Cortex

CrossMark

Note

Sleep preserves original and distorted memory traces

Scott A. Cairney [a], Shane Lindsay [b], Ken A. Paller [c] and M. Gareth Gaskell [a,*]

[a] Department of Psychology, University of York, York, UK
[b] Psychology, School of Life Sciences, University of Hull, Hull, UK
[c] Department of Psychology, Northwestern University, Evanston, IL, USA

ARTICLE INFO

ABSTRACT

Retrieval facilitates the long-term retention of memories, but may also enable stored representations to be updated with new information that is available at the time of retrieval. However, if information integrated during retrieval is erroneous, future recall can be impaired: a phenomenon known as retrieval-induced distortion (RID). Whether RID causes an "overwriting" of existing memory traces or leads to the co-existence of original and distorted memory traces is unknown. Because sleep enhances memory consolidation, the effects of sleep after RID can provide novel insights into the structure of updated memories. As such, we investigated the effects of sleep on memory consolidation following RID. Participants encoded word locations and were then tested before (T1) and after (T2) an interval of sleep or wakefulness. At T2, the majority of words were placed closer to the locations retrieved at T1 than to the studied locations, consistent with RID. After sleep compared with after wake, the T2-retrieved locations were closer to *both* the studied locations *and* the T1-retrieved locations. These findings suggest that RID leads to the formation of an additional memory trace that corresponds to a distorted variant of the same encoding event, which is strengthened alongside the original trace during sleep. More broadly, these data provide evidence for the importance of sleep in the preservation and adaptive updating of memories.

1. Introduction

Retrieval practice can greatly benefit long-term retention, but it may also enable stored representations to be updated with new information that is available at the time of retrieval (Dudai & Eisenberg, 2004). Memory updating via retrieval provides a means of adapting to changes in the external environment. However, if retrieved information is erroneous, its integration within an existing representation can cause distortion and impair future recall (Schacter, Guerin, & St Jacques, 2011). Multiple bouts of remembering in the

absence of suitable feedback and correction may therefore lead to substantial memory inaccuracies.

Such retrieval-induced distortion (RID) was demonstrated by Bridge and Paller (2012). Participants learned a set of object locations and were then tested on all objects immediately after training, on a subset of objects the following day, and again for all objects another day later. RID was observed as locations recalled in the final test were preferentially closer to the locations recalled the prior day than to the locations that were originally learned. Moreover, brain potentials at the time of intervening retrieval predicted later memory distortion. Thus, rather than simply strengthening memory for the studied information, retrieval promoted the storage of retrieved information, which was often inaccurate.

Standard models of systems consolidation propose that episodic memories are initially dependent on both the hippocampus and neocortex, but become gradually independent of the hippocampus as consolidation unfolds (Marr, 1971; McClelland, 2013; McClelland, McNaughton, & O'Reilly, 1995). The multiple trace-transformation (MTT) account of memory, by contrast, proposes that the hippocampus is involved in the retrieval of episodic memories for as long as they exist (Moscovitch & Nadel, 1998; Nadel & Moscovitch, 1997; Winocur, Moscovitch, & Bontempi, 2010; Winocur, Moscovitch, & Sekeres, 2013). A central tenet of this model is that each time an episodic memory is retrieved it is re-encoded within the hippocampus as a new trace with unique contextual properties. Thus, the more often a memory is retrieved, the greater the number of corresponding hippocampal traces that will exist. Neocortical networks may therefore draw upon these traces to support a gist-like, decontextualised version of the original memory. From an MTT perspective, RID may lead to the presence of an additional hippocampal memory trace that corresponds to a distorted variant of the same encoding event. Subsequent retrieval operations would then draw upon both the original and distorted traces, resulting in a blend of the two. Alternatively, if the hippocampus does not re-encode distinct episodic traces during retrieval, then memory updating and distortion may result from some form of "overwriting" within relevant neural networks. Retrieval operations in this framework would then utilise only a single, distorted trace during memory recall.

There is now robust evidence that sleep facilities the consolidation of hippocampal-dependent, episodic spatial memories (Rasch, Buchel, Gais, & Born, 2007; Rudoy, Voss, Westerberg, & Paller, 2009; Wilhelm, Diekelmann, & Born, 2008). How sleep influences originally-learned and distorted spatial memories can thus provide novel insights into the neurocognitive mechanisms of memory updating. Accordingly, we examined the effects of sleep on memory consolidation following RID. Participants encoded word locations and were then tested before (T1) and after (T2) a period of sleep or wakefulness. RID was indicated when T2-recalled locations were closer to T1-recalled locations than to the studied locations. Thus, the distance between locations recalled at T1 and T2 provided an index of distorted memory content. However, the change in word-location error (distance between the studied and recalled locations) from T1 to T2 provided an index of original memory content.

We predicted three possible outcomes. First, if memory updating is achieved by an "overwriting" of existing information at retrieval, then sleep after RID should strengthen distorted but not original memory content (i.e., better recall of T1 locations but not studied locations after sleep vs wake). Second, if RID leads to an additional memory trace for the distorted location, then sleep should strengthen this and the original location trace to similar extents (i.e., better recall of both T1 and studied locations after sleep vs wake). Third, if post-retrieval memory traces are in some way subsidiary to original memory traces, then sleep should preferentially strengthen the original location memories (i.e., better recall of the studied locations than the T1 locations after sleep vs wake).

2. Materials and methods

2.1. Participants

Sixty healthy males were randomly assigned to a sleep group ($n = 30$, mean $\pm$ SD age = 19.77 $\pm$ 1.33 years) or a wake group ($n = 30$, mean $\pm$ SD age = 20.10 $\pm$ 1.49 years). Participants had no history of sleep, psychiatric or neurological disorders, were medication-free and had not consumed alcohol/caffeine within 24 h of the study.

2.2. Procedure

Two sessions were separated by a 120-min interval (see Fig. 1A). The first session commenced at 9pm (enabling the sleep group to have a normal bed time), and began with a word-location task. For each 3 sec trial of an initial passive viewing phase, participants viewed a rectangular box containing one of 50 words from Maki, McKinley, and Thomson (2004). Word boxes were presented on a grid background and appeared in randomised screen locations. After two passive viewing rounds, participants carried out an active learning phase. For each trial, one of the 50 words was presented centrally and, using the mouse, participants moved it to the location studied at passive viewing. The word then reappeared in the studied (i.e., correct) location for 3 sec. After two active learning rounds, words that had been consecutively placed within a standardised distance of 150 pixels (4.8 cm on our 27″ 1920 × 1080 display) from the studied location were dropped from the task and active learning continued until this criterion was met for all remaining words (Bridge & Paller, 2012; Cairney, Lindsay, Sobczak, Paller, & Gaskell, 2016). There was a marginal group difference in the number of rounds required to reach the criterion for all 50 words [sleep group mean $\pm$ SD = 10.00 $\pm$ 2.95 rounds; wake group mean $\pm$ SD = 12.00 $\pm$ 4.50 rounds; $t(58) = 2.04$, $p = .05$]. Importantly, however, there were no significant group differences in memory performance at baseline (see Section 3). The first of two tests (T1) for all 50 word locations was then carried out. This followed the same procedures as one round of active learning, but without location feedback. As a secondary assessment of declarative memory, participants then completed a word pair task. Details of this task and the results are available in the Supplementary Materials. Afterwards (~11pm), sleep group participants went to bed for 90 min. Sleep was confirmed with polysomnography. After waking, these individuals took a 30-min break to recover from sleep inertia. Participants in the