In recent years, compelling evidence has been found to support a role of sleep in the consolidation of non-declarative memory (Stickgold et al. 2000; Walker et al. 2002; Fenn et al. 2003; Wagner et al. 2004). A coherent body of theoretical (Crick and Mitchison 1983; Alvarez and Squire 1994; McClelland et al. 1995; Meeter and Murre 2005) and physiological (Wilson and McNaughton 1994; Buzsaki 1996; Izquierdo et al. 1997) evidence suggests that sleep might also have an important role in consolidation of hippocampus-dependent episodic memories. According to these studies, sleep may be used to recode these memories from a highly plastic store, the hippocampus, to a more stable one with larger capacity, involving the neocortex. Such recoding would protect memories from fast degradation through interference and passive forms of decay.

While a significant body of work indirectly supports the above notion (e.g., Peigneux et al. 2004; Schmidt et al. 2006; Takashima et al. 2006; for review, see Stickgold 2005), direct evidence for a causal role of sleep in episodic memory consolidation is scarce. Indeed, many studies that tapped directly into the effect of sleep are confounded by unspecific effects of sleep deprivation and circadian effects on memory performance. Moreover, beneficial effects of sleep on episodic memory have been attributed, by some, to a lack of interference during sleep, rather than to active consolidation (Wixted 2004).

One possible way to control for the time of day effect is to make use of a daytime nap. A recent study based on this approach demonstrated sleep-related declarative memory benefits (Tucker et al. 2006). When using this approach, however, the difference in amount of sleep and interference between the nap and the control group might still introduce a confound. A few other recent studies have shown beneficial effects on memory of a full night’s sleep. These studies addressed the mentioned confounds through the use of extensive control groups for time of day effects (Ellenbogen et al. 2006; Gais et al. 2006) or by enhancing slow wave sleep through electrical stimulation (Marshall et al. 2006). Taken together, these studies suggest enduring effects of sleep on associative memory in tasks that rely to a large extent on the language system.

Here, we investigate whether these sleep effects on episodic/declarative memory generalize to other types of associations. We chose a task in which faces have to be associated to spatial locations (Fig. 1A). The acquisition of object–location associations has been shown to depend on the hippocampal formation (Miller et al. 1997; Stepankova et al. 2004; Piekema et al. 2006). An additional objective was to investigate the mechanism through which sleep enhances recall. We thus assessed how the order of sleeping and waking after learning affects retention and whether there are “carry-over” effects from sleep-time processing onto wake-time retention and vice versa.

The experimental procedure involves subjects undergoing natural sleep–wake cycles. Each subject is assigned to one of four retention interval conditions (Fig. 1B): a 12-h interval that starts in the morning, between 8 and 11 AM (wake group); a 12-h interval that starts in the evening, between 8 and 11 PM (sleep group); a 24-h interval that starts in the morning (wake–sleep group); and a 24-h interval that starts in the evening (sleep–wake group). Recall of face–location associations in each subject is tested 10 min after learning (immediate recall) and following the retention interval (delayed recall).

Using this approach, confounds of abnormal sleep–wake patterns are avoided, while concerns regarding circadian effects are addressed in two manners: First, recall rates immediately after learning are compared between the morning and evening learners. Second, possible influences of time of day on retrieval would act orthogonally to the length of the retention interval and should thus not exert a systematic influence on the interpretation of sleep–wake effects over all groups (for similar procedures, see Fenn et al. 2003; Ellenbogen et al. 2006; Gais et al. 2006). Importantly, the 24-h groups allow us to examine whether any beneficial effects of sleep are due merely to a lack of interference. In that case, retention over the two 24-h intervals should be similar, as they contain similar amounts of sleep and waking activity. Conversely, if forgetting in the two groups differs, this implies that there are enduring effects of either diurnal or nocturnal memory processing that “carry over” into the subsequent phase. Both possibilities were evaluated.
were trained in three rounds: two rounds of male and female faces (10 of each). Twenty face consisted of grayscale, front-facing, photographs of unfamiliar location task (adapted from Takashima et al. 2006). Task stimuli – previous 24 h. Subsequently, they started training on the face well as sleep quality and psychoactive substance use in the pre-
ationnaire regarding general selection criteria (specified earlier), as
viation 3.0 yr; 61 females/16 males).

Selection of participants was based on a questionnaire and interview. Exclusion criteria were current use of psychoactive medication or illicit drugs, history of drug abuse, head trauma, neurological or psychiatric illness, diagnosed sleep abnormalities, self-perceived sleep problems, or unusual sleep patterns (<5 h of sleep in a regular night; sleeping outside the 10 PM–10 AM window; frequent interruptions of sleep, etc.). Incidental sleep disturbances, psychoactive drug use, and excessive alcohol intake within the 24 h before the start of the experiment and throughout the experiment were also registered and were cause for exclusion. Written informed consent was obtained according to the local medical ethics committee, and subjects received either a monetary fee or study credits for their participation.

Eighty-six subjects completed the experiment. Six of these were disqualified from analysis based on noncompliance with criteria regarding sleep quality and psychoactive substance use during the experiment. Three more were excluded because they performed below chance level (12.5% hit rate) on immediate and/or delayed recall. Analyses were, thus, performed on data from 77 subjects (mean age 22 yr, range 18–35 yr, standard deviation 3.0 yr; 61 females/16 males).

Upon arrival at the laboratory, subjects completed a questionnaire regarding general selection criteria (specified earlier), as well as sleep quality and psychoactive substance use in the previous 24 h. Subsequently, they started training on the face-location task (adapted from Takashima et al. 2006). Task stimuli consisted of grayscale, front-facing, photographs of unfamiliar male and female faces (10 of each). Twenty face-location pairings were trained in three rounds: two rounds of “passive” training followed by a round of “active” training. During passive training, subjects focused on a fixation cross surrounded by 8 gray target dots (1000 ms; Fig. 1A, left). After the central presentation of a face (1100 ms), the appropriate target turned green (Fig. 1A, middle). This prompted the subject to move the cursor to the target using a joystick. Then, the face reappeared centrally and moved to the target where it remained for 2000 ms before the next trial started (Fig. 1A, right). Using this procedure the 20 face-location pairings were presented twice.

In the active training that followed, a face was presented centrally for 1000 ms, prompting the subject to move the cursor to one of the targets. For correctly chosen targets, the face reappeared in the center, surrounded by a green box. After 1000 ms it moved to the target location where it remained for 500 ms. If an incorrect target location was chosen, a red box appeared in the center for 1000 ms. After the box disappeared the correct target turned green, prompting the subject to move the cursor to this target. Then the face reappeared centrally and moved to the target location where it remained for 500 ms before the next trial started.

Following a 10-min break, during which a nonstrenuous, word-categorization task (filler task) was performed, subjects were tested for recall of the learned material. The test procedure was identical to the active training phase, except that no feedback on performance was given: When subjects selected a location through a joystick movement, this location became white for 200 ms and then the next trial started.

After the recall test, subjects left the laboratory and returned either 12 or 24 h later. Upon arrival, they completed another questionnaire, regarding sleep quality and psychoactive substance use during the retention interval. Then recall of the 20 face-location pairs was tested a second time. The order of presentation of the faces was randomized over subjects and over the rounds of training and testing.

Table 1 presents immediate recall performance and performance after the retention intervals for the four experimental groups. Recall scores are given as the percentage of correctly remembered face-location associations. There were no statistically significant differences in immediate recall scores between the four groups (analysis of variance [ANOVA] with the four groups as between-subject factor: F(3, 73) = 2.01, P = 0.12; post-hoc Tukey honestly significant difference [HSD] tests comparing each group with all others: all P-values > 0.05).

Effects of the sleep-wake pattern (retention starting with sleep; retention starting with waking) and the length of the retention interval (12 h; 24 h) on recall were evaluated through a repeated measures ANOVA. In line with the classic memory literature, forgetting increased with the length of the retention interval (F(1,73) = 6.08; P = 0.016). More importantly, performance was better when learning was followed by sleep than when it was followed by day-time waking (F(1,73) = 18.44; P < 0.0005). Thus, there appears to be a benefit of sleeping shortly (a few hours) after encoding.

In Figure 2, forgetting over the retention interval in each group is expressed as the difference between immediate and delayed recall scores (see also the last column of Table 1). Direct comparison of the two 12-h groups showed significantly less forgetting in the sleep group than in the wake group (t = −2.56, df = 37, P = 0.015). Moreover, comparison of the 24-h groups showed that forgetting was less for the sleep–wake group than for the wake–sleep group (t = −2.46, df = 36, P = 0.019). These analyses were also performed with forgetting expressed as a proportion of immediate recall (forgetting = delayed/initial recall).

Figure 1. (A) Screen view of the passive training of the face–location association task. During training the subject is presented with a screen with eight locations, marked by gray circles, and a fixation cross in the center (left). A face is presented on the fixation cross (middle) and the subject had to move a joystick-controlled cursor to an indicated location marked by one of the peripheral circles becoming green (white in the figure). Subsequently, the face moved to the indicated location (right). (B) Schematic outline of experimental procedure for the 4 groups (12-h wake group, 12-h sleep group, 24-h wake–sleep group, 24-h sleep–wake group). (T) Training session; (R1) immediate recall test; (R2) recall test after retention interval.

<table>
<thead>
<tr>
<th>Group</th>
<th>N (%)</th>
<th>Imm. correct</th>
<th>Del. correct</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wake</td>
<td>19</td>
<td>52 ± 20</td>
<td>45 ± 19</td>
<td>−7 ± 10</td>
</tr>
<tr>
<td>Sleep</td>
<td>20</td>
<td>41 ± 13</td>
<td>42 ± 16</td>
<td>1 ± 8</td>
</tr>
<tr>
<td>Wake–sleep</td>
<td>20</td>
<td>52 ± 21</td>
<td>41 ± 16</td>
<td>−12 ± 12</td>
</tr>
<tr>
<td>Sleep–wake</td>
<td>18</td>
<td>50 ± 17</td>
<td>47 ± 19</td>
<td>−3 ± 8</td>
</tr>
</tbody>
</table>
The main findings remain unaltered (sleep group vs. wake group: \( t = 2.15, df = 37, P = 0.038 \); sleep–wake group vs. wake–sleep group: \( t = 2.31, df = 36, P = 0.027 \)).

We further tested whether there was significant forgetting in each of the four groups. Forgetting was significantly different from zero in the 12-h wake group (\( t = 2.93, df = 18, P = 0.009 \)) and in the 24-h wake–sleep group (\( t = 5.94, df = 19, P < 0.0005 \)), but not in the 12-h sleep (\( t = -0.57, df = 19, P = 0.58 \)) and 24-h sleep–wake (\( t = 1.57, df = 17, P = 0.14 \)) groups.

These findings confirm the beneficial effect of sleep on memory. Moreover, findings in the 24-h group indicate that the order of sleeping and waking after learning affects memory retention. This suggests that differential effects of sleeping and waking on memory are not due solely to reduced interference during sleep. Rather, some persisting effect of post-encoding processes on later phases of retention may be occurring.

To assess the nature of these post-encoding processes, forgetting over the second half of the wake–sleep interval was compared with forgetting in the 12-h sleep condition and forgetting in the second half of the sleep–wake interval was compared with forgetting in the 12-h wake condition (Fig. 3). We inferred how memory decay progressed over the second half of the retention interval in each 24-h group, assuming that forgetting over the first half was, on average, the same as in the respective 12-h groups (see legend of Fig. 3). This plausible assumption relies on the fact that subjects in all groups are derived from the same population and, thus, demonstrate similar memory characteristics. In accordance with such an assumption, there are no statistically significant differences between the four groups with respect to the mean of immediate recall scores (see above). Also, there are no significant differences between groups in the variance of any of the memory measures (Levene test for homogeneity of variances: forgetting (R1 – R2), \( L = 0.74, df(3, 73), P = 0.53 \); forgetting rate (R2/R1), \( L = 0.12, df(3, 73), P = 0.95 \); immediate recall, \( L = 1.9, df(3, 73), P = 0.13 \); delayed recall, \( L = 0.35, df(3, 73), P = 0.79 \)). Also, note that forgetting is calculated relative to immediate recall performance in each subject to further minimize group differences.

Statistical analysis showed that inferred decay during sleep in the wake–sleep group was significantly larger than observed decay during sleep in the 12-h sleep group (\( t = -7.47, df = 38, P < 0.0005 \)). This implies that the amount of time spent waking between encoding and subsequent sleep significantly reduces the beneficial effects of sleep on retention. Conversely, the difference between inferred memory decay during the waking period in the sleep–wake group and observed decay in the 12-h wake group (Fig. 3b) was not significant (\( t = 1.57, df = 35, P = 0.13 \)). This indicates that sleep does not substantially reduce forgetting during the subsequent day, that is, we were not able to establish that sleep protects the memory trace against interference.

To test for possible circadian effects on encoding and immediate retrieval, we performed an ANOVA on the immediate retrieval data, with “time of learning” (morning/evening) as between-subject factor. No significant effect was found, that is, the immediate recall performance was similar in the morning and in the evening. Hence, confounds from time-of-learning effects on the interpretation of the retention data are unlikely.

In summary, the results of this study show a significant difference in retention of novel face–location associations over 12-h sleep and wake intervals, with less forgetting occurring during sleep than during waking. Retention is also significantly better over a 24-h retention interval when sleep follows learning directly than when sleep is delayed by a day of waking activity.

It would be tempting to explain these findings in terms of a sleep consolidation process that stabilizes memories against day-time decay. However, our analyses do not support this notion, as diurnal forgetting directly following learning and after a night of intervening sleep do not differ significantly. Rather, we observe differential forgetting during sleep, dependent on whether sleep...
follows learning immediately or after a day of interference: If the sleep phase follows shortly after learning, the memory trace does not deteriorate at all. However, if the memory trace has started to erode over the day, it continues eroding during the following night.

One way to explain these findings is that the fate of memories during sleep may depend on their relative strength or stability at sleep onset. After all, the longer the time between learning and sleep, the more a memory trace will have decayed. Thus, in our experiment, memories that have been encoded in the evening might have stronger representations at sleep onset than memories encoded in the morning. This interpretation is consistent with a PET study by Peigneux et al. (2004) showing that cerebral reactivation during sleep is modulated by the strength of the memory traces.

We will here outline two physiological hypotheses accounting for a different fate of strong and weak memories during sleep. The first hypothesis was implemented in a computational neural network model from our group (Meeter and Murre 2005) and is in line with neurophysiological findings from several groups (Wilson and McNaughton 1994; Buzsaki 1996; Isquierdo et al. 1997). According to this hypothesis, consolidation involves the spontaneous activation of subsets of neurons during sleep. Occasionally, a representation will be brought above the critical activation value for pattern completion, leading to full reactivation of that memory pattern and further strengthening of the memory trace by synaptic plasticity (in the model this involves recoding to the cortex). The model shows that neurons belonging to strong memories have a higher chance of being reactivated during sleep than weaker ones. The model also shows that strengthening of any one memory occurs at the expense of other overlapping representations, which tend to be eroded. This leads to a consolidation trajectory for a given memory that is positively related to its original encoding strength, negatively influenced by a long wake period between encoding and sleep, and negatively influenced by temporally proximal encoding of other materials. Thus, memories are weakened or strengthened during sleep, depending on recent “memory history.”

A second working hypothesis does not involve actual reactivation of recently learned representations during the subsequent sleep. The synaptic homeostasis hypothesis claims that synaptic potentiations, acquired during learning, are downscaled during the following sleep period, so that the overall synaptic strength in the network is brought back to baseline (Tomon and Cirelli 2006). According to this hypothesis, weak memories may be lost during sleep because they do not survive the downscaling process. Importantly, the first hypothesis entails a mechanism that modifies individual memory traces and their relative strength over time. Conversely, the second hypothesis predicts static maintenance of the relative strength of memories that persist in the network. However, according to both hypotheses, weak memories could be lost during sleep.

In interpreting these findings in the context of everyday life, it should be considered that, in our experimental setup, group differences in the representational strength of task-related material at sleep onset are only due to the amount of wake time (and thus interference) between learning and sleep onset. However, in a real-life setting, initial encoding strength of stimuli (for instance, in relation to affective and motivational circumstances) will co-determine memory strength at sleep onset.

Beneficial effects of sleep on declarative memory have recently been found in a few other carefully controlled studies. One of these shows that boosting of slow oscillations during sleep potentiates episodic memory (Marshall et al. 2006). Another reports beneficial effects of sleep on foreign language learning (Gais et al. 2006). The latter study uses a setup very similar to our own and, in line with our findings, suggests that sleep is particularly beneficial when it follows shortly after learning. As suggested above, one explanation for these findings is that sleep-time processing of memories is related to their relative representational strength at sleep onset. This notion could be tested more directly through experiments in which the strength of memory traces is manipulated experimentally.

A third study suggests a possible mechanism underlying consolidation, showing that sleep following learning renders memory traces more resistant to subsequent associative interference (Ellenbogen et al. 2006). Notably, this study adopts an AB–AC paradigm to test the resistance of the material learned before sleep against interference from closely related material. Conversely, our setup tests memory under circumstances of low interference (from material unrelated to the task). Under such circumstances, beneficial effects of sleep on subsequent retention appear to be modest. In fact, we found only a trend in this direction. On the other hand, there is a robust negative effect of diurnal trace deterioration on subsequent nocturnal retention. Hence, our findings suggest an additional consolidation mechanism, which also acts under circumstances of low interference. This mechanism might be viewed as a “memory filter” that retains strong memories while weak ones are sifted out.

It might also be noted that the aforementioned study (Ellenbogen et al. 2006) did not report a significant difference in retention over 12-h wake and 12-h sleep intervals. This may be due to a combination of insufficient power and ceiling effects, as subjects learned the task to a 100% performance criterion. To uncover the subtle effects of sleep–wake patterns on retention, our own study used a larger number of subjects (N = 77) and trained subjects to an initial performance well below 100%.

As a final consideration, it should be taken into account that the effects on memory in our study could also be due to factors that co-vary with the sleep–wake cycle or the environmental light–dark cycle. It is therefore important that the findings are confirmed in ulterior studies using different approaches. In conclusion, our results support a beneficial effect of sleep on episodic memory, showing that such effects also apply to nonverbal face–location associations. Importantly, our findings suggest that sleep does not lead to an indiscriminate strengthening of all memories present at sleep onset. Rather, they are in line with a consolidation process, in which the fate of a memory depends, at least in part, on its relative strength at sleep onset. Strong memories tend to be preserved, while weaker memories erode still further.

This kind of process could serve to “clean up” memory, ridding weakly encoded memories, while selectively preserving more strongly encoded ones. Such a mechanism is in line with neural network perspectives on hippocampal memory processing and consolidation, which emphasize limited capacity, competition, and overwriting (Treves and Rolls 1994; McClelland et al. 1995; Meeter et al. 2004; Meeter and Murre 2005; Talanini et al. 2005). It might receive further support from behavioral experiments in which encoding strength is experimentally manipulated.

Acknowledgments

This study was supported by grants 051-04-102, 051-04-100, and 451-06-006 from the Dutch National Science Foundation (NWO). We thank Martin Paassen for his assistance on the project.

References


