safe	4	58
scale	5	60
search	6	66
seek	4	69

send	4	74
sharp	5	72
shook	5	57
shop	4	63
shore	5	61
sick	4	51
site	4	64
sleep	5	65
slight	6	53
slow	4	60
smile	5	58
snow	4	59
soft	4	61
soil	4	54
song	4	70
sought	6	55
spite	5	56
stone	5	58
store	5	74
sweet	5	70
tall	4	55
task	4	60
taste	5	59
taught	6	50
term	4	79
theme	5	55
thick	5	67
throat	6	51
till	4	50
tiny	4	50
title	5	77
tone	4	78
tree	4	59
truck	5	57
twice	5	74
upper	5	72
valley	6	73
vast	4	61
vital	5	56
vote	4	75
wage	4	56
warm	4	67
weather	7	69

were	5	61
wheel	5	56
wild	4	56
wind	4	63
wine	4	72
wood	4	55
wore	4	65
worry	5	55
worse	5	50
writer	6	73
yellow	6	55

Appendix B: Demographics Questionnaire

Please provide a response to the questions below. You may choose to decline to provide a response to any question that you do not wish to answer. This information will only be associated with the data you provided in today's session and will not be associated with your personal identity.

1. Gender: ________________
2. Sex: _____________
3. Age: ______________
4. Do you currently use any medications that may impair cognitive functioning? YES / NO
5. Do you currently take any over-the-counter or prescribed medications to help you sleep? YES / NO
6. Do you currently take any substances (e.g. over-the-counter medications, prescribed medications, caffeine, or energy drinks) to help you stay awake? YES / NO

Appendix C: Kessler Psychological Distress Scale

Please tick the answer that is correct for you:	All of the time (score 5)	Most of the time (score 4)	Some of the time (score 3)	A little of the time (score 2)	None of the time (score 1)
1. In the past 4 weeks, about how often did you feel tired out for no good reason?					
2. In the past 4 weeks, about how often did you feel nervous?					
3. In the past 4 weeks, about how often did you feel so nervous that nothing could calm you down?					
4. In the past 4 weeks, about how often did you feel hopeless?					
5. In the past 4 weeks, about how often did you feel restless or fidgety?					
6. In the past 4 weeks, about how often did you feel so restless you could not sit still?					
7. In the past 4 weeks, about how often did you feel depressed?					
8. In the past 4 weeks, about how often did you feel that everything was an effort?					
9. In the past 4 weeks, about how often did you feel so sad that nothing could cheer you up?					
10. In the past 4 weeks, about how often did you feel worthless?					

Appendix D: Sleep Condition Indicator

Item	4	3	2	1	0
Thinking about a typical night in the last month…					
1…. how long does it take you to fall asleep?	0-15 min	16-30 min	31-45 min	46-60 min	≥ 60 min
2…. if you then wake up during the night… how long are you awake for in total? (add all the wakenings up)	0-15 min	16-30 min	31-45 min	46-60 min	≥ 60 min
3…. how many nights a week do you have a problem with your sleep?	0-1	2	3	4	5-7
4…. how would you rate your sleep quality?	Very good	Good	Average	Poor	Very Poor
Thinking about the past month, to what extent has sleep…					
5…. affected your mood, energy, or relationships?	Not at all	A little	Somewhat	Much	Very Much
6…. affected your concentration, productivity, or ability to stay awake?	Not at all	A little	Somewhat	Much	Very Much
7…. troubled you in general?	Not at all	A little	Somewhat	Much	Very Much
Finally…					
8…. how long have you had a problem with your sleep?	I don't have a problem/ <1 mo	1-2 mo	3-6 mo	7-12 mo	> 1 yr

Appendix E: PROMIS-SF

In the past 7 days…		**Not at all**	**A little bit**	**Somewhat**	**Quite a bit**	**Very much**
Sleep 108	My sleep was restless……	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
Sleep 115	I was satisfied with my sleep……	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
Sleep 116	My sleep was refreshing……	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
Sleep 44	I had difficulty falling asleep……	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

In the past 7 days…		**Never**	**Rarely**	**Sometimes**	**Often**	**Always**
Sleep 87	I had trouble staying asleep……	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
Sleep 90	I had trouble sleeping……	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
Sleep 110	I got enough sleep……	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1

In the past 7 days…		**Very poor**	**Poor**	**Fair**	**Good**	**Very Good**
Sleep 109	My sleep quality was……	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1

In the past 7 days…		Not at all	A little bit	Somewhat	Quite a bit	Very much
Sleep 10	I had a hard time getting things done because I was sleepy……	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
Sleep 119	I felt alert when I woke up……	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
Sleep 18	I felt tired……	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
Sleep 25	I had problems during the day because of poor sleep……	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
Sleep 27	I had a hard time concentrating because of poor sleep……	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
Sleep 30	I felt irritable because of poor sleep……	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
Sleep 6	I was sleepy during the daytime……	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
Sleep 7	I had trouble staying awake during the day……	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5