BRIEF REPORT

What drives sleep-dependent memory consolidation: Greater gain or less loss?

Kimberly M. Fenn · David Z. Hambrick

Abstract When memory is tested after a delay, performance is typically better if the retention interval includes sleep. However, it is unclear what accounts for this well-established effect. It is possible that sleep enhances the retrieval of information, but it is also possible that sleep protects against memory loss that normally occurs during waking activity. We developed a new research approach to investigate these possibilities. Participants learned a list of paired-associate items and were tested on the items after a 12-h interval that included waking or sleep. We analyzed the number of items *gained* versus the number of items *lost* across time. The sleep condition showed more items gained and fewer items lost than did the wake condition. Furthermore, the difference between the conditions (favoring sleep) in lost items was greater than the difference in gain, suggesting that loss prevention may primarily account for the effect of sleep on declarative memory consolidation. This finding may serve as an empirical constraint on theories of memory consolidation.

Keywords Memory · Sleep · Consolidation · Decay

Introduction

Nearly a century ago, Jenkins and Dallenbach (1924) had participants learn lists of nonsense syllables and then tested their memory for the lists at various points. Jenkins and

K. M. Fenn (✉) · D. Z. Hambrick
Department of Psychology, Michigan State University,
East Lansing, MI 38824, USA
e-mail: kfenn@msu.edu

Dallenbach found that memory performance was better after retention intervals that included sleep. This basic pattern of results has since been replicated by a number of researchers. For example, consistent with the results of a number of other studies (Ellenbogen, Hulbert, Jiang & Stickgold, 2009; Ellenbogen, Hulbert, Stickgold, Dinges & Thompson-Schill, 2006; Gais & Born, 2004; Gais, Lucas & Born, 2006; Gais, Molle, Helms & Born, 2002; Molle, Marshall, Gais & Born 2004; Plihal & Born, 1997, 1999), we found that participants' memory for paired associates was better after a retention interval that included sleep than after one that only including waking activity (Fenn & Hambrick, 2012). Together with findings from research on nondeclarative memory (cf. Brawn, Fenn, Nusbaum & Margoliash, 2010; Fenn, Nusbaum & Margoliash, 2003; Huber, Ghilardi, Massimini & Tononi, 2004; Karni, Tanne, Rubenstein, Askenasy & Sagi, 1994; Stickgold, James & Hobson, 2000), this pattern of results suggests that a critical function of sleep is to consolidate new memories.

While the finding that sleep benefits declarative memory performance is well established, the question of what accounts for this fact remains in dispute in the literature. One possibility, which has been mentioned in media accounts of research on sleep and memory, is that a period of sleep enhances retrieval of information from long-term memory—that information that was not remembered before sleep becomes more accessible during sleep, and thus easier to remember after sleep. We alluded to this possibility to explain our finding that memory for paired associates was better after sleep than after being awake (Fenn & Hambrick, 2012).

Another possibility is that sleep protects against memory loss that normally occurs during waking activity, due to interference. In fact, one prominent theory of forgetting has argued that encoding of information during waking interferes with newly formed memory, and sleep benefits

memory by providing a period of time when memory is safe from retroactive interference (Wixted, 2004). Consistent with this possibility, several studies have shown that when participants learn declarative information in the morning and remain awake for a full day prior to sleep, they show more forgetting at a 24-h test than participants who are trained in the evening and go to sleep a few hours after training (Benson & Feinberg, 1977; Gais et al., 2006; Nesca & Koulack, 1994; Payne, Tucker, Ellenbogen, Wamsley, Walker, Schacter & Stickgold, 2012).

Present study

Here, we used a new research approach to investigate these possible explanations for the finding of better memory performance after sleep than waking. The study took place in two sessions, following the procedure in our earlier study (Fenn & Hambrick, 2012). In Session 1, which was conducted in either the morning or the evening, we presented participants with a list of paired associates and tested their memory. In Session 2, which occurred after a 12-h retention interval that included being either awake or asleep, we gave participants a final memory test.

We then used item data to compute subscores reflecting different types of change in memory performance across sessions. Our major analyses focused on two subscores. *Memory gain* was the number of items that a participant recalled in the final test (in Session 2) that had not been recalled at any point during training (in Session 1), whereas *memory loss* was the number of items that a participant recalled at the final training test (in Session 1) that were not recalled in the final test (in Session 2).

Our research question was whether the sleep and wake conditions would differ in their average levels of memory gain and memory loss. The finding of a higher level of memory gain in the sleep condition than in the wake condition, combined with no difference between the conditions in memory loss, would suggest that the better memory performance after sleep is due primarily to enhanced retrieval of information. By contrast, the finding of a lower level of memory loss for the sleep than for the wake condition, combined with the lack of a difference between the conditions in memory gain, would suggest that this pattern is due primarily to protection against loss during sleep (see the supplementary online materials for more detailed discussion of this point).

As a secondary analysis, we tested for correlations between the memory gain and memory loss subscores. A finding that the subscores correlated significantly with each other would suggest that a common mechanism may contribute to both gain and loss. In contrast, a finding that the subscores correlated near zero with each other would suggest that the two subscores may reflect independent mechanisms.

Method

Participants

The participants were 495 native English speakers who reported no history of sleep or memory disorders. A number of the participants were excluded from all analyses because they did not complete the experiment ($n = 38$), because they reported diagnosed sleep or psychological disorders or habitual difficulty sleeping ($n = 10$), or because of experimenter error during the test ($n = 4$). An additional 89 participants were excluded because they reported napping during a waking retention interval;[1] naps of even very short duration can result in consolidation in this task (Lahl, Wispel, Willigens & Pietrowsky, 2008). The remaining participants[2] were 354 undergraduate students between the ages of 18 and 35. Demographic information was not obtained from 72 of the participants; of the 282 who did report this information, 203 were female and 79 were male.

Stimuli

The stimulus set consisted of 48 pairs of semantically related nouns (from Fenn & Hambrick, 2012). The word pairs were adapted from Gais and Born (2004) and were matched for frequency, imagery, and concreteness (Francis & Kučera, 1982).

Procedure

We conducted two experimental sessions, separated by a 12-h retention interval. For the wake condition ($n = 165$), the first session occurred at 9:00 and the second session occurred at 21:00. For the sleep condition ($n = 189$), the first session began at 21:00 and the second session began at 9:00 the following morning, after a regular sleep phase. The experimental sessions began within a 30-min window of these times.

In Session 1, participants studied 48 pairs of semantically related words and were tested on the word pairs. The stimuli were presented for 4,000 ms, with a 1,500-ms intertrial interval. Immediately after study, the participants were given a cued-recall test on 40 of the word pairs. The first four

---

[1] The participants who napped also showed significantly less sleep on the night of the study than did either the wake or the sleep condition. A one-way ANOVA revealed a significant effect of condition (wake, sleep, or nap) on sleep duration $F(2, 433) = 11.41$, $p < .001$, and Scheffe's post-hoc tests revealed that participants who napped showed significantly less sleep on average than did either the wake ($p < .001$) or the sleep ($p < .001$) condition. We therefore excluded participants who napped from all of the analyses.

[2] All participants were recruited for an experiment requiring three separate sessions. In the third session, the participants returned to complete a cognitive test battery, but the results concerning those tests are to be described in a separate report.

and the final four items presented during study never appeared on any of the tests, to control for primacy and recency effects on memory performance, as is standard in the literature. Testing on these items could have artificially increased immediate memory performance, particularly for items that remained in working memory. During the test, the first word of each pair was presented, and participants were given unlimited time to type the second word. After each response, participants were given two forms of feedback: They first were told whether their response was correct or incorrect, and then were shown the correct words in the pair. Participants were trained to a criterion of 60 % correct. If criterion was not met, the entire cued-recall test, including feedback, was repeated, until criterion was achieved. It should be noted that the final test during training always included feedback, so this test was also a learning trial. Thus, we expected performance to improve on the delayed test in Session 2. In Session 2, participants were given a final cued-recall test, without feedback. As in the previous test, the first word in each pair was presented, and the participants had unlimited time to respond. Items were presented randomly in the study phase and in all tests.

Memory measures

As was already mentioned, we computed subscores reflecting change in memory performance from Session 1 to Session 2. The two outcomes that were critical to the present investigation were memory gain and memory loss (Table 1). To reiterate, memory gain was the number of items that a participant recalled in the final test (in Session 2) that had not been recalled at any point during training (in Session 1), whereas memory loss was the number of items that a participant had recalled at the final training test (in Session 1) that were not then recalled in the final test (in Session 2).

Table 1 Outcomes for each item: Gained, lost, recovered, or lost during training

	Early Training Tests	Final Training Test	Final Test
Gain	No	No	Yes
Loss	No	Yes	No
Loss	Yes	Yes	No
Recovered	Yes	No	Yes
Lost During Training	Yes	No	No

"Yes" indicates correct recall during that particular test, and "No" indicates that recall was not correct. "Early Training Tests" refers to performance on all tests prior to the final training test. An item was scored as being correct if it was correctly recalled on *any* of the early tests; the item did not need to be recalled on every test. The "Final Training Test" was the last test during training, when criterion was achieved, and the "Final Test" was the last test of the experiment, during Session 2

(Note that lost items may have also been correctly recalled during early training tests.)

We also computed subscores reflecting two other possible outcomes. The *recovered* subscore was the number of items that had been correctly recalled during any of the early training tests but not on the final training test, and that were then recalled correctly during Session 2. The *lost-during-training* subscore was the number of items that had been correctly recalled during at least one early training test, but that were not recalled during the final training test (and not recalled during Session 2).

Results

To determine whether the sleep condition showed greater improvement from the final training test (Session 1) to the final test (Session 2) than the wake condition, as in previous studies, we performed a repeated measures analysis of variance (ANOVA) with Condition (wake or sleep) as a between-subjects factor and Test (final training test or final test) as a within-subjects factor. We found main effects of condition [$F(1, 352) = 9.46$, $p < .01$] and test [$F(1, 352) = 291.13$, $p < .001$], as well as a Test × Condition interaction [$F(1, 352) = 28.66$, $p < .001$], indicating that the sleep condition improved more across the retention interval than did the wake condition (Fig. 1).

Thus, we replicated the finding of greater improvement in memory performance after a retention interval that includes sleep than after one that includes only waking activity. With this established, we performed analyses on the subscores to better understand the source of this effect. First, we performed a repeated measures ANOVA with Condition (wake or sleep) as a between-subjects factor and Subscore Type (gain or loss) as a within-subjects factor.[3] Main effects emerged of both condition, $F(1, 352) = 24.73$, $p < .001$, and subscore type, $F(1, 352) = 1,721.86$, $p < .001$. The effect of subscore type indicates a larger amount of gain than of loss across the conditions. We also found a significant Condition × Subscore Type interaction, $F(1, 352) = 7.26$, $p < .01$. The difference between the sleep and wake conditions in loss was greater than the difference between the conditions in gain (Fig. 2). Indeed, the sleep-versus-wake effect size was more than four times larger for loss ($d = 0.69$) than for gain ($d = 0.17$).[4]

---

[3] We transformed the loss variable by multiplying each value by −1, so that better performance would be in the same direction for this subscore and the gain subscore.

[4] One potential concern about this finding is that some participants scored near ceiling on the final training test, and therefore had little room for change in memory performance across the retention interval, thereby artificially producing this interaction. However, when participants who were at ceiling on the final training test were removed from the analysis, the interaction remained significant (see the supplemental online materials).
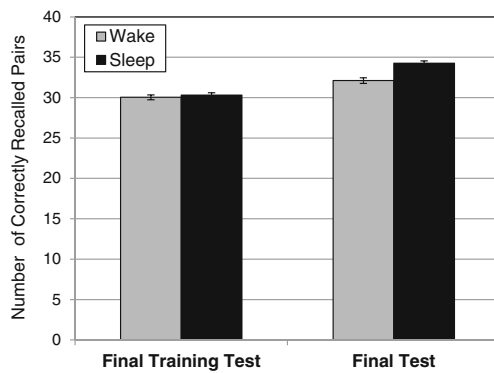
Fig. 1 Numbers of correctly recalled word pairs (out of 40) on the final training test (during Session 1) and on the final test (during Session 2) for the wake and sleep conditions. Error bars represent ±1 standard error of the means

We also compared the conditions in average number of items *recovered* and *lost during training* for participants who took two or more tests to reach criterion ($n = 222$; 62.7 % of our sample). These values were quite low for both conditions (Table 2) and were not significantly different across conditions (ts < 1, ps > .4).

Correlations

The correlation between the gain and loss subscores approached statistical significance in the wake condition but was quite small ($r = .14$, $p = .07$), and was near zero in the sleep condition ($r = .03$, $p = .42$). Thus, there was some evidence that memory gain and memory loss may reflect independent mechanisms. Considering all four subscores, the average correlations were $r = -.07$ in the wake condition and $r = .06$ in the sleep condition (see the supplementary online materials).

Circadian and sleep analyses

Because participants in the wake condition were trained in the morning and tested in the evening, while participants in
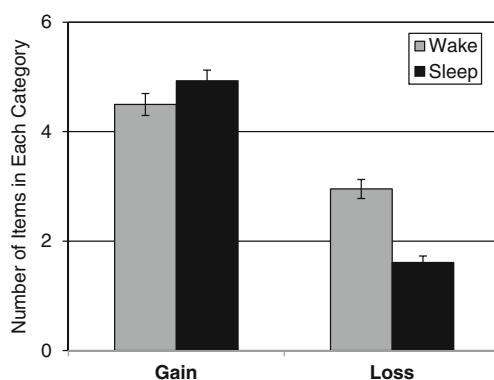

Fig. 2 Numbers of items gained and lost for the wake and sleep conditions. Error bars represent ±1 standard error of the means

Table 2 Numbers of items *gained*, *lost*, *recovered*, and *lost during training* for the wake and sleep conditions

	Gain	Loss	Recovered	Lost During Training
Wake	4.49 (2.5)	2.95 (2.2)	0.87 (1.0)	0.33 (0.6)
Sleep	4.93 (2.7)	1.60 (1.6)	0.98 (1.1)	0.34 (0.6)

Values in parentheses indicate standard deviations

the sleep condition were trained in the evening and tested in the morning, it is possible that diurnal or circadian differences might explain our finding of less loss and greater gain in the sleep condition. To investigate this possibility, we compared the wake and sleep conditions in terms of performance on both the initial and the final training test, as well as on the number of tests to reach criterion. If one of our conditions performed better on these measures than the other did, this might suggest that paired-associate learning was better in either the morning or the evening. However, we found no evidence of circadian variation in this experiment: Recall on the first training test was not significantly different between the two conditions, $t(352) = 0.09$, $p = .92$, and the conditions also showed similar performance on their final training tests, $t(352) = 0.62$, $p = .53$ (Table 3). Furthermore, no significant difference was apparent in the numbers of tests to reach criterion, $t(352) = 1.4$, $p = .14$. Thus, our results cannot be attributed to performance differences based on time of day.

We also compared the amounts of self-reported sleep on the night of the study between our conditions. The wake condition reported an average of 6.96 ± 1.3 h (mean ± *SD*) of sleep, whereas the sleep condition reported an average of 6.79 ± 1.6 h.[5] These values were not significantly different, $t(345) = 1.14$, $p = .25$, and the trend was for the wake condition to report having slept more than the sleep condition. Thus, longer duration of sleep cannot explain the superior performance in the sleep condition.

Discussion

The finding that memory performance improves more after a retention interval that includes sleep than after one that includes only waking activity is well replicated, but still not well understood. Participants studied paired associates and were tested immediately after training and after a 12-h interval that was composed either entirely of waking or included sleep. We analyzed the numbers of items gained

---

[5] Four participants in the wake condition and three participants in the sleep condition did not report sleep duration.

Table 3 Average numbers of word pairs correctly recalled on the first and on the final test during training, as well as the average numbers of tests to reach criterion for participants in the wake condition and the sleep condition

	First Training Test	Final Training Test	Tests to Criterion
Wake	21.32 (7.3)	30.03 (3.8)	1.71 (0.7)
Sleep	20.46 (8.3)	30.30 (4.1)	1.83 (0.8)

Values in parentheses indicate standard deviations

and lost across time and found that the sleep condition showed both greater gain of individual items and less loss of individual items than did the wake condition. The difference between the conditions in the numbers of items lost was greater than the difference in the numbers of items gained.

Although we have discussed gain in terms of an improvement in memory performance, it is important to note that participants received feedback on their final training test. That is, on each trial, they were given the correct word pair after they had responded. Thus, we can assume that memory improved as a result of the feedback given on this test. However, we did not test the participants after this criterion test. Therefore, our measure of memory performance on the final training test likely underestimated performance. This means that our measure of memory gain may not reflect actual memory *improvement*. Regardless, using the exact same task and same design that has been used to show memory enhancement across sleep, we have shown that the strongest effect of sleep on declarative memory is in protection against loss, not in memory gain.

The finding that sleep both increases gain and reduces loss raises the question of whether these two effects reflect the same underlying mechanism. One possibility is that both effects reflect protection against loss in the sleep condition. The smaller number of items lost in the sleep condition, as compared to the wake condition, likely reflects protection from interference and loss of memory. The increased gain in the sleep condition, however, may also reflect protection against loss. Both conditions were given feedback on the final training test. We therefore expected both conditions to improve between the final training test and the final test. If both groups acquired the same number of new items on the final training test, the reduced gain in the wake condition may actually reflect *loss* of some of the training-related gain in memory. For example, both groups may have acquired five new word pairs on the final training test. However, the wake group may have lost some of this initial gain, potentially due to interference. Therefore, the gain observed in the sleep group may reflect maintenance of memory from the final training test, whereas the lower amount of gain in the wake group may reflect some loss of memory that had been acquired on the final training test. Thus, it is possible that

both the gain and loss scores reflect the same underlying mechanism: prevention against loss. This is consistent with early accounts of the role of sleep in memory performance (cf. Jenkins & Dallenbach, 1924) and with a more recent theory that has argued that sleep benefits memory by protecting it against retroactive interference, and thus, forgetting (Wixted, 2004).

If the gain and loss subscores do in fact reflect the same underlying mechanism, we would expect the subscores to be correlated with each other. However, this correlation was nonsignificant in both conditions. This finding may suggest that two mechanisms contribute to consolidation. We speculate that the mechanism underlying the reduced memory loss in the sleep condition is protection against interference. By contrast, the mechanism underlying increased memory gain in the sleep condition may be enhanced retrieval. Several explanations are possible for how sleep might enhance retrieval ability. For example, several studies have shown that neurons that are active during task activity are subsequently reactivated during sleep (Dave & Margoliash, 2000; Ji & Wilson, 2007; Louie & Wilson, 2001; O'Neill, Senior, Allen, Huxter & Csicsvari, 2008; Wilson & McNaughton, 1994), and this reactivation of a memory trace may enhance retrieval processes. Thus, it is possible that two separate consolidation mechanisms operate during sleep: a passive mechanism of protection from interference, and an active mechanism of information processing. While we cannot definitively argue for an active consolidation process, it is clear that even if sleep works to actively enhance retrieval processing, the strongest effects in this task are in protection against loss.

In conclusion, we have adopted a new approach to the study of declarative memory consolidation. To better understand the exact effect that sleep has on memory, future studies would benefit from employing the analytic approach described here. The pattern of results that we report also represents an important empirical constraint on theorization about sleep-related effects on declarative memory, in that proposed explanations for consolidation must be able to account not only for greater gain following sleep, but also for diminished loss.

References

Benson, K., & Feinberg, I. (1977). Beneficial effect of sleep in an extended Jenkins and Dallenbach paradigm. *Psychophysiology, 14,* 375–384.

Brawn, T. P., Fenn, K. M., Nusbaum, H. C., & Margoliash, D. (2010). Consolidating the effects of waking and sleep on motor-sequence learning. *Journal of Neuroscience, 30,* 13977–13982.

Dave, A. S., & Margoliash, D. (2000). Song replay during sleep and computational rules for sensorimotor vocal learning. *Science, 290,* 812–816.

Ellenbogen, J. M., Hulbert, J. C., Jiang, Y., & Stickgold, R. (2009). The sleeping brain's influence on verbal memory: Boosting resistance to interference. *PLoS ONE, 4,* e4117. doi:10.1371/journal.pone.0004117

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

Fenn, K. M., & Hambrick, D. Z. (2012). Individual differences in working memory capacity predict sleep-dependent memory consolidation. *Journal of Experimental Psychology: General, 141,* 404–410.

Fenn, K. M., Nusbaum, H. C., & Margoliash, D. (2003). Consolidation during sleep of perceptual learning of spoken language. *Nature, 425,* 614–616.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston, MA: Houghton Mifflin.

Gais, S., & Born, J. (2004). Low acetylcholine during slow-wave sleep is critical for declarative memory consolidation. *Proceedings of the National Academy of Sciences, 101,* 2140–2144.

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

Gais, S., Molle, M., Helms, K., & Born, J. (2002). Learning-dependent increases in sleep spindle density. *Journal of Neuroscience, 22,* 6830–6834.

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

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

Ji, D., & Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocamus during sleep. *Nature Neuroscience, 10,* 100–107.

Karni, A., Tanne, D., Rubenstein, B. S., Askenasy, J. J. M., & Sagi, D. (1994). Dependence on REM sleep of overnight perceptual skill. *Science, 265,* 679–682.

Lahl, O., Wispel, C., Willigens, B., & Pietrowsky, R. (2008). An ultra short episode of sleep is sufficient to promote declarative memory performance. *Journal of Sleep Research, 17,* 3–10.

Louie, K., & Wilson, M. A. (2001). Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron, 29,* 145–156.

Molle, M., Marshall, L., Gais, S., & Born, J. (2004). Learning increases human electroencephalographic coherence during subsequent slow sleep oscillations. *Proceedings of the National Academy of Sciences, 101,* 13963–13968.

Nesca, M., & Koulack, D. (1994). Recognition memory, sleep and circadian rhythms. *Canadian Journal of Experimental Psychology, 48,* 359–379.

O'Neill, J., Senior, T. J., Allen, K., Huxter, J. R., & Csicsvari, J. (2008). Reactivation of experience-dependent cell assembly patterns in the hippocampus. *Nature Neuroscience, 11,* 209–215.

Payne, J. D., Tucker, M. A., Ellenbogen, J. M., Wamsley, E. J., Walker, M. P., Schacter, D. L., & Stickgold, R. (2012). Memory for semantically related and unrelated declarative information: The benefit of sleep, the cost of wake. *PLoS ONE, 7,* e33079. doi:10.1371/journal.pone.0033079

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

Plihal, W., & Born, J. (1999). Memory consolidation in human sleep depends on inhibition of glucocorticoid release. *Learning and Memory, 10,* 2741–2747.

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

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. doi:10.1146/annurev.psych.55.090902.141555