The role of sleep in false memory formation

Jessica D. Payne [a,c,*], Daniel L. Schacter [c], Ruth E. Propper [d], Li-Wen Huang [c], Erin J. Wamsley [a], Matthew A. Tucker [a], Matthew P. Walker [b], Robert Stickgold [a,*]

[a] Center for Sleep and Cognition, Department of Psychiatry, Harvard Medical School and Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02215, United States
[b] Sleep and Neuroimaging Laboratory, University of California Berkeley, Department of Psychology, 3331 Tolman Hall, Berkeley, CA 94720, United States
[c] Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, United States
[d] Merrimack College, North Andover, MA, United States

ARTICLE INFO

ABSTRACT

Memories are not stored as exact copies of our experiences. As a result, remembering is subject not only to memory failure, but to inaccuracies and distortions as well. Although such distortions are often retained or even enhanced over time, sleep's contribution to the development of false memories is unknown. Here, we report that a night of sleep increases both veridical and false recall in the Deese–Roediger–McDermott (DRM) paradigm, compared to an equivalent period of daytime wakefulness. But while veridical memory deteriorates across both wake and sleep, false memories are preferentially preserved by sleep, actually showing a non-significant improvement. The same selectivity of false over veridical memories was observed in a follow-up nap study. Unlike previous studies implicating deep, slow-wave sleep (SWS) in declarative memory consolidation, here veridical recall correlated with *decreased* SWS, a finding that was observed in both the overnight and nap studies. These findings lead to two counterintuitive conclusions – that under certain circumstances sleep can promote false memories over veridical ones, and SWS can be associated with impairment rather than facilitation of declarative memory consolidation. While these effects produce memories that are less accurate after sleep, these memories may, in the end, be more useful.

© 2009 Elsevier Inc. All rights reserved.

0. Introduction

Growing evidence suggests that sleep plays an important role in memory consolidation (Payne, Ellenbogen, Walker, & Stickgold, 2008b; Rasch & Born, 2007; Smith, 1995; Stickgold, 2005; Walker & Stickgold, 2006). While sleep's benefit was once thought to apply mainly to procedural forms of memory, it has recently been shown to benefit declarative memory as well (see Marshall & Born, 2007; Payne et al., 2008b for review). Memory consolidation is often conceptualized as a time-dependent, off-line process that stabilizes memories against interference and decay, allowing them to persist over time (McGaugh, 2000). This notion of memory stabilization implies that memories are solidified in high fidelity, true to their original form. Yet substantial evidence shows that memories can become increasingly distorted with time (Bartlett, 1932; McDermott, 1996; Payne, Elie, Blackwell, & Neuschatz, 1996; Seamon

et al., 2002), suggesting that the process of consolidation does not always yield veridical representations of our experiences.

A large body of research has focused on the formation of false memories, in which people recollect events that never occurred (Brainerd & Reyna, 2005; Gallo, 2006; Roediger & McDermott, 2000; Schacter & Slotnick, 2004). Yet, while a growing number of studies support a role for sleep in the consolidation of veridical information, it is unknown whether sleep also influences the development of false memories. Understanding whether sleep affects the formation of false memories is important because it is directly related to questions about how memories are consolidated and stored, how memory representations change over time, and whether these changes can be useful and adaptive.

Here, we tested whether sleep influences false recall, using a list learning task known as the Deese–Roediger–McDermott (DRM) paradigm (e.g. Roediger & McDermott, 1995). This declarative memory task reliably produces high rates of confident false memories for unstudied "critical" words (e.g. *window*) that are semantically associated to studied wordlists (e.g. *door, glass, pane, shade, ledge, sill, house, open, curtain,* etc.). Previous research has demonstrated that long-term memory for critical words actually exceeds veridical memory for studied words (McDermott, 1996; Payne et al., 1996; Seamon et al., 2002; Toglia, Neuschatz, & Goodwin,

* Corresponding authors. Address: Center for Sleep and Cognition, Department of Psychiatry, Harvard Medical School and Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02215, United States. Fax: +1 617 667 8498.
E-mail addresses: jdpayne@bidmc.harvard.edu (J.D. Payne), rstickgold@hms.harvard.edu (R. Stickgold).

1999). For example, McDermott (1996) demonstrated that a 2-day delay between study and test produced levels of false recall that exceeded levels of veridical recall, noting that, unlike many DRM studies of immediate memory where veridical and false recall tend to increase together, over longer delays false memories persist over veridical ones. Thus, in addition to the encoding and retrieval factors known to influence false memory (Brainerd & Reyna, 2005; Gallo, 2006), these studies raise the possibility that slow, offline memory consolidation processes influence false memory development as well. This prediction seems particularly plausible given growing evidence that sleep-based consolidation does more than just stabilize memories in veridical form, but also transforms them in ways that render memories less accurate in some respects, but perhaps more useful in the long run (Ellenbogen, Hu, Payne, Titone, & Walker, 2007; Payne, Stickgold, Swanberg, & Kensinger, 2008a; Wagner, Gais, Haider, Verleger, & Born, 2004).

There is a growing consensus in the literature that the consolidation of hippocampus-dependent memories is modulated by deep, slow-wave sleep (SWS) (Marshall & Born, 2007). SWS is characterized by slow (1–4 Hz), high amplitude brain waves in the EEG and is associated with hippocampal sharp wave-ripples (SPW-Rs), events that may provide a means of communication between hippocampal and neocortical memory stores as memories undergo the process of consolidation (Buzsaki, 1996, 1998). Spatial navigation studies in rodents and humans have shown that hippocampal networks involved in spatial memory acquisition can be reactivated during sleep – particularly SWS (Peigneux et al., 2004; Wilson & McNaughton, 1994), and that this reactivation is linked to improved performance the following day in humans (Peigneux et al., 2004). SWS appears to play a similar role in the veridical consolidation of hippocampus-dependent declarative memories (Marshall & Born, 2007 for review; Rasch, Buchel, Gais, & Born, 2007). For example, Rasch et al. (2007) exposed human subjects to an odor cue (a rose scent) while they learned object-location pairings in the memory game 'concentration' during the evening. fMRI revealed increased hippocampal activation in response to the odor when presented during SWS the following night, and this led to improved declarative memory retention the following morning. Accurate performance on this task, which requires good memory for objects, as well as the ability to correctly bind objects to their specific locations, requires the highly specific relational contextual processing known to depend on the hippocampus (Cohen & Eichenbaum, 1995; Davachi & Wagner, 2002; Giovanello, Schnyer, & Verfaellie, 2004; O'Keefe and Nadel, 1978). These studies and others (e.g. Takashima et al., 2006) strongly suggest that SWS plays a role in the consolidation of hippocampus-dependent forms of memory.

The DRM task differs from these tasks, however, in that it draws on both of the major components of declarative memory – episodic (context-specific event memory), and semantic (context-independent conceptual knowledge).[1] Remembering detailed information about the experimental context, such as the sound of the words as they were presented and characteristics of the speaker's voice, are episodic memory components (i.e. specific to the experimental context or episode), whereas knowing that all of the words in a list are related in meaning is a semantic memory component (i.e. based on pre-existing knowledge of the shared meaning among the words).

While false memory of critical words is thought to rely solely on semantic processing (because there is no contextual information available for non-presented words), correct memory for studied words relies on both context-specific episodic processing and, perhaps to a greater degree, on context-independent semantic pro-

cessing (simply knowing the theme of a word list allows some accurate retrieval). Consistent with this notion, recent neuroimaging studies have demonstrated that both false and veridical memory formation in the DRM task rely heavily on regions associated with semantic processing, such as the left ventrolateral prefrontal cortex and left lateral temporal cortex, while veridical memory formation also relies on medial temporal regions, including the hippocampus (Dennis, Kim, & Cabeza, 2007; Kim & Cabeza, 2007a; Kubota et al., 2006). Thus, although performance on spatial and episodic memory tasks benefit from SWS, accurate performance on the DRM task, with its strong semantic component, may draw on a different complex of neural resources and thus different sleep-stages than the strictly hippocampus-dependent tasks described in the sleep and memory literature to date.

1. Experiment 1

1.1. Methods

1.1.1. Participants

Healthy, medication-free college students (mean age = 20.5) from two Boston area colleges participated for payment or course credit. We initially conducted this experiment at Merrimack College in N. Andover, MA (n = 101 total), and subsequently repeated it at Harvard University (n = 84 total). The Harvard study served to replicate the sleep/wake differences observed in the Merrimack subjects (see Results), and to provide a matched baseline for subsequent sleep polysomnographic (PSG) experiments using the Harvard population (see Experiments 2 and 3). Because performance patterns in the two colleges were virtually identical, all analyses in the main text reflect their combined performance. Individual college statistics can be found in the Supplementary Information online. Given that Merrimack and Harvard colleges represent different populations, the similarities across schools increase our confidence in the robustness and reliability of the results.

All subjects provided informed consent, which was approved by local IRBs, and were screened for self-reported sleep and mental health disorders, irregular sleep habits, and medication use. Subjects maintained their normal sleep schedule for two days prior to the experiment, and were required to sleep for at least 6 h each night. Subjects reported mean bedtimes of 12:28AM, rise times of 8:12AM, and sleep times of 7.4 h. In addition, participants abstained from caffeine and alcohol for two days before and throughout the experiment.

1.1.2. Procedures

All subjects listened to a recording of eight DRM wordlists (Roediger & McDermott, 1995), and later attempted to recall them. Subjects were randomly assigned either to study the lists at 9AM, returning for testing at 9PM that evening ("Wake" group, n = 29 at Merrimack; n = 43 at Harvard), or to study the lists at 9PM, returning for testing at 9AM the next morning ("Sleep" group, n = 27 at Merrimack; n = 41 at Harvard). Two additional Merrimack College groups studied the wordlists at either 9AM (n = 24) or 9PM (n = 21) and were tested for recall just 20 min later ("AM" and "PM Control" groups, respectively), in order to obtain baseline measures of memory recall after a short delay, and also to rule out potential circadian influences on encoding and retrieval. Note that because the AM and PM control groups were run at Merrimack College only, all analyses comparing Sleep and Wake performance to these 20 min delay baselines (e.g. Fig. 2) were performed using Merrimack subjects.

Subjects were tested in small groups. They were told that they were participating in a memory test, and that they should listen carefully to the words they were about to hear because they would be tested on them later. They were then presented with eight DRM

---

[1] It should be noted that while many theorists agree that episodic and semantic memories represent separate memory systems, with episodic memories relying more on hippocampal processing than semantic memories (Moscovitch et al., 2005), not all researchers agree with this idea (Manns, Hopkins, & Squire, 2003).

lists, corresponding to the following critical words: *window*, *doctor*, *chair*, *rough*, *anger*, *soft*, *cup*, and *mountain*. Each list consisted of the twelve associated words with the highest relatedness ratings for that critical word (Stadler, Roediger, & McDermott, 1999). Thus, a total of 96 words were presented. Words were presented aurally, in descending strength of association, at a rate of one word every 2 s. Following the final word of each list, there were 12 s of silence, followed by a 1 s tone, followed by 2 s of silence, followed by the first word of the next list. Words were recorded in an unfamiliar male voice. Subjects heard the lists only one time. They were then released to go about their normal activities until the time of the recall test.

At recall, subjects were given a blank piece of paper and asked to recall as many words as possible from the lists they heard previously. They were informed that they had 10 min to recall as many words as they could remember. After 8 min had passed, they were told they had 2 min remaining.

1.1.3. Memory measures

Recalled words were categorized as studied words (those heard during the initial session), critical words (the central, unstudied word associated with each list), or intrusions (other non-studied words reported at recall). Results for studied words are presented both as (1) overall recall – the total number of studied words recalled and (2) corrected recall – the number of studied words minus the number of intrusions – which was used to correct for possible recall bias.

1.2. Results

Overall, sleep led to greater recall of both studied words and unstudied critical words (but not intrusions) than did wake. A repeated measures ANOVA, comparing performance in the Sleep and Wake groups across the three categories of recall, revealed a highly significant interaction $[F(2, 276) = 12.1$, $P < 0.0001$, $\eta_p^2 = 0.08]$. Recall of studied words was significantly better in the Sleep group than the Wake group, both for overall recall $(21.9 \pm 1.2$ *vs.* 15.7 ± 0.9 (mean words recalled ± s.e.m.), $t(138) = 4.1$, $P < 0.0001$, $d = 0.7$) and for corrected recall $(16.2 \pm 1.4$ *vs.* 9.5 ± 1.1, $t(138) = 3.8$, $P = 0.0002$, $d = 0.6$; Fig. 1A; see Supplementary Information and Fig. S1 for individual college statistics and Fig. S2a for results depicted as proportion correct). Subjects in the Sleep group also falsely recalled more critical words (27%) than subjects in the Wake group $(3.6 \pm 0.2$ *vs.* 2.9 ± 0.2, $t(138) = 2.8$, $P = 0.005$, $d = 0.5$; Fig. 1B; see Supplementary Information for more details about critical word recall and Fig. S2b for results depicted as proportion correct). In contrast, intrusion errors

(false recall of other non-studied words) were non-significantly lower in the Sleep group $(5.6 \pm 0.9$ *vs.* 6.2 ± 0.7, $P = 0.60$; Fig. 1B). This finding, in addition to the corrected recall differences, rules out a general output bias after sleep.

Interestingly, similar to other studies investigating recall in the DRM task across long delay intervals (McDermott, 1996; Payne et al., 1996; Seamon et al., 2002), subjects in the Sleep and Wake groups recalled 23% and 16% of studied words, respectively, but falsely recalled 46% and 36% of the critical words; therefore, in both groups, recall of critical words exceeded recall of studied words following a 12 h delay.

Recall performance was similar in the AM and PM control groups, with no significant differences between them emerging in overall recall of studied words $(22.9 \pm 1.5$ *vs.* 25.6 ± 1.8, respectively; $P = 0.24$), corrected recall of studied words $(19.1 \pm 1.6$ *vs.* 22.0 ± 2.0, respectively; $P = 0.27$), recall of critical words $(3.4 \pm 0.32$ *vs.* 3.7 ± 0.41, respectively; $P = 0.57$), or recall of intrusions $(4.1 \pm 0.83$ *vs.* 4.1 ± 0.93, respectively; $P = 0.99$). These findings argue that circadian processes did not significantly affect encoding or recall of the words at these specific times. Standard measures of subjective sleepiness, acquired using the Stanford Sleepiness Scale (Hoddes, Zarcone, Smythe, Phillips, & Dement, 1973), also were not significantly different between the AM and PM control groups $(3.2 \pm 0.2$ *vs.* 3.0 ± 0.2; $P = 0.72$), suggesting further that circadian differences in cognitive performance or general alertness do not account for the recall differences seen between the Sleep and Wake groups.

To explore how memory changed across the 12 h retention interval, we next compared recall 20 min after study (for the combined AM and PM control groups) to recall 12 h after study. Recall of studied words deteriorated significantly from the 20 min baseline in both the Wake $[t(72) = 6.2$, $P < 0.0001$, $d = 1.1]$, and the Sleep $[t(70) = 3.9$, $P = 0.0002$, $d = 1.0]$ groups (Fig. 2, left). However, the deterioration was significantly more pronounced in the Wake group than in the Sleep group $(-42.9 \pm 4.6\%$ *vs.* $-28.7 \pm 0.5.2\%$, $t(54) = 2.1$, $P = 0.04$, $d = 0.5$), suggesting that sleep protects memories against deterioration over time (Jenkins & Dallenbach, 1924).

Recall of critical words also decreased significantly from baseline, but only in the Wake group $[-22.4 \pm 8.6\%$; $t(72) = 2.2$, $P = 0.03$, $d = 0.4$; Fig. 2, right, open bar]. In contrast, recall of critical words actually increased, albeit non-significantly, after sleep $(+2.1 \pm 9.0\%)$, and showed better recall than after wake $[t(54) = 2.0$, $P = 0.05$, $d = 0.5$; Fig. 2, right, solid bar]. In support of other DRM studies examining delayed retrieval (e.g. McDermott, 1996; Payne et al., 1996; Seamon et al., 2002), this finding demonstrates a divergence between studied words and critical words
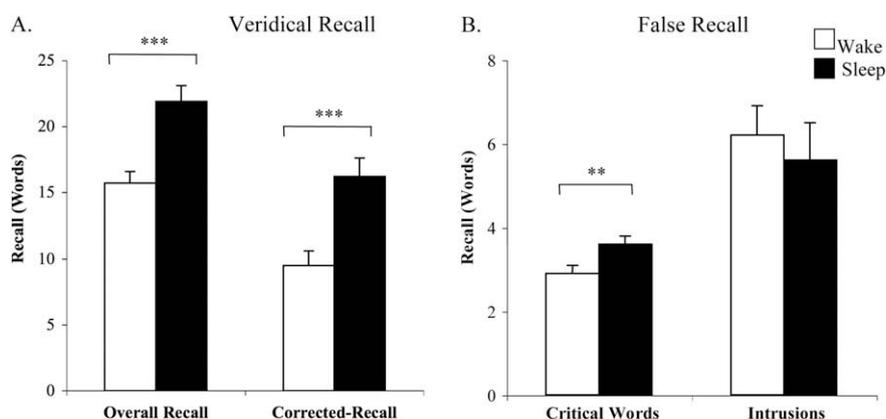
Fig. 1. Sleep increases both veridical and false recall. Results represent combined performance in both college populations (see Supplementary Information for individual college statistics). (A) Increased recall of studied words (overall and corrected recall) and (B) increased recall of critical words (strongly semantically associated, but unstudied words). Note that sleep did not increase false recall of words other than the critical word (intrusions). Error bars = s.e.m. ***$P < 0.001$; **$P < 0.01$.
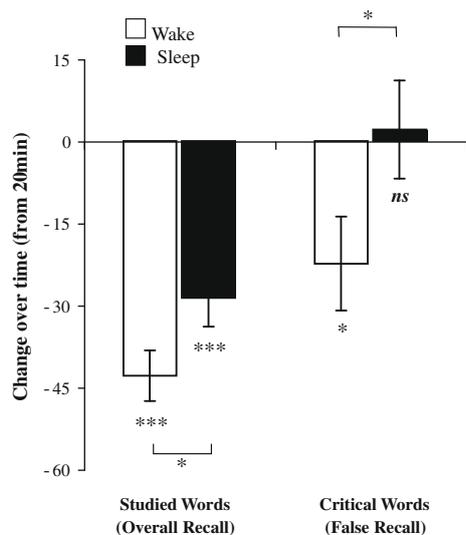
Fig. 2. Change in recall performance across 12 h, relative to the 20 min baseline, in the Wake and Sleep groups. Values have been converted to percents to equate recall of studied words (out of 96 possible words) and critical words (out of eight possible words). While veridical recall of studied words deteriorated significantly from baseline in both the Wake and Sleep groups, false recall of critical words deteriorated from baseline only in the Wake group. False recall in the Sleep group non-significantly *increased* from baseline. Error bars = s.e.m. ***$P < 0.001$; *$P < 0.05$.

over time, and suggests that the observed increase in false recall following sleep cannot easily be explained as a byproduct of increased veridical recall. It also suggests that sleep may be required for the preferential recall of false over veridical memories observed across long delays (e.g. McDermott, 1996; Payne et al., 1996; Seamon et al., 2002).

2. Experiment 2

To determine which specific sleep stages are correlated with the increased recall seen in the Sleep group, we polysomnographically (PSG) monitored the nocturnal sleep of an additional group of subjects. As before, subjects were trained at 9PM and tested at 9AM the following morning. But these subjects spent the intervening night in the sleep laboratory.

2.1. Methods

2.1.1. Participants and procedures

Subjects recruited at Harvard University (n = 22) arrived at the sleep laboratory at Beth Israel Deaconess Medical Center at approximately 8:30PM. They provided informed consent and then listened to the DRM wordlists in a quiet testing room. They were then moved to the sleep laboratory where they were wired for PSG recording – a process that typically took about an hour. During this time, subjects were engaged in conversation with the experimenter and assistants in order to prevent intentional rehearsal of the words. Subjects were allowed to read until they were ready to go to sleep, and sleep recording began at lights out. Subjects were awakened following a 9 h sleep opportunity (typically between 7 and 8AM), at which point the electrodes were removed, and a shower and breakfast were provided. Subjects then returned to the testing room for the recall test at approximately 9AM.

Sleep was recorded with an Embla A1 digital system. The montage included EOG (electrooculography), EMG (electromyography), and EEG leads (O1, O2, C3, Cz, C4), with each electrode referenced to the contralateral mastoid. Sleep data were scored according to the standards of Rechtschaffen and Kales. Data from one subject had to be discarded, due to a failure of the PSG equipment. A sum-

mary of sleep measures is provided in Table 1. Subjects slept an average of 7.7 h, with a mean sleep latency of 27 min from lights out (see Supplementary Information). Average times spent in different sleep stages did not differ from established norms.

2.2. Results

Notably, recall scores for the overnight PSG sleep subjects were comparable to the subjects who slept at home, with no significant differences in overall recall of studied words (25.7 ± 1.7 *vs.* 21.9 ± 1.2, respectively; $P = 0.11$), corrected recall of studied words (19.0 ± 2.4 *vs.* 16.2 ± 1.4, respectively; $P = 0.33$), recall of critical words (4.0 ± 0.2 *vs.* 3.6 ± 0.2, respectively; $P = 0.36$), or recall of intrusions (6.7 ± 1.1 *vs.* 5.6 ± 0.9, respectively; $P = 0.54$). The similarities in performance between the Sleep groups minimize concerns about environmental differences between the sleep laboratory and home.

In contrast to a wealth of evidence suggesting a beneficial effect of slow wave sleep (SWS) on standard declarative memory tasks (Marshall & Born, 2007), overnight recall of studied DRM words showed a significant but *negative* correlation with time spent in SWS, whether calculated for overall time spent in SWS [$r(21) = -0.47$, $P = 0.03$; Fig. 3A] or percent of total sleep time spent in SWS (SWS%) [$r(21) = -0.55$, $P = 0.009$; Fig. 3B]. Similar correlations were observed between corrected recall and SWS [$r(21) = -0.60$, $P = 0.004$; Fig. S3], and SWS% [$r(21) = -0.53$, $P = 0.015$; Fig. S4]. Power in the delta (1–4 Hz) band was assessed via spectral analysis of all artifact- and arousal-free NREM epochs (see Supplementary Information). Spectral power density was calculated using Welch's method, applied to successive 4 s epochs (Hanning window, 50% overlap). Spectral analyses revealed no significant correlations between recall performance and relative delta power either within all of NREM sleep or within SWS specifically (all Ps > 0.20).

There was also a positive correlation between the percent of Stage 2 NREM sleep obtained during the night (Stage 2%) and recall (overall recall: $r(21) = 0.49$, $P = 0.02$; Fig. S5; corrected recall: $r(21) = 0.49$, $P = 0.02$). However, correlations between total time spent in Stage 2 sleep and both overall and corrected recall failed to reach significance. Moreover, when subjected to a stepwise regression analysis, the Stage 2% correlation with recall was rejected, suggesting that this positive correlation was secondary to the negative SWS correlation. Correlations between sleep stages and recall of critical words could not be meaningfully analyzed, because of the narrow range of critical words recalled (80% of overnight subjects recalled either four or five of the eight possible critical words).

3. Experiment 3

To further clarify the relationship between sleep and memory performance in the DRM task independent of nocturnal period

Table 1
Sleep parameters for subjects in overnight sleep study (Section 2).

Sleep parameter	Mean time (min) ± s.e.m	% Total sleep time ± s.e.m
Total sleep time	461 ± 12	
Wake after sleep onset	43 ± 10	
Sleep latency	27 ± 7	
Stage 1	33 ± 3	7.2 ± 0.7
Stage 2	251 ± 9	54.4 ± 2.0
Stage 3	40 ± 3	8.7 ± 0.7
Stage 4	45 ± 4	9.8 ± 0.9
SWS (Stages 3 + 4)	86 ± 4	18.7 ± 0.9
REM	90 ± 5	19.5 ± 1.1

Note: All measures are in minutes. Sleep Latency = latency to sleep onset (first epoch of sleep). Figs. S1–S4, Stages 1–4; SWS, slow wave sleep; REM, Rapid Eye Movement sleep.
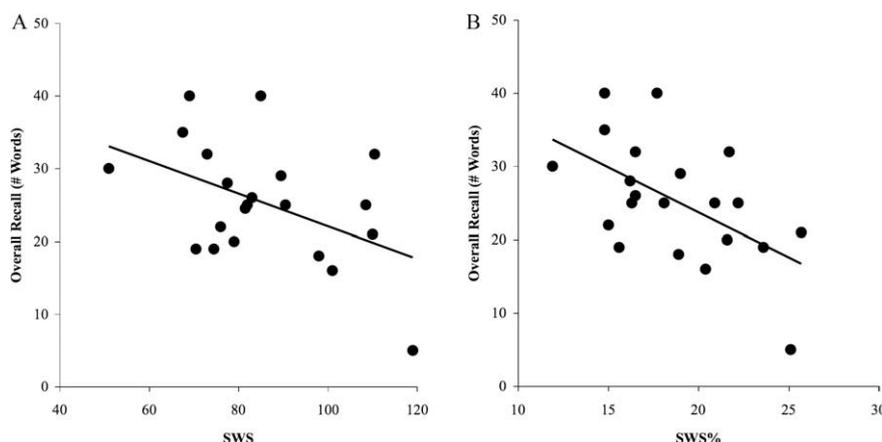
Fig. 3. Negative correlation between overall recall and SWS. (A) Correlation between overall recall and total minutes spent in SWS, $r = -0.47$, $P = 0.03$ and (B) Correlation between overall recall and percent of total sleep time spent in SWS, $r = -0.55$, $P = 0.009$. See Fig. S2 for corrected recall correlations.

and better controlling for circadian and interference factors, and also to replicate the negative correlation between SWS and correctly recalled words found in Section 2, we conducted an additional PSG experiment across a daytime nap.

3.1. Methods

3.1.1. Participants and procedures

Subjects ($n = 30$) were recruited from Harvard University, trained on the wordlists at noon and tested at 6:30PM after either napping for an average of 88 min ($n = 16$; see Table 2 for sleep parameters), or remaining awake ($n = 14$). On each testing day, two subjects reported to the laboratory at 12:00PM to listen to the wordlists. Both subjects were then wired for PSG, after which they were randomly assigned, one to the Nap and one to the Wake condition. The Wake subject watched an emotionally neutral movie during the other subject's nap, which began at approximately 1:15PM. Wake subjects were monitored to ensure they remained awake during this interval. Both subjects then remained in the laboratory listening to books on tape until testing at 6:30PM. One nap subject had to be excluded from EEG analyses due to a power failure.

3.2. Results

Similar to the overnight Sleep groups (Figs. 1 and S1), subjects in the Nap group recalled significantly more critical words than those in the Wake group [4.3 ± 0.4 *vs.* 2.9 ± 0.4, $t(28) = 2.4$, $P = 0.02$, $d = 0.9$; Fig. 4B; see Fig. S2b for results depicted as proportion correct]. In contrast, recall of studied words did not differ between the Nap and Wake groups [26.6 ± 3.0 words in the Nap

group *vs.* 26.4 ± 4.2 words in the Wake group; $t(28) = 0.05$, $P > 0.90$; Fig. 4A], corrected recall [18.6 ± 3.3 *vs.* 17.8 ± 5.9; $t(28) = 0.13$, $P > 0.80$; Fig. 4A; see Supplementary], nor did the number of intrusions [8.0 ± 1.3 *vs.* 8.6 ± 2.5, $t(28) = 0.23$, $P > 0.80$; Fig. 4B]. These findings again suggest that sleep may selectively promote the recall of critical over studied words in the DRM task.

Importantly, we again observed the significant negative correlation between SWS and recall [$r(15) = -0.54$, $P = 0.037$; Fig. 5]. This time, the correlation with Stage 2 NREM sleep did not emerge, providing further support for our earlier finding that the Stage 2 NREM correlation with overnight recall may be secondary to the SWS correlation. As in Section 2, spectral analyses failed to reveal significant correlations between recall performance and relative delta power, either within all of NREM sleep or within SWS specifically (all Ps > 0.30).

4. Discussion

Several studies using the DRM task have demonstrated that false memories for semantic associates are more likely to persist than veridical memories across long time delays (McDermott, 1996; Payne et al., 1996; Seamon et al., 2002; Toglia et al., 1999). These studies raise the possibility that slow, offline processes, many of which are preferentially active during sleep (Ellenbogen, Payne, & Stickgold, 2006; Payne et al., 2008b for review), may influence false memory development. Our data suggest that sleep indeed plays a role in the evolution of false memories, with sleep increasing recall of critical words over studied words compared to equivalent periods of wakefulness. The nap study (Section 3) is most convincing in this regard, showing that sleep led to nearly 50% greater recall of critical words compared to wake controls, but did not lead to increased recall of veridical memories (26.6 *vs.* 26.4 words recalled in the nap and wake groups, respectively). Similarly, in Section 1, in spite of substantial deterioration from baseline of veridical memories across wake and sleep (though deterioration was significantly greater across wake), subjects showed a marked reduction in false memories across wakefulness, but no decrease at all after a night of sleep. Taken together, these findings suggest that while sleep influences the consolidation of both veridical and false memories in the DRM task, it has its biggest impact on the latter when assessed via recall memory (see Diekelmann, Landolt, Lahl, Born, and Wagner (2008) for different results with recognition memory).

Our results raise several questions. First, why were critical words preferentially recalled across periods of sleep? That this is not simply a global enhancement of recall after sleep is clearest

Table 2
Sleep parameters for the nap study (Experiment 3).

Sleep parameter	Mean + s.e.m	% Total sleep time ± s.e.m
Total sleep time	88.1 ± 7.2	
Wake after sleep onset	12.6 ± 3.7	
Sleep latency	13.8 ± 3.9	
Stage 1	9.2 ± 0.8	15.6 ± 5.5
Stage 2	42.2 ± 4.7	44.9 ± 4.0
Stage 3	12.4 ± 2.9	13.2 ± 2.8
Stage 4	8.9 ± 2.7	10.4 ± 3.2
SWS (Stages 3 + 4)	21.2 ± 3.6	23.6 ± 4.3
REM	15.5 ± 3.0	15.8 ± 2.9

Note: All measures are in minutes. Sleep Latency = latency to sleep onset (first epoch of sleep), Figs. S1-S4, Stages 1–4; SWS, slow wave sleep; REM, Rapid Eye Movement sleep.
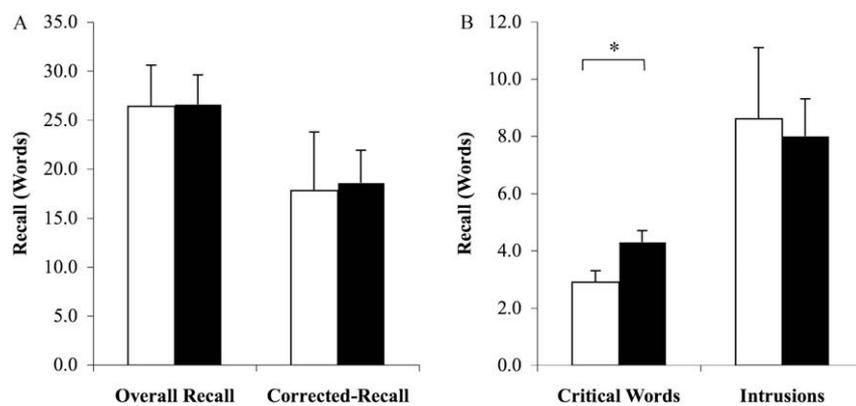
Fig. 4. An afternoon nap selectively increases false recall. (A) equivalent recall of studied words in the Nap and Wake groups (overall and corrected recall) and (B) increased recall of semantically associated, but unstudied words (critical words) in the nap group. Error bars = s.e.m. *$P = 0.02$.
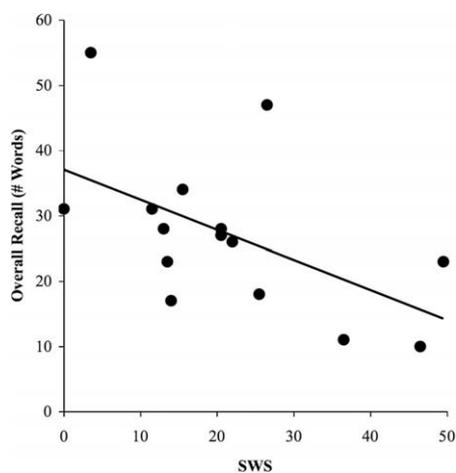
Fig. 5. Negative correlation between overall recall and total minutes spent in SWS in the nap study, $r = -0.54$, $P = 0.037$ (Section 3).

in Section 3, where recall of studied words was virtually identical in the nap and wake conditions, yet recall of critical words was substantially greater in the nap condition. A similar effect was seen in Section 1, where recall of critical words deteriorated significantly from the 20-min baseline in the Wake, but not the Sleep group. Thus, sleep appears to have produced an enrichment in the "recall" of these false memories.

Several findings lessen concerns about circadian or interference accounts of these results. Since both the nap and wake groups encoded the words at approximately 12PM and retrieved them at 6:30PM in Section 3, any circadian effects would necessarily be identical for both groups. Similarly, the preferential overnight retention of critical words seen in Section 1 cannot simply reflect a circadian influence of nighttime on retention, as sleep in the afternoon produces similar effects. Concerns about circadian confounds are further allayed as recall of critical words did not differ between the AM and PM 20 min delay control groups.

It also seems unlikely that interference effects can solely account for the increase in recall of critical words following the nap. Wake subjects experienced only an average of 88 min more wakefulness than nap subjects in Section 3. While this time interval could allow significant interference to occur, it is not obvious why increasing the time available for interference from 270 to 360 min would provoke such a large (50%) increase in critical word recall. More importantly, however, a strict interference account would predict equal protection of all memories during sleep, yet

this was clearly not the case. Only memory for critical words was enriched by sleep in the nap study, and sleep preferentially promoted critical words over studied words in the overnight study as well. This selective enhancement of critical words, with no effect on studied words in the nap study, is difficult to explain in terms of interference factors alone.

A second question raised by our results is why, if subjects correctly recalled more studied words after a night of sleep than across a day of wakefulness in Section 1, did the nap and no-nap groups produce virtually identical recall of studied words in Section 3? One possibility is that sleep produces an active strengthening of the memories, and the longer overnight sleep period provided greater strengthening than the briefer nap. But given that recall was negatively correlated with SWS in *both* groups, it seems unlikely that this was a major factor in the difference between the full night and nap groups.

An alternative possibility is that sleep provides merely passive protection against the weakening of memories by waking interference, and the greater disparity in time spent asleep in the 12 h Sleep and Wake groups compared to the Nap and No-nap groups leads to a greater disparity in recall in the 12 h condition. If this were the case, total sleep time (TST) and performance should be positively correlated in the overnight study, yet neither overall recall ($r = 0.10$, $P = 0.67$) nor corrected recall ($r = -0.12$, $P = 0.59$) of studied words correlated with TST. Instead, strong correlations between amounts of SWS and post-sleep recall of studied words ($r = -0.47$ and -0.55) and similar regression slopes (0.44 and 0.30 words/min SWS) emerged for overnight and nap studies. An interference account cannot easily account for these correlations, and would actually predict that time spent in SWS should correlate positively with recall, as there is minimal exposure to potentially interfering mentation during SWS compared to the intense mental activity seen in other sleep states (i.e. the dreams of Rapid Eye Movement (REM) sleep). The SWS correlations also argue that circadian influences cannot explain the beneficial influence of sleep on veridical recall. This argument is further supported by the lack of differences in veridical recall between the AM and PM 20 min delay control groups, which suggests that circadian influences do not affect encoding or retrieval processes.

While these arguments do not rule out at least some role for interference and circadian influences in the sleep effects seen here (Schmidt, Collette, Cajochen, & Peigneux, 2007; Wixted, 2004, 2005), they strongly suggest that sleep plays an active, rather than merely passive, role in memory consolidation (Ellenbogen et al., 2006; Payne et al., 2008b for review).

Our findings raise two final questions that are more difficult to address, and about which we can only speculate. First, how do we

understand the consequences of a system that preferentially enriches false memories? One possibility is that when faced with a large amount of similar or related material, an adaptive memory system may preserve only what is most relevant to future needs (Bartlett, 1932). Thus, in the DRM task, an efficient system might preferentially extract and retain the general theme or gist of information over the specific details, unless subjects are instructed otherwise (which was not the case in this study) (Brainerd & Reyna, 2001, 2005; Reyna & Brainerd, 1998).

Because there are several theoretical accounts of the DRM false memory effect (see Gallo (2006) for review), it is important to acknowledge that false memories for critical words may arise from memory for gist (Brainerd & Reyna, 2005), activation of associated words (Roediger, Watson, McDermott, & Gallo, 2001), or some combination of both. In the latter account (Roediger, Balota, & Watson, 2001; Underwood, 1965), activation in a semantic network spreads from representations of presented words to critical words, and associative (potentially summative) activation of the critical word is subsequently misattributed to real experience. Both views suggest that critical words receive considerable processing, which could lead to the tagging of the word for subsequent sleep-dependent consolidation (Buzsaki, 1998; Sejnowski and Destexhe, 2000). While our results clearly demonstrate that sleep plays a role in the long-term recall of critical words, they do not distinguish between the alternative explanations of false memory formation, nor do they rule out an effect of sleep on the retrieval (*vs.* consolidation) of critical over veridical words. Moreover, false recall of semantically related materials is only one type of memory distortion, and it is unclear whether sleep's influence would extend to other forms of false memory (e.g. falsely accepting misleading information). These are important questions for future research to resolve.

The second question is how to explain the finding that across a night of sleep, as well as across an afternoon nap, recall of studied words in the DRM paradigm correlates *negatively* with the amount of SWS obtained. Given that SWS is sensitive to sleep deprivation and associated with rebound effects, one possibility is that the poor recall associated with abundant SWS is an artifact of poor sleep (that is, the abundant SWS in subjects with poor recall is a signature of prior sleep deprivation, and such subjects may be impaired during encoding of the lists). If this hypothesis has merit, veridical recall should be correlated with sleep duration the night prior to encoding, and/or the night of the experiment (in between training and test) in the Sleep group, but this was not the case (Night prior to encoding; overall recall, $r = 0.01$, $P = 0.97$; corrected recall, $r = -0.02$, $P = 0.93$; Night of the experiment; overall recall, $r = 0.09$, $P = 0.55$; corrected recall, $r = 0.14$, $P = 0.38$). Moreover, recall did not correlate with ratings on the Stanford Sleepiness Scale (all Ps > 0.40).

So how do we understand this result, particularly in light of the *beneficial* influence of SWS on the consolidation of hippocampus-dependent spatial and episodic memories (Marshall & Born, 2007; Peigneux et al., 2004; Rasch et al., 2007; Takashima et al., 2006)? Although performance on the DRM task can benefit from detailed, context-specific episodic memory, it is much more reliant on general, semantic memory processing. In fact, several studies of associative false memory have shown that semantic processing enhances not only false memory for critical words, but also true memory for studied words (Gallo, Roediger, & McDermott, 2001; Kim & Cabeza, 2007a, 2007b; Rhodes & Anastasi, 2000; Toglia et al., 1999). Kim and Cabeza (2007a) found a significant correlation between recall of studied words and recall of critical words, which they suggest provides evidence that semantic processing contributes to both false and veridical memory formation (Kim & Cabeza, 2007a, p. 2145).

We found the same correlation between studied and critical words in our study (Section 1), both in the wake ($r = 0.55$,

$P < 0.0001$) and Sleep groups ($r = 0.34$, $P = 0.001$), although interestingly a Fisher's r–z transformation revealed that the correlation was stronger following wake than following sleep ($P = 0.05$ for the comparison of the two correlation coefficients), likely because sleep has a proportionally greater influence on false than on true memory as indicated above (see Fig. 2). These correlations suggest that while recall of studied and critical words can be dissociated by various manipulations (Toglia et al., 1999), including the sleep/wake manipulation used here, they may nevertheless draw on some of the same memory resources – namely those used for semantic processing (Kim & Cabeza, 2007a, 2007b). While we were unable to investigate the relation of critical word recall to specific sleep stages due to low variability in recall of these words (of which there were only eight in total), given that both veridical and false recall stem from semantic processing of related word lists (Toglia et al., 1999), we would predict critical word recall, like veridical recall, to be negatively related to SWS. Studies are currently underway to test this prediction.

In light of the differences between the DRM task and the spatial and episodic tasks described in the sleep literature to date (Marshall & Born, 2007), the negative correlations found here may suggest that SWS, while facilitating the consolidation of detailed contextual and episodic memories, may impair subsequent performance on tasks that benefit from general semantic knowledge. All aspects of performance on the DRM task rely much more on semantic processing than other tasks used in the sleep and declarative memory literature. Because many researchers agree that episodic and semantic memory systems rely on different brain systems (e.g. Tulving, 2002), and at times perform antagonistic functions (i.e. storing veridical details to keep memories separate *vs.* extracting semantic regularities to emphasize what memories share in common), there may be good reason for the two types of memory to utilize SWS in different ways. Nonetheless, our results reflect correlations and do not demonstrate causation, and further studies are required to clarify the nature of this relationship.

An alternative explanation is that the negative correlations reflect trait-like differences in SWS, with subjects who habitually obtain less SWS possibly suffering from deficient consolidation of item-specific memory and/or relying more on pre-existing semantic knowledge. The relationship between individual differences in sleep architecture and strengths and weaknesses in different forms of memory processing has received little attention but is a promising avenue for future research.

Our results suggest that sleep plays a role in the preservation and persistence of false memories. On the face of it, this seems strikingly maladaptive. Why would one want to falsely recall information that was never encountered? In most cases this would certainly be undesirable. But the DRM task, with its reliance on semantic processing, appears to generate a special kind of false memory, where words representing the semantic meaning of a list are remembered even though they were never presented. Given that our brains cannot possibly store every detail we encounter, it may sometimes be beneficial to remember the meaning or 'gist' of information at the expense of the details, even if that information is represented by a word that was never studied.

Our findings add to a small but growing body of work suggesting that sleep does more than simply consolidate memories in veridical form, additionally transforming and restructuring them so that insights and abstractions can be made (Gomez, Bootzin, & Nadel, 2006; Wagner et al., 2004), inferences can be drawn (Ellenbogen et al., 2007), integration can occur (Dumay and Gaskell, 2007), and emotionally salient aspects of information can be preferentially remembered over neutral aspects (Payne et al., 2008a). Susceptibility to memory distortion might be the price we pay for the flexible use of our memories – memories that are less accurate, but more useful, following sleep.

Acknowledgments

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.nlm.2009.03.007.

References

Bartlett, F. C. (1932). *Remembering*. Cambridge University Press.

Brainerd, C. J., & Reyna, V. F. (2001). Fuzzy-trace theory: dual processes in memory, reasoning, and cognitive neuroscience. *Advances in Child Development and Behavior, 28*, 41–100.

Brainerd, C. J., & Reyna, V. F. (2005). *The Science of false memory*. Oxford: Oxford University Press.

Buzsaki, G. (1996). The hippocampo-neocortical dialogue. *Cerebral Cortex, 6*, 81–92.

Buzsaki, G. (1998). Memory consolidation during sleep: a neurophysiological perspective. *Journal of Sleep Research, 7*(Suppl. 1), 17–23.

Cohen, N. J., & Eichenbaum, H. (1995). *Memory, amnesia, and the hippocampal system*. Cambridge: The MIT Press.

Davachi, L., & Wagner, A. D. (2002). Hippocampal contributions to episodic encoding: insights from relational and item-based learning. *Journal of Neurophysiology, 88*, 982–990.

Dennis, N. A., Kim, H., & Cabeza, R. (2007). Effects of aging on true and false memory formation: an fMRI study. *Neuropsychologia, 45*, 3157–3166.

Diekelmann, S., Landolt, H. P., Lahl, O., Born, J., & Wagner, U. (2008). Sleep loss produces false memories. *PLoS ONE, 3*(10), e3512.

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

Ellenbogen, J. M., Hu, P. T., Payne, J. D., Titone, D., & Walker, M. P. (2007). Human relational memory requires time and sleep. *Proceedings of the National Academy of Sciences USA, 104*, 7723–7728.

Ellenbogen, J. M., Payne, J. D., & Stickgold, R. (2006). The role of sleep in declarative memory consolidation: passive, permissive, active or none? *Current Opinion in Neurobiology, 16*, 716–722.

Gallo, D. (2006). *Associative illusions of memory: Research on false memory for related events*. New York: Psychology Press.

Gallo, D., Roediger, H. L., & McDermott, K. B. (2001). Associative false recognition occurs without strategic criterion shifts. *Psychonomic Bulletin and Review, 8*, 579–586.

Giovanello, K. S., Schnyer, D. M., & Verfaellie, M. (2004). A critical role for the anterior hippocampus in relational memory: evidence from an fMRI study comparing associative and item recognition. *Hippocampus, 14*, 5–8.

Gomez, R. L., Bootzin, R. R., & Nadel, L. (2006). Naps promote abstraction in language-learning infants. *Psychological Science, 17*, 670–674.

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., & Dement, W. C. (1973). Quantification of sleepiness: a new approach. *Psychophysiology, 10*, 431–436.

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

Kim, H., & Cabeza, R. (2007a). Differential contributions of prefrontal, medial temporal, and sensory-perceptual regions to true and false memory formation. *Cerebral Cortex, 17*, 2143–2150.

Kim, H., & Cabeza, R. (2007b). Trusting our memories: dissociating the neural correlates of confidence in veridical versus illusory memories. *Journal of Neuroscience, 27*, 12190–12197.

Kubota, Y., Toichi, M., Shimizu, M., Mason, R. A., Findling, R. L., Yamamoto, K., et al. (2006). Prefrontal hemodynamic activity predicts false memory – a near-infrared spectroscopy study. *Neuroimage, 31*, 1783–1789.

Manns, J. R., Hopkins, R. O., & Squire, L. R. (2003). Semantic memory and the human hippocampus. *Neuron, 38*, 127–133.

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

McDermott, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory and Language, 35*, 212–230.

McGaugh, J. L. (2000). Memory – a century of consolidation. *Science, 287*, 248–251.

Moscovitch, M., Rosenbaum, R. S., Gilboa, A., Addis, D. R., Westmacott, R., Grady, C., et al. (2005). Functional neuroanatomy of remote episodic, semantic and spatial memory: a unified account based on multiple trace theory. *Journal of Anatomy, 207*, 35–66.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Oxford University Press.

Payne, D. G., Elie, C. J., Blackwell, J. M., & Neuschatz, J. S. (1996). Memory illusions: recalling, recognizing, and recollecting events that never occurred. *Journal of Memory and Language, 35*, 261–285.

Payne, J. D., Ellenbogen, J. M., Walker, M. P., & Stickgold, R. (2008b). The role of sleep in memory consolidation. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference*. New York: Elsevier.

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

Peigneux, P., Laureys, S., Fuchs, S., Collette, F., Perrin, F., Reggers, J., et al. (2004). Are spatial memories strengthened in the human hippocampus during slow wave sleep? *Neuron, 44*, 535–545.

Rasch, B., & Born, J. (2007). Maintaining memories by reactivation. *Current Opinion in Neurobiology, 17*, 698–703.

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

Reyna, V. F., & Brainerd, C. J. (1998). Fuzzy-trace theory and false memory: new frontiers. *Journal of Experimental Child Psychology, 71*, 194–209.

Rhodes, M. G., & Anastasi, J. S. (2000). The effects of a levels-of-processing manipulation on false recall. *Psychonomic Bulletin and Review, 7*, 158–162.

Roediger, H. L., Balota, D. A., & Watson, J. M. (2001). Spreading activation and the arousal of false memories. In H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95–115). Washington DC: American Psychological Association.

Roediger, H. L., & McDermott, C. M. (1995). Creating false memories: remembering words not presented in lists. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 21*, 803–814.

Roediger, H. L., & McDermott, C. M. (2000). Distortions of memory. In E. Tulving & F. Craik (Eds.), *The oxford handbook of memory* (pp. 149–162). New York: Oxford University Press.

Roediger, H. L., 3rd, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: a multiple regression analysis. *Psychonomic Bulletin and Review, 8*, 385–407.

Schacter, D. L., & Slotnick, S. D. (2004). The cognitive neuroscience of memory distortion. *Neuron, 44*, 149–160.

Schmidt, C., Collette, F., Cajochen, C., & Peigneux, P. (2007). A time to think: circadian rhythms in human cognition. *Cognitive Neuropsychology, 24*, 755–789.

Seamon, J. G., Luo, C. R., Kopecky, J. J., Price, C. A., Rothschld, L., Fung, N. S., et al. (2002). Are false memories more difficult to forget than accurate memories? The effect of retention interval on recall and recognition. *Memory and Cognition, 30*, 1054–1064.

Sejnowski, T. J., & Destexhe, A. (2000). Why do we sleep? *Brain Research, 886*(1–2), 208–223.

Smith, C. (1995). Sleep states and memory processes. *Behavioral Brain Research, 69*, 137–145.

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

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

Takashima, A., Petersson, K. M., Rutters, F., Tendolkar, I., Jensen, O., Zwarts, M. J., et al. (2006). Declarative memory consolidation in humans: a prospective functional magnetic resonance imaging study. *Proceedings of the National Academy of Sciences USA, 103*, 756–761.

Toglia, M. P., Neuschatz, J. S., & Goodwin, K. A. (1999). Recall accuracy and illusory memory: when more is less. *Memory, 7*.

Tulving, E. (2002). Episodic memory: from mind to brain. *Annual Review of Psychology, 53*, 1–25.

Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology, 70*, 122–129.

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

Walker, M. P., & Stickgold, R. (2006). Sleep, memory, and plasticity. *Annual Review of Psychology, 57*, 139–166.

Wilson, M. A., & McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science, 265*, 676–679.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology, 55*, 235–269.

Wixted, J. T. (2005). A theory about why we forget what we once knew. *Current Directions in Psychological Science, 14*, 6–9.