Neurobiology of Learning and Memory 137 (2017) 123–133

Contents lists available at ScienceDirect

Neurobiology of Learning and Memory

journal homepage: www.elsevier.com/locate/ynlme

The impact of sleep on true and false memory across long delays

Enmanuelle Pardilla-Delgado *, Jessica D. Payne *

University of Notre Dame, Department of Psychology, United States

ARTICLE INFO

Article history:
Received 4 June 2016
Revised 21 October 2016
Accepted 25 November 2016
Available online 27 November 2016

Keywords:
False memory
Sleep
Gist
Episodic memory
Semantic memory
Slow-wave sleep

ABSTRACT

While the influence of sleep on memory has a long history, sleep's role in the formation of false memories is less clear. Moreover, virtually nothing is known about the development of false memories beyond delays of about 12 h. Here, for the first time, we assess post-sleep development of true and false memories across longer delay intervals of 24 and 48 h. Although technically a false memory, remembering information that is related to the theme, or gist, of an experience can be considered an adaptive process. Some evidence suggests that sleep, compared to a wake period, increases both true and gist-based false memories in the Deese-Roediger-McDermott (DRM) task, but not all studies have returned this result, and most studies cannot rule out the possibility that sleep is merely protecting the information from interference, as opposed to actively aiding its consolidation. Here, to equate amount of time spent awake and asleep across groups, we assess how the positioning of sleep relative to memory encoding impacts retention across longer delays of 24 and 48 h. Participants encoded 16 DRM lists in the morning (WAKE 1st Groups) or evening (SLEEP 1st Groups), and were tested either 24 or 48 h later at the same time of day. Results demonstrate that true memory is better when participants sleep soon after learning. Sleeping first also increased false memory, but only in low performers. Importantly, and similar to previous studies, we found a negative correlation between slow-wave sleep (SWS) and false memory, suggesting that SWS may be detrimental for semantic/gist processing.

1. Introduction

Memory has long been considered to involve reconstructive processes that fill in information that has been lost over time, rather than simply involving a complete and unchanging crystallized entity. Early examples of work that illustrate memory's reconstructive nature include Bartlett's (1932) schema studies and Brewer's (1977) inference work, among others (see Loftus, 2005; Schacter, 1995). Roediger and McDermott revived the investigation of the reconstructive nature of memory by adapting a simple word-list memorization task earlier reported by Deese (1959). In this research paradigm, known as the Deese-Roediger-McDermott (DRM) task, participants study a list of semantically related items (e.g., nurse, hospital, medicine, etc.) and often falsely remember an unpresented, but semantically related critical "lure" word (in this case, doctor) with the same level of confidence as their correctly remembered studied words (Roediger & McDermott, 1995; Seamon et al., 2002).

1.1. The effect of sleep on false memory formation is not clear

Although the effects of sleep on mnemonic processes have been studied for over a century (Ebbinghaus, 1885; Jenkins & Dallenbach, 1924; see Payne, 2011; Payne, Ellenbogen, Walker, & Stickgold, 2008), it was only recently that the relationship between sleep-related consolidation processes and false memory development began to be assessed. While the effects of sleep on episodic memory are mostly beneficial (see Rasch, Büchel, Gais, & Born, 2007; Diekelmann & Born, 2010 for review), the relationship between sleep and false memory remains unclear.

Diekelmann, Landolt, Lahl, Born, and Wagner (2008) first assessed the question of whether sleep would affect performance on the DRM task. They compared recognition memory performance between three groups: nocturnal sleep, nocturnal sleep deprivation, and daytime wakefulness. Sleep deprivation, but not nocturnal sleep, increased the rate of false memories recognized, while leaving true memories unchanged compared to daytime wakefulness. These findings were attributed to the idea that sleep deprivation impairs prefrontal cortex (PFC) activity, which is crucial for successful source monitoring and has been related to decreased false memory recognition (Curran, Schacter, Johnson, & Spinks, 2001).

* Address: University of Notre Dame, Haggar Hall, Room B24, United States.
E-mail addresses: epardill@nd.edu (E. Pardilla-Delgado), jpayne7@nd.edu (J.D. Payne).

http://dx.doi.org/10.1016/j.nlm.2016.11.016
1074-7427/

Fenn, Gallo, Margoliash, Roediger, and Nusbaum (2009) found that a 12 h period filled with sleep, when compared to a similar period filled with daytime wakefulness, had no effect on true recognition memory and reduced false recognition memory when stimuli were presented both aurally (Experiment 1) and visually (Experiment 2). The authors argued that sleep might have benefited consolidation of verbatim, item-specific details, therefore enhancing source-monitoring processes, which in turn results in decreased false memories. This interpretation is in line with the activation/monitoring theory of false memory formation (Roediger, Balota, & Watson, 2001; Roediger & McDermott, 1995, 2000), which argues that presentation of the studied words (e.g., nurse, hospital, etc.) triggers spreading activation of semantic networks, which in turn produces summative activation of the unstudied critical word (e.g., doctor). Later, during retrieval, participants fail to monitor the source of their memories and become susceptible to false memories. Fenn and colleagues argued that sleep increased source-monitoring processes, resulting in reduced false memories. However, this reduction of false memory post-sleep may be more easily obtained with recognition memory testing, as opposed to recall memory testing, because the presentation of a word to be recognized reactivates the sensory details of the study words and reinstates the encoding context, which aids source monitoring (Cabeza, Rao, Wagner, Mayer, & Schacter, 2001). Nevertheless, these results are not in agreement with those of Diekelmann and colleagues (2008) who also used recognition testing after a normal night of sleep, but found no effect on memory performance (compared to a wake group).

Payne and colleagues (Payne et al., 2009) who were the first to test the impact of sleep on free recall, as opposed to recognition memory in the DRM task, reported that sleep *increased* false recall of critical words, compared to an equivalent period of daytime wakefulness. Interestingly, a full night of sleep also benefited correct recall of studied words. These findings support the idea that sleep may generally benefit memory performance on a task that relies heavily on gist processing. According to the Fuzzy Trace Theory of false memory formation (Reyna & Brainerd, 1995; Reyna & Kiernan, 1994), memories are composed of both verbatim traces, which represent specific, contextual details of experience, and gist traces, which are 'fuzzy' representations of the general meaning of the information. In the DRM paradigm, it has been argued that studied words ('true memories') are dependent on both verbatim and gist traces, while critical words ('gist-based false memories') are dependent solely on gist because they were never actually presented and thus are not associated with contextual details (Brainerd & Reyna, 2002). Payne and colleagues (2009) argued that because memory for gist is utilized when recalling both true and false memories in the DRM task (Reyna & Brainerd, 1998), sleep may be advantageous to memory overall in such a gist-heavy, semantic-type task, resulting in increased true *and* false memory, at least when a free recall task is used to assess memory performance. Payne et al.'s results are also consistent with the fact that a majority of studies reporting beneficial effects of sleep on episodic memory have used recall to probe memory, whereas those using recognition are not as consistent (see for review Diekelmann, Wilhelm, & Born, 2009).

In addition, Payne et al. (2009) also found that an afternoon 90-min nap, compared to a 90-min wake period, resulted in an increase in false recall. To our knowledge, this is the only experiment to date that can argue for a lack of circadian and interference effects in the sleep and false memory literature, as participants in the nap and wake groups are trained and tested at the same time of day, and the nap, unlike a full night of sleep, provides only a very brief buffer against potentially interfering content that might be encountered during wakefulness.

In 2010, Diekelmann, Born, and Wagner repeated their original experiment comparing a full night of sleep, a full night of sleep deprivation, and a day spent awake. However, instead of using recognition to test DRM performance, they used free recall, as did Payne et al. (2009). Their results were consistent with those of Payne et al. (2009), as the sleep group resulted in increased false recall when compared to the wake group, although only for "low" performers. No differences were found in true recall. Diekelmann and colleagues defined low performers by applying a median split to the corrected true recall scores (i.e., studied words minus intrusions). Comparing true memory performance between the Payne et al. (2009) and Diekelmann, Born, and Wagner (2010) studies, it is clear that true memory performance in the Payne et al. study (true recall M = 29.75) was more similar to "low" (true recall M = 23.29) performers than "high" (true recall M = 47.05) performance of the Diekelmann et al. study. Another study by McKeon, Pace-Schott, and Spencer (2012), in which low recall rates (lower than Payne et al., 2009) were also reported, provides similar evidence to support this idea. The concurrence of results in these three studies strongly suggests that sleep increases false recall of critical words in the DRM paradigm, and is also in agreement with previous studies indicating a greater sleep benefit for weak, rather than strong, memory traces (Drosopoulos, Schulze, Fischer, & Born, 2007; Kuriyama, Stickgold, & Walker, 2004).

1.2. A possible role of slow-wave sleep (SWS) in false memory formation

Although humans go through the cyclic process of rapid eye movement (REM) and non-rapid eye movement (NREM) sleep, which includes deep, slow-wave sleep (SWS) and lighter Stages 1 and 2 throughout the night, the only sleep stage to be implicated thus far in the formation of gist-based false memories is SWS (e.g., Lo, Sim, & Chee, 2014; Payne et al., 2009), and this evidence is tenuous at best. For example, in the Payne et al. (2009) study, the authors reported a negative correlation between the recall of studied words and SWS (both SWS duration and SWS percentage). This is surprising given substantial research suggesting that SWS is beneficial for declarative memory consolidation (see Gais & Born, 2004). The authors suggested that, unlike other exclusively episodic (context-specific) memory tasks used in the sleep and memory literature (e.g., word pairs), performance on the DRM task relies largely on gist or semantic (context-independent) processing. This idea is supported by studies that have found a positive correlation between true and false memory performance (e.g., Kim & Cabeza, 2007; Payne et al., 2009). As stated above, according to fuzzy trace theory, both studied and critical words rely to some extent on gist-based, semantic processing (Reyna & Brainerd, 1998). Because episodic and semantic memory processing may rely on different neural systems (Tulving, 2002), one explanation for the negative correlation in the Payne et al. study is that, while SWS benefits memory for the contextually-linked details (verbatim traces) of episodic memories, it impairs memory for context-independent semantic/gist information that dominates a semantic memory task like the DRM (i.e., SWS impairs the ability to see the forest for the trees).

The idea suggested by Payne et al. (2009), that SWS might impair semantic information processing, demands explicit testing. If one argues that the strong, but not exclusive, semantic processing recruited while consolidating the studied DRM words (true memories) drives the negative correlation with SWS, then one would expect a significantly greater negative correlation between SWS and memory for the critical lures (which, as mentioned previously, are processed exclusively by semantic/gist-based processing). However, false recall did not correlate with SWS (or any sleep stage) in the Payne study, likely due to a power issue created by the fact that only eight total DRM wordlists were used, and thus there were only eight possible critical words available for recall. In the present study we doubled the number of studied lists from 8 to 16. By using more lists we increased the range of the critical-word recall and recognition performance, improving the variability and therefore the chance of reliably detecting a correlation between SWS and false memory.

Additionally, there is only a single nap experiment (experiment 3, Payne et al., 2009) to date that suggests an active role of sleep in false memory performance (i.e., sleep's benefits on memory go above and beyond protection from interference). In the classic sleep vs. wake 12-h design used by most authors to date, the wake group spends the 12-h delay filled with everyday activities that could function as interference to the memory task. Therefore, in the current study, we equated the time spent awake and asleep with delays of 24 and 48 h. These longer-delay groups allowed us to determine whether sleep has an active or merely (passive) protective effect on DRM performance by equating time spent awake and asleep in all conditions (only the timing of sleep varied) while simultaneously examining, for the first time, going beyond 12 h delays to determine longer-term effects of sleep on true and false memory performance in the DRM task.

Based on the findings discussed above (Diekelmann et al., 2010; McKeon et al., 2012; Payne et al., 2009) we hypothesized that sleeping shortly following encoding, when compared to having a full day of waking activity first (~16 h before going to sleep), would result in increased false and true memory. We expected that this increase should be observed regardless of retention delay between encoding and testing, because sleeping shortly after learning has been shown to benefit memory (see Payne, Chambers, & Kensinger, 2012; Payne, Tucker, et al., 2012). We measured both free recall and recognition memory in order to assess memory discriminability and bias separately (e.g., Hu, Stylos-Allan, & Walker, 2006), and to further explore the apparent differences in recall versus recognition performance in the DRM task across sleep-filled delays.

In line with Payne et al. (2009), we further predicted that both true and false memory performance would be negatively correlated with SWS. Furthermore, we hypothesized that the correlation between false memory and SWS would be stronger than the one between true memory and SWS. This expectation was developed under the idea that the negative correlation between true recall and SWS that Payne et al. (2009) observed was a result of the semantic nature of the DRM task. That is, this task relies heavily on semantic/gist traces even for the processing of studied words and even more so for processing critical lures. Because there are no episodic/contextual features for critical words (because they were never presented), memory for these (false) words is thought to depend completely on semantic/gist processing (Reyna & Brainerd, 1998). Thus we expect a stronger negative correlation between false memories and SWS than between true memories and SWS.

2. Method

2.1. Participants

One hundred and thirty-four college students (78 females; age 18–28 yrs) were recruited from the University of Notre Dame to participate in this study, and they were compensated with either course credit or cash payment. After providing informed consent, subjects were screened for self-reported sleep and mental health disorders, irregular sleep habits, and use of medications that affect the central nervous system. Participants were asked to refrain from alcohol and caffeine for 24 h before and throughout the experiment. Three participants failed to follow these instructions and therefore were not included in the analyses. Eight participants failed to return for the second (testing) session of the experiment, one participant reported previous participation in a different experiment using the DRM paradigm, and technical issues prevented data analysis of five participants. Thus, final data analysis was performed on 117 participants [69 females; mean age: 19.5 (range 18–28)].

2.2. Procedure

Participants listened to recordings of 16 DRM wordlists (Stadler, Roediger, & McDermott, 1999) through headphones after being instructed to pay close attention to the words because they would be tested on them in the next session. Sixteen word lists with the highest probability of false recall (Stadler et al., 1999) were chosen for this experiment. Each list consisted of 12 semantically associated words, and thus a total of 192 words were presented. The first eight lists range from a probability of false recall of 65–42%; the second eight lists range from a probability of 61–44%. Words were presented in an unfamiliar male voice, in descending strength of association, at a rate of one word every 2 s. At the end of each list there was a 12 s period of silence, a 1 s tone, 2 s of silence, and then the start of the next list. Participants heard the lists only one time each.

Fig. 1 shows the experimental procedure. Participants were pseudo-randomly assigned to complete the encoding phase of the task at 9:30 AM (WAKE 1st groups) and returned 24 h (W24 group; n = 30) or 48 h (W48 group; n = 30) later (9:30 AM) for the retrieval tests, or at 9:30 PM (SLEEP 1st groups) and returned 24 h (S24 group; n = 29) or 48 h (S48 group; n = 28) later (9:30 PM) for the retrieval tests. Participants were asked to avoid napping during the retention interval, which was confirmed by a daytime-activities questionnaire in the retrieval session.

Following the encoding session, the SLEEP 1st groups were wired for polysomnography (PSG) recording, put to bed at approximately 11:30 PM and given an 8.5 h sleep opportunity. The PSG montage included electroencephalography (EEG; F3, F4, C3, C4, Cz), electromyography (EMG), and electrooculography (EOG) with each electrode referenced to the contralateral mastoid. In the morning, the electrodes were removed and participants were reminded of whether they needed to come back that same night (24 h) or the following night (48 h), depending on their experimental condition. After completing the encoding phase in the morning, the WAKE 1st groups went about their everyday activities and returned for the retrieval tests 24 or 48 h later.

The retrieval session involved two separate tasks: a free recall task immediately followed by a recognition task. Because we were more interested in free recall memory (Diekelmann et al., 2010; McKeon et al., 2012; Payne et al., 2009), free recall testing always came first. A blank piece of paper was given to the participants and the experimenter asked them to write any words they remembered from the previously heard lists. Participants also wrote how confident they felt about each response using a 1–4 scale (1 being not sure at all, 4 being completely sure; Diekelmann et al., 2008). They had 10 min to complete this test and were notified when they had 2 min remaining. To more fully probe memory, participants then completed a recognition test in which words were presented visually on a computer. Each word was presented until the participant made an old or new judgment. This was followed by a self-paced confidence judgment that was made for each word using the same 1–4 scale as in the recall test. The recognition test included 96 words randomly presented in black font with a white background. It consisted of 48 study words taken from serial positions 1, 8, and 10 in the DRM studied lists (Fenn et al., 2009;

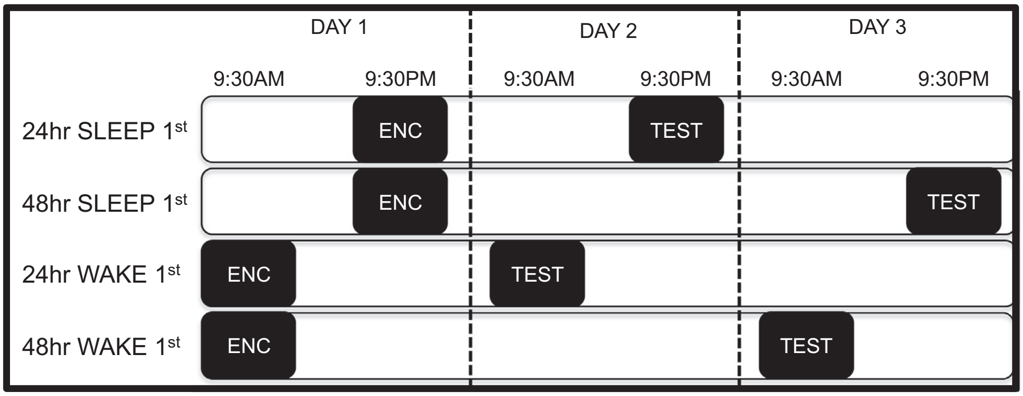

Fig. 1. Experimental protocol. SLEEP 1st groups encoded (ENC) the DRM word lists in the evening and were tested 24 (S24, n = 29) or 48 h (S48, n = 28) later. WAKE 1st groups encoded the DRM word lists in the morning and were tested 24 (W24, n = 30) or 48 h (W48, n = 30) later. The testing session (TEST) consisted of a free recall test followed by a recognition test.

Roediger & McDermott, 1995), the 16 critical lures, and 32 unrelated foils taken from eight other unpresented DRM lists.

2.3. Memory measures

For the free recall test, words were categorized as studied list words (previously heard words; true memory), critical words (semantically-related unpresented words; false memory), and intrusions (other unpresented words; intrusions). For the recognition test, true rate was defined as hit rate; "old" responses given to study words divided by the total number [48] of presented study words (e.g., nurse, hospital). False rate was defined as false alarm rate to critical lures (FAC); "old" responses given to related lures (e.g., doctor) divided by the total number [16] of related lures presented. Foil rate was defined as false alarm rate to foils (FAF); "old" responses given to unrelated foils divided by the total number [32] of unrelated foils presented. These raw memory measures were not used for analysis. Instead we used the non-parametric signal detection measure A' as an independent measure of discriminability [hit rate (H) and foil rate (FAF) for true recognition, and false rate (FAC) and foil rate (FAF) for false recognition], along with its corresponding independent measure of response bias B'', to take into account individual differences in the baseline propensity to accept both presented (hits) and non-presented (false alarms) items (Boice & Gardner, 1988; Donaldson, 1992; Snodgrass & Corwin, 1988). The formulas for true recognition are: $A' = 1/2 + [(H - FAF)(1 + H - FAF)]/[4H(1 - FAF)]$ when $H \geqslant FAF$ and $A' = 1/2 + [(FAF - H)(1 + FAF - H)]/[4FAF(1 - H)]$ when $FAF \geqslant H$. The formulas for false recognition are: $A' = 1/2 + [(FAC - FAF)(1 + FAC - FAF)]/[4FAC(1 - FAF)]$ when $FAC \geqslant FAF$ and $A' = 1/2 + [(FAF - FAC)(1 + FAF - FAC)]/[4FAF(1 - FAC)]$ when $FAF \geqslant FAC$. The formulas for response bias to true words (B'') are: $B'' = [H(1 - H) - FAF(1 - FAF)]/[H(1 - H) + FAF(1 - FAF)]$ when $H \geqslant FAF$ and $B'' = [FAF(1 - FAF) - H(1 - H)]/[FAF(1 - FAF) + H(1 - H)]$ when $FAF > H$. The formulas for response bias to false words (B'') are: $B'' = [FAC(1 - FAC) - FAF(1 - FAF)]/[FAC(1 - FAC) + FAF(1 - FAF)]$ when $FAC \geqslant FAF$ and $B'' = [FAF(1 - FAF) - FAC(1 - FAC)]/[FAF(1 - FAF) + FAC(1 - FAC)]$ when $FAF > FAC$. For B'' 0 = neutral bias, while negative values indicate a liberal bias and positive values indicate a conservative bias.

2.4. Statistical analyses

Mixed ANOVAs were used for most comparisons between groups. Further, the Greenhouse-Geisser correction was used when the sphericity assumption was violated in these mixed ANOVAs, as measured by Mauchly's Test. Also, appropriate change in degrees of freedom was performed when Levene's test for equality of variance was significant in independent *t*-tests (only used in true recall analysis).

3. Results

3.1. The impact of sleeping soon after learning on free recall memory

A 2 (order: SLEEP 1st, WAKE 1st) × 2 (time: 24 h, 48 h) × 3 (memory type: true, false, intrusions) mixed ANOVA, with memory type as a within-subjects factor and order and time as between-subjects factors, was performed to assess performance on free recall memory.

Main effects of order ($F(1,113) = 18.25$, $p < 0.001$, $\eta_p^2 = 0.14$) and memory type ($F(1,159) = 132.37$, $p < 001$, $\eta_p^2 = 0.54$) were found, while the main effect of time was not significant ($F(1,113) = 0.56$, $p = 0.46$, $\eta_p^2 < 0.01$). The memory type × time interaction was not significant ($F = 1.01$, $p = 0.34$, $\eta_p^2 < 0.01$). Because there were no significant main effects or interactions with time in this or any of the subsequent analyses, all main analyses were conducted on the combined S24 and S48 groups (SLEEP 1st) and the W24 and W48 groups (WAKE 1st); see Table 1 for a complete breakdown of memory performance in all groups across each of the memory types (true, false, intrusions). Importantly however, the memory type × order interaction was significant ($F(1,159) = 4.13$, $p = 0.03$, $\eta_p^2 = 0.04$). In order to examine our *a priori* prediction that the SLEEP 1st groups would have higher recall of both studied ('true') words and critical lure ('false') words than the WAKE 1st groups, independent *t*-tests were used to compare performance between the combined SLEEP 1st and WAKE 1st groups. As predicted, participants who slept first had significantly increased true memory, ($t(93) = 3.29$, $p = 0.001$, Cohen's $d = 0.61$), and significantly increased false memory ($t(115) = 2.62$, $p = 0.01$, Cohen's $d = 0.49$) compared to those who remained awake first (see Fig. 2).

3.2. The impact of sleeping soon after learning on recognition memory

Results from the recognition test (measured with A': Donaldson, 1992; Snodgrass & Corwin, 1988) support the idea that sleeping first benefits true memory and extend our free recall findings. A 2 (order) × 2 (time) × 2 (A' memory type: true, false) mixed ANOVA, with 'memory type' as a within-subject factor, was performed to investigate discriminability in the recognition test. Only a main effect of memory type ($F(1,113) = 11.86$, $p = 0.001$,

Table 1
Memory performance descriptive statistics. Means (standard deviations).

	S24	W24	S48	W48
TRUE recall[a]	22.58 (11.85)	15.37 (6.87)	19.50 (12.46)	14.37 (8.34)
FALSE recall[b]	4.27 (2.62)	3.27 (1.96)	4.11 (1.87)	3.07 (2.02)
INTRUSION recall	9.31 (7.54)	6.03 (5.24)	9.57 (8.63)	6.60 (5.09)
TRUE recognition[b]	0.63 (0.09)	0.61 (0.07)	0.64 (0.06)	0.59 (0.06)
FALSE recognition	0.65 (0.12)	0.64 (0.07)	0.66 (0.08)	0.63 (0.08)
TRUE response bias[a]	0.00 (0.08)	0.04 (0.09)	−0.02 (0.07)	0.01 (0.06)
FALSE response bias	−0.05 (0.12)	0.00 (0.12)	−0.07 (0.09)	−0.02 (0.10)

[a] p < 0.05 between S24 and W24.
[b] p < 0.05 between S48 and W48.

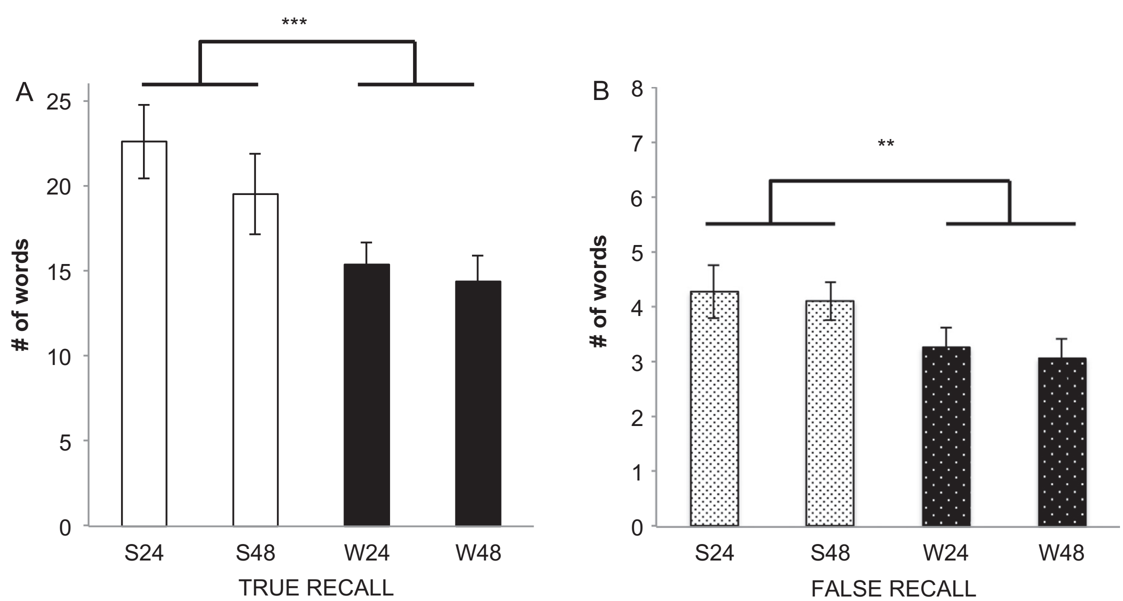

Fig. 2. Number of words recalled. (A) Recall of studied words (TRUE RECALL). (B) Recall of critical lures (FALSE RECALL). Asterisks represent p-values from individual *t*-tests on the SLEEP 1st groups vs. WAKE 1st groups, collapsed across time. Error bars denote SEM. ** p = 0.01. *** p = 0.001.

$\eta_p^2 = 0.10$) and a main effect of order ($F(1,113) = 4.48$, $p = 0.04$, $\eta_p^2 = 0.04$) were found, showing that false words were better recognized than true words. Although no interactions were significant (all F's < 1.17 all p's > 0.28), *a priori* comparisons on the collapsed groups showed that sleeping first resulted in significantly higher true ($t(115) = 2.73$, $p = 0.007$, Cohen's $d = 0.51$), but not false recognition ($t(115) = 1.26$, $p = 0.21$, Cohen's $d = 0.23$) (see Fig. 3A and B).

A 2 (order) × 2 (time) × 2 (B'' memory type: true, false) mixed ANOVA, with 'memory type' as a within-subject factor, was performed to investigate memory bias in the recognition test. A main effect of memory type was found ($F(1,113) = 47.84$, $p < 0.001$, $\eta_p^2 = 0.23$) showing that participants were more liberal in their false memory judgments than in their true memory judgments. There was also a main effect of order ($F(1,113) = 6.93$, $p = 0.01$, $\eta_p^2 = 0.06$), showing that sleeping soon after learning led to more liberal responding in the recognition test (see Fig. 3C and D). Follow-up independent *t*-tests on the collapsed groups showed that sleeping first resulted in significantly lower bias for both true ($t(115) = -2.74$, $p = 0.007$, Cohen's $d = 0.51$) and false recognition ($t(115) = -2.31$, $p = 0.02$, Cohen's $d = 0.43$). No significant interactions were found (all F's < 0.41, all p's > 0.52). Whereas sleeping soon after learning increased memory discriminability (A') for only studied (true) words, it decreased bias (B'') for both studied and non-studied, critical (false) words.

3.3. The impact of sleeping soon after learning is dependent on memory performance

False memory performance has been shown to depend on general memory performance (Diekelmann et al., 2010). Because Diekelmann et al. (2010) found that sleep increased false memory only in low performers, we divided participants into low and high performers using a median split for true recall, resulting in four groups: high SLEEP 1st, low SLEEP 1st, high WAKE 1st, and low WAKE 1st performers. One-way ANOVAs revealed increased false recall in the SLEEP 1st groups compared to the WAKE 1st groups ($F = 10.79$, $p < 0.001$, $\eta^2 = 0.22$), while no significant differences were found for intrusions ($F = 2.17$, $p = 0.096$, $\eta^2 = 0.05$). Consistent with Diekelmann et al. (2010), planned follow-up comparisons showed that sleeping first increased false recall only in low performers ($t(60) = 3.67$, $p = 0.001$), but not in high performers ($t(53) = 1.31$, $p = 0.20$; Fig. 4).

The fact that we observed increased false memory after sleep in low performers not only supports previous findings (Diekelmann et al., 2010), but also suggests that such participants rely more heavily on non-contextual, gist-based processing to retrieve the DRM words (i.e., both true and false). If semantic/gist processing is recruited for overall memory performance in this task, one might expect true and false memory performance to be positively correlated in this study, as it has been in several others (e.g., Kim & Cabeza, 2007; Payne et al., 2009). Indeed, true recall was positively

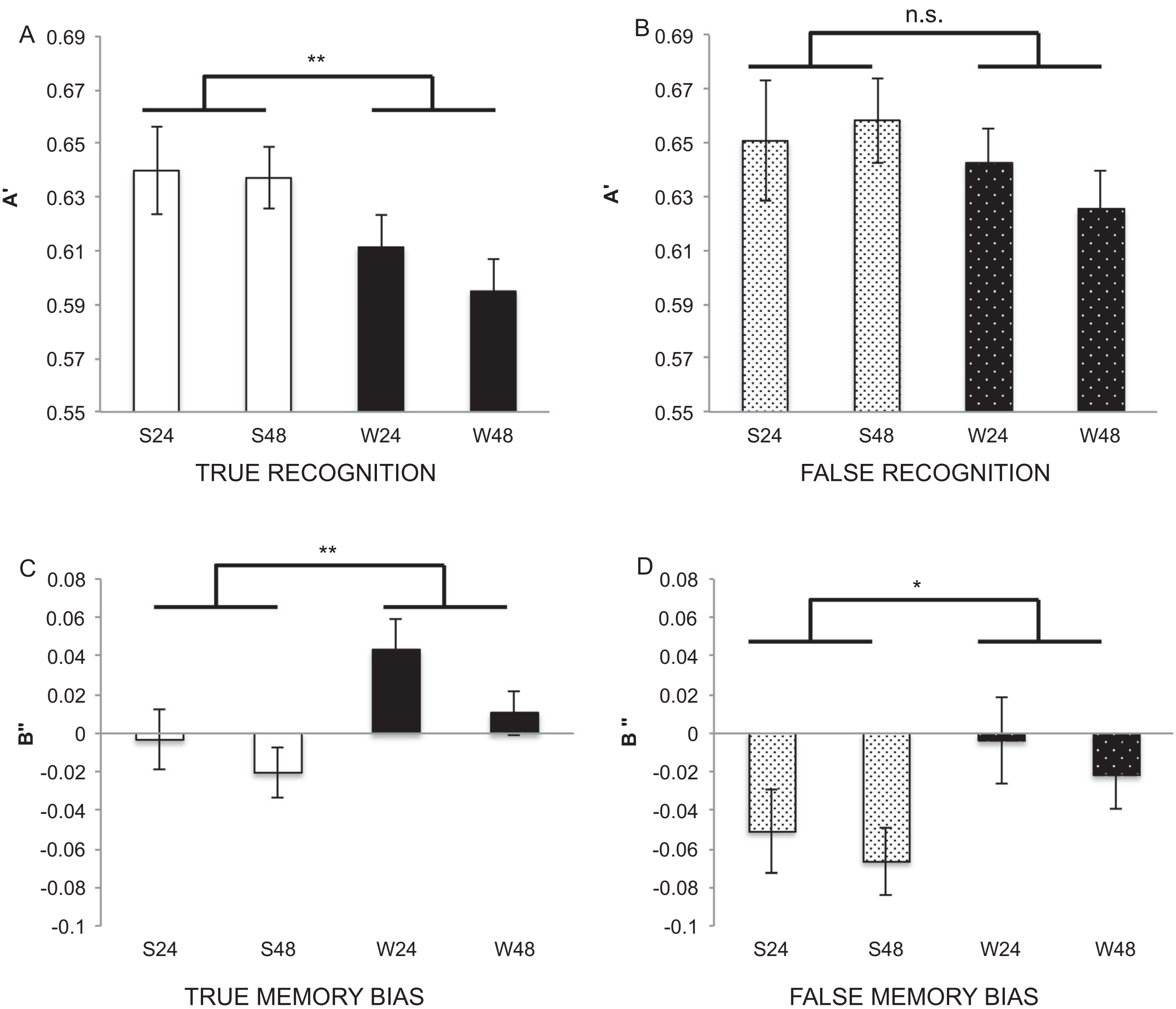

Fig. 3. Recognition memory, measured by A', for (A) studied words (TRUE) and (B) non-studied critical lures (FALSE). Memory bias in the recognition test, measured by B″, for (C) TRUE and (D) FALSE words. Asterisks represent p-values from individual *t*-tests on the SLEEP 1st groups vs. WAKE 1st groups, collapsed across time. Error bars are SEM. * p = 0.02. ** p = 0.007.

correlated with false recall in the current study, but importantly, this was only the case in low performers (r = 0.38, p = 0.035 for low, r = −0.02, p = 0.94 for high). Thus, sleeping soon after learning might increase false memory, but only in those participants who encoded the material poorly to begin with.

3.4. The relationship between slow-wave sleep (SWS) and false memory

Table 2 shows the descriptive statistics for PSG measures for the SLEEP 1st groups. Of note, the S24 group obtained more minutes in SWS and SWS% than the S48 group [t(54) = 2.73, p = 0.008 for SWS min, t(54) = 2.57, p = 0.01]. This is considered further in the Discussion (see Section 4.2). Payne and colleagues (2009) argued that SWS, if detrimental for semantic/gist processing, should be negatively correlated to false memory performance. In their study, no such correlation was found due to lack of power, whereas in the current study, false recognition memory was indeed negatively correlated with SWS%, r = −0.33, p = 0.01, and SWS duration, r = −0.33, p = 0.01, although true recognition memory was not correlated with either measure of SWS (SWS% r = −0.11, p = 0.44; SWS duration r = −0.07, p = 0.60). Moreover, measures of true and false recall memory were not significantly correlated with SWS duration or SWS% (see Table 3).

We next explored whether the relationship between recognition memory and SWS was modulated by performance level. SLEEP 1st participants were divided into low and high performers using a median split for true A'. False memory was negatively correlated with SWS, but only in low performers (r = −0.54, p = 0.003 and r = −0.49, p = 0.008 for SWS duration and percentage of total sleep time, respectively; see Fig. 5A solid black line). Although the correlations for high performers were also in the predicted negative direction, they did not reach significance (r = −0.29, p = 0.13 and r = −0.30, p = 0.13 for SWS duration and percent, respectively; Fig. 5A dotted gray line). Interestingly, a similar negative correlation between SWS and true recognition memory was apparent in low, but not in high, performers (r = −0.47, p = 0.012 and r = −0.46, p = 0.013 for SWS duration and percent, respectively; see Fig. 5B). This is consistent with previous findings using free recall measures for studied words (Payne et al., 2009).

It is possible that SWS is indirectly detrimental to the formation of gist-based false memories in the DRM task by supporting the

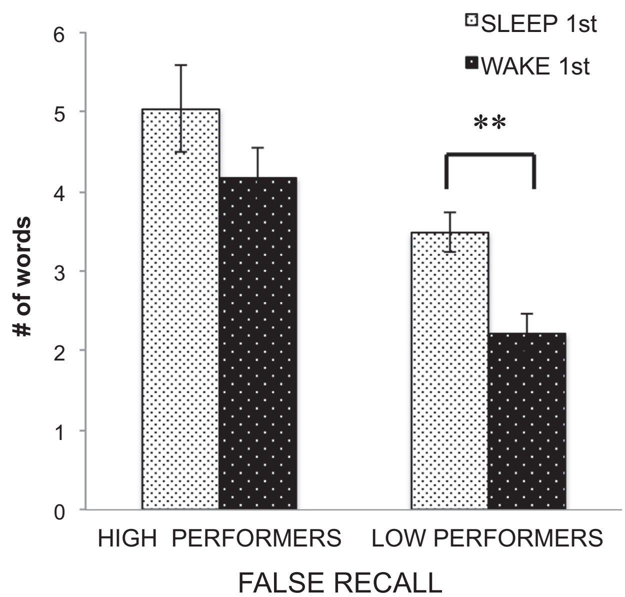

Fig. 4. FALSE RECALL in high and low performers. Asterisks represent p-values from individual *t*-tests on the SLEEP 1st groups vs. WAKE 1st groups, collapsed across time. Error bars denote SEM. $^{***}p = 0.001$.

Table 2
PSG descriptive statistics. TRT: total recorded time, TST: total sleep time, SOL: sleep onset latency, WASO: wake after sleep onset, S1: stage 1, S2: stage 2, SWS: slow-wave sleep, REM: rapid eye movement, Wake: minutes spent awake, Movement: Movement in minutes during REM sleep. Percentage values are related to total sleep time. Parentheses denote SD.

	S24	S48	AVG
TRT	510.85 (15.45)	510.92 (5.69)	510.89 (13.53)
TST	467.64 (29.21)	451.48 (36.85)	459.56 (33.95)
SOL	9.68 (10.43)	13.77 (18.59)	11.72 (15.08)
REM Latency	105.41 (44.26)	97.77 (33.01)	101.59 (38.92)
WASO (min)	28.53 (15.53)	37.75 (30.59)	33.14 (24.49)
% Efficiency	91.15 (4.17)	88.48 (6.95)	89.82 (5.84)
S1 (min)	45.58 (16.58)	51.12 (30.38)	46.86 (24.63)
S1%	9.2 (3.69)	11.58 (7.25)	10.39 (5.83)
S2 (min)	236.73 (30.23)	229.94 (34.20)	233.34 (32.17)
S2%	50.97 (5.80)	50.81 (5.98)	50.89 (5.84)
SWS (min)	93.08 (26.86)	75.64 (20.44)	84.36 (25.24)
SWS%	19.88 (5.17)	16.68 (4.08)	18.29 (4.89)
REM (min)	92.67 (19.58)	95.41 (27.38)	94.05 (23.63)
REM%	19.32 (4.13)	29.90 (5.02)	20.42 (4.58)
Wake (min)	38.21 (20.27)	51.51 (37.24)	44.87 (30.46)
Movement (min)	4.92 (3.42)	6.19 (3.75)	5.56 (3.61)

Table 3
Pearson correlations for SWS measures (SLEEP 1st groups combined) and memory measures.

	Pearson's r	p-value
SWS%		
True recall	0.10	0.46
False recall	0.08	0.58
True recognition	−0.11	0.44
False recognition	−0.33	0.01
SWS minutes		
True recall	0.13	0.34
False recall	−0.01	0.92
True recognition	−0.07	0.60
False recognition	−0.33	0.01

consolidation of contextual and sensory (i.e., episodic) details of the DRM study words; that is, by increasing the participants' ability to discriminate between a presented and a non-presented, critical (theme) word. Some have argued that true memory retrieval in the DRM task recruits both verbatim and gist memory traces (Brainerd & Reyna, 2002). If so, using the classic hit vs. foil comparison may make it difficult to identify whether an experimental manipulation influences true memory by affecting verbatim, gist, or both types of memory traces. Therefore, in the following exploratory analysis, we examined the impact of sleep's positioning using a more detail-based measure of recognition memory, *verbatim* recognition, which provides a measure of how well participants discriminate between presented items versus semantically related material that was never presented. Conceptually, this measure assesses contextual and/or sensory features of the studied words (e.g., memory of the sound of a speaker's voice during the presentation of a specific word). For this *verbatim* recognition measure, we used H = hit rate and FAC = false rate (e.g., Arndt & Reder, 2003; Curran et al., 2001) for the A' formula (see Memory Measures in Method section).

Interestingly, verbatim recognition was *positively* correlated with SWS% and SWS duration. Once again, however, this correlation was only observed in low performers ($r = 0.37$, $p = 0.05$ for SWS%; $r = 0.42$, $p = 0.03$ for SWS duration), but not in high performers ($r = 0.20$, $p = 0.29$ for SWS%; $r = 0.19$, $p = 0.34$ for SWS duration). Together, these findings suggest that SWS is detrimental for gist-based memory consolidation, directly, by affecting semantic processing and, indirectly, by supporting contextual/detail memory, particularly in low performers.

4. Discussion

The current study investigated how sleeping shortly after learning, compared to after a day filled with waking activity, affects the development of true and false memories across longer delays than those examined in previous experiments (e.g., Fenn et al., 2009; Payne et al., 2009). Our study extends previous findings (e.g., Payne et al., 2009) that investigated the development of true and false memories across a 12 h delay, filled with either a full night of sleep or a full day of wakefulness. It could be claimed that the enhancing effect of nocturnal sleep reported by previous studies (McKeon et al., 2012; Payne et al., 2009) was due to a lack of outside interference during the period of sleep, instead of sleep-dependent processes (Wixted, 2004). Here, however, all participants in the 24 h and 48 h conditions had equivalent amounts of sleep and waking interference between encoding and retrieval. Yet sleeping soon after learning increased true memories and to a lesser extent, gist-based false memories. Sleeping soon after learning also resulted in a bias shift; participants were more liberal in their recognition memory (i.e., more prone to endorse items), for both true and false words, if they slept first. Furthermore, slow-wave sleep (SWS duration and percent) was negatively related to false memory, indicating a possible detrimental role of SWS on gist/semantic processing. Importantly, false memory performance, and its relationship with SWS, might depend on true memory performance (e.g., Diekelmann et al., 2010). Only low performers recalled more false words after sleeping first, and only in these performers did we observe a negative correlation between SWS and false memory, and between SWS and true memory (e.g., Payne et al., 2009).

Sleep may increase both true and false memory performance in the DRM task, as was the case in previous studies (McKeon et al., 2012; Payne et al., 2009). However, while the strongest effect in the Payne et al. study was the selective protection of gist-based false memories, this does not seem to be the case in the current study. Sleeping first, compared to being awake, resulted in a significant increase of both true and false *recall*, but a significant increase

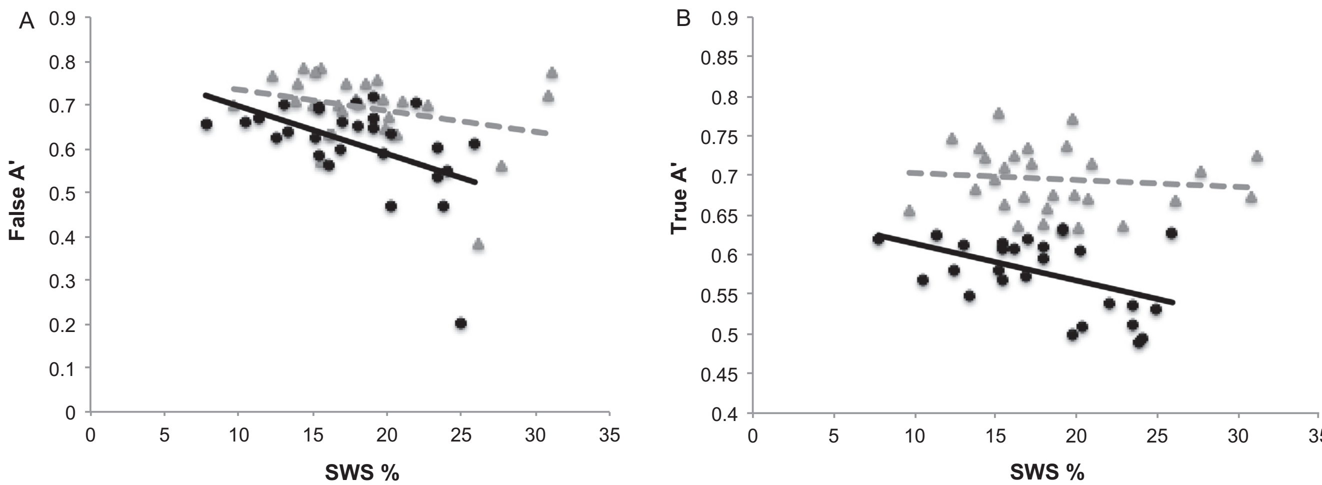

Fig. 5. (A) The negative correlation between SWS and FALSE recognition memory is only observed in low performers (solid black line; $r = -0.49$, $p = 0.008$). (B) A negative correlation between SWS and TRUE recognition memory is apparent only in low performers (solid black line; $r = -0.46$, $p = 0.013$). Black circles represent low performers, Gray triangles represent high performers.

in true *recognition* only. A possible explanation for the lack of a significant difference in false recognition might arise from lack of statistical power with the DRM critical words. Although we attempted to increase power in the current study by using 16 DRM word lists (as opposed to 8 in the Payne et al., 2009 study), which resulted in 16 possible critical words to falsely recall, this nevertheless may have been heavily countered by the use of longer delay intervals of 24 and 48 h (as opposed to 12 h). An inspection of the recall data from the Payne et al. (2009) study supports this idea. The average false recall in the 12 h-delay sleep groups in Payne et al. is approximately 4 critical words (3.6 in exp. 1, and 4.0 in exp. 2), which is quite similar to the average false recall in the current study (4.2 for S24 and 4.1 for S48). Another factor, along with a lack of power, could also explain the nonsignificant difference in false recognition performance: the use of a recognition task to test memory is thought to aid source monitoring. Cabeza and colleagues (2001) had participants complete a recognition test in an fMRI scanner and found that while hippocampal activity was similar for true and false words, the parahippocampal gyrus was more activated by true than false words. This differentiation in the parahippocampal gyrus suggests reactivation of associated perceptual and contextual details of the studied (true) words, likely supporting source monitoring, which has been robustly shown to decrease false memory (e.g., Neuschatz, Benoit, & Payne, 2003). Therefore, it is possible that a recognition test made it more difficult to observe significant effects in false recognition performance in the current study. Nonetheless, we observed significant differences in false recall performance, which, interestingly, were modulated by true memory performance, as in Diekelmann et al. (2010).

The current study seems to support the idea that weak traces are the ones targeted by sleep, i.e., only participants that weakly encoded the material received a benefit from sleep (Drosopoulos et al., 2007; Kuriyama et al., 2004; but see Tucker & Fishbein, 2008 for opposite findings). For example, Drosopoulos and colleagues showed that individuals who learned word-pair lists to a lower criterion received a greater sleep-dependent benefit than those who learned the lists to a higher criterion. Similarly, in the current study, only low performers had increased false recall after sleep. It is possible that our participants, particularly low performers, relied more heavily on gist/semantic processing, due to a sleep-dependent benefit in gist extraction. This is supported by the fact that only in low performers did we observe a positive correlation between true and false memory performance, consistent with several previous studies (e.g., Kim & Cabeza, 2007; Payne et al., 2009), suggesting shared mechanisms for true and false memory formation.

We had predicted a negative correlation between false memory performance and SWS. Similar to what Payne and colleagues (2009) found, we also predicted a negative correlation between true memories and SWS. Most likely because their study was statistically underpowered, Payne and colleagues did not find a correlation between SWS and false memories. With only eight DRM word lists, there was not enough variability to detect a significant correlation. In the current study, by doubling the number of DRM word lists to 16, we did observe the SWS-false memory negative correlation that Payne et al. predicted. It should also be noted that another recent study found a similar correlation in an older population (Lo et al., 2014). Importantly, we found that SWS was negatively correlated to false and true recognition performance, but in low performers only.

At first, a negative correlation between SWS and true memories seems to be in conflict with a mass of evidence suggesting a *beneficial* effect of SWS on episodic memory (e.g., Alger, Lau, & Fishbein, 2012; Plihal & Born, 1997; see Rasch & Born, 2013 for review). However, as Payne and colleagues (2009) argued, because of the highly semantic nature of the DRM paradigm, and the role that gist-processing likely plays in the retrieval of both false (critical) and true (studied) words, this task is not as context-dependent (and thus hippocampus-dependent, Kim & Cabeza, 2007) as other episodic tasks used in the sleep and memory literature (e.g., paired-associates task, card locations in the memory game 'concentration'). In turn, the DRM paradigm relies heavily on context-independent gist-based processing for memory formation of both the semantically related studied (true) words in the DRM lists and the unpresented, yet highly related (false) critical lures (Brainerd & Reyna, 2002). Our data provide further support for the idea that semantic/gist processing may be negatively affected by SWS, particularly in low performers.

Another possible explanation for the detrimental role of SWS on false memory formation may be that SWS strengthens the available (albeit few) contextual details associated with the DRM words that were actually studied (i.e., memory for the sound of the voice speaking a given studied word, memory for its serial position in a given DRM list). In this way, SWS may be indirectly detrimental to false memory formation by benefiting the sensory/contextual traces during consolidation, possibly by reactivating and reinforcing the individual traces of the studied words (and their features). This idea is supported by the fact that, in low performers only, we found a positive correlation between SWS and our measure of *verbatim* recognition, i.e., a measure that provides a stricter estimate of participants' ability to discriminate between item-specific, detail-oriented memories for study words (e.g., Curran et al., 2001) and highly similar critical lure words. Detailed, contextually specific memories may be reactivated in the hippocampus during SWS, which would in turn lead to facilitated consolidation of these item-specific features, and perhaps also impede the consolidation (and/or retrieval) of gist traces. Thus, our findings suggest that low performers may rely more on gist/semantic processing after sleep overall, and that SWS may be detrimental for gist extraction in the DRM task by hindering gist traces and reinforcing contextual/verbatim details of the study words.

We had predicted that the negative correlation between false memory and SWS would be stronger than the one between true memory and SWS. This hypothesis was developed under the idea that the previously reported (Payne et al., 2009) true recall-SWS negative correlation was mainly due to the semantic nature of the entire DRM task (i.e., semantic gist plays a role in both true and false memory formation in the DRM task). False memory, as opposed to true memory, in this task relies *completely* on semantic processing and therefore, a false memory-SWS negative correlation should be significantly stronger than a true memory-SWS correlation. In the current study the correlation between false recognition and SWS% ($r = -0.49$) was indeed stronger, albeit not significantly so, than the correlation between true recognition and SWS% ($r = -0.46$ in low performers; see solid lines in Fig. 5). A plausible explanation for this is that, although the current study was successful at increasing the statistical power to capture this correlation (by doubling the encoded DRM lists from 8 to 16), some power might have been lost by increasing the time delay between encoding and testing (from 12 h to 24/48 h), and even more word lists may have been necessary to detect the predicted finding.

4.1. Sleeping soon after learning results in a shift towards liberal bias

Most sleep and declarative memory studies that use recognition to probe memory do not separate discriminability from bias. Merely a handful of studies have measured memory in terms of both discrimination and bias, only to find that sleep can increase (Hu et al., 2006; Sterpenich et al., 2007), decrease (Atienza & Cantero, 2008) or have no effect on bias (Baran, Pace-Schott, Ericson, & Spencer, 2012; Sterpenich et al., 2009; Wagner, Kashyap, Diekelmann, & Born, 2007). For example, Hu et al. (2006) found that sleep increased recognition bias (i.e., shifted participants to a conservative bias) of both neutral and emotional (negative) stimuli. The authors argue that, although there is no neural mechanism formally stated for bias, the high levels of acetylcholine (ACh) during REM sleep might drive this effect. It is also of note that diseases (e.g., Alzheimer's disease) that exhibit low levels of cortical ACh are associated with abnormally liberal (low) recognition bias (Fuld, Katzman, Davies, & Terry, 1982), which suggests that the process of memory bias can be neurochemically mediated (Hu et al., 2006). During REM sleep, ACh concentrations in the hippocampus reach levels similar to those during waking (Hasselmo, 1999). Hu et al. (2006) suggested that this ACh elevation during sleep is responsible for the conservative bias shift. It is important to note that in the Hu et al. study, sleep increased bias more so for emotional than neutral stimuli, another reason to argue the importance of REM sleep in this process, since this stage is known to be selectively important for emotional memory consolidation (see Diekelmann & Born, 2010). Nonetheless, the authors did not report a correlation (or lack of) between REM sleep and their bias measure.

If high levels of ACh during REM sleep shift bias to a more conservative behavior, especially for arousing stimuli, then we might speculate that the low(est) levels of ACh during SWS may shift bias to the liberal side of the spectrum, particularly in the case where no arousing stimuli are present as in the current study. Nonetheless, in the present study no correlations between SWS and bias measures were found (nor between SWS and foil rate, or total recall of words). More importantly, the negative correlation between SWS and false recognition found here suggests that a more liberal bias would result in better memory performance for critical words. However, response bias and discriminability are independent from each other, therefore, it is possible to affect bias without affecting discriminability (Green & Swets, 1966), which is what our findings report; sleep did not increase false recognition, but it did decrease false memory bias.

4.2. Limitations and future directions

The current study is not without limitations. This study was a follow-up to Payne et al. (2009), where only recall memory was tested. Our recall test was always followed by a recognition test, which has been shown to increase recognition rates (e.g., Roediger & McDermott, 1995). Thus, recognition rates may be inflated in the current study and future experiments should counterbalance the order of the retrieval tests, or exclusively test recognition memory to clarify the recognition memory findings. This might also explain why we did not detect significant differences in the recognition test; previous free recall could have increased recognition performance to (functional) ceiling levels, resulting in nonsignificant differences between the groups.

In the current study, we observed some differences in memory across the delay intervals (see Table 1). That is, increased true recall was significantly increased only in the S24 group whereas false recall was significantly increased only in the S48 group (although note that all effects were in the predicted direction). Non-sleep studies (e.g., Seamon et al., 2002) have found that memory for studied words decreases at a faster rate than memory for critical words. This could explain why our 48 h between-group comparison (S48 vs. W48) for true recall, albeit in the predicted direction, did not reach significance. Additionally, even with our large sample size, the effect sizes of the presented recall results can be considered medium-sized, particularly for false memory (Cohen's $d = 0.49$). This may help explain the dissociation for false memory in the current study, and in the field at large.

Additionally, the SLEEP 1st groups had significantly different amounts of SWS during the first post-encoding night. This difference, albeit coincidental and unrelated to experimental manipulation, may help explain the differences presented in Table 1. It is possible that the decreased amount of SWS (which we suggest is detrimental for gist processing) obtained by the S48 group in that first post-encoding night might have resulted in increased false memory two days after during retrieval. Nonetheless, this is unlikely because in both groups, the non-significant patterns were again in the predicted direction. That is, true recall was higher in the S48 group, compared to the W48 group, and false recall was higher in the S24 group, compared to the W24 group (see Table 1 and Fig. 2). It is also possible that this difference in SWS may, in part, explain the lack of significant interactions with delay intervals (i.e., our 'time' variable). However, a more likely explanation for this lack of interaction is simply that our difference of 24 h between the groups was not long enough to detect significant effects of delay.

Furthermore, due to time and resource constraints, only the first night was recorded with PSG in this study, and only the SLEEP 1st conditions. To better understand the role of sleep in true and false memory performance across longer delays, all nights should be recorded with PSG to address long-term changes in the relationship between sleep and true and false memories. Nonetheless, we argue that our study targets the most important time period (i.e., the first night) post-encoding, during which sleep-dependent consolidation processing has been found to have its greatest impact (e.g., Fischer, Hallschmid, Elsner, & Born, 2002; Gais, Lucas, & Born, 2006; Payne, Chambers, et al., 2012; Payne, Tucker, et al., 2012; Stickgold, James, & Hobson, 2000).

Diekelmann et al. (2010) argued that a possible explanation for their different patterns depending on memory performance was that high and low performers had inherently different mnemonic strategies (e.g., high performers could have encoded words more literally, which might prevent formation of the gist of the DRM wordlists, therefore decreasing false recall). Another possibility is that low performers were paying less attention to the words at encoding. Additional goals of future research should be to measure attention and to assess mnemonic strategies used throughout the study to account for the variables as possible explanations for differences in high and low performers.

In the current study, participants had similar absolute levels of interference during the retention interval (e.g., both 24 h-delay groups had ~16 h of interference before retrieval, only the timing of sleep was different). Although this can be considered as evidence supporting an active role of sleep in memory consolidation, it may alternatively support an "opportunistic" view of sleep (Mednick, Cai, Shuman, Anagnostaras, & Wixted, 2011), such that sleep is useful for memory consolidation by providing protection from interference during a critical time window that is close to the encoding event.

5. Conclusion

The objective of this study was to investigate the role of sleep in the formation and long-term consolidation of true and false memories in the DRM paradigm. Sleeping soon after learning resulted in enhanced true memory and, to a lesser extent, false memory. Further, SWS was detrimental to false memory and, in low performers, to true memory formation. Due to the nature of the task (i.e., lists of semantically related words), in the DRM, gist/semantic processing is arguably more recruited than sensory/contextual processing. The current findings suggest that SWS negatively affects semantic/gist processing, particularly in low performers. Most sleep and memory research focuses in retrieval after a 12 h period of time filled with sleep or wake activity. Some argue that sleep passively shields memory processing from interference (Wixted, 2004). In the current study, all participants slept, but only those that slept soon after learning had higher memory for both studied and non-studied, critical words, supporting an active role of sleep in the formation of true and false memories.

Acknowledgements

The authors thank everyone in the Sleep, Stress, and Memory Lab, particularly Tony J. Cunningham and Sara E. Alger for their helpful feedback on the manuscript. This work was supported by the National Science Foundation Graduate Research Fellowship Program (DGE-1313583).

References

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

Arndt, J., & Reder, L. M. (2003). The effect of distinctive visual information on false recognition. *Journal of Memory and Language, 48*(1), 1–15.

Atienza, M., & Cantero, J. L. (2008). Modulatory effects of emotion and sleep on recollection and familiarity. *Journal of Sleep Research, 17*(3), 285–294.

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

Bartlett, M. (1932). Remembering: A case in experimental and social psychology.

Boice, R., & Gardner, R. M. (1988). A computer program to generate parametric and nonparametric signal-detection parameters. *Bulletin of the Psychonomic Society, 26*(4), 365–367.

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

Brewer, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition, 5*(6), 673–678.

Cabeza, R., Rao, S. M., Wagner, A. D., Mayer, A. R., & Schacter, D. L. (2001). Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proceedings of the National Academy of Sciences, 98*(8), 4805–4810.

Curran, T., Schacter, D. L., Johnson, M. K., & Spinks, R. (2001). Brain potentials reflect behavioral differences in true and false recognition. *Journal of Cognitive Neuroscience, 13*(2), 201–216.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*(1), 17.

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

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

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

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

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General, 121*(3), 275.

Drosopoulos, S., Schulze, C., Fischer, S., & Born, J. (2007). Sleep's function in the spontaneous recovery and consolidation of memories. *Journal of Experimental Psychology: General, 136*(2), 169.

Ebbinghaus, H. (1885). *Memory*. New York: Columbia University.

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

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

Fuld, P. A., Katzman, R., Davies, P., & Terry, R. D. (1982). Intrusions as a sign of Alzheimer dementia chemical and pathological verification. *Annals of Neurology, 11*(2), 155–159.

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

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

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. John Wiley and Sons.

Hasselmo, M. E. (1999). Neuromodulation: Acetylcholine and memory consolidation. *Trends in Cognitive Sciences, 3*(9), 351–359.

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

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

Kim, H., & Cabeza, R. (2007). Differential contributions of prefrontal, medial temporal, and sensory-perceptual regions to true and false memory formation. *Cerebral Cortex, 17*(9), 2143–2150.

Kuriyama, K., Stickgold, R., & Walker, M. P. (2004). Sleep-dependent learning and motor-skill complexity. *Learning & Memory, 11*(6), 705–713.

Lo, J. C., Sim, S. K., & Chee, M. W. (2014). Sleep reduces false memory in healthy older adults. *Sleep, 37*(4), U247–U665.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*(4), 361–366.

McKeon, S., Pace-Schott, E. F., & Spencer, R. M. (2012). Interaction of sleep and emotional content on the production of false memories. *PLoS ONE, 7*(11), e49353.

Mednick, S. C., Cai, D. J., Shuman, T., Anagnostaras, S., & Wixted, J. T. (2011). An opportunistic theory of cellular and systems consolidation. *Trends in Neurosciences, 34*(10), 504–514.

Neuschatz, J. S., Benoit, G. E., & Payne, D. G. (2003). Effective warnings in the Deese-Roediger-McDermott false-memory paradigm: The role of identifiability. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(1), 35.

Payne, J. D. (2011). Sleep on it!: Stabilizing and transforming memories during sleep. *Nature Neuroscience, 14*(3), 272.

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

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

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

Payne, J. D., Tucker, M. A., Ellenbogen, J. M., Wamsley, E. J., Walker, M. P., Schacter, D. L., & Stickgold, R. (2012b). Memory for semantically related and unrelated declarative information: The benefit of sleep, the cost of wake. *PLoS ONE, 7*(3), e33079.

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

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

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

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

Reyna, V. F., & Brainerd, C. J. (1998). Fuzzy-trace theory and false memory: New frontiers. *Journal of Experimental Child Psychology, 71*(2), 194–209.

Reyna, V. F., & Kiernan, B. (1994). Development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology, 30*(2), 178.

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

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

Roediger, H. L., & McDermott, K. B. (2000). Tricks of memory. *Current Directions in Psychological Science, 9*(4), 123–127.

Schacter, D. L. (1995). Memory distortion: History and current status. *Memory distortion: How minds, brains, and societies reconstruct the past*, 1–43.

Seamon, J. G., Luo, C. R., Kopecky, J. J., Price, C. A., Rothschild, L., Fung, N. S., & Schwartz, M. A. (2002). Are false memories more difficult to forget than accurate memories? The effect of retention interval on recall and recognition. *Memory & Cognition, 30*(7), 1054–1064.

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

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

Sterpenich, V., Albouy, G., Boly, M., Vandewalle, G., Darsaud, A., Balteau, E., ... Rauchs, G. (2007). Sleep-related hippocampo-cortical interplay during emotional memory recollection. *PLoS Biology, 5*(11), e282.

Sterpenich, V., Albouy, G., Darsaud, A., Schmidt, C., Vandewalle, G., Vu, T. T. D., ... Collette, F. (2009). Sleep promotes the neural reorganization of remote emotional memory. *The Journal of Neuroscience, 29*(16), 5143–5152.

Stickgold, R., James, L., & Hobson, J. A. (2000). Visual discrimination learning requires sleep after training. *Nature Neuroscience, 3*(12), 1237–1238.

Tucker, M. A., & Fishbein, W. (2008). Enhancement of declarative memory performance following a daytime nap is contingent on strength of initial task acquisition. *Sleep, 31*(2), 197.

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

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

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