Sleep-Associated and Retrieval-Associated Memory Consolidation

Inaugural-Dissertation zur Erlangung der Doktorwürde
der Philosophischen Fakultät II
(Psychologie, Pädagogik und Sportwissenschaft)
der Universität Regensburg

vorgelegt von

MAGDALENA ABEL

aus Regensburg

Regensburg 2011

Erstgutachter: Prof. Dr. Karl-Heinz T. Bäuml

Zweitgutachter: Prof. Dr. Klaus W. Lange

Acknowledgement

This work would never have been finished without the support that I received from many different people over the years.

First and foremost, I would like to thank Prof. Karl-Heinz Bäuml, the initiator and supervisor of this thesis. He provided me not only with the infrastructure necessary for a project like this one, but also gave me professional and scientific input as well as motivating words, whenever I needed them.

Second, I want to thank everybody at the Bäuml lab for sharing their thoughts and helpful comments, but even more so for sharing coffee breaks and good times. My (current and former) colleagues Alp Aslan, Bernhard Pastötter, Ina Dobler, Oliver Kliegl, Andreas Schlichting, Christoph Holterman, Margit Frimberger, Anuscheh Samenieh, Tobias Staudigl, Maria Wimber, and Simon Hanslmayr clearly made university a more enjoyable place. Moreover, without the lab's research assistants (e.g., Doro Bodmann, Elisabeth Rose, Karin Aumüller, Leon Kroczek, Lisa Kugler, Marie Fellner, Michael Wirth, Sabine Krug, and Sandra Kaltner) the data collections this thesis is built on would never have been possible.

Last but not least, props go out to my friends and relatives, who kept me company throughout the years. I would like to dedicate this work to my 'extended family', and especially to my parents. Your love and support never cease to amaze me, I owe you everything.

Preface

When somebody is facing a tough decision, a suggestion frequently made is to advise the person to sleep on the respective problem. As this phrase is part of our conventional language use, most people seem to agree or to have made the experience themselves that things can fall into place during the night, and that one may rise with a new perspective the next day. Several anecdotes provided by scientists and artists support the idea that sleep can add something to our cognition; the chemist Kekule, for instance, reported that the potential ring-shaped form of the benzene molecule occurred to him in a dream in which he saw a snake bite into its own tail (for a list of similar anecdotes, see Maquet & Ruby, 2004). However, the proposal that the contents of our minds are further processed during periods of sleep does not rely solely on introspective feelings and reports. As a matter of fact, a considerable amount of experimental research has been devoted to disentangling the effects that sleep may exert on cognition.

A strong focus has thereby been put on the effect of sleep on memory consolidation. In general, the process of memory consolidation is assumed to stabilize memory contents and, thus, to be beneficial for later remembering. Consolidated memories are further assumed to show less time-dependent forgetting and to be less susceptible to the detrimental effects of retroactive interference (e.g., Wixted, 2004). During the last two decades, corresponding evidence has been brought forward that links sleep, and more specifically, certain sleep stages and sleep parameters to the effective consolidation of memories (for a review, see Diekelmann & Born, 2010). While we sleep, memory contents are assumed to be reactivated, and thereby to be stabilized and to be made more resistant to forgetting. Such and similar processes have been suggested to underly the familiar phenomenon of insight after sleep (e.g., Cai, Mednick, Harrison, Kanady, & Mednick, 2009; Wagner, Gais, Haider, Verleger, & Born, 2004).

The first part of this thesis will be dedicated to investigating sleep-associated memory consolidation in more detail. Indeed, previous research indicates that sleep may not benefit all memories equally. For instance, it has been found that sleep-related benefits are more pronounced for memories that are harder to access (Drosopoulos, Schulze, Fischer, & Born, 2007). If so, sleep may not only counteract time-dependent forgetting by stabilizing memory contents, but may also counteract experimentally induced forms of forgetting that are assumed to rely on a reduced accessibility of memories. By investigating interference effects and list-method directed forgetting, the prior work is to be extended to different conditions and a different paradigm.

Irrespective of the vast amount of research on sleep-associated memory consolidation, there are other factors as well that impact upon the contents of our minds. While sleep is mostly regarded as a pleasant state connected with rest and inactivity, active retrieval from memory constitutes quite the opposite: It takes place during wakefulness and requires deliberate efforts. Yet, research on the so-called testing effect has previously shown that such retrieval from memory is also beneficial for long-term retention. In particular, less time-dependent forgetting has been observed after retrieval in comparison to restudy or distractor conditions (for a review, see Roediger & Butler, 2011). As reduced time-dependent forgetting is seen as one criterion for memory consolidation (e.g., Wixted, 2004), one could speculate that retrieval from memory might be connected to the consolidation of memory contents, too.

Nevertheless, another branch of research shows that retrieval can also entail detrimental effects for related, but nonretrieved material. The term retrieval-induced forgetting circumscribes the finding that retrieval of a subset of material prompts forgetting of related, nonretrieved material relative to control material (Anderson, Bjork, & Bjork, 1994). Such forgetting is investigated by applying the retrieval-practice paradigm and assumed to rely on inhibitory processes that are initiated during retrieval in order to resolve the interference caused by automatic activation of related material (e.g., Bäuml, 2008). Concerning the longevity of the effect conflicting results have been reported; while some studies found the effect to be lasting, others found it to be vanishing with time (e.g., Garcia-Bajos, Migueles, & Anderson, 2009; MacLeod & Macrae, 2001).

The second part of this thesis will be dedicated to investigating whether retrieval may also be connected to processes of memory consolidation. If so, retrieved material should show less time-dependent forgetting and less susceptibility to retroactive interference than nonretrieved control material (e.g., Wixted, 2004). Applying the retrieval-practice paradigm allows not only to replicate and extend previous research on the testing effect by examining time-dependent forgetting and interference susceptibility of practiced memories, it also allows to investigate how related material is affected by retrieval practice. As active retrieval has not only been reported to strengthen the respective memories, but also to weaken related and interfering memories, possible effects of retrieval-associated memory consolidation could also be of interest for the perspective on retrieval-induced forgetting.

Finally, the third part of this thesis will summarize the obtained data on effects of sleep and retrieval. The idea that both sleep and retrieval could be related to memory consolidation may beforehand seem like a rather counterintuitive proposition. Sleep, on the one hand, is usually perceived as a passive state and assumed to recur solely because our bodies need rest. Whatever exactly mediates the mnemonic benefits associated with sleep seems to come without the need to actively work for it. Memory retrieval, on the other hand, requires an awake and active mind as well as cognitive efforts. If sleep and retrieval could be connected because they entail the similar basic effects associated with memory consolidation, this could spark new research and deepen our understanding of memory consolidation in a more general way.

Contents

Abstract

After encoding, memory contents need to undergo a phase of stabilization in order to be remembered in the long-term. Such stabilization is referred to as memory consolidation, and is assumed to be observable in less time-dependent forgetting and less susceptibility to interference. Evidence for a role of sleep in the effective consolidation of memories has previously been provided. In addition, research on the so-called testing effect has also indicated a contribution of retrieval practice to long-term memory, as the active retrieval from memory has been shown to be able to boost retention.

This thesis investigated effects of sleep and retrieval on memory consolidation. Effects of sleep-associated memory consolidation were examined by inducing interference in paired-associate learning and by applying list-method directed forgetting: Replicating previous work, sleep was found to be of preferential benefit for memories that are hard to access; thereby, sleep counteracted both incidental and intentional forms of forgetting. Applying the retrieval-practice paradigm, effects of retrieval-associated consolidation were observed: Retrieval was found to stabilize directly retrieved and related memories, thereby making them less susceptible to both time-dependent forgetting and retroactive interference.

The data presented here indicate that both sleep and retrieval may be associated with memory consolidation. Possible differences and parallels between sleep-associated and retrieval-associated memory consolidation as well as implications of such a proposal are discussed.

Part I

Sleep-Associated Memory Consolidation

Chapter 1

A Current Perspective

Humans and most other mammals spend a significant part of their lives asleep (Cirelli & Tononi, 2008; Siegel, 2009). Observed from the outside, sleep does not appear to be a very exciting affair: Responsiveness is reduced and hardly any visible activity takes place. In addition, behavioral control and consciousness are lost (not considering dream mentation, which is a debatable topic of its own; see for example Hobson, Pace-Schott, & Stickgold, 2000; Schredl & Erlacher, 2011). Maybe because of sleep's seemingly dull nature, the general assumption prevailed for long that the brain was completely switched off during times of sleep and that no important processes occurred during these time windows (e.g., Payne, Ellenbogen, Walker, & Stickgold, 2008; for a historical review, see Dement, 2003). However, especially the loss of consciousness that comes with sleep struck scientists as a fact that demanded explanation. Rechtschaffen (1998) phrased this line of thought as follows:

> *"Sleep has persisted in evolution even though it is apparently maladaptive with respect to other functions. While we sleep we do not procreate, protect or nurture the young, gather food, earn money, write papers, etc. It is against the logic of natural selection to sacrifice such important activities unless sleep serves equally or more important functions."*

Indeed, research of the past decades has shown that sleep supports more than one essential function (e.g., immune function, Imeri & Opp, 2009; or the regulation of affect, Walker & van der Helm, 2009; see also Siegel, 2005) and that the concomitant inactivity of the body is contrasted by considerable activity of the brain that may mediate these functions (e.g., Yoo, Gujar, Hu, Jolesz, & Walker, 2007). Yet, most of the scientific attention in recent years was paid to a further function attributed to sleep. The general proposal is that sleep consolidates memories, leaving them less prone to detrimental influences. In particular, sleep-dependent memory consolidation has been referred to as the most important function of sleep, as it can account for why we lose consciousness every night (e.g., Born & Wilhelm, in press).

1.1 Memory Consolidation and its Relation to Sleep

The term memory consolidation refers to the assumption that memory contents need to be stabilized in order to persist. Directly after encoding, memories are assumed to initially remain labile and vulnerable (Alberini, Milekic, & Tronel, 2006), which is why they have to undergo a phase of stabilization in order to last. Early evidence for such a time-dependent phase of memory consolidation came from studies showing that memory for recent events can both be impaired or enhanced by interfering treatments directly after encoding (e.g., by new learning, electroconvulsive shock, or stimulant drugs; for a review, see McGaugh, 2000). Thus, memory consolidation is assumed to transform initially fragile memories into more stable memory representations (e.g., Dudai, 2004) that are, as a consequence of this stabilization, less susceptible to retroactive interference (Stickgold, 2005; Wixted, 2004). Yet, such consolidated memory content may again be modified and updated by a process referred to as reconsolidation (for reviews, see Nader & Einarsson, 2010, or Sara, 2008): In humans, after longer delays, a reminder of previously encoded information has been shown to be sufficient to shift memories from a stable into a labile state again so that they may anew be subject to change (e.g., Hupbach, Gomez, Hardt, & Nadel, 2007).

The hypothesis that sleep is beneficial for memory and protects from forgetting has almost been around since the first experimental studies on human memory (e.g., Jenkins & Dallenbach, 1924). By now, the link between sleep and memory consolidation is well established (for reviews, see Diekelmann & Born, 2010; or Payne et al., 2008). As the same neuronal networks are involved in both the encoding and the consolidation of memory contents, and because no new incoming information is loaded on these networks during sleep, memory consolidation is assumed to work more effectively during sleep than during wakefulness - thereby making the loss of consciousness an adaptive feature of sleep that benefits the stabilization of recently encoded information (Born, Rasch, & Gais, 2006; Diekelmann, Wilhelm, & Born, 2009).

There are many experimental findings that document the importance of sleep for consolidation processes. Beyond others, Plihal and Born (1997) showed that sleep benefits both declarative and procedural memory contents. Participants studied paired associates and were trained on the mirror-tracing task before they either were allowed to sleep or had to stay awake. After sleep, recall of paired associates was found to be improved, both in comparison to baseline performance and in comparison to performance after wake. In parallel, participants needed less time after sleep to trace figures in the mirror-tracing task, both in comparison to baseline performance and in comparison to performance after wake. Such sleep-related benefits have by now been shown to emerge across different tasks (e.g., finger-tapping tasks, Fischer, Hallschmid, Elsner, & Born, 2002; saccade-learning tasks, Gais et al., 2008; or verbal list learning, Ficca, Lombardo, Rossi, & Salzarulo, 2000). In addition, they have also been found in various settings that may be closer to daily life (e.g., when having to remember to execute a goal, Scullin & McDaniel, 2010; when perceptually learning a language, Fenn, Nusbaum, & Margoliash, 2003; or when having to retain information on navigation in a spatial environment, Ferrara et al., 2008).

Ellenbogen, Hulbert, Stickgold, Dinges, and Thompson-Schill (2006a) furthermore reported evidence belaying that sleep renders memories less susceptible to retroactive interference. Participants studied a list of paired associates before going to sleep or staying awake. After 12 hours, the previously encoded paired associates were either tested immediately or after additional learning of new and similar paired associates. Overall, recall of the previously studied material was better after sleep compared to wake. However, a significant difference between the respective sleep and wake conditions only emerged after interference was induced: The additional learning impaired performance for the original material to a significantly higher degree in the wake compared to the sleep conditions, leading to overall higher recall rates in the sleep conditions. The authors concluded that sleep protected memories from retroactive interference, a finding that fulfills one of the criteria required for effective memory consolidation (e.g., Wixted, 2004).

While many studies have focused on rather immediate effects of sleep on memory contents, only few studies investigated the long-term effects of such consolidation. Nevertheless, it could be shown that the benefits entailed by sleep can persist over prolonged delay intervals. Applying a visual paired-associate task, Tucker, Tang, Uzoh, Morgan, and Stickgold (2011) for instance showed that less time-dependent forgetting occurs over a 24-h delay when subjects slept closely after encoding in comparison to when they stayed awake for 12 hours before going to sleep. Similar results were also obtained by Talamini, Nieuwenhuis, Takashima, and Jensen (2008), who reported sleep-associated benefits for the recall of face-location associations to persist over a delay interval of 24 hours, and by Gais, Lucas, and Born (2006), who found a benefit of sleep over sleep deprivation for vocabulary learning after a delay of 48 hours. Thus, the sequence of sleep and wake intervals seems to be of importance beyond the previously described effects found immediately after sleep and wake manipulations. Nevertheless, another finding by Wagner, Hallschmid, Rasch, and Born (2006) seems to be notable: Participants read several neutral and several emotional text passages before they were either allowed to sleep or had to stay awake. After four years, participants' recall of text contents was still found to be superior in the sleep condition compared to the wake condition, although only for the emotional but not for the neutral texts. This finding indicates that the impact of sleep may, at least under specific conditions, be serious and long-lasting. All in all, sleep stabilizes memory contents and reduces time-dependent forgetting, another criterion of memory consolidation (e.g., Wixted, 2004).

A further line of experimental evidence points out that not even a whole night of sleep is necessary in order to be able to observe effects of sleep-associated memory consolidation. Napping studies, in particular, have investigated the impact of short episodes of sleep during daytime on cognition, thereby conveniently eliminating time of day or sleep deprivation as potentially confounding factors. In such studies, subjects are typically asked to memorize material and are then either allowed to sleep for a specific duration of a few minutes to a few hours or are asked to stay awake for the same amount of time. While most napping studies converge on the finding that 1 to 1.5 hours of sleep are enough to prompt memory consolidation (e.g., Durrant, Taylor, Cairney, & Lewis, 2011; Mednick, Nakayama, & Stickgold, 2003; Nishida & Walker, 2007), astonishing results were reported by Lahl, Wispel, Willigens, and Pietrowsky (2008). The authors found that even very brief naps of about 6 minutes were sufficient to induce the better retention of a word list in comparison to a wake condition. Similarly, a 10-min nap was reported to be enough to consolidate motor memories (Debarnot, Castellani, Valenza, Sebastiani, & Guillot, 2011). Such findings give further drive to the debate about what exactly triggers sleep-associated memory consolidation.

Plihal and Born (1997), for instance, did not only show that declarative as well as procedural memories benefit from sleep, they also provided first evidence for a possible dissociation between the sleep-associated consolidation of declarative and procedural memories: While declarative memories where shown to profit from early sleep during the night, procedural memories mainly profited from late sleep. This indicates that different sleep stages may support the sleep-associated consolidation of different contents of memory. By now, considerable evidence indicates that declarative memories indeed benefit the most from slow-wave sleep that predominantly occurs early during the night (see also Daurat, Terriet, Foret, & Tiberge, 2007; Drosopoulos, Wagner, & Born, 2005; for a review, see Born, 2010). Other evidence, however, does link REM sleep to memory stabilization (e.g., Rauchs et al., 2004) and to other aspects said to be connected with consolidation, for instance to the abstraction and generalisation of memory contents (for a review, see Walker & Stickgold, 2010). The idea that the consolidation of declarative memory contents is primarily dependent on slow-wave sleep is additionally called into question by the previously described finding that effects of sleep-associated memory consolidation are detectable after very brief episodes of sleep. As stages of deep sleep are rarely or only very shortly reached in corresponding napping studies, it is argued that these stages cannot ultimately be tied to the observed effects. Alternatively, it has been suggested that sleep parameters occuring already in lighter sleep stages could trigger processes of sleep-associated memory consolidation. Sleep spindles during stage 2 sleep, for instance, have been proposed to act as such a trigger (for a review, see Fogel & Smith, 2011). Another alternative account states that the intact cycling between sleep stages observed during regular nights of sleep is beneficial for memory (for a comparison of the accounts, see also Diekelmann et al., 2009). To date, however, no final consensus has been reached yet concerning the question which component of sleep mediates effective memory consolidation.

Nevertheless, for declarative memory, the most convincing experimental evidence explaining effects of sleep-dependent consolidation relies on the assumption that memories are reactivated during slow-wave sleep, thereby being strengthened and stabilized (Rasch & Born, 2007). Early evidence for the existence of such replay of previous experiences during sleep came from studies on rodents. Wilson and McNaughton (1994), for instance, showed that cells in the rat hippocampus that were found to be simultaneously active during a spatial learning task were more likely to be co-activated again during post-learning sleep than they had been during pre-learning sleep (for a review on related research on rodents, see O'Neill, Pleydell-Bouverie, Dupret, &

Csicsvari, 2010). In recent years, research on humans succeeded in establishing a clear link between the reactivation of memory contents during sleep and their sleep-associated consolidation. The next paragraphs will provide a brief outline on corresponding studies and their theoretical implications.

In a study by Rasch, Büchel, Gais, and Born (2007), participants initially encoded object locations. During the presentation of every stimulus, a rose odor was presented as well, thereby being linked to the encoded material. After encoding, subjects were allowed to go to bed and were, subliminally in their sleep, either again presented with the same odor cue as during encoding or were presented a neutral vehicle instead. On a final delayed test during which no cue was present, it could be observed that the repeated presentation of the odor cue during slow-wave sleep had led to better memory performance in comparison to when the vehicle had been presented instead. Notably, similar results were recently reported by Rudoy, Voss, Westerberg, and Paller (2009), who initially paired each to-be-encoded stimulus with an individual auditory cue. Half of the applied cues were then again presented during slow-wave sleep. On a final test, such cueing during sleep was again found to have led to better memory performance for the related stimuli, this time in a within-subjects comparison to recall of the other half of stimuli that were connected to different cues not presented during sleep. It was argued that the cues triggered additional or amplified reactivation of the originally encoded memories related to the cues, thereby further improving memory performance.

In line with this reasoning, no unspecific benefit was observed by Rasch et al. (2007) when the odor cue was only presented during sleep, but had beforehand not been associated with the encoded material. Furthermore, Rasch et al. (2007) reported that the mnemonic benefit was only present when the odor cue was administered during slow-wave sleep, but not when it was presented during REM sleep or equal periods of wakefulness. In addition, no corresponding effect could be obtained for a procedural task relying to a higher degree on other brain structures than the hippocampus. Consistently, research on rodents suggests that the hippocampus, a brain structure that is known to be essential for declarative memory (Squire, 1992), coordinates the replay and reactivation of memory traces, thereby interacting with cortical areas, transfering recently encoded memories to sites of long-term storage, and binding memory traces stored in different brain areas together (O'Neill et al., 2010). All in all, these and similar findings suggest that declarative memories are reactivated in the hippocampus during slow-wave sleep, which may explain a range of the observed effects of sleep-associated consolidation. Note, however, that alternative accounts for the benefits of sleep exist and may be found elsewhere (e.g., Axmacher, Draguhn, Elger, & Fell, 2009; Tononi & Cirelli, 2006).

A point of critique frequently expressed when discussing active mechanisms of sleep-dependent memory consolidation is the following: During wake, subjects are exposed to continuous input, which, it is reasoned, could itself interfere with memories acquired earlier and harm performance in experimental wake groups. Therefore, so the idea, the better performance observed after sleep would not have to be caused by more effective consolidation that is actively mediated by sleep, but could just as well be caused by the simple fact that sleep passively protects memory from interfering learning (e.g., Wixted, 2004). Sleep, according to this view, would not lead to more effective consolidation and better memory performance because of the induced stabilization. Rather, wakefulness would lead to impaired performance due to interference, without the engagement of any consolidation processes in experimental sleep groups. However, if this were the case, memory contents should not be stabilized at all during sleep, as they would only be passively protected from new learning. They should therefore, during subsequent intervals of wakefulness, be subject to the same time-dependent decay as in experimental wake groups.

Inconsistent with this prediction, Ellenbogen et al. (2006a) showed that memories are less susceptible to interference after sleep than after wake, suggesting that memories are indeed stabilized during sleep. In addition, several studies indicate that effects of sleep-dependent memory consolidation basically stay the same when performance of wake and sleep groups is compared after prolonged intervals of 24 hours or longer (Ellenbogen et al., 2006a; Gais et al., 2006; Tucker et al., 2011). Such intervals include similar amounts of sleep, wake and, consequently, of new and potentially interfering learning experiences in both groups. Performance in the sleep groups is still found to be better compared to wake groups, the only remaining difference between groups now being the sequence of sleep and wake intervals (with sleep groups obtaining sleep with only little delay after learning, while wake groups do not go to sleep until after about 12 hours). In additon, polysomnographical studies have established links between specific sleep stages or sleep parameters and effects of sleep-dependent memory consolidation (e.g., Fogel & Smith, 2011; Marshall, Kirov, Brade, Mölle, & Born, 2011; Poe, Walsh, & Bjorness, 2010), which is regarded as an indication of active contributions of sleep to processes of memory consolidation (for a detailed discussion of the possible passive or active role of sleep, see Ellenbogen, Payne, & Stickgold, 2006b).

1.2 Forms of Sleep-Dependent Memory Consolidation

As research on the topic increases, more and more different types of sleep-dependent memory consolidation are distinguished. On the one hand, it is differentiated between forms of synaptic consolidation and forms of system consolidation (Dudai, 2004; Born & Wilhelm, in press). On the other hand, it has been found that sleep-dependent memory consolidation may manifest itself in a row of behavioral effects (Diekelmann et al., 2009; Payne, 2011).

Synaptic consolidation is assumed to occur in the short term, within minutes or hours after new information is encoded. This form of consolidation is supposed to be achieved through molecular remodeling processes on the synaptic level, which rapidly stabilize internal memory representations by changing synaptic connections. Synaptic consolidation may occur both during sleep and wakefulness. System consolidation, in contrast, is assumed to occur in the long term, i.e., within days to years. The standard model assumes that, in the course of this form of consolidation, memory traces are reorganized, thereby becoming less dependent or even independent of the hippocampus, and more dependent on neocortical areas. Recently, experimental evidence for a link between sleep and system consolidation was brought forward: Comparing brain activity of a sleep condition and a sleep deprivation condition, the hippocampus was found to be more active after two days in the sleep condition; moreover, this hippocampal activity was functionally related to activity in the medial prefrontal cortex in the sleep condition only. However, after a delay of six months, the correct recall of memories in the sleep condition was preferentially related to activity in the medial prefrontal cortex and less to activity of the hippocampus, while a different pattern of activity emerged in the sleep deprivation condition (Gais et al., 2007). Taken together, sleep is assumed to induce a shift in memory representations' long-term location within the brain, from one brain system to another (for alternative models, see Moscovitch, Nadel, Winocur, Gilboa, & Rosenbaum, 2006; Nadel & Moscovitch, 1997; Redondo & Morris, 2011).

Moreover, increasing evidence indicates that sleep's benefit for memory performance may itself be versatile and observable in varying patterns of behavioral results. Until now, it has been found that sleep can stabilize, enhance, or transform memories. The stabilization of memory contents may manifest itself in less time-dependent forgetting (e.g., Talamini et al., 2008) as well as in less susceptibility to interference (Ellenbogen et al., 2006a) after sleep compared to wake. Sleep-dependent enhancement has mostly been reported for procedural memory (e.g., Debarnot et al., 2009; Gais et al., 2008), but also for declarative memory (e.g., Tucker & Fishbein, 2008; Wilhelm et al., 2011): In comparison to baseline performance before sleep, performance after sleep has been found to be significantly improved. The sleep-associated transformation of memory content is assumed to rely on unbinding and restructuring of memory traces during sleep. By combining recent with remote memories, integrating associations, enlarging semantic networks, and through processes of schematization, qualitative changes in performance are supposed to be induced (Payne, 2011). For instance, it could be shown that sleep changes associative memory (Stickgold, Scott, Rittenhouse, & Hobson, 1999), improves creative problem solving (Cai, Mednick, Harrison, Kanady, & Mednick, 2009), facilitates insight into hidden rules (Wagner, Gais, Haider, Verleger, & Born, 2004), and accelerates the formation of explicit knowledge (Drosopoulos, Harrer, & Born, 2011). For further thoughts on how sleep might mediate transformation processes, see Tse et al. (2011), or Lewis & Durrant (2011).

Nevertheless, it remains unclear what determines which of the three forms of sleep-dependent memory consolidation described above ultimately takes effect. Currently, research starts to identify modulating factors of sleep-dependent memory consolidation, as for instance the strength (Drosopoulos, Schulze, Fischer, & Born, 2007), emotionality (Payne & Kensinger, 2010), or future relevance of memory contents (Wilhelm et al., 2011). However, it remains to be investigated whether the different forms of sleep-dependent memory consolidation are affected differently by these and similar factors, or whether they might even be triggered by specific events and circumstances. Alternatively, synaptic and system consolidation might be expressed differently on a behavioral level, or, more elementary and therefore more likely, different tasks and methodic modifications might alter the effects.

Although the first part of this thesis is dedicated to investigating effects of sleep-associated memory consolidation, the following chapter will, in a first step, deal with another prominent theory in memory research, namely with interference theory. Previous findings and their implications will be reviewed and discussed, both for basic research on interference effects and for research on the role of sleep in this respect. In a second step, the results of two fresh experiments will be reported that build upon the prior work. In particular, these experiments will deal with the question whether sleep-associated consolidation counteracts experimentally induced forgetting due to interference, and whether sleep-associated consolidation is modulated by memory strength.

Chapter 2

Interference

2.1 Background

Interference has often been assumed to constitute one, if not *the*, major cause of forgetting (for a review of interference theory, see Wixted, 2004). The general idea is that target information is recalled less well on a postponed test if additional learning has taken place. Proactive interference, on the one hand, refers to the finding that memory performance for target information is reduced by previous encoding of other information in comparison to a control condition, in which only the target information was encoded. The term retroactive interference, on the other hand, is used when performance for target information is negatively affected by the subsequent encoding of other information in comparison to a control condition, in which only the original information was encoded. Interference has been examined in list learning and paired-associate learning; both of these experimental approaches and their major outcomes shall be described in the following passages.

When applying list learning, subjects are usually asked to encode either only one target list or several additional lists of unrelated, single items. Recall performance for the target list is then analyzed with regard to the number of lists studied before or after this list and is usually found to be impaired by previous and subsequent learning (e.g., Underwood, 1957). McGeoch and McDonald (1931), for instance, let their subjects learn a list containing adjectives and manipulated what kind of material was to be studied afterwards. While some of the subjects encoded additional lists that differed profoundly from the original list (e.g., lists containing numbers instead of adjectives), other subjects encoded material more similar to the original list (e.g., lists comprising synonyms for the previously studied adjectives). In comparison to a single-list control condition, recall performance for the first list of adjectives was reduced in all interference conditions, but this reduction was the more pronounced the more similar the interfering material was to the original list. It was concluded that forgetting due to interference increases if the additional information is highly similar to the target information.

Tulving and Psotka (1971) applied list learning to investigate whether retroactive interference affected the accessibility of information or, alternatively, had an impact on its general representation and availability in memory. Participants studied either a single list or up to six lists. The lists were categorized, i.e., the items of each list belonged to four distinct semantic categories; the category names were not provided during study. On a free recall test, the expected memory impairment due to interfering learning was present; however, a closer look at the data revealed that interference affected mainly the number of categories recalled and less the number of items recalled within each category. Indeed, when participants were asked to take another test and were provided with category cues, the impairment was no longer evident. The authors concluded that, in this experiment, interference affected the accessibility of higher order units, whereas the items within these units remained accessible. On a more general level, it was concluded that retroactive interference did not cause forgetting because it directly affected memories, but because these memories could not be accessed due to insufficient retrieval cues.

In contrast to list learning, a typical interference experiment applying paired-associate learning is conducted in the following manner: Subjects are presented a first list of several unrelated word pairs (e.g., door - cherry) and are asked to memorize them. After a specific learning criterion has been reached for this list over repeated study-test cycles, a second list is encoded by the subjects in the same way. Critically, the paired associates of the second list (e.g., door - glasses) possess the same first words as the pairs from the previously studied list. The shared first words are, in a subsequent test phase, used as cues (e.g., door - ?) to elicit recall of paired associates from both lists (i.e., cherry, glasses). This testing procedure, during which the cue word is presented and subjects are asked to name both target words that have previously been paired with the cue word, is called the modified modified free recall procedure (MMFR; Barnes & Underwood, 1959). The major finding emerging from studies on paired-associate learning is that additional and interfering learning of another list impairs memory performance in comparison to a control condition, in which subjects are required to study only one single list (for a detailed description and further variations of the paired-associate task, see Crowder, 1976).

During past decades, several different accounts have been discussed in order to explain how exactly interference causes forgetting. For retroactive interference, three prominent accounts can be distinguished. For instance, it was suggested that gradual associative unlearning of the first target word took place, when a second target word was newly linked to a specific cue during repeated study-test cycles (Melton & Irwin, 1940). However, the unlearning assumption is not well in line with the report by Tulving & Psotka (1971), showing that retroactive interference does not depend on an actual weakening of the original memory traces, but rather on their inaccessibility. Another account proposed that the effect might be due to stronger target words blocking recall of weaker target words at test, because items related to a common cue were assumed to compete for recall (McGeoch, 1942). Alternatively, a suppression of response set was suggested to explain retroactive interference; it was assumed that the later a response to a specific stimulus was learned, the more dominant it was and, thus, suppressed response sets acquired earlier (Postman & Stark, 1969; Postman, Stark, & Fraser, 1968).

For proactive interference, basically two contrasting proposals exist that were made to explain the effect. On the one hand, it was suggested that previous encoding might impair subsequent encoding, leading to weaker memory traces and worse recall. This idea was based on several reports of a release from proactive interference if the critical items were not similar to previous ones but differed, for instance, with regard to category affiliation (e.g., Keppel & Underwood, 1962; Wickens, 1970). However, the assumption that these reports mirrored encoding effects was called into question by Gardiner, Craik, and Birtwistle (1972) who could show that they were mainly a matter of retrieval cues. On the other hand, it was suggested that proactive interference might arise because, as more and more information is encoded, it becomes more and more difficult to distinguish the recently encoded bit of information from other bits encoded before. The temporal discrimination account (e.g., Baddeley, 1990; Wixted & Rohrer, 1993) proposed that search for a target item at recall can, under conditions of proactive interference, not be restricted to the last bit of information that was presented, but includes and considers all the previously encoded units of information as well. According to this account, forgetting arises because, within such an enlarged search set, specific information is harder to find. Note that the proposal is also compatible with the finding by Keppel and Underwood (1962), because affiliation of target items to a new semantic category may also enhance the discriminability of the respective items. However, a more detailed discussion of various accounts aimed at explaining either proactive or retroactive intereference may be found elsewhere (e.g., Anderson & Neely, 1996; Crowder, 1976; Wixted, 2004).

In the course of time, many studies were carried out that examined effects of interference. In the following, only a brief outline shall be given on results concerning immediate interference effects in paired-associate learning and the durability of the phenomena. Initially, on an immediate test without further delay, it was found that effects of retroactive interference were more pronounced than effects of proactive interference, when the standard anticipation-plus-study method was applied (Melton & von Lackum, 1941). This procedure usually encompasses the execution of several study-test cycles after initial learning of each list, until participants reach a specific learning criterion defined beforehand. However, it was shown that the immediate pattern of retroactive and proactive interference effects could be reversed by omitting the execution of such study-test cycles; if the acquisition phase consisted of pure study trials without any intermediate testing, proactive interference was initially more pronounced than retroactive interference (Tulving & Watkins, 1974).

When investigating recall performance across prolonged retention intervals, an overall complementary process to forgetting due to retroactive interference was found. Underwood (1948) asked his subjects to study two paired-associate lists and tested recall performance after either 5 hours or after 48 hours. Interestingly, performance for the first list was stable between the 5-hour and the 48-hour delay, while performance for the second list decreased significantly with delay. After 5 hours, performance for the second list had been superior to performance for the first list; but after 48 hours, performance was found to be equal for both lists, as memory for the first list remained stable over time.

Underwood (1948) attributed this finding of stable memory performance for the first list to a process of spontaneous recovery from retroactive interference, that occurred, in parallel to recovery from extinction in animal conditioning, over time. Since then, the finding has been replicated several times for retroactive interference (for a review, see Brown, 1976), while no such effect was reported for proactive interference. Moreover, according to Brown (1976), two forms of recovery from retroactive interference over time can be distinguished. Absolute recovery, on the one hand, is found when recall performance for a first list of paired associates significantly improves across a delay interval. Relative recovery, on the other hand, does not refer to an actual improvement in memory performance for the first list; instead, it refers to the finding that recall performance declines significantly less for the first of two lists than for a single-list condition. Nevertheless, the phenomenon of spontaneous recovery from retroactive interference is controversial. While some researchers argued that evidence for the effect was, if anything, merely inconclusive (e.g., Keppel, 1968), others stated that recovery from retroactive interference was a reliable phenomenon reported in plenty of studies (e.g., Wheeler, 1995). According to Roediger and Guynn (1996) “the phenomenon does seem to exist, but it remains poorly understood” (p.230).

Ekstrand (1967) was the first to investigate whether delay intervals filled with sleep or wake affected phenomena of interference differently. His subjects underwent the paired-associate task: They studied two lists of paired associates and, subsequently, stayed awake during the day or slept regularly during the night. After eight hours, it was found that sleep facilitated recall of target items from both lists; however, this facilitation was found to be larger for the first list than for the second list. Ekstrand (1967) concluded that sleep led to recovery from retroactive interference. To test the hypothesis that the effect might be related to specific sleep stages, Ekstrand, Sullivan, Parker and West (1971) conducted a replication study that included, beyond others, also a short delay condition. Intriguingly, after 20 minutes, the same effect of recovery from retroactive interference emerged (albeit insignificant). Because of this finding, Ekstrand et al. (1971) reasoned that the release from interference observed in the previous study (Ekstrand, 1967) must have had occurred already before subjects went to sleep, and therefore dropped the whole research topic.

In 2007, Drosopoulos et al. reconsidered the problem and reexamined the results reported by Ekstrand et al. (1971). Except for introducing several additional control groups, the experimental design stayed the same: Subjects studied two lists of paired associates and were either tested immediately, after 20 minutes, or went to sleep or stayed awake before taking the test. Again, better recall performance after sleep was found for the first list of paired associates, while no such effect was found for the second list. In contrast to the results obtained by Ekstrand et al. (1971), recall performance did not differ between groups that were immediately tested or tested after a short delay of 20 minutes; i.e., no recovery from interference was evident after 20 minutes. However, Drosopoulos et al. (2007) examined interference by comparing performance for the first and second list, and not by comparing performance of the first or second list to adequate single-list conditions (although data for such single-list conditions had been collected). They argued that recall of paired associates after no and after a short delay was lower for the first list in comparison to the second list because of retroactive interference, and that sleep recovered first-list associations, because only in the sleep group recall of first-list paired associates was not inferior to recall of second-list associations. In all other groups, differences between first-list and second-list recall were evident and comparable in magnitude.

This first experiment led Drosopoulos et al. (2007) to develop the hypothesis that sleep boosts associative strength, and that this could preferentially be the case for weaker associations (as was found for the presumably 'weaker' first-list items, but not for the 'stronger' second-list items). In a second experiment, this hypothesis was put to the test. Subjects studied two lists of paired associates, but these lists were independent of each other and did not share the first cue word. Both lists were studied consecutively, until either a rather intense learning criterion of 90 % correct or a rather weak learning criterion of 60 % correct was reached. After intervals filled with sleep or wakefulness, Drosopoulos et al. (2007) found a benefit of sleep for both lists - but only in the weak encoding condition, not in the intense encoding condition. All in all, the conclusion was drawn that sleep preferentially benefits memories of rather weak associative strength, no matter what the cause of the reduced strength is (e.g., retroactive interference, or shallow encoding).

Without a doubt, the results reported by Drosopoulos et al. (2007) and, partly, as well by Ekstrand (1967), are interesting. If sleep really "nullifies interference" (as stated by Drosopoulos et al., 2007, p.179), it would not only counteract normal time-dependent forgetting in the future by stabilizing memory contents, but would even thwart an experimental manipulation applied before sleep onset that has been shown to cause forgetting. However, this critical conclusion should be based on a different analysis than the one applied by Drosopoulos et al. (2007). More precisely, a comparison of double-list performance to single-list performance is necessary in order to be able to state whether or not influences of additional learning (i.e., of proactive or retroactive interference) impair memory performance after shorter or longer delays. As mentioned above, Drosopoulos et al. (2007) did collect data for such single-list conditions; however, the corresponding analysis was not reported. In addition, the authors assumed that retroactive interference had somehow 'weakened' first-list associations. However, this interpretation is at odds with classic interference literature (e.g., Tulving & Psotka, 1971), showing that interference does not so much affect the general strength of memories, but rather their accessibility.

Aside from the missing analysis and the unsteady interpretation, details of the experimental procedure applied by Drosopoulos et al. (2007) may have had an additional impact on the reported results. In particular, three points seem worth considering. First, in parallel to classic interference experiments, the anticipation-plus-study method was applied to ensure equivalent encoding in all experimental conditions. After initial encoding, subjects underwent several study-test cycles until they reached a specific learning criterion (of 90 % correct in the first experiment, and of 90 % or 60 % correct in the second experiment). However, recent evidence indicates that testing itself has a beneficial impact on long-term retention (for a review of literature on the so-called testing effect, see Roediger & Butler, 2011). As paired associates were repeatedly tested by Drosopoulos et al. (2007) before the delay intervals, this procedure may have influenced consolidation in additon to the actual manipulation of sleep vs. wake. Second, the item material used by Drosopoulos et al. (2007) was semantically related. Items from the same semantic categories were not used within pairs, but were nevertheless spread across the first and second list. As has been shown, partial retrieval from a semantic category may cause forgetting of the rest of the category (for a review on retrieval-induced forgetting, see Anderson, 2003; or Bäuml, 2008). Therefore, the possibility cannot be excluded that such retrieval-induced forgetting influenced memory performance in addition to retroactive interference. Third, there were only ten subjects in each of the experimental groups. This small sample size, criticized by the authors themselves, appears to be too small to draw general conclusions.

All in all, a remake of the study seems necessary that rules out possibly confounding factors and, by applying the appropriate analysis, investigates whether sleep really counteracts retroactive interference (and, if so, in what way). In the following section, the details and outcomes of such a study are to be described; in addition to retroactive interference, effects of proactive interference were investigated as well.

2.2 Experiment 1a: Sleep and Interference (Strong Encoding)

Although the results reported by Drosopoulos et al. (2007) are striking, recent literature on possible effects of retrieval suggests that the study-plus-anticipation method may have acted as a confounding variable that could have biased the results. Here, likewise applying the paired-associate task, strong encoding in both groups was ensured by repeated presentation of the pairs instead of by repeated study-test cycles. In addition to measuring baseline performance for both a single-list and a double-list condition after 20 minutes, sleep and wake groups' recall levels were assessed after twelve hours. By comparing memory performance of sleep and wake groups that initially studied two lists of paired associates to memory performance of sleep and wake groups that only had to encode one list before the delay interval, the influence of sleep on effects of both retroactive and proactive interference was examined. In addition, time-dependent forgetting of lists was assessed by comparing the short-delay control to the wake condition. Sleep-associated stabilization was analyzed by comparing sleep to wake conditions.

Method

Participants

120 subjects participated in the experiment, either for course credit or a small compensatory amount of money. The sample consisted of 26 male and 94 female subjects, mean age was 23.4 years (range 18-30 years). All participants completed a screening questionnaire and interview prior to selection (Ellenbogen et al., 2006a). This approach was chosen to ensure that no subject in the final sample suffered from any neurological, psychiatric, or sleep disorders, or was under the influence of drugs or medication affecting the central nervous system. All subjects spoke German as their native language, reported to have regular sleep-wake cycles, and were compliant with the instructions provided by the investigators. Subjects were randomly assigned to one of the experimental conditions and were tested either individually or in pairs of two. Between experimental conditions, no differences with regard to age, habitual sleep duration, subjective ratings of sleep quality, or a rough estimate of intelligence (as assessed by the connect-the-numbers test; Oswald & Roth, 1987) were evident (all $ps > .05$).

Material

Item material consisted of five separate lists of fifteen single items. Items were taken from different semantic categories out of the category norm provided by Van Overschelde, Rawson, and Dunlosky (2004) and translated into German. Hence, items were unrelated, both within and between lists. Two of the five item lists were randomly chosen; items from these two lists always served as cue words. The sequence of items within all lists was randomized; then, in order to create paired-associate lists, the three single-item lists were merged with the two cue lists to pairs of two or three, respectively. Thus, one out of the remaining three single-item lists was combined with one of the previously chosen cue lists and used as a single paired-associate list. The remaining two single-item lists were combined with the remaining list of cues and were used as double paired-associate lists. Paired associates were created by sequentially combining items from the respective lists; i.e., by combining the respective first items, the second items, and so forth. All possible pairings of cue and item lists were equally often used as single and double paired-associate lists. To control for possible time of day confounds, participants used the Stanford Sleepiness Scale to indicate how alert and activated they felt at the beginning of each session (Hoddes, Zarcone, Smythe, Phillips, & Dement, 1973).

Design

The experiment had a 3 x 2 mixed factorial design. The factor CONDITION (20-min control, 12-h wake, 12-h sleep) was manipulated between subjects. In the sleep and wake conditions, the factor INTERFERENCE (single list, double list) was manipulated between subjects as well; in the 20-min control condition, INTERFERENCE was manipulated within subjects. The experiment started at 9 a.m. or 9 p.m., respectively. Subjects either studied one list or two lists of paired-associates (only in the 20-min control condition, both interference conditions were absolved successively). Recall performance was tested either after a short delay of 20 minutes or after a long delay of 12 hours that included either diurnal wakefulness or nocturnal sleep. All participants in the 12-h wake condition reported to have stayed awake and not to have taken any naps during the day, whereas all participants in the 12-h sleep group reported to have slept regularly during the night (mean sleep duration: 7.5 hours; range 5-10 hours); none of the participants consumed alcohol between the two sessions.

Procedure

Study Phase. In double-list conditions, both paired-associate lists were presented on three consecutive study cycles to ensure robust encoding. Before presentation of the first list started, subjects were informed that they would have to memorize paired associates. Item pairs were then presented in random order and at a rate of 4 sec each centrally on a computer screen. When all 15 paired-associates of the first list had been presented, a new study cycle began and the same pairs of the same list were presented again in random order. With a third corresponding study cycle, encoding of the list was completed. A short distractor phase of about 10 minutes followed, during which subjects absolved a cognitive test (the d2 test of attention; Brickenkamp, 2002) and engaged in problem solving tasks. Afterwards, subjects encoded the second list of paired associates that all shared the first words with the previously encoded list of paired associates. Subjects were instructed to memorize the following word pairs in addition to the ones they had already studied. The encoding procedure for the second list was identical to that for the first list; item pairs were shown for 4 sec each, and the list was presented on three consecutive study cycles. After encoding of the second list, another distractor phase of about 10 minutes followed. During this distractor phase, subjects absolved another cognitive test (the connect-the-numbers test; Oswald & Roth, 1987) and again engaged in problem solving tasks. Afterwards, subjects from the short-delay control condition took the final memory test assessing recall for both lists of paired associates; subjects from the long-delay conditions left the laboratory and returned after 12 hours to take the same test.

In single-list conditions, subjects only studied one list of paired-associates. The general encoding procedure was the same as in double-list conditions; item pairs were presented for 4 sec each and on three consecutive study cycles. Critically, the amount of time spent in the study phase was held constant between single-list and double-list conditions. This was done by combining the two short distractor phases described above for the double-list conditions to one long distractor phase, and by increasing the time that subjects were allowed to engage in the respective problem solving tasks to account for the time it took the other subjects to encode the second list. Subjects from the short-delay control condition took the memory test assessing recall of the single list after both encoding and distractor phase had been absolved; subjects from the long-delay conditions left the laboratory and returned after 12 hours to take the same test. Note that, in order to reduce the already considerable sample size, half of the subjects from the single-list conditions first encoded the paired associates and subsequently engaged in the distractor tasks, thereby forming the adequate control condition to assess retroactive interference. The other half of the subjects first engaged in the distractor tasks and encoded the list of paired associates at the same point in time, when subjects from the double-list conditions encoded the second list - thereby forming the adequate control condition to assess proactive interference. It was speculated that recall performance for one single list should be about equal - no matter at what exact point of time during the study phase the list had been encoded.

Importantly, subjects in the short-delay condition absolved both single-list and double-list conditions; their sequence was balanced between subjects. As subjects engaged in the two previously described short distractor phases when absolving the double-list condition, a further and similarly long distractor was needed to fill the remaining time after and/or before encoding of the single list. Therefore, subjects in the short-delay condition additionally engaged in another cognitive test (Standard Progressive Matrices; Raven, 1999) after and/or before encoding of the single list; time was held constant between single-list and double-list conditions (see Figure 2.1 for an illustration of the experimental procedure and conditions).

Test Phase. In the double-list conditions, testing took place in the form of a MMFR test (Barnes & Underwood, 1959). Subjects were confronted with a cue word that appeared centrally on a computer screen and were asked to write down both target words they had previously studied in relation to the respective cue. In addition, subjects were asked to indicate, which of the two target words had been studied first and which second. After 15 seconds, the next cue word appeared on the screen. In the single list conditions, testing took place in the same way. However, as only one target item had to be remembered, presentation time of each cue word was reduced to 10 seconds. After the final test phase, all subjects were debriefed and thanked for their participation. Note that only items recalled in connection to the right cue were counted as correctly recalled for the later analysis of memory performance.

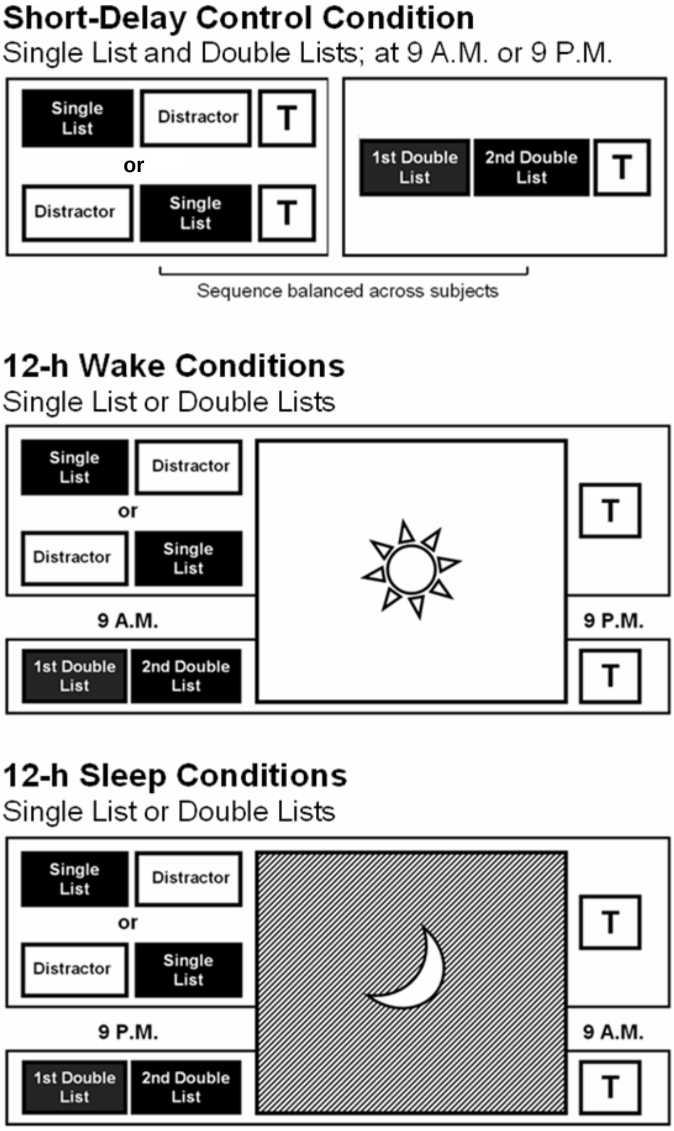

Figure 2.1: Procedure and conditions employed in *Experiment 1a*. In the short-delay control condition, subjects absolved two blocks, successively encoding and recalling the single and the double lists (T = Test of the list(s) encoded in the respective block). Sequence of blocks was balanced across subjects; the experiment was conducted at either 9 a.m. or 9 p.m. Subjects in the 12-h wake conditions started the experiment at 9 a.m., whereas subjects in the 12-h sleep conditions started it at 9 p.m. In both long-delay conditions, subjects encoded either the single list or the double lists, and were tested after 12 hours that were either filled with diurnal wakefulness or nocturnal sleep.

Results

Prerequisites

Ratings of Alertness. Ratings on the Stanford Sleepiness Scale (Hoddes et al., 1973) did not differ between morning and evening groups in the first session, $F(1, 118) = 3.40$, $MSE = 0.71$, $p > .05$. There was also no difference between morning and evening groups concerning their ratings of alertness in the second session, i.e., after the 12-h delay, $F(1, 94) < 1.0$.

Morning and Evening Short-Delay Control Conditions. To further confine sample size, half of the subjects from the 20-min control condition were tested in the morning, while the other half was tested in the evening. Indeed, the two groups did not differ with regard to performance for the single list, $U(23) = 72.0$, $Z < 0.001$, $p > .95$, the first double list, $U(23) = 67.5$, $Z = 0.26$, $p > .75$, or the second double list, $U(23) = 64.0$, $Z = 0.47$, $p > .60$. Thus, the morning and evening short-delay conditions will be combined to one general short-delay control condition for all further analyses.

Sequence of Lists in the Short-Delay Control Condition. Moreover, half of the subjects from the short-delay control condition started the experimental session with the single-list condition, while the other half began with the double-list condition. However, sequence did not influence memory performance for the single list, $U(23) = 68.0$, $Z = 0.25$, $p > .80$, the first double list, $U(23) = 69.5$, $Z = 0.15$, $p > .85$, or the second double list, $U(23) = 68.5$, $Z = 0.20$, $p > .80$. Hence, for all further analyses, data from the short-delay control condition will be merged without further considering sequence of lists.

Single Lists in the Short-Delay Control Condition. Additionally, half of the subjects from the short-delay condition encoded the single list to control for retroactive interference, while the other half encoded it to control for proactive interference. As memory performance for the single list did not differ with respect to this procedural difference, $U(23) = 71.0$, $Z = 0.06$, $p \geq .95$, data are combined for all further comparisons.

Single Lists in the Sleep and Wake Conditions. As in the short-delay control condition, half of the subjects from single-list wake and sleep groups encoded the list to control for retroactive interference, while the other half encoded it to control for proactive interference. Again, memory performance for the single list did not differ between these two approaches in either group, $Us(23) \geq 47.0$, $Zs \leq 1.45$, $ps \geq .15$. Therefore, data are combined to form one single-list conditon in each of the groups.

Effects of Interference

Retroactive Interference. For the short-delay condition, a paired-samples t-test confirmed that memory performance differed significantly for the single list and the first double list (91.7 % correct vs. 85.6 % correct), $t(23) = 3.60$, $SEM = 1.70$, $p = .01$, which indicates that retroactive interference was induced after a short 20-min delay.

To assess retroactive interference in the 12-h delay conditions, a 2 x 2 ANOVA with the factors of INTERFERENCE (single list, first double list) and CONDITION (12-h wake, 12-h sleep) was calculated. No significant main effect of INTERFERENCE was found, $F(1,92) = 1.84$, $MSE = 383.26$, $p > .15$, suggesting that retroactive interference did not affect memory performance across both conditions. A significant main effect of CONDITION emerged, $F(1,92) = 4.50$, $MSE = 383.26$, $p < .05$, reflecting superior memory performance in the 12-h sleep group. Moreover, a significant interaction of the two factors was found, $F(1,92) = 5.42$, $MSE = 383.26$, $p < .03$, which suggests that CONDITION affected memory performance for the two lists differently. While sleep in comparison to wake did not affect single-list performance (75.3 % correct vs. 76.1 % correct), $t(46) = 0.15$, $p > .85$, it led to better memory performance for the first double list (79.2 % correct vs. 61.4 % correct), $t(46) = 3.04$, $p < .01$. Consistently, retroactive interference was evident in the 12-h wake condition, $t(46) = 2.36$, $p < .03$, but not in the 12-h sleep condition, $t(46) = 0.78$, $p > .40$ (see Figure 2.2 for a plot of the results).

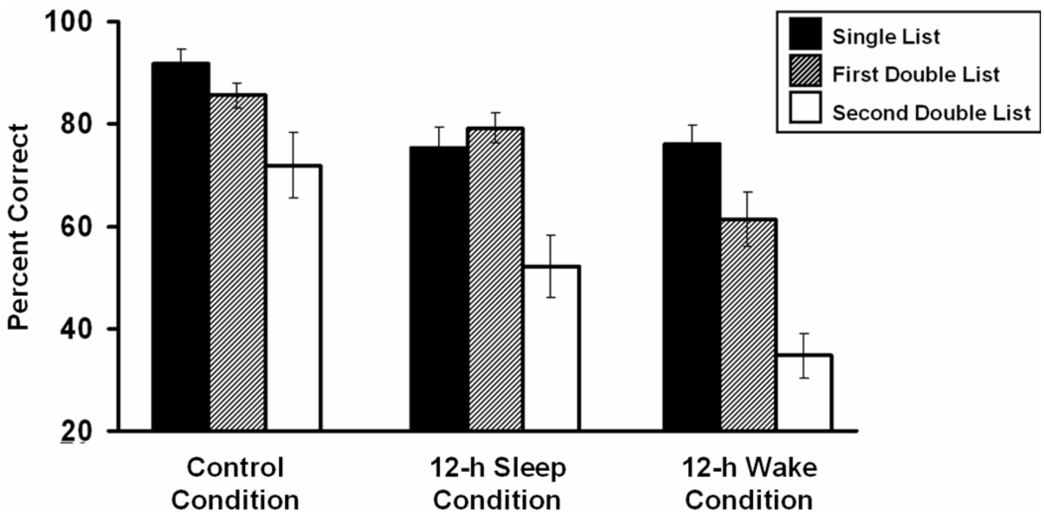

Figure 2.2: Results of *Experiment 1a.* Mean recall performance for single list and double lists is shown separately for control, sleep and wake conditions (error bars represent standard errors of the mean).

Proactive Interference. A paired-samples t-test confirmed that memory performance differed significantly for the single list and and the second double list in the short-delay condition (91.7 % correct vs. 71.9 % correct), $t(23) = 4.09$, $SEM = 4.82$, $p = .001$; proactive interference was induced after a short 20-min delay.

To assess proactive interference in the 12-h delay conditions, a 2 x 2 ANOVA with the factors of INTERFERENCE (single list, second double list) and CONDITION (12-h wake, 12-h sleep) was run. A significant main effect of INTERFERENCE was found, $F(1, 92) = 51.59$, $MSE = 485.10$, $p = .001$, indicating that proactive interference affected memory performance across conditions. The ANOVA also revealed a marginally significant main effect of CONDITION, $F(1, 92) = 3.38$, $MSE = 485.10$, $p = .07$, reflecting the tendency for better memory performance in the sleep group. Moreover, a significant interaction of the two factors emerged, $F(1, 92) = 4.09$, $MSE = 485.10$, $p = .05$, which suggests that CONDITION again affected memory performance for the two lists differently. As described above, sleep in comparison to wake did not affect single-list performance (75.3 % correct vs. 76.1 % correct),

$t(46) = 0.15$, $p > .85$; however, it produced better memory performance for the second double list (52.1 % correct vs. 34.7 % correct), $t(46) = 2.43$, $p = .02$. In contrast to retroactive interference, proactive interference was evident in both 12-h delay conditions, $ts(46) \geq 3.20$, $ps < .01$.

Time-Dependent Forgetting and Sleep-Associated Stabilization

To investigate time-dependent forgetting, three one-way ANOVAS with the factor of CONDITION (20-min control, 12-h wake) were calculated; separately for the single list, the first double list, and the second double list. For all three lists, significant differences in memory performance between the two conditions emerged, $Fs(1, 46) > 11.25$, $MSEs < 670.50$, $ps < .01$. As memory performance was higher in the 20-min control condition, time-dependent forgetting across 12 hours of wakefulness was evident for all three lists.

To investigate sleep-associated consolidation, performance for the single list and the two double lists was analyzed separately. For the single list, a one-way ANOVA with the factor of CONDITION (12-h sleep, 12-h wake) revealed no significant difference in memory performance between sleep and wake, $F(1, 46) < 1.0$. As in the wake condition, memory performance for the single list got worse across the 12-h sleep delay, $F(1, 46) = 11.29$, $MSE = 285.47$, $p < .01$. For the double lists, a 2 x 2 ANOVA with the factors of LIST (first double list, second double list) and CONDITION (12-h sleep, 12-h wake) was calculated. Significant main effects of both LIST, $F(1, 46) = 61.21$, $MSE = 283.18$, $p < .001$, and CONDITION were found, $F(1, 46) = 10.00$, $MSE = 740.92$, $p < .01$, reflecting better memory performance for the first than for the second double list and for the sleep than for the wake condition. However, no significant interaction of the factors was observed, $F(1, 46) < 1.0$, indicating that sleep stabilized memory performance for the two lists equally. Time-dependent forgetting across the 12-h sleep delay was only evident for the second double list, $F(1, 46) = 5.05$, $MSE = 936.57$, $p < .05$, but did not reach significance for the first double list, $F(1, 46) = 2.70$, $MSE = 181.60$, $p > .10$.

Intrusion Errors

Rate of intrusion errors was analyzed separately. For single-list conditions, this rate included incorrectly ‘recalled’ items that were completely new words, paired with the wrong cue, or items that had actually been presented as cues. For double-list conditions, it additionally included falsely ‘recalled’ items that had been presented in the respective other list. Because intrusion rates were different for single and double lists anyway, no corresponding comparisons were calculated. Possible effects of condition on overall intrusion errors were investigated separately for single list and double lists.

For the single list, a one-way ANOVA revealed that intrusion rate did not differ between sleep, wake, and control conditions, $F(2, 69) < 1.0$. Furthermore, the overall intrusion rate was generally low ($M = 0.46$, $SD = 0.84$). For the double lists, a 2 x 3 ANOVA with the factors of LIST (first double list, second double list) and CONDITION (20-min control, 12-h wake, 12-h sleep) was run. A significant main effect of LIST emerged, $F(1, 69) = 3.86$, $MSE = 1.04$, $p = .05$, reflecting a lower intrusion rate for the second than for the first double list. No significant main effect of CONDITION was found, $F(1, 69) < 1.0$, and also no significant interaction of the two factors, $F(2, 69) < 1.0$, indicating that intrusion rate did not differ between conditions. Again, the overall intrusion rates were rather low (first double list: $M = 1.18$, $SD = 1.30$; second double list: $M = 0.85$, $SD = 1.34$).

Discussion

The results replicate previous findings by Ekstrand (1967) and Drosopoulos et al. (2007). Retroactive interference was present after a short delay and persisted after 12 hours filled with wakefulness, but was abrogated if sleep followed encoding. The effect seems to depend on selective sleep-associated stabilization of the first double list. While sleep did not benefit the retention of the single list, it provided a benefit for the double list, leading to relative recovery from retroactive interference after sleep. Indeed, no time-dependent forgetting was found after sleep for the first double list, whereas equal time-dependent forgetting was evident for the single list both after sleep and wake. Thus, the results suggest that the previously reported elimination of retroactive interference after sleep does not depend on confounded variables, as for instance prior retrieval or semantically related material. Instead, and by providing the crucial comparison of single-list to double-list performance, the data further underpin the conclusion that sleep may not only counteract normal (time-dependent) forgetting, but also experimentally induced forms of forgetting.

Extending previous findings, effects of sleep on proactive interference were additionally investigated. In line with results provided by Ekstrand (1967), memory performance for the second double list was better after sleep compared to wake. Although significant time-dependent forgetting of this list was present both after sleep and wake, the comparison of performance after sleep and wake indicates that sleep stabilized memory performance for the first and second double list to an equal degree. Consistent with classic interference studies (for a review, see Brown, 1976), no recovery from proactive interference was observed: Forgetting due to proactive interference was not only present after a short delay, but was also evident after both sleep and wake, irrespective of the better memory performance in the double-list sleep group compared to the double-list wake group. Therefore, sleep did counteract proactive interference as well, but failed to abrogate it because the numerical distance to the single list was greater for the second than for the first double list. An analysis of intrusions additionally straigthened out that none of the observed differences between conditions were due to elevated or reduced error rates.

In parallel to the study by Drosopoulos et al. (2007), no benefit of sleep emerged for the single list. Drosopoulos et al. (2007) argued that this finding might be due to sleep not providing any further benefits for memory contents that are already strongly positioned after encoding. As the learning criterion in the study by Drosopoulos et al. (2007) was set to 90 % correct, the single list was probably very robustly encoded and, thus, memory performance for this list did not additionally profit from sleep. In the double list conditions, however, retroactive interference reduced recall performance for the respective memories, which could explain why they, in contrast, did show effects of sleep-associated stabilization. This line of argumentation can be applied to the present results as well: As a consequence of three consecutive study cycles for all paired-associate lists, memory performance for the single list was close to ceiling in the short-delay condition. Although time-dependent forgetting was observed in both 12-h delay groups and, thus, a ceiling effect seems rather unlikely, sleep did not stabilize memory performance for the single list in comparison to wake. In contrast, a benefit of sleep compared to wake was evident for both double lists in the interference conditions, suggesting that competition between memories and consequently reduced performance could somehow have triggered sleep-associated stabilization. The data presented here point in this direction, but, in parallel to previously reported data, do not provide any more information on how such triggering might actually work and what exact factor could be responsible.

Drosopoulos et al. (2007) argued that the observed pattern of results was basically an effect of reduced memory strength caused by retroactive interference. In a second experiment, they indeed provided evidence for the hypothesis that memory strength may affect sleep-associated memory consolidation. Subjects studied two non-interfering lists of paired associates to a criterion of either 90 % correct or 60 % correct. A benefit of sleep emerged for both lists - but only in the weaker encoding condition. Drosopoulos et al. (2007) concluded that sleep preferentially profits memories of weaker associative strength, irrespective of the exact cause of the reduced strength (according to their reasoning, such a reduction in strength could either be due to retroactive interference or weak encoding). However, the finding that memory strength (manipulated by changed encoding conditons) can influence sleep-associated stabilization might also be important when trying to figure out whether or not sleep abrogates effects of retroactive interference only under specific conditions (e.g., under conditions of strong encoding) or in a more general way (e.g., also under conditions of weak encoding). Following the results provided by Drosopoulos et al. (2007), one would expect to find a benefit of sleep for single-list performance, too, if items were weakly encoded. If, however, under weak encoding conditions the benefit of sleep remained the same for the double lists, retroactive interference might persist after sleep - because performance for none of the lists would be at ceiling and all lists should, therefore, profit from subsequent sleep.

In the next section, a follow-up experiment will be reported that strongly resembles Experiment 1a. To investigate whether the abrogation of retroactive interference after sleep depends on specific study conditions, memory strength was reduced at encoding in all experimental sleep and wake groups.

2.3 Experiment 1b: Sleep and Interference (Weak Encoding)

Results reported in the previous section replicate findings by Drosopoulos et al. (2007) and provide further evidence for the claim, that sleep can eliminate experimentally induced forgetting by supporting relative recovery from retroactive interference. However, such recovery from retroactive interference due to sleep has until now only been reported if lists were encoded robustly. Therefore, it remains unclear whether the finding is connected to specific encoding conditions or whether sleep counteracts retroactive interference in general. Here, the same experimental procedure as in Experiment 1a was used, but encoding strength was reduced by omitting the repeated presentation of paired-associate lists during study. Again comparing memory performance of sleep and wake groups that initially studied two double lists of paired associates to memory performance of sleep and wake groups that encoded a single list before the delay interval, the influence of sleep on interference effects was examined under weak encoding conditions.

Method

Participants

96 subjects participated in the experiment, either for course credit or payment. The sample consisted of 25 male and 71 female subjects, mean age was 22.5 years (range 19-35 years). All participants completed a screening questionnaire prior to selection (Ellenbogen et al., 2006a) to ensure that no subject in the final sample suffered from any neurological, psychiatric, or sleep disorders, or was under the influence of drugs or medication affecting the central nervous system. All subjects spoke German as their native language, reported to have regular sleep-wake cycles, and were compliant with the instructions provided by the investigators. Subjects were randomly assigned to one of the experimental conditions and were tested either individually or in pairs. Between the wake and sleep conditions, no differences with regard to age, subjective ratings of sleep quality, or a rough estimate of intelligence (as assessed by the connect-the-numbers test; Oswald & Roth, 1987) were evident (all $ps > .50$).

Material

Item material was the same as in Experiment 1a. It consisted of five separate lists of fifteen single items that were combined to pairs of two or three lists, respectively. The same paired associates as in Experiment 1a were used. Again, subjects filled out the Stanford Sleepiness Scale to indicate how alert and activated they felt at the beginning of each session (Hoddes et al., 1973).

Design

The experiment had a 2 x 2 between-subjects design. The factors CONDITION (12-h wake, 12-h sleep) and INTERFERENCE (single list, double list) were both manipulated between subjects. The experiment started at 9 a.m. or 9 p.m., respectively. Subjects either studied one list or two lists of paired associates.

Recall performance was tested after a delay of 12 hours, that was filled with either sleep or wakefulness. All participants in the 12-h wake condition reported to have stayed awake and not to have taken any naps during the day, whereas all participants in the 12-h sleep group reported to have slept regularly during the night (mean sleep duration: 7.7 hours; range 5-11 hours); none of the participants consumed alcohol between the two sessions.

Procedure

The general experimental procedure was the same as in Experiment 1a. However, in the study phase, lists were not repeated; paired associates were only presented once for 4 sec each and in random order to cause comparatively weak encoding. Subjects engaged in the same distractor tasks as in Experiment 1a; again, time spent in the study phase was held constant between single-list and double-list conditions. In parallel to Experiment 1a, half of the subjects from the single-list conditions first encoded the paired associates and subsequently engaged in the distractor tasks, thereby forming the adequate control condition to assess retroactive interference. The other half of the subjects first engaged in the distractor tasks and encoded the list of paired associates at the same point in time, when subjects from the double-list conditions encoded the second list - thereby forming the adequate control condition to assess proactive interference.

After the study phase, all subjects left the laboratory and returned after 12 hours to complete the experiment. The test phase was conducted in the same manner as in Experiment 1a: Subjects were confronted with a cue word that appeared centrally on a computer screen and were asked to write down the one or, respectively, two target words they had previously studied in relation to the presented cue. Again, subjects had 10 or, respectively, 15 sec per trial to recall the paired associates. After the final test phase, all subjects were debriefed and thanked for their participation.

Results

Prerequisites

Ratings of Alertness. Ratings on the Stanford Sleepiness Scale (Hoddes et al., 1973) did not differ between the morning and evening groups in the first session, $F(1, 94) < 1.0$, or in the second session, i.e., after the 12-h delay, $F(1, 94) = 3.33$, $MSE = 0.80$, $p > .07$.

Single Lists. Half of the subjects from single-list wake and sleep groups encoded the list as a control for retroactive interference, while the other half encoded it as a control for proactive interference. As in Experiment 1a, memory performance for the single list did not differ between these two approaches in either group, $Us(23) \geq 64.5$, $Zs < .45$, $ps > .65$. Therefore, data will be combined to form one single-list conditon in each of the two groups.

Effects of Interference

Retroactive Interference. To assess retroactive interference in the 12-h delay conditions, a 2 x 2 ANOVA with the factors of INTERFERENCE (single list, first double list) and CONDITION (12-h wake, 12-h sleep) was calculated. A significant main effect of CONDITION was found, $F(1, 92) = 34.22$, $MSE = 350.70$, $p < .001$, reflecting superior memory performance in the 12-h sleep group. However, the ANOVA revealed no significant main effect of INTERFERENCE, $F(1, 92) < 1.0$, but a significant interaction of the two factors, $F(1, 92) = 8.24$, $MSE = 350.70$, $p < .01$, which suggests that retroactive interference varied in the two conditions. In the wake condition, memory performance for the single list was better than for the first double list (28.9 % correct vs. 19.4 % correct), $t(46) = 2.00$, $p = .05$; i.e., significant retroactive interference was found after 12 hours of wakefulness. In contrast, recall of the first double list was better than recall of the single list in the sleep condition (51.9 % correct vs. 39.4 % correct), $t(46) = 2.20$, $p < .04$; i.e., retroactive interference was abrograted after 12 hours of sleep. Consistently, a benefit of sleep in comparison to wake was evident for recall of the single list, $t(46) = 2.04$, $p = .05$, but was more pronounced for recall of the first double list, $t(46) = 6.37$, $p < .001$.

Proactive Interference. To assess proactive interference in the 12-h delay conditions, a 2 x 2 ANOVA with the factors of INTERFERENCE (single list, second double list) and CONDITION (12-h wake, 12-h sleep) was run. A significant main effect of INTERFERENCE was found, $F(1, 92) = 61.78$, $MSE = 256.46$, $p < .001$. This indicates that proactive interference affected memory performance across both conditions. The ANOVA also revealed a significant main effect of CONDITION, $F(1, 92) = 13.66$, $MSE = 256.46$, $p < .001$, reflecting better memory performance in the sleep group. However, no significant interaction of the two factors emerged, $F(1, 92) < 1.0$, which suggests that INTERFERENCE did not affect memory performance differently in the two conditions. Indeed, significant proactive interference was evident in both 12-h delay conditions, $ts(46) \geq 4.90$, $ps < .001$. As described above, sleep in comparison to wake did benefit single-list performance (39.4 % correct vs. 28.9 % correct), $t(46) = 2.04$, $p = .05$, and did also benefit memory performance for the second double list (15.3 % correct vs. 2.5 % correct), $t(46) = 3.75$, $p = .001$ (see also Figure 2.3 for a plot of the results).

Sleep-Associated Stabilization

To examine sleep-associated memory stabilization, a 2 x 2 ANOVA with the factors of LIST (first double list, second double list) and CONDITION (12-h sleep, 12-h wake) was calculated. Significant main effects of both LIST, $F(1, 46) = 122.07$, $MSE = 145.70$, $p < .001$, and CONDITION were found, $F(1, 46) = 39.60$, $MSE = 322.18$, $p < .001$, reflecting better memory performance for the first than for the second double list and for the sleep than for the wake condition. A significant interaction of the factors was observed as well, $F(1, 46) = 17.40$, $MSE = 145.70$, $p < .001$, indicating that sleep-associated mnemonic benefits were greater for the first than for the second double list.

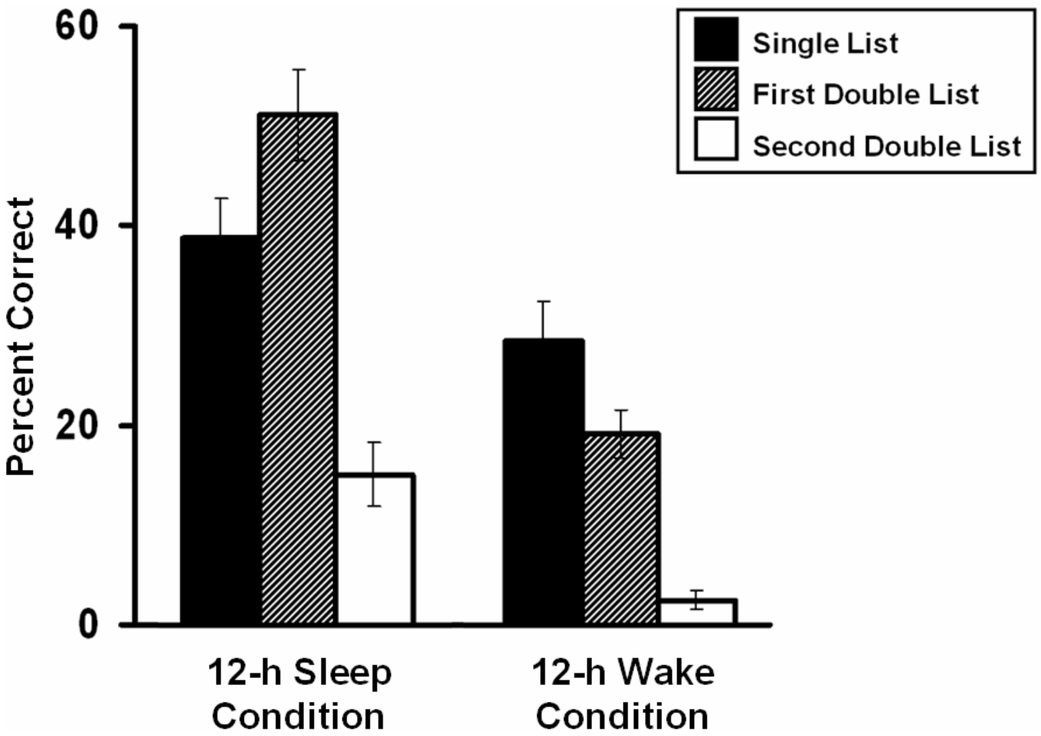

Figure 2.3: Results of *Experiment 1b*. Mean recall performance for the single list and the double lists is plotted separately for sleep and wake conditions (error bars represent standard errors of the mean).

Intrusion Errors

In parallel to Experiment 1a possible effects of condition on overall intrusion errors were investigated separately for single list and double lists. For the single list, a one-way ANOVA revealed that intrusion rate did not differ between sleep and wake, $F(1,46) = 2.54$, $MSE = 1.61$, $p > .10$. The overall intrusion rate was still rather low ($M = 1.04$, $SD = 1.29$).

For the double lists, a 2 x 3 ANOVA with the factors of LIST (first double list, second double list) and CONDITION (12-h wake, 12-h sleep) was calculated. A significant main effect of LIST emerged, $F(1,46) = 18.95$, $MSE = 1.72$, $p < .001$, reflecting less intrusions for the second than for the first double list. No significant main effect of CONDITION was found, $F(1,46) < 1.0$, but a significant interaction of the two factors, $F(1, 46) = 4.09$, $MSE = 1.72$, $p = .05$, indicating that CONDITION affected intrusion rates differently for the two lists. Indeed, there was no difference between conditions for the first double list, $F(1, 46) < 1.0$, but for the second double list, $F(1, 46) = 4.34$, $MSE = 1.23$, $p < .05$. For this list, the intrusion rate was higher in the sleep compared to the wake condition ($M = 1.13$ vs. $M = 0.46$). Again, the overall intrusion rates were rather low (first double list: $M = 1.96$, $SD = 1.97$; second double list: $M = 0.79$, $SD = 1.15$).

Discussion

The results basically replicate the general pattern that was observed in Experiment 1a, but this time under conditions of weak encoding: Retroactive interference was present after 12 hours of wakefulness, but was abrogated after 12 hours of sleep. Proactive interference was again present in both wake and sleep conditions. As expected, under weak encoding conditions memory performance for the single list was found to be subject to sleep-associated consolidation, while it had been rendered unaffected by sleep under strong encoding conditions in Experiment 1a. Here, sleep benefited memory performance for the single list and the second double list equally, while it entailed a more pronounced benefit for recall of the first double list in comparison to recall of the single list.

Therefore, Experiment 1b shows that the finding of sleep counteracting forgetting due to retroactive interference does not hinge on strong encoding conditions (or on no sleep-associated benefits for strongly encoded single lists), but more general on the fact that sleep benefits first double lists more under both strong and weak encoding conditions. In Experiment 1a, however, sleep did also benefit memory performance for the second double list more than memory performance for the single list. Here, the benefits of sleep were about the same for single list and second double list, and clearly more pronounced for the first double list. This finding, though, is probably due to a floor effect concerning memory performance for the second double list in the 12-h wake condition, which differed barely from zero and couldn't have dropped any further. Similarly, the lower rate of intrusion errors for the second list in the wake compared to the sleep condition was very likely tied to this floor effect as well; rate of intrusion errors thus cannot explain the observed differences between sleep and wake.

To be able to evaluate the extent of sleep-associated stabilization for first and second double lists under this condition of weak encoding, comparisons to a short-delay control condition (as in Experiment 1a) would be more appropriate. However, the corresponding data are still to be collected. Such a control condition would also allow to investigate statistically, whether the finding of no retroactive interference in the sleep conditions again arose because of a relative recovery from retroactive interference (e.g., Brown, 1976). Anyway, as memory performance for the first double list was even enhanced above memory performance for the single list in the sleep conditions, and because the regular pattern of results one might expect to observe in short-delay conditions would be better memory performance for a single list than for both double lists, it seems very likely that sleep again lead to a relative recovery from retroactive interference.

2.4 Conclusions

In this chapter, the data of two experiments were reported that investigated the influence of sleep on effects of interference in paired-associate learning. The results clarify an unresolved issue, but also generate further questions. They illustrate that sleep seems to counteract effects of retroactive interference in general, and that this role of sleep does not depend on strong encoding conditions as one could possibly have assumed based on previous data sets. Sleep very likely provokes a relative recovery from retroactive interference by entailing more distinct benefits for the first of two lists than for a single list, thereby counteracting not only time-dependent forgetting but also an experimentally induced form of forgetting. In line with previous research (for a review, see Brown, 1976), no such recovery was observed for proactive interference; nevertheless, sleep counteracted proactive interference as well, but could not close the greater numerical distance to the single list.

Thus, the present data are both replication and extension to a precedent study by Drosopoulos and colleagues (2007), who hypothesized that the observed pattern of results might have arisen due to the fact that sleep especially benefits memories that are, for one reason or another, reduced in strength. Though only partially, the results presented here speak in support of such a suggestion. Sleep-associated consolidation effects were far more pronounced under the weak encoding conditions of Experiment 1b than under the strong encoding conditions of Experiment 1a, both for single lists and for first double lists, which tempts one to conclude that, as far as sleep-associated consolidation is concerned, 'the weaker' might indeed be 'the better'. Moreover, larger sleep-related benefits for item lists affected by retroactive interference in comparison to single lists were observed across both Experiments 1a and 1b. However, according to classic interference literature (e.g., Gardiner et al., 1972; Tulving & Psotka, 1971) interference effects do probably not rely on reduced memory strength in double-list conditions, but rather on a reduced accessibility due to competition between memories that share the same cue; this view is also widely favored by the contemporary field (e.g., Baddeley, 1990; Nairne & Pandeirada, 2008). Therefore, the conclusions drawn by Drosopoulos et al. (2007) should be extended to include the proposition that not only weakly encoded memories, but also memories that are harder to access might preferentially benefit from sleep-associated memory consolidation.

Clearly, broader and more multifaceted research is required to further investigate sleep-associated memory consolidation. Paired-associate learning has been applied frequently in the past when the impact of sleep on declarative memory performance was to be investigated. However, there are other paradigms prevalent in modern memory research that could help to further elucidate processes of sleep-associated memory consolidation which may be influenced by both associative memory strength and competition between memory contents. For instance, while interference in the paired-associate task causes incidental forgetting, a cue to forget previously encoded material has been found to cause intentional forgetting (Bjork, LaBerge, & LeGrande, 1968). More precisely, the list-method directed forgetting task seems to be a suitable candidate to further examine the proposal that sleep may also counteract experimentally induced forms of forgetting. Little is known yet about sleep's impact on voluntary forms of forgetting. However, as a cue to forget is assumed to render items less accessible by reducing their associative binding to the current list context (Bäuml, Pastötter, & Hanslmayr, 2010), one could argue that, parallel to the findings on interference presented in this chapter, especially items reduced in accessibility might benefit from sleep in the directed forgetting task. The next chapter introduces list-method directed forgetting and reports data on effects of sleep-associated memory consolidation in this paradigm.

Chapter 3

Directed Forgetting

Many people wish they could voluntarily forget specific information and erase it from their minds, be it because the information is emotionally straining, annoying, or simply irrelevant. Memory research indicates that the latter may indeed be possible: Obsolete and outdated memories can be deliberately forgotten, at least under specific circumstances. Especially one paradigm, the list-method directed forgetting task (Bjork et al., 1968), has been applied to investigate the issue.

3.1 Background

In the list-method directed forgetting task, subjects usually study two lists of unrelated items. Between lists, they are presented a cue, indicating whether the first list will be relevant for an upcoming test or not. A forget cue after list 1 indicates that the list will not be tested later and may be forgotten; subjects are asked to focus on the second list instead. In contrast, a remember cue indicates that the first list will be tested later and has to be remembered; subjects are asked to additionally encode the second list. After list-2 encoding, recall performance for both lists is assessed, irrespective of which cue was given between lists. Memory performance for list 1 has been found to be impaired when subjects received a forget cue in comparison to when subjects received a remember cue; i.e., the first list can intentionally be forgotten when subjects are cued to do so. In addition, memory performance for list 2 has been found to be improved in the forget compared to the remember condition; i.e., forgetting of list 1 also results in enhanced memory performance for list 2. Therefore, such forgetting of outdated memories has been assumed to be adaptive for our memory system, as it frees up memory capacity and helps to remember current and relevant information more effectively (e.g., Bjork, 1989). Forgetting of list 1 has also been termed the ‘costs’ of directed forgetting, whereas the gain for list-2 performance has also been labelled the ‘benefits’ of directed forgetting (for reviews, see Bäuml et al., 2010; MacLeod, 1998).

Several accounts have been proposed to explain both costs and benefits of directed forgetting. One-mechanisms accounts attribute both findings to a single mechanism. For instance, it was suggested that a forget cue after the first list might stop rehearsal of the first list and induce selective rehearsal of the second list instead, which would explain both costs and benefits of the forget cue if one assumed that subjects in the remember condition simultaneously engaged in rehearsal of both lists (Bjork, 1970). However, Geiselman, Bjork, and Fishman (1983) found forgetting also for incidentally encoded material, thereby eliminating selective rehearsal as a potential explanation. Instead, they suggested that the forget cue triggers an active inhibitory process, that reduces accessibility of list 1 and that, because the first list can no longer interfere with list-2 recall, can also explain the benefits for list 2. Alternatively, the context change account (Sahakyan & Kelley, 2002) proposes that the forget cue causes an internal context change, and that because the context at retrieval is different from the one present at list-1 encoding, forgetting of this list occurs, while list 2 can be remembered more efficiently.

Two-mechanism accounts attribute costs and benefits of directed forgetting to separate mechanisms, thereby following studies reporting that costs and benefits do not necessarily occur together (e.g., Benjamin, 2007; Zellner & Bäuml, 2006). For instance, while assuming that an internal context change can account for list-1 forgetting, Sahakyan and Delaney (2003) suggest that the forget cue could additionally lead to an evaluation of the previous learning and, consequently, to a switch of encoding strategies, which could explain why list 2 is remembered better in the end. Another proposal is that the costs of list 1 arise due to inhibitory processes, but that the benefits are caused by a reset of encoding after the forget cue (Bäuml, Hanslmayr, Pastötter, & Klimesch, 2008). In any case, the directed forgetting effect cannot be explained by assuming that it simply is a result of demand characteristics; subjects have been found to be unable to recall list-1 items, no matter how highly motivated they are to do so (MacLeod, 1999).

Evidence for the assumption that forgetting of obsolete memories is an active process comes from a study showing that the engagement of controlled strategies is necessary for forgetting to arise (Foster & Sahakyan, 2011). Moreover, it has been reported that postcue encoding is crucial for the forgetting effect to emerge (Pastötter & Bäuml, 2007, 2010), suggesting that the processes mediating forgetting operate during list-2 encoding. Directed forgetting is reliably found in free recall tests (e.g., Geiselman et al., 1983; MacLeod, 1998). However, the forgetting effect has not been observed in recognition tests (e.g., Geiselman et al., 1983; Sego, Golding, & Gottlob, 2006), indicating that inhibition does not affect items directly. Rather, it is assumed that the whole list as a unit is affected (see Bäuml, 2008), as for instance indicated by the finding that memory for all items is compromised, irrespective of whether they were learned intentionally or incidentally (Geiselman et al., 1983). The accessibility of list 1 seems to be impaired, possibly by a reduction of associative links between the two lists' study contexts (Bäuml et al., 2010).

However, to date little is known about the durability and persistence of the directed forgetting effect. In most studies that applied the paradigm, distractor phases taking less than two minutes or no distractor phases at all were placed after encoding of the second list. Two studies reported regular effects of directed forgetting after distractor phases of 5 min (Conway, Harries, Noyes, Racsmany, & Frankish, 2000; Racsmany, Conway, Garab, & Nagymate, 2008), but to my knowledge there are no studies that employed longer delay intervals. Therefore, it remains unclear whether directed forgetting is still present after a prolonged period of time, or whether the effect fades away as time passes.

An experiment was conducted to fill this and another empirical gap at once: First, the experiment was to answer the question whether directed forgetting persists after prolonged retention intervals. Second, as described earlier, there is evidence for a role of sleep in the longevity of memories, and other evidence for the assumption that both strength and accessibility of memories may modulate this role of sleep. Therefore, a distinction was made between retention intervals filled with nocturnal sleep or diurnal wakefulness, to further address the question whether sleep and wake affect directed forgetting differently, that is assumed to rely on a reduction in associative strength. The corresponding experiment is described in the following section.

3.2 Experiment 2: Sleep and Directed Forgetting

There are hardly any published data available that could guide expectations when it comes to the durability of voluntary forms of forgetting and, in particular, of directed forgetting. Directed forgetting has been shown to persist for a few minutes (e.g., Racsmany et al., 2008), which might indicate that the effect is incorporated in long-term memory and should persist after longer retention intervals as well. However, it could also be argued that the forget cue may lose its salience with time, and that access to list-1 items might in general be recovered after prolonged retention intervals. More possibilities arise if one assumes that intervals filled with sleep or wakefulness could additionally influence the effect. For instance, following the findings on interference effects presented in the previous chapter as well as reports by other researchers (e.g., Drosopoulos et al., 2007; Ellenbogen, Hu, Payne, Titone, & Walker, 2007), especially items reduced in associative accessibility might benefit from sleep in the directed-forgetting task. As an instruction to forget is assumed to render items less accessible by reducing their associative binding to potential retrieval cues (Bäuml et al., 2010), one could argue that sleep in comparison to wake might mainly entail benefits for items in the forget condition, and less for items in the remember condition. If so, sleep might revive outdated memories, thereby undermining people's goal to forget obsolete memories and eliminating directed forgetting.

Here, the results of a list-method directed forgetting experiment are reported that was conducted to clarify the issue. By comparing memory performance of sleep and wake groups for list 1 after either a forget or a remember cue had been presented, the durability of the effect and the role of sleep for its persistence were to be investigated.

Method

Participants

256 subjects participated in the experiment in return for a small compensatory amount of money. The sample consisted of 85 male and 171 female subjects, mean age was 22.7 years (range 18-35 years). All participants completed a screening questionnaire and interview prior to selection (Ellenbogen et al., 2006a). Accordingly, no subject in the final sample suffered from any neurological, psychiatric, or sleep disorders, or was under the influence of drugs or medication affecting the central nervous system. All subjects spoke German as their native language, reported to have regular sleep-wake cycles, and were compliant with the instructions provided by the investigators. Subjects were randomly assigned to one of the experimental conditions and were tested either individually or in pairs. Between experimental conditions, no differences with regard to age, habitual sleep duration, subjective ratings of sleep quality, or a rough estimate of intelligence (as assessed by the connect-the-numbers test; Oswald & Roth, 1987) were evident (all *ps* $> .05$).

Material

Item material consisted of four lists, each containing sixteen concrete German nouns. Items were taken from different semantic categories out of the norm provided by Van Overschelde et al. (2004) and translated into German. Hence, items were unrelated, both within and between lists. In addition, all items within one list had unique initial letters. The four lists were divided into two sets of two lists that were equally often used across conditions. Within sets, the sequence of lists was balanced across subjects. For each list, eight items were defined as targets that participants were always asked to recall first on the later memory test. This was done because previous research indicates that directed forgetting is only present for the first half of a 16-item list on an immediate test (Bäuml & Samenieh, 2010). To get relatively 'pure' measures of the forgetting effect, the analysis of participants' memory performance was restricted to these eight target items of the first list.

Design

The experiment had a 2 x 2 design with the between-participants factors of CONDITION (12-wake, 12-h sleep) and CUE (forget, remember). In the remember condition, the first list was followed by a cue to remember the list for an upcoming test, whereas in the forget condition, it was followed by a cue to forget the list (e.g., Bjork, 1989). In the wake condition, participants studied the two lists of items at 9 a.m., and returned for the test after a 12-h waking retention interval; in the sleep condition, participants studied the same material at 9 p.m., and returned for the test after a 12-h interval that included regular sleep (see also Figure 3.1 for an illustration of procedure and experimental conditions). All participants in the 12-h wake condition reported to have stayed awake and not to have taken any naps during the day, whereas all participants in the 12-h sleep group reported to have slept regularly during the night (mean sleep duration: 7.5 hours; range 5-10 hours); none of the participants consumed alcohol between the two sessions.

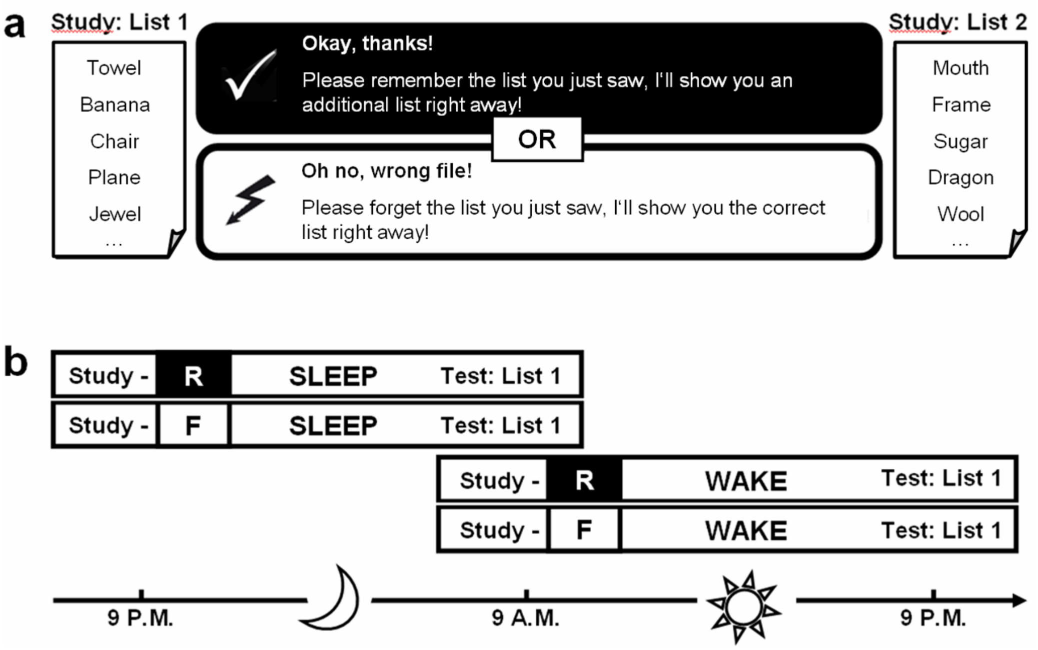

Figure 3.1: a) Procedure applied in *Experiment 2*. Between study of two lists, subjects were either asked to keep on remembering the first list (Remember Condition) or to forget the first list and to focus on the second list instead (Forget Condition). b) Experimental groups. Half of the subjects received a remember cue for list 1, the other half received a forget cue for this list; half of the subjects started the experiment at 9 a.m. and stayed awake before taking the final test after 12 hours, whereas the other half started the experiment at 9 p.m. and slept regularly in the meantime; R = Remember Condition, F = Forget Condition.

Procedure

Study Phase. Before presentation of the first list started, subjects were asked to try to memorize as many items as possible from the following list. Items were then presented in random order and at a rate of 4 sec each centrally on a computer screen. Two consecutive study cycles were conducted in this manner to ensure sufficient encoding. After presentation of the first list, subjects in the remember condition were simply asked to additionally encode a second list that was then presented in the same way as list 1. In the forget condition the procedure was different. To make sure that participants believed the cue to forget the first list, a computer crash was simulated: After the last item of the first list had been presented, the presentation shut down. The instructors acted surprised, apologized, and told the subjects that they had obviously opened a wrong, broken file. They asked the participants to try to forget the list they had just seen and to focus on another list that was enclosed in the correct, undamaged file instead. After this cover story, list 2 was presented in the same way as list 1.

Immediately after study, participants engaged in several distractor tasks to prevent rehearsal of any of the lists. First, they had to absolve the d2 test of attention (Brickenkamp, 2002), which was followed by the connect-the-numbers test (Oswald & Roth, 1987). In addition, subjects had to read through a number of moral dilemmata that required moral decision making (Greene, Nystrom, Engell, Darley, & Cohen, 2004). Approximately fifteen minutes after the study phase, subjects were allowed to leave the laboratory, and returned after 12 hours to complete the experiment.

Test Phase. As the main interest of this study was to investigate directed forgetting of the first list after 12 hours, list 1 was always tested before list 2 during the test phase. Before testing started, subjects in the forget condition were debriefed and asked to try to remember as many items of list 1 irrespective of the previously simulated computer crash. Then, the target items of the first list were tested before the rest of the list. Recall sequence was controlled through presentation of the items' unique initial letters. The item cues were presented successively and, both within the target set and the rest of the items, in random order for 10 sec each. Participants were asked to recall a studied list-1 item that fit the initial-letter cue. After all item cues for list 1 had been presented, subjects were asked to recall list 2 in the same way; nevertheless, only recall performance for list 1 will be included in the following analysis, as recall sequence may have biased performance for list 2.

Results

Ratings of Alertness. Ratings on the Stanford Sleepiness Scale (Hoddes et al.,

1973) did not differ between morning and evening groups in the first session, $F(1, 254) < 1.0$. There was also no difference between morning and evening groups concerning their ratings of alertness in the second session, i.e., after the 12-h delay, $F(1, 254) = 3.49$, $MSE = 0.88$, $p > .05$.

Directed Forgetting. To assess memory performance for the first list, a 2 x 2 ANOVA with the factors of CUE (forget, remember) and CONDITION (12-h wake, 12-h sleep) was calculated. Significant main effects of CONDITION, $F(1, 252) = 39.07$, $MSE = 332.98$, $p < .001$, and CUE were found, $F(1, 252) = 10.59$, $MSE = 332.98$, $p = .001$, as well as a significant interaction of the two factors, $F(1, 252) = 5.75$, $MSE = 332.98$, $p < .02$. The main effect of CONDITION reflects better memory performance in the sleep compared to the wake groups. Post-hoc t-tests confirmed that a benefit of sleep was evident irrespective of which cue had been given after list 1, $ts(126) > 2.60$, $ps < .01$. While the main effect of CUE suggests that the instruction to forget the first list was effective across both conditions, the significant interaction indicates that target recall was differently affected by CUE depending on whether the participants slept or stayed awake during the delay. Compared to the remember cue the forget cue impaired target recall when participants stayed awake (41.0 % correct vs. 28.1 % correct), $t(126) = 4.17$, $p < .001$, but did not affect recall when participants slept during the retention interval (49.8 % correct vs. 47.9 % correct), $p > .55$ (see also Figure 3.2 for a plot of the results).

Discussion

The results clarify two important issues that have been neglected so far. On the one hand, they suggest that directed forgetting is a powerful effect that may be found after prolonged retention intervals. Subjects in the wake conditions showed reliable forgetting of information they considered as outdated and irrelevant after 12 hours. On the other hand, the results provide evidence for a whole different role of sleep in this respect. Indeed, successful forgetting of outdated memories seems to depend on whether

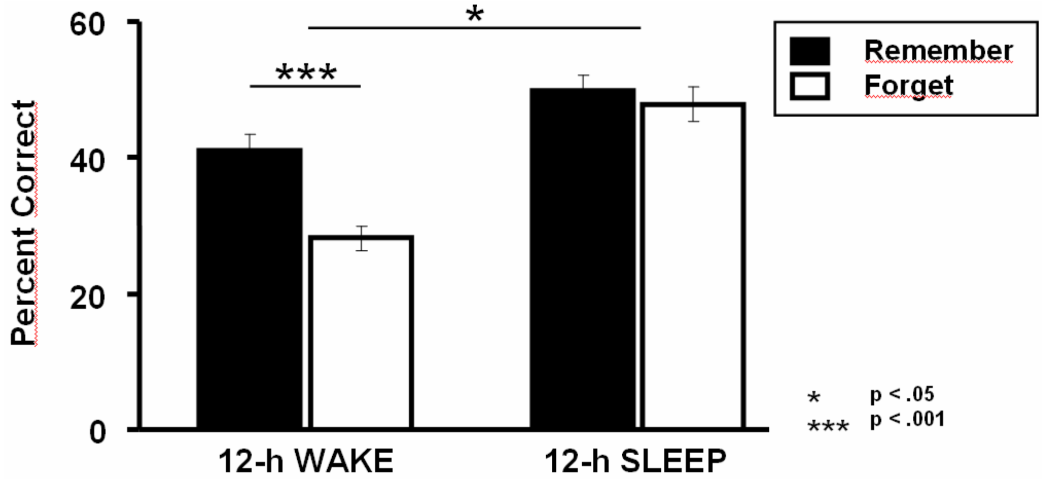

Figure 3.2: Results of *Experiment 2*. Mean recall performance for list 1 is plotted as a function of condition and cue (error bars represent standard errors of the mean).

sleep or wakefulness follows encoding. When wakefulness follows encoding of to-be-forgotten information, forgetting of outdated information is successful; when sleep follows the encoding, no forgetting of the information arises. Numerous previous studies demonstrated successful forgetting of outdated information after short retention intervals (for reviews, see Bäuml et al., 2010; MacLeod, 1998). The finding of no such forgetting after a 12-hour interval with sleep, therefore, indicates that sleep does not stabilize this forgetting, but rather revives outdated memories.

Interestingly, an effect of sleep was not only found when subjects were cued to remember the first list, but the effect was even more pronounced when subjects were cued to forget it. The finding that beneficial effects of sleep are not restricted to memories that people wish to maintain, but are exceedingly present for memories people wish to forget, has theoretical importance. It supports the view that sleep preferentially strengthens weak associative links (e.g., Drosopoulos et al., 2007; Ellenbogen et al., 2007) that may underlie the forgetting of outdated information in list-method directed forgetting (e.g., Bäuml et al., 2010). As forgetting, in one way or another, is frequently assumed to be related to memories being reduced in strength or accessibility (e.g., Anderson, 2003; Bäuml, 2008; Levy & Anderson, 2008; Wixted, Mickes, & Squire, 2010; Yonelinas, Aly, Wang, & Koen, 2010), sleep may counteract different forms of forgetting, e.g., incidental and intentional forms, in a more general way. This hypothesis is also in line with a recent report on abrogation of another form of voluntary forgetting in the think / no think paradigm (Fischer, Diekelmann, & Born, 2011), which specifically linked REM sleep to such a sleep-associated abrogation of forgetting.

Moreover, the finding that memories considered as outdated and irrelevant profit from sleep-associated memory consolidation could also be of practical relevance, because it suggests that sleep may only be partially beneficial for effective learning. Effective learning often requires successful remembering of relevant information and successful forgetting of out-of-date information (e.g., Bjork, 1989). By reviving outdated memories, sleep could therefore counteract effective learning. However, the debate about whether remembering as much information as possible or forgetting specific information is more adaptive and valuable in the long-term, is an old one and has, in different contexts, been conducted before (e.g., Anderson & Schooler, 2000; Nairne & Pandeirada, 2008). Therefore, the recovery of outdated information after sleep could also be interpreted as adaptive, because sleep more or less provides the possibility of a reset: After a night of sleep, memories may again be equally accessible and, thus, can be newly weighted and highlighted according to the demands arising during the following day. Future research will have to address which of the two interpretations is more adequate.

As noted earlier, there are no published data available on the durability and persistence of list-method directed forgetting. Recently, however, two reports emerged on effects of sleep and wake on a different version of the directed forgetting task. In item-method directed forgetting, namely, a cue to either forget or remember is given right after the initial presentation of every single item during study. This manipulation typically results in a similar effect as the manipulation applied in list-method directed forgetting: On a later test, items immediately cued to be forgotten are recalled significantly worse than items immediately cued to be remembered.

Using the item-method directed forgetting procedure, Rauchs et al. (2011) had their subjects encode stimulus material and tested recognition performance after either regular sleep or sleep deprivation. While memory did not differ for items cued to be remembered, the sleep deprivation group recognized significantly more items cued to be forgotten, thereby decreasing the item-method directed forgetting effect. Similarly, Saletin, Goldstein, and Walker (2011) found decreased forgetting of items cued to be forgotten for a wake in comparison to a napping condition. However, in this case, recall of to-be-forgotten items did not differ between wake and napping conditions, but napping selectively profited recall of to-be-remembered items. Although the results reported by Rauchs et al. (2011) and Saletin et al. (2011) are at least consistent in showing intact item-method directed forgetting after sleep and reduced item-method directed forgetting after wake, they do not contradict the present results on list-method directed forgetting. While item-method directed forgetting is probably caused by differential encoding of to-be-forgotten and to-be-remembered information, information in list-method directed forgetting is restructured after initial processing is already complete (MacLeod, 1999). Therefore, as the mechanisms underlying item-method and list-method directed forgetting are argued to be not the same, it is plausible to assume that sleep may influence the two effects differently.

Another recent report may appear to be contradicting the present results on list-method directed forgetting. Although Wilhelm et al. (2011) did not apply a directed forgetting task but, instead, paired-associate learning, they concluded from their data, that sleep selectively benefits memories one considers as being relevant in the future. After repeated study-test cycles, subjects were either told that the just encoded material was relevant for a final test coming up the following morning, or that it would be no longer needed, because a different task would have to be absolved the next day. Sleep-related benefits were only observed if subjects were informed about the test, but not if subjects didn't believe that the material would still be relevant.

Indeed, if the perceived relevance of memories impacted upon their sleep-associated memory consolidation, one would expect sleep in list-method directed forgetting to benefit mostly memories in the remember condition, and less in the forget condition. Although sleep-associated memory consolidation was observed in the remember conditions as well, sleep was found to entail more pronounced benefits in the forget conditions. However, in comparison to the procedure applied by Wilhelm et al. (2011), list-method directed forgetting is a paradigm that comes with a long history of research and many theoretical assumptions. As a direct request to forget the previously encoded material is assumed to impede access to the information (e.g., Bäuml et al., 2010), and because no such active reduction of accessibility should have taken place in the study by Wilhelm and colleagues (2011), sleep-associated memory consolidation in list-method directed forgetting may to a higher degree benefit to-be-forgotten information that is harder to access (e.g., Drosopoulos et al., 2007), while still providing gains for information considered as relevant for the future.

All in all, the just presented data further support the claim that the accessibility of memories affects sleep-associated consolidation. In list-method directed forgetting, this leads to a revival of outdated, to-be-forgotten information after sleep, whereas intact directed forgetting is observed across wake. Still, a regular sleep-associated benefit is observed for to-be-remembered information that is regarded as relevant for the future, thereby reconciling the finding with previous ones.

3.3 Conclusions

In the last two chapters evidence was provided for sleep-associated memory consolidation. In both paired-associate learning and list-method directed forgetting, memory performance was repeatedly found to be better after sleep compared to wake. In particular, however, and in line with previous research (e.g., Drosopoulos et al., 2007), sleep-associated memory consolidation was shown to benefit memories of rather weak associative strength. In Experiment 1a and 1b, more pronounced sleep-associated benefits were observed for a first double list than for a single list, while, in Experiment 2, sleep-associated benefits were larger for information cued to be forgotten than for information cued to be remembered. By providing higher gains for initially inaccessible material, sleep counteracted both retroactive interference and directed forgetting, thus abolishing both incidental and intentional forms of experimentally induced forgetting.

As outlined before, sleep-associated memory consolidation is assumed to rely on a reactivation of memory contents during sleep that ultimately stabilizes the respective memories and leaves them less prone to forgetting (e.g., Ellenbogen et al., 2007, or Rasch et al., 2007; see also Diekelmann & Born, 2010). Recent research on memory retrieval indicates that there may be interesting parallels between effects of sleep and retrieval. Indeed, conscious retrieval from memory has previously been shown to entail an awake reactivation of memories as well (e.g., Carr, Jadhav, & Frank, 2011), and stabilizing long-term effects of such memory retrieval have been reported repeatedly in the literature on the testing effect, showing reduced normal forgetting after retrieval in comparison to reexposure or distractor conditions (for a review, see Roediger & Butler, 2011). Thus, it could be hypothesized that retrieval may consolidate memories as well.

The next part of this thesis will be dedicated to investigating this proposal. To begin with, previous findings on the effects of retrieval on long-term retention will be reviewed. Then the outcomes of three experiments will successively be reported that were conducted to clarify whether retrieval stabilizes memories, thereby assuring less time-dependent forgetting and less susceptibility to interference, and whether such retrieval-associated consolidation during wake could have an impact upon later sleep-associated consolidation.

Part II

Retrieval-Associated Memory Consolidation

Chapter 4

A Current Perspective

Frequently and across many professional domains, administering or taking a test is regarded as a mere means to assess performance. Consequently, in educational contexts, students usually rely on familiar learning techniques when preparing for exams, such as rote learning and cramming, or rereading the relevant paragraphs over and over again. However, classic as well as recent research indicates that active retrieval from memory is not only a way to test performance, but may, at least in the long-term, be even more effective in boosting memory performance than additional study. Although this finding's potential relevance for all kinds of educational settings lies at hand, its implications have widely been neglected by practicioners, and have not been incorporated into general knowledge so far. The next few paragraphs shall therefore give a brief outline on research on the so-called testing effect.

4.1 The Testing Effect: Basic Findings

Among the first to investigate effects of testing on memory performance was Gates (1917). Based on extensive data collections he concluded that repeated testing may be beneficial for memory, both in the short-term and in the long-term. Since then, his findings have been replicated and refined many times (for reviews, see Roediger & Karpicke, 2006a; Roediger & Butler, 2011). Basically, what has been observed repeatedly across several decades, is that less time-dependent forgetting occurs after testing in comparison to restudy or distractor conditions (e.g., Thompson, Wenger, & Bartling, 1978; Wheeler, Ewers, & Buonanno, 2003), with repeated recall tests being most efficient in retarding forgetting (Wheeler & Roediger, 1992). While, initially after study, repeated restudy seems to be more beneficial for memory performance than repeated testing (e.g., Hogan & Kintsch, 1971), the pattern is reversed after longer delays, and a profound mnemonic benefit of testing due to less time-dependent forgetting becomes evident (e.g., Allen, Mahler, & Estes, 1969; see also Roediger & Karpicke, 2006b).

Further research has focused on identifying limitations and boundary conditions of the testing effect. However, this research essentially documents what powerful an effect is entailed by active retrieval from memory. For instance, it could be shown that the testing effect is not caused by any specific kind of test, but can be found for a variety of tests, including multiple choice testing (Kang, McDermott, & Roediger, 2007; Marsh, Roediger, Bjork, & Bjork, 2007) and completion tests (Hinze & Wiley, 2011). The benefits of testing have further been shown not to be restricted to the exact context present at retrieval. Rather, it has been reported that testing seems to promote the transfer of knowledge from one context to another (Butler, 2010; Rohrer, Taylor, & Sholar, 2010). In addition, retrieval has not only been shown to be superior to restudy (e.g., Karpicke & Roediger, 2007), but has also been observed to entail substantially bigger benefits for long-term retention than elaborative study techniques, such as concept mapping (Karpicke & Blunt, 2011).

One could speculate that success at retrieval is necessary for the testing effect to emerge, and indeed, providing corrective feedback after retrieval has been shown to further increase the positive effects of testing by diminishing the possibility to potentially stick to incorrectly retrieved information (e.g., Butler & Roediger, 2008; Pashler, Rohrer, Cepeda, & Carpenter, 2007).

Yet, unsuccessful retrieval attempts seem to come with their own benefit by improving the effectiveness of subsequent learning (e.g., Kornell, Hays, & Bjork, 2009; Richland, Kornell, & Kao, 2009). Moreover, testing of previously learned information has been shown to facilitate the consecutive encoding of new information, presumably by a reset of encoding and by reducing the extent to which the previously studied material interferes with the newly encoded material (e.g., Pastötter, Schicker, Niedernhuber, & Bäuml, 2011; Szpunar, McDermott, & Roediger, 2008; Weinstein, McDermott, & Szpunar, 2011). Thus, the impact of active retrieval from memory is not limited to past experiences, but also affects learning events taking place in the future.

Importantly, there is extensive evidence by now supporting the claim that active retrieval from memory is not only beneficial for retention in a laboratory setting, but also in realistic educational contexts. Many articles have been published that report successful replications of the testing effect in classroom and real-life study settings (e.g., Butler & Roediger, 2007; Larsen, Butler, & Roediger, 2008; McDaniel, Roediger, & McDermott, 2007; McDaniel, Agarwal, Huelser, McDermott, & Roediger, 2011). Moreover, testing effects have reliably been observed for a variety of stimulus materials more relevant in daily life than the stimulus material usually employed in most laboratory studies on human memory. For instance, retention for prose passages (e.g., Agarwal, Karpicke, Kang, Roediger, & McDermott, 2008), visuospatial maps (Carpenter & Pashler, 2007), or natural concepts (Jacoby, Wahlheim, & Coane, 2010) was reported to be superior after subjects had been tested on it. More recently, a role for long-term effects of retrieval in eyewitness reports has been established, suggesting that testing can on the one hand promote the accuracy of reports, but on the other hand also increase suggestibility - as long as no warning about potentially misleading information is provided (Chan & Langley, 2011; Thomas, Bulevich, & Chan, 2010).

Given all the experimental evidence underscoring the importance of testing for long-term retention, the notion that additional ungraded tests have not been implemented more in educational systems as a consequence seems kind of queer. However, another branch of research shows that people are hardly ever aware of the benefits that come along with taking tests. Although some students report to turn to self-testing when occupied with studying, only roughly a quarter of them does so, because they perceive to gain additional input from tests - rather, students report to test themselves in order to be able to evaluate their learning progress (e.g., Kornell & Bjork, 2007), i.e., to assess their performance. Another survey among students found that only few employ retrieval practice as a learning strategy, and that the most applied technique to promote learning still was rereading and restudying (Karpicke, Butler, & Roediger, 2009). In line with these findings, subjects have been reported to make wrong predictions about their future success in retaining repeatedly studied or repeatedly tested information: While subjects overestimated the probability that they would be able to remember information they had studied several times, they underestimated the probabilty that they would be able to remember material they had retrieved several times (Agarwal et al., 2008). All in all, there seems to be a lack of awareness for the mnemonic gain that is caused by recalling studied information.

4.2 How Does Retrieval Practice Benefit Retention?

Concerning the underlying mechanisms responsible for the long-term effects of retrieval practice, not much progress has been made yet. Recently, Carpenter (2011) stated that “we currently know very little about what this process involves, however, and how it benefits retention” (p.1547). Nevertheless, several proposals have been made to account for why testing is so advantageous for long-term memory (Roediger & Butler, 2011). For instance, it was suggested that memory traces might, as a consequence of retrieval, become more elaborate, possibly because such retrieval establishes additional retrieval routes (e.g., Carpenter, 2009; McDaniel & Masson, 1985), enhances organizational processes (e.g., Zaromb & Roediger, 2010), or because it triggers the generation of more potent mediators between retrieval cue and to-be-remembered target (e.g., Carpenter, 2011; Pyc & Rawson, 2010). In addition, it was proposed that the effort necessary to retrieve specific information from memory might mirror the magnitude of information processing, as, with more effort at retrieval, more pronounced benefits were found in the long-term (e.g., Gardiner, Craik, & Bleasdale, 1973; Pyc & Rawson, 2009). The term 'desirable difficulties' (Bjork, 1994) was coined to circumscribe this finding.

Clearly, the theoretical view on the effect that testing exerts on memory has changed over time. While Tulving (1967) still likened the overall effect of recall tests to that of additional study trials, the present perspective is that retrieval adds more to memory than mere restudy does (e.g., Roediger & Butler, 2011). Indeed, a link between retrieval practice and processes of reconsolidation has been suggested (see also Lasry, Levy, & Tremblay, 2008). More precisely, based on multiple trace theory (e.g., Moscovitch & Nadel, 1998), it was hypothesized that each retrieval from memory induced reconsolidation, thereby creating an additional memory trace which supports recall in the future. As work on rodents identified a molecular basis of reconsolidation (for a review, see Nader & Einarsson, 2010), this rather new approach to explain the benefits of testing could be regarded as more mechanistic and a little less psychological than previous ones.

However, there is one crucial difference between most studies on the testing effect and studies on reconsolidation in humans. On the one hand, the usual procedure when investigating effects of testing on memory is to implement the critical manipulation of either restudy or testing cycles right after initial learning (e.g., Butler & Roediger, 2007; Carpenter, Pashler, Wixted, & Vul, 2008; Karpicke & Roediger, 2007) - or, put differently, incorporated into the initial learning phase. Studies on reconsolidation in humans, on the other hand, typically place delays of at least 24 hours between initial study and the critical manipulation that is assumed to induce reconsolidation, as well as between this manipulation and a final memory test (e.g., Hupbach, Hardt, Gomez, & Nadel, 2008; Hupbach, Gomez, & Nadel, 2011; note, however, that spacing between learning and retrieval as well as between repeated testing events has been subject to experimental studies on the testing effect, too; e.g., Karpicke & Bauernschmidt, 2011; Karpicke & Roediger, 2010). Because consolidation is regarded as a multilayer process lasting longer than just a few seconds (e.g., Dudai, 2004), the proposal that retrieval right after encoding induces reconsolidation is rather astonishing. How can contents undergo a phase of reconsolidation if they have not even had a chance to be fully consolidated?

As outlined before, consolidation has been defined as a stabilization of memory contents that becomes evident in less time-dependent forgetting and less susceptibility to interfering treatments (e.g., Stickgold, 2005; Wixted, 2004). Because the mnemonic benefit of testing relies on less time-dependent forgetting after retrieval in comparison to restudy (e.g., Allen, Mahler, & Estes, 1969; Roediger & Karpicke, 2006b) this could be seen as an interesting parallel between memory retrieval and memory consolidation. In the following, it will be examined whether retrieval gives rise to more effective consolidation. As the retrieval-practice paradigm (Anderson, Bjork, & Bjork, 1994) was applied to investigate the issue, the next chapter will begin with a brief introduction of the paradigm and basic findings. In a first step, it will then be examined if retrieval attenuates time-dependent forgetting and if this attenuation is further modulated by sleep. In a second step, it will be investigated if retrieval practice also protects from retroactive interference.

Chapter 5

Retrieval Practice Effects and Normal Forgetting

5.1 Background

In the retrieval-practice paradigm (Anderson et al., 1994), participants typically study categorized item material (e.g., FRUIT-*Mango*, FRUIT-*Apple*, SPORTS-*Tennis*, SPORTS-*Soccer*) and afterwards are repeatedly asked to retrieve a subset of the items from a subset of the categories in an intermediate retrieval-practice phase (e.g., FRUIT-*Man*____). Such retrieval practice creates three item types: Practiced items from practiced categories (i.e., FRUIT-*Mango*), unpracticed items from practiced categories (i.e., FRUIT-*Apple*), and control items from unpracticed categories that did not appear during the retrieval-practice phase at all (i.e., SPORTS-*Tennis*, SPORTS-*Soccer*).

When memory for all three item types is assessed in an ultimate test, the typically observed results are the following: Due to retrieval practice, memory performance for practiced items is usually found to be enhanced in comparison to control items. At the same time, however, memory performance for unpracticed items is found to be reduced compared to memory performance for control items. This reduction in recall of unpracticed items has been termed retrieval-induced forgetting (RIF); the facilitating effect of retrieval on practiced items will in the following be referred to as retrieval-induced enhancement (RIE).

Retrieval-induced forgetting is often explained by the assumption that inhibitory processes operate during retrieval practice. The proposal is that, during retrieval attempts, a category's not-to-be-retrieved items interfere and, to overcome the interference, are inhibited and reduced in strength through the involvement of executive control processes (for reviews, see Anderson, 2003; Bäuml, 2008; Bäuml, Pastötter, & Hanslmayr, 2010). Behavioral evidence for the hypothesis that such an inhibitory mechanism directly affects item representations, comes from studies that report intact RIF over a variety of testing procedures, as for example word-stem completion tests (e.g., Anderson et al., 1994; Bäuml & Aslan, 2004), recognition tests (e.g., Hicks & Starns, 2004; Spitzer & Bäuml, 2007), and independent probe tests (e.g., Anderson & Spellman, 1995; Saunders & MacLeod, 2006). Consistent with the assumption of the involvement of executive control processes, it could be shown that retrieval-induced forgetting is attenuated by dual task performance during the retrieval-practice phase (Roman, Soriano, Gomez-Ariza, & Bajo, 2009). Stress (e.g., Koessler, Engler, Riether, & Kissler, 2009) or changes in mood (e.g., Bäuml & Kuhbandner, 2007) during retrieval practice can also abrogate the effect. Furthermore, neurocognitive findings provide first insights into how interference during retrieval is reflected in the brain (Kuhl, Dudukovic, Kahn, & Wagner, 2007; Staudigl, Hanslmayr, & Bäuml, 2010) and into how inhibition modulates the neural activity observed at test (Wimber et al., 2008).

An alternative account attributes RIF to blocking of the relatively weaker, unpracticed items caused by preferential recall of the relatively stronger, practiced items at test (for this noninhibitory account of RIF, see Camp, Pecher, & Schmidt, 2007; Jakab & Raaijmakers, 2009; Williams & Zacks, 2001). It has been found, however, that the strengthening of practiced items during retrieval practice is not necessary for RIF to emerge: Even when retrieval is made impossible in this intermediate phase by providing implausible retrieval cues, forgetting of related unpracticed items is observed (e.g., Storm, Bjork, Bjork, & Nestojko, 2006; Storm & Nestojko, 2010). This finding further adds to the evidence in favor of an inhibitory account, that does not predict a dependency of RIF on retrieval success; rather, it suggests that any effort to retrieve specific target memories irrespective of its success should be sufficient to stimulate the inhibition of irrelevant and interfering material. A further finding arguing against blocking as the origin of RIF is that no RIF emerges if a noncompetitive restudy phase instead of a competitive retrieval-practice phase follows upon encoding (Anderson, Bjork, & Bjork., 2000; Bäuml, 2002); i.e., RIF is a phenomenon specific to retrieval. While restudied items are enhanced above control items similar to the enhancement of practiced items mediated by retrieval practice, no forgetting of the related material emerges after restudy, suggesting that RIF does not rely on the preferential recall of enhanced items at test. Moreover, RIF is found to be present when recognition tests are applied (e.g., Hicks & Starns, 2004; Spitzer & Bäuml, 2007); as, on such tests, all items themselves are directly presented, blocking by comparatively stronger items can not have caused the observed RIF effect.

The basic effects of partial retrieval practice have so far been replicated many times. In addition, RIE and RIF have been observed for other types of stimulus material than categorized word lists (e.g., more episodic material, Ciranni & Shimamura, 1999; Macrae & MacLeod, 1999; Spitzer & Bäuml, 2009; or information socially shared in conversations, Coman, Manier, & Hirst, 2009; Cuc, Koppel, & Hirst, 2007) and have also been found in more applied settings closer to daily life (e.g., in eyewitness reports, MacLeod, 2002; Shaw, Bjork, & Handal, 1995; during the acquisition of a foreign language, Levy, McVeigh, Marful, & Anderson, 2007; in person perception and stereotypes, Dunn & Spellman, 2003). Moreover, the adaptive nature of forgetting due to retrieval of a related target memory has been emphasized. For instance, RIF has been reported to enable future learning (Storm, Bjork, & Bjork, 2008), to be linked to a lower rate of everyday cognitive failures (Groome & Grant, 2005), and to reduce neural processing demands (Kuhl et al., 2007).

Concerning the durability of RIF, inconsistent results have been reported. The results from some previous studies suggest that RIF is a transient phenomenon that diminishes as time between retrieval practice and final test passes. For instance, MacLeod and Macrae (2001) reported no RIF effect after a delay interval of 24 hours; while recall of practiced and control items was reduced with delay, no such reduction was found for unpracticed items. Similarly, across two experiments, Chan (2009) observed no RIF after a delay of 24 hours; while recall of control items was reduced after the delay, recall of unpracticed items was stable. For the practiced items results were less unequivocal, as memory performance for this item type was reduced in one experiment but not in the other. In contrast to this line of evidence, results from other previous studies suggest that RIF can also be lasting. Garcia-Bajos, Migueles, and Anderson (2009) showed that RIF was present after a delay of one week (for similar results, see also Tandoh & Naka, 2007). In addition, Garcia-Bajos et al. found that memory performance for all three item types was reduced across the one-week delay (see also Chan, 2010, for the possible role of delay after 24 hours). Thus, there is no consistency in the data currently available. Whether RIF is a rather longlasting or a rather transient phenomenon remains unclear, just as the question whether there are factors that might modulate the effect's durability.

Recently, Racsmany, Conway, and Demeter (2010) reported evidence indicating that sleep might modulate RIF: Racsmany et al. (2010) found RIF to be absent after a 12-hour wake interval, but to be present after a 12-hour sleep interval, suggesting that sleep after retrieval practice may be important for the persistency of RIF. In a second experiment, the authors also included a short-delay control condition and provided data that numerically point to a reduction of memory performance for control items, but not for practiced and unpracticed items across 12 hours of wakefulness; in comparison, memory performance for all three item types seems to have been stable across the sleep interval. It should be noted, however, that Racsmany et al. (2010) did not address the question whether delay affects item types differently and, as a result, did also not report statistical analyses. Nevertheless, the presented pattern of results suggests that sleep may affect RIF through different effects on practiced and unpracticed items versus control items.

Drawing firm conclusions on the role of sleep and wake delays for RIF and the single item types from this prior work (Racsmany et al., 2010) might be premature, however: Regular RIF after a short delay was not examined (Exp. 1) or not evident (Exp. 2), which makes it difficult to evaluate the effect of the wake delay on RIF. In addition, no appropriate control conditions of circadian effects were included, which complicates the evaluation of a possible sleep effect, because time of day may have additionally affected the results. Finally, the prior work did not analyze the impact of delay on the single item types, an analysis which could further elucidate when and why delay may be a boundary condition for RIF.

In the following section, a fresh experiment will be reported that was run to examine RIF for short and long delays, to analyze effects of wake and sleep for both RIF and the single item types, and to control for circadian effects at the same time. In order to go beyond a 'simple' replication of the finding presented by Racsmany et al. (2010), an additional goal was to examine the above issues in a purely episodic RIF task. Following prior work by Ciranni and Shimamura (1999) and Spitzer and Bäuml (2009), perceptual instead of semantic categories were used to examine how wake and sleep delays influence RIF and the single item types, i.e., practiced, unpracticed and control items. Method and outcomes of the experiment will be described in the following section.

5.2 Experiment 3: Retrieval Practice Effects, Sleep and Normal Forgetting

Here, the effects of wake delay on RIF were analyzed by comparing recall performance between short-delay morning (20-min wake) and long-delay wake groups (12-h wake). The effects of sleep delay on RIF were investigated by comparing recall performance of short-delay evening (20-min sleep) and long-delay sleep groups (12-h sleep). Generally, intact and comparable RIF and RIE in the two short-delay control conditions (20-min wake, 20-min sleep) were expected. Following Racsmany et al. (2010), one would further expect to find abrogated RIF after 12 hours of wakefulness, but intact RIF after 12 hours of sleep. If, on the one hand, the present experiment replicated findings by Chan (2009, 2010) and the general pattern of results presented by Racsmany et al. (2010), it could moreover be reasoned that a difference in the normal (time-dependent) forgetting across sleep (20-min sleep, 12-h sleep) and wake (20-min wake, 12-h wake) should be observed mainly for control items, and less for practiced and unpracticed items. If, on the other hand, there was no difference in forgetting of the single item types, as for example suggested by Garcia-Bajos et al. (2009), the RIF effect might persist over time and be present after both sleep and wake.

Method

Participants

Ninety-six undergraduates ($M = 23.5$; range 19-33 years) took part in the experiment voluntarily and in return for financial reimbursement. All participants completed a screening questionnaire and interview prior to selection (Ellenbogen et al., 2006a). This approach was chosen to ensure that no participant in the final sample suffered from any neurological, psychiatric, or sleep disorders, or was under the influence of drugs or medication affecting the central nervous system. All participants spoke German as their native language, reported to have regular sleep-wake cycles, and were compliant with the instructions provided by the investigators. Participants were randomly assigned to one of the four experimental conditions and were tested individually. There was no difference between groups with reference to age, a rough estimate of intelligence (as assessed by speed of cognitive processing; Oswald & Roth, 1987), and habitual sleep duration, all $ps > .15$.

Material

Item material consisted of 24 semantically unrelated German nouns divided into sets of eight items. Within sets, all items had unique initial letters. Each set was equally often assigned to the font colors red, blue, and yellow, and served equally often as practiced and unpracticed category. Each item was equally often used as practiced and unpracticed word (see also Spitzer & Bäuml, 2009; see Figure 5.1 for an illustration of the applied color categories and the different item types). To control for possible time of day confounds, participants used the Stanford Sleepiness Scale to indicate how activated they felt at the beginning of each session (Hoddes et al., 1973).

Design

The experiment had a 2 x 2 x 3 mixed factorial design: CONDITION (wake, sleep) and DELAY (20-min delay, 12-h delay) were manipulated between participants, and ITEM TYPE (practiced items, unpracticed items, control items) was varied within participants. For each participant, the experiment started at 9 a.m. or 9 p.m. Testing occurred either 20 min after the retrieval-practice phase (20-min wake and 20-min sleep) or after an additional 12-h delay (12-h wake and 12-h sleep conditions). No participant in the 12-h wake condition took a nap during the day, whereas all participants in the 12-h sleep group slept regularly during the night (mean sleep duration: 7.6 hours; range 6.0-9.3 hours); none of the participants consumed alcohol between the two sessions.

Procedure

Study Phase. Initially, participants were instructed to memorize the 24 items in their respective font colors. Items were presented for 5 sec each and appeared centrally on a computer screen. They were shown individually, and in a pseudorandomized order with no two items of the same color category following

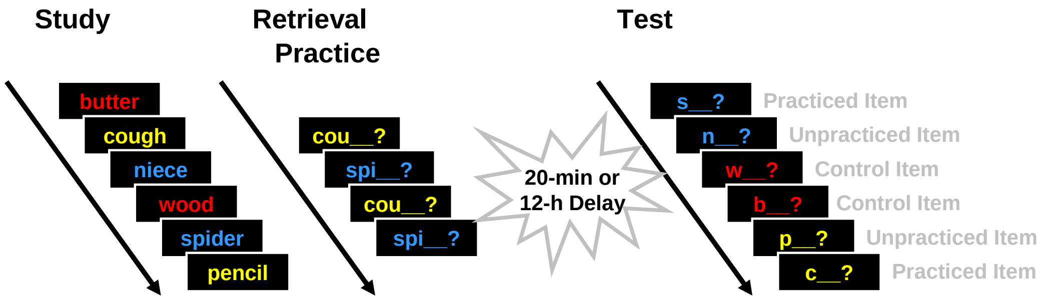

Figure 5.1: Depiction of the experimental procedure applied in *Experiment 3*. Partial retrieval practice after initial study created three different item types (i.e., practiced and unpracticed items from practiced categories, and control items from unpracticed categories). All initially studied items were tested on a final test that was administered either 20 minutes or 12 hours after retrieval practice.

each other. Promptly following the first study cycle, a second study cycle was conducted in exactly the same way.

Retrieval-Practice Phase. Immediately after study, participants were asked to recall half of the words from two of the three color categories in two successive retrieval cycles, providing the words' font colors and unique word stems as retrieval cues. Retrieval practice created three item types: Practiced items from practiced categories, unpracticed items from practiced categories, and control items from unpracticed categories. After retrieval practice, all participants engaged in a distractor task for 20 min. In particular, all participants were informed that all items, regardless of whether they had been practiced in this phase or not, would be tested later.

Test Phase. For the two control groups (20-min wake, 20-min sleep), testing occured immediately after the distractor task. The participants of the 12-h wake and 12-h sleep groups left the laboratory and returned after 12 hours to complete the experiment (see Figure 5.2 for an illustration of the experimental conditions). At test, all originally studied words were tested; the words' font colors and unique first letters were provided as retrieval cues. The cues were presented successively in a blocked, randomized manner: The sequence of categories was randomly chosen, but all items of one color category were tested successively. Within practiced categories, practiced and unpracticed items were blocked and it was randomly chosen which block was tested first. Participants were asked to give their responses orally without any time constraint. The investigators coded per key press whether the given answers were correct or not, which prompted the next cue to appear on the screen.

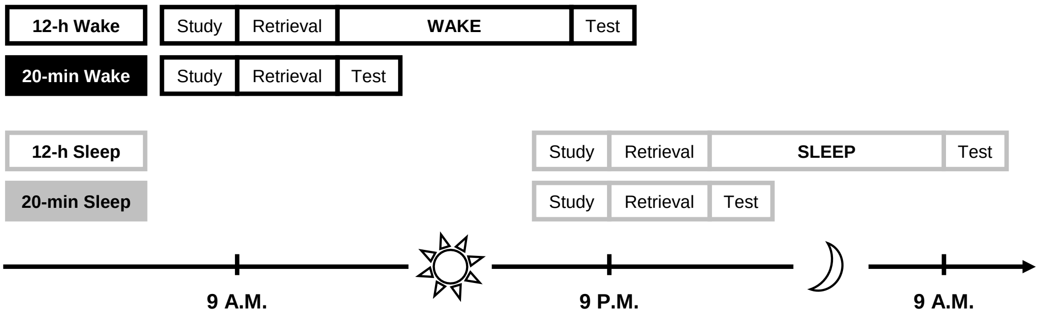

Figure 5.2: Experimental conditions in *Experiment 3.* Half of the subjects started the experiment at 9 a.m., while the other half started it at 9 p.m. Subjects were either tested shortly after retrieval practice (20-min wake, 20-min sleep) or after an additional 12-h delay (12-h wake, 12-h sleep).

Results

Prerequisites

Ratings of Alertness. Ratings on the Stanford Sleepiness Scale (Hoddes et al., 1973) did not differ between the four groups in the first session, $F(3, 92) = 1.44$, $MSE = 0.57$, $p > .20$. There was also no difference between the 12-h wake and 12-h sleep groups concerning their ratings of alertness in the second session, i.e., after the 12-h delay, $t(46) = 1.48$, $p = .15$.

Retrieval Success. Mean recall success rate in the retrieval-practice phase was 95.8 %, $SD = 7.3$, and it was unaffected by experimental group, $F(3, 92) < 1.0$. A comparison between the two groups that underwent retrieval practice in the morning (20-min wake, 12-h wake) and the two groups that underwent retrieval practice in the evening (20-min sleep, 12-h sleep) also yielded no significant difference, $F(1, 94) < 1.0$.

Circadian Effects. When comparing the 20-min wake and 20-min sleep control groups, no circadian memory effects arose for any of the three item types, $ts(46) < 1.35$, $ps > .15$.

Delay-Induced Forgetting and Sleep-Associated Consolidation

A 2 x 2 x 3 ANOVA with the factors of CONDITION (sleep, wake), DELAY (20-min delay, 12-h delay), and ITEM TYPE (practiced, unpracticed, or control items) revealed significant main effects of DELAY, $F(1, 92) = 7.36$, $MSE = 471.96$, $p < .01$, and ITEM TYPE, $F(1, 184) = 313.41$, $MSE = 63281.79$, $p < .001$, but no significant main effect of CONDITION, $F(1, 92) < 1.0$. The main effect of ITEM TYPE reflects the pattern of better recall for practiced items than for control items, and of better recall for control items than for unpracticed items (see below, for details); the main effect of DELAY reflects the decrease in recall with delay. In addition, a significant interaction between the three factors was found, $F(2, 184) = 3.49$, $MSE = 201.92$, $p < .05$, suggesting that CONDITION affected the role of delay for item type. No significant two-way interactions emerged, all $ps > .20$ (see also Figure 5.3 for a plot of the results).

To further investigate the effect of sleep on delay-induced forgetting for the three item types, two post-hoc 2 x 3 ANOVAs with the factors of DELAY (20-min delay, 12-h delay) and ITEM TYPE (practiced, unpracticed, or control items) were calculated, separately for the wake and sleep conditions. For the wake conditions, significant main effects of ITEM TYPE, $F(2, 92) = 144.26$, $MSE = 211.26$, $p < .001$, and DELAY, $F(1, 46) = 7.61$, $MSE = 479.88$, $p < .01$, were found, as well as a significant interaction between the two factors, $F(2, 92) = 4.50$, $MSE = 211.26$, $p < .02$. The interaction reflects the fact that reliable delay-induced forgetting was present for control items (54.7 % correct vs. 34.4 % correct), $t(46) = 3.94$, $p < .001$, but was absent for both practiced items (85.4 % correct vs. 79.7 % correct) and unpracticed items (37.0 %

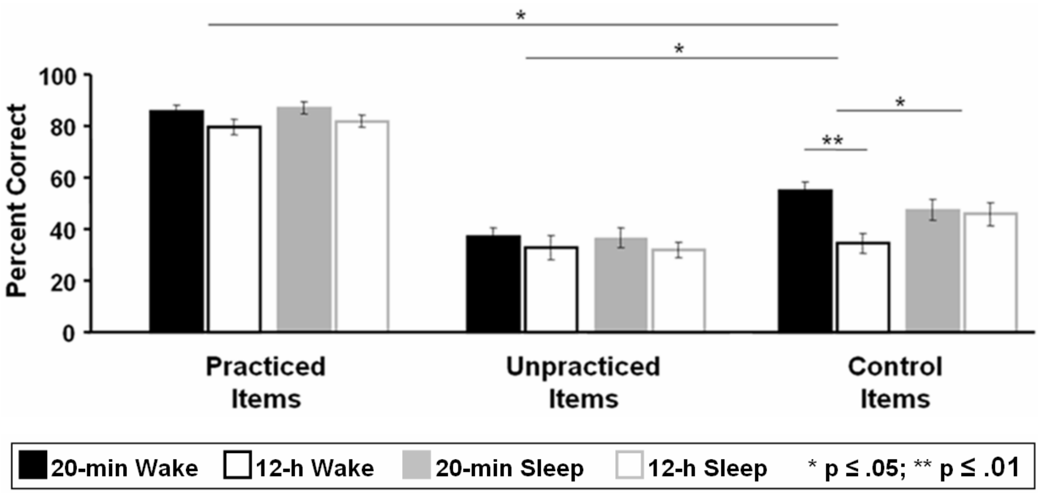

Figure 5.3: Results of *Experiment 3*. Mean recall performance for the three item types is plotted as a function of condition and delay (error bars represent standard errors of the mean).

correct vs. 32.8 % correct), $ts(46) < 1.41$, $ps > .15$. For the sleep conditions, a significant main effect of ITEM TYPE was found, $F(2, 92) = 170.66$, $MSE = 192.58$, $p < .001$, but there was no main effect of DELAY, $F(1, 46) = 1.13$, $MSE = 464.03$, $p > .25$, and no significant interaction between the two factors, $F(2, 92) < 1.0$. Consistently, for none of the three item types delay-induced forgetting emerged, $ts(46) < 1.65$, $ps > .10$. Memory performance for practiced items was stable across the sleep delay (87.0 % correct vs. 82.8 % correct), just as for control items (47.4 % correct vs. 45.8 % correct), and unpracticed items (36.5 % correct vs. 31.8 % correct).

Retrieval-Induced Enhancement and Retrieval-Induced Forgetting

Retrieval-induced enhancement (RIE) refers to better recall for practiced items than control items. A 2 x 4 ANOVA with the factors of ITEM TYPE (practiced or control items) and CONDITION (20-min wake, 20-min sleep, 12-h wake, 12-h sleep) revealed significant main effects of ITEM TYPE, $F(1, 92) = 341.64$, $MSE = 201.72$, $p < .001$, and CONDITION, $F(3, 92) = 4.43$, $MSE = 340.56$,

$p < .01$, and a marginal significant interaction between the two factors $F(3, 92) = 2.24$, $MSE = 201.72$, $p = .09$. The main effect of ITEM TYPE reflects RIE; the marginal interaction reflects the tendency for larger RIE in the 12-h wake group than the other three groups, although RIE was present in all four groups, $ts(23) > 7.45$, $ps < .001$.

RIF refers to inferior recall for unpracticed items compared to control items. A 2 x 4 ANOVA with the factors of ITEM TYPE (control or unpracticed items) and CONDITION (20-min wake, 20-min sleep, 12-h wake, 12-h sleep) showed a significant main effect of ITEM TYPE, $F(1, 92) = 28.65$, $MSE = 205.26$, $p < .001$, a marginally significant main effect of CONDITION, $F(3, 92) = 2.53$, $MSE = 507.71$, $p = .06$, and a significant interaction between the two factors, $F(3, 92) = 2.80$, $MSE = 205.26$, $p < .05$. The main effect of ITEM TYPE reflects RIF; the interaction reflects the tendency for reduced RIF in the 12-h wake group. Indeed, post-hoc analyses showed that RIF was absent in the 12-hour wake group, $p > .70$, but was present in the three remaining groups, $ts(23) > 2.64$, $ps < .02$.

Discussion

The results show that RIF is present after a short delay, but eliminated after 12 hours of wakefulness: Memory performance for control items was found to be reduced, but was maintained for practiced and unpracticed items across the wake delay. Additionally, the data provide evidence that RIF is maintained after a 12-hour interval filled with sleep: Memory performance for all item types was stable across the sleep delay. As the described effects are not attributable to circadian effects, the pattern of results indicates that wake but not sleep delay can reduce or eliminate RIF. Furthermore, the results suggest that sleep modulates RIF by differently affecting control items, because the normal (time-dependent) forgetting of control items differed between sleep and wake delays whereas it did not differ for practiced and unpracticed items.

Prior work on the role of delay on RIF was inconsistent in the critical conclusion whether RIF was present or absent after longer delay intervals. The present results agree with the finding of no RIF after 12 hours of wakefulness that was reported by Racsmany et al. (2010). Other previous studies, however, examined longer delay intervals and did not differentiate between the sequence of sleep and wake upon retrieval practice. Nevertheless, the present pattern of results observed after 12 hours of wakefulness is consistent with previous reports of abrogated RIF after delays of 24 hours (Chan, 2010; MacLeod & Macrae, 2001). In contrast, the results somewhat differ from previous reports of intact RIF after a delay of one week (Garcia-Bajos et al., 2009; Tandoh & Naka, 2007).

While the present data apparently cannot resolve the inconsistency in results between the two conflicting lines of previous studies, they help identifying a first factor that can modulate the effect of delay on RIF, i.e., sleep. The finding that RIF is present after a delay interval filled with sleep, but not after a delay interval filled with wake, and the finding that the difference in results is mainly driven by a consolidating effect of sleep on the control items is basically consistent with the prior work by Racsmany et al. (2010). However, the finding goes beyond this prior work by showing RIF also for short delays, by controlling for circadian effects, and by analyzing effects of sleep for the single item types. Also, it generalizes the prior work by demonstrating the effects of sleep in a purely episodic RIF task.

In any case, a recommendation for future RIF studies can be derived from the present data. In typical experiments that investigate memory performance after delays, some subjects are tested in the morning while others are tested in the evening. Among those subjects tested in the evening, some may have worked hard during the whole day whereas others may have liked to enjoy a nap somewhen in the afternoon, etc. While the present results suggest that time of day per se does not affect RIF and does not affect performance differently for the three item types, they indicate that, across a longer delay, it can make a difference whether sleep closely follows upon encoding and retrieval practice or not. If sleep follows after retrieval practice, it is more likely that RIF will be observed on a later test than if a longer wake-delay interval follows instead. Thus, sleep should be controlled for in future RIF experiments examining the impact of delay in order to get 'pure' measures of the effects of delay on RIF. More consistent research is needed to provide a reliable answer to the question whether the RIF effect is transient or persistent over prolonged delays.

The present study examined effects of retrieval practice under 'standard' conditions, for which retrieval practice has usually been found to cause forgetting. However, as mentioned above, there are boundary conditions for RIF, as for instance mood (Bäuml & Kuhbandner, 2007) or stress (e.g., Koessler et al., 2009). Also, no RIF has been reported for interconnected material after short delays (e.g., Anderson & McCulloch, 1999; Bäuml & Hartinger, 2002), while even the opposite effect of retrieval-induced facilitation was observed for such material after longer delays (e.g., Chan, McDermott, & Roediger, 2006; Chan, 2009). Interestingly, however, similar effects of delay on the single item types were reported for low- and high-integration conditions (see Chan, 2009). Therefore, the question arises whether the effects of sleep found in the present and prior work generalize from low- to high-integration conditions, or if sleep affects the two types of materials differently. As the degree of integration was not manipulated in this study, future work is required to address the issue.

Ultimately, the data may not only bear interesting features for research on RIF, but also for research on sleep-associated memory consolidation. Prior work has shown that sleep does not influence all types of memories equally (see also Diekelmann & Born, 2010). For instance, sleep-related benefits were found to be bound to the strength or accessibility of memories (e.g., Drosopoulos et al., 2007), their emotional tone (e.g., Payne, Stickgold, Swanberg, & Kensinger, 2008), or their relevance for the future (e.g., Wilhelm et al, 2011). The present experiment suggests a further condition that could restrict sleep-associated memory consolidation, namely the previous testing of memorized material (e.g., Roediger & Butler, 2011). The finding that both practiced and unpracticed items show no time-dependent forgetting suggests effects of retrieval practice that go beyond influences on the practiced and unpracticed items' memory strength. Indeed, normal forgetting is typically largely unaffected by item strength (e.g., Slamecka & McElree, 1983). Therefore, normal forgetting should not only have been present for the control items but should have been present for the (stronger) practiced and the (weaker) unpracticed items as well. Rather, the finding supports recent literature on the testing effect, which reports reduced forgetting of memories after retrieval, both of the retrieved material itself and of related nonretrieved items (e.g., Chan, 2009; Roediger & Karpicke, 2006b). It indicates that sleep-associated consolidation may be present mainly for not-yet-retrieved items, and may be reduced after retrieval practice.

Why exactly testing protects the practiced and related unpracticed items against forgetting is an interesting issue. One possibility could be that retrieval triggers consolidation processes (e.g., Spear & Müller, 1984), for instance by directly reactivating practiced materials (e.g., Carr et al., 2011) and by incidentally reactivating related and interfering, but unpracticed materials at the same time (e.g., Kuhl et al., 2007; Staudigl et al., 2010). As outlined before, other proposals to explain the benefits related to retrieval have previously been made, as for instance the proposal that retrieval could mediate more effective cues (e.g., Pyc & Rawson, 2010), or that there might be a link between retrieval and processes of reconsolidation (e.g., Roediger & Butler, 2011; for another proposal, see also Kornell, Bjork, & Garcia, 2011). More research is needed to elucidate what underlies the mnemonic benefits of retrieval and to shed further light on the question whether (and if so, how) prior retrieval may indeed modulate subsequent sleep-associated memory consolidation. In sum, the results presented here demonstrate that both sleep and retrieval can render memories less susceptible to delay-induced forgetting. The retrieval-associated effect occured for practiced material and generalized to related unpracticed items, whereas the sleep-associated effect was evident only for not-yet-retrieved material unrelated to practiced material.

5.3 Conclusions

Experiment 3 was carried out to test whether delay may constitute a boundary condition for RIF, and whether sleep modulates the role of delay. The results replicated a previous finding (Racsmany et al., 2010) by showing that intact RIF is found after sleep, but not after wake; thus, whether delay is a boundary condition for RIF seems to depend on whether sleep or wake follow upon retrieval practice. In addition, the results replicated other previous findings by showing that retrieval reduced the time-dependent forgetting of retrieved memories (e.g., Roediger & Butler, 2011), and further generalized this finding by showing that also related but unpracticed material was stabilized (see also Chan, 2009, 2010). Moreover, it was shown that retrieved and related items in comparison to control items did not experience any further gains due to sleep-associated memory consolidation, possibly because the items had already been stabilized as a consequence of retrieval.

Notably, this ostensive influence of prior retrieval practice on later sleep-associated memory consolidation also modified sleep's influence on forgetting. While, in the first part of this thesis, it was shown across three experiments and across two different paradigms that sleep can counteract experimentally induced forms of forgetting by preferentially stabilizing memory contents that are initially harder to access, a different pattern was observed in Experiment 3. As retrieval practice had presumably already stabilized the unpracticed, related items, no additional (and thus, preferential) sleep-associated benefit could emerge. Because control items, in contrast, did benefit from sleep, RIF was maintained after sleep, but not after wake. Thus, by influencing subsequent processes of sleep-associated consolidation, retrieval-associated stabilization may also be decisive for the question whether experimentally induced forms of forgetting are evident in the long-term or not.

However, the idea that retrieval could be related to a stabilization of memory contents can be tested in one more way. Indeed, consolidation is assumed to be evident not only in reduced time-dependent forgetting, but also in reduced susceptibility to interference (e.g., Stickgold, 2005). For instance, sleep-dependent consolidation has been shown to make memories less susceptible to retroactive interference (Ellenbogen et al., 2006a). Whether retrieval protects from the detrimental effects of retroactive interference as well remains to be investigated. The following chapter will deal with this question.

Chapter 6

Retrieval Practice Effects and Interference

As pointed out in one of the previous chapters, interference theory attributes a great deal of daily phenomena of forgetting to the fact that interfering material is encoded previously or subsequently to the encoding of later to-be-retrieved target information (e.g., Wixted, 2004). Notably, there are parallels between time and interference: Both have been discussed to cause forgetting and to be somewhat interrelated (e.g., Jonides et al., 2008); moreover, both have been employed as a means to probe memories and to assess whether they have been stabilized or not. It is argued that stabilized or consolidated memories should be somewhat protected from the detrimental effects of both time and interference, whereas unstabilized memories should suffer from regular time-dependent or interference-induced impairment (see also Dudai, 2004).

In the previous chapter, data were presented showing that retrieval protects both retrieved and related material from time-dependent forgetting. The following sections will report details and outcomes of two experiments that were conducted in order to test, whether retrieval in comparison to another kind of reprocessing stabilizes memories and makes them less susceptible to retroactive interference. If this were the case, the finding would mirror results obtained for sleep in comparison to wake intervals (Ellenbogen et al., 2006a) and further strengthen the idea, that awake memory retrieval could be associated with consolidation.

6.1 Experiment 4a: Interference after Retrieval Practice

Again applying the retrieval-practice paradigm, this experiment was aimed at examining whether partial retrieval practice protects both practiced and related unpractied items from retroactive interference. If practiced and unpracticed items were indeed stabilized through retrieval-associated processes, memory performance for these items should be significantly less impaired by an interference manipulation than memory performance for control items that were not retrieved or reactivated and, as a consequence, not stabilized.

Effects of interference on the different item types were studied by conducting two highly similar experimental blocks: Subjects underwent one standard block of retrieval-induced forgetting and one block in which retroactive interference was additionally elicited between retrieval practice and final test. Impairment due to interference was assessed by comparing memory performance for the item types across the two experimental blocks.

Method

Participants

32 healthy and drug-free students participated in the experiment, either for partial course credit or a small financial reimbursement. Mean age was 22.3 years (range 19-28 years). All participants spoke German as their native language.

Material

Item material was taken from the two category norms provided by Van Overschelde et al. (2004) and Scheithe and Bäuml (1995). Two item sets were created that consisted of twelve exemplars of six semantic categories each. The twelve exemplars were further divided into six target items and six lures. Target items were used for the standard conduction of the retrieval-practice paradigm; lures were presented for additional study between retrieval practice and test to induce interference in one of the two experimental blocks. Within categories, all items possessed unique initial letters. The two item sets were equally often assigned to blocks with and without interference, respectively.

Design

The experiment had a 3 x 2 design. The two factors of ITEM TYPE (practiced items, unpracticed items, control items) and INTERFERENCE (with interference, without interference) were both manipulated within subjects. Each participant successively underwent two highly similar experimental blocks. Retrieval practice always followed upon initial study, creating the three item types (practiced items, unpracticed items, control items). However, before a final test of all items was administered, subjects either studied additional items from the same semantic categories (experimental block with interference) or engaged in a distractor task for the same duration (experimental block without interference). The sequence of blocks with and without interference was balanced across subjects. Between the two blocks, subjects were given a short break of about 5 minutes (see Figure 6.1 for an illustration of the two experimental blocks).

Procedure

Study Phase. During each of the two experimental blocks, 36 items belonging to six different semantic categories were studied. Initially, participants were

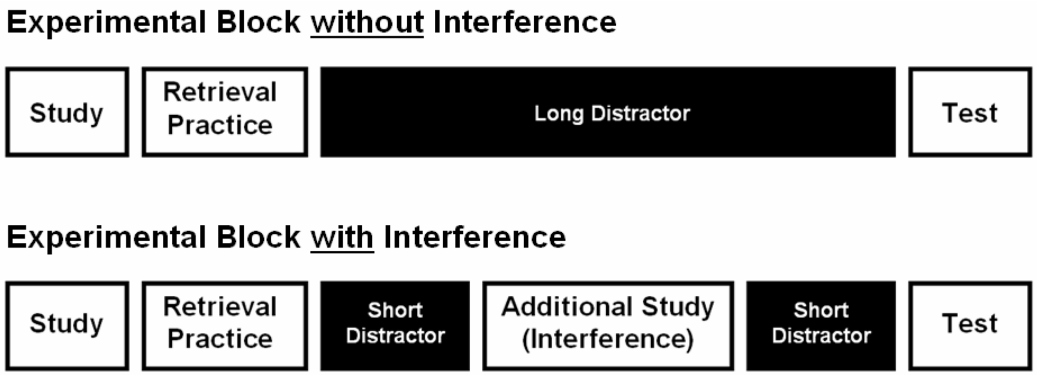

Figure 6.1: Depiction of the two experimental blocks applied in *Experiment 4a*. Each Subject absolved both experimental blocks; sequence of blocks was balanced across subjects.

instructed to memorize the following items in relation to their categories. Items then appeared together with their respective categories centrally on a computer screen and were presented for 3 sec each. They were shown individually and in a pseudorandomized order with no two items of the same category following each other.

Retrieval-Practice Phase. Immediately after study, participants were asked to recall half of the words from four of the six semantic categories in two successive retrieval cycles. The words' categories and unique word stems were provided as retrieval cues. Subjects had 8 sec to recall the respective items and to write down their answers. Thus, retrieval practice created three item types: Practiced items from practiced categories, unpracticed items from practiced categories, and control items from unpracticed categories.

Interference Phase. In the experimental block with interference, subjects studied 36 additional items after retrieval practice that belonged to the same six semantic categories encoded during the initial study phase. Subjects were asked to additionally memorize as many of the following items as possible. Again, items and their respective categories were presented for 3 sec each and in a pseudorandomized manner with the single restriction that no two items of the same category followed each other. Right before and after the interference manipulation, subjects counted backwards in steps of three for 30 sec to control for working memory effects. All in all, the interference phase took three minutes. In the experimental block without interference, subjects instead solved simple arithmetic problems between retrieval practice and test for three minutes to rule out time as a confounding variable.

Test Phase. For each of the experimental blocks, testing occurred immediately after either the interference phase or the long distractor task had been absolved. The words' semantic categories and unique first letters were provided as retrieval cues and were presented successively in a blocked, randomized manner for 8 sec each centrally on a computer screen: The sequence of categories was randomly chosen, but all items of one category were tested successively. Within practiced categories, practiced and unpracticed items were blocked and it was randomly chosen which block was tested first. Subjects had 8 sec to recall each item and were asked to write down their respective answers. After 8 sec, the next retrieval cue appeared on the screen. In the experimental block with interference, all initially studied items were tested first; the additionally encoded items were tested second. After the final test phase, all subjects were debriefed and thanked for their participation.

Results

Prerequisites

Retrieval Success. Mean recall success rate in the retrieval-practice phase was 90.8 % (SD = 7.8) in the experimental block without interference and 90.9 % (SD = 8.2) in the experimental block with interference. There was no difference between the two blocks, $t(31) = 0.10$, $p > .91$.

Interference Manipulation. Mean recall of the additionally encoded items in the experimental block with interference was 62.8 % ($SD = 13.3$).

Effects of Interference on the Different Item Types

A 3 x 2 ANOVA with the factors of ITEM TYPE (practiced, unpracticed, or control items) and INTERFERENCE (with interference, without interference) revealed significant main effects of ITEM TYPE, $F(2, 62) = 60.37$, $MSE = 197.74$, $p < .001$, and INTERFERENCE, $F(1, 31) = 12.77$, $MSE = 167.93$, $p < .01$. The main effect of ITEM TYPE reflects the pattern of better recall for practiced items than for control items, and of better recall for control items than for unpracticed items (see below, for details); the main effect of INTERFERENCE reflects a decrease in recall in the experimental block with interference. In addition, a significant interaction between the two factors was found, $F(2, 62) = 4.52$, $MSE = 177.58$, $p < .02$, suggesting that INTERFERENCE affected the three item types differently (see Figure 6.2 for a plot of the results).

To further investigate the effect of interference on items from practiced and unpracticed categories, two separate post-hoc 2 x 2 ANOVAs with the factors of INTERFERENCE (with interference, without interference) and ITEM TYPE (practiced, control items; or control items, unpracticed items) were calculated. For practiced and control items, significant main effects of ITEM TYPE, $F(1, 31) = 103.24$, $MSE = 164.64$, $p < .001$, and INTERFERENCE, $F(1, 31) = 22.13$, $MSE = 103.61$, $p < .001$, were found, as well as a significant interaction between the two factors, $F(1, 31) = 9.30$, $MSE = 139.99$, $p < .01$. The interaction indicates that RIE was more pronounced in the block with interference than in the block without interference, although it was present in both blocks, $ts(31) > 5.80$, $ps < .001$. This conclusion is further supported by the fact that recall of control items was impaired by the interference manipulation (69.0 % correct vs. 54.2 % correct), $t(31) = 4.62$, $p < .001$, while recall of practiced items was not (85.7 % correct vs. 83.6 % correct), $t(31) = 0.94$, $p > .35$.

For control and unpracticed items, a significant main effect of INTERFERENCE, $F(1, 31) = 12.07$, $MSE = 103.61$, $p < .01$, as well as a significant interaction between the two factors was found, $F(1, 31) = 4.76$,

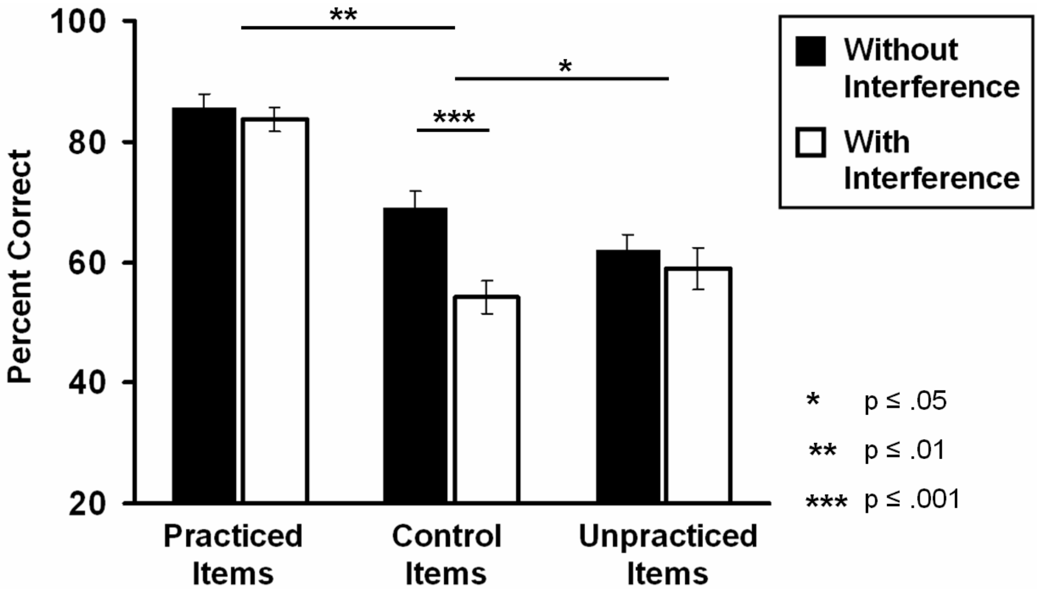

Figure 6.2: Results of *Experiment 4a*. Mean recall performance is shown for the three item types in the two experimental blocks. Error bars represent standard errors of the mean.

$MSE = 1098.34$, $p < .04$, while the main effect of ITEM TYPE was not significant, $F(1, 31) < 1.0$. The interaction suggests that RIF was more pronounced in one of the two blocks; indeed, a difference in recall between control and unpracticed items was present in the block without interference, $t(31) = 3.04$, $p < .001$, but not in the block with interference, $t(31) = 0.92$, $p > .35$. This difference in RIF arises due to the fact that recall of control items was impaired by additional learning (69.0 % correct vs. 54.2 % correct), $t(31) = 4.62$, $p < .001$, whereas recall of unpracticed items was not (62.0 % correct vs. 58.9 % correct), $t(31) = 0.75$, $p > .45$.

Discussion

In the previous chapter, the outcomes of an experiment were reported, showing that retrieval practice protects from time-dependent forgetting. The results displayed here extend this finding by showing that retrieval practice also shelters memories from interference-induced forgetting. Memory performance for both practiced and related unpracticed items showed no impairment due to retroactive interference. In contrast, memory performance for control items that did not appear during the retrieval-practice phase, and that were consequently not subject to any stabilization, showed pronounced impairment due to retroactive interference. As stated earlier, memory consolidation is assumed to manifest itself in less time-dependent forgetting and less susceptibility to interference (e.g., Stickgold, 2005; Wixted, 2004). If these are regarded as the measures by which the presence or absence of consolidation can be assessed, the results presented here suggest that retrieval practice consolidates practiced and unpracticed items, while control items remain unaffected.

Recently, Halamish and Bjork (2011) examined the impact of test difficulty on occurrence and shape of the testing effect. In a paired-associate task, subjects were asked to read through a second list of interfering paired associates, after they had either restudied a first list of paired associates several times, or after they had repeatedly been tested on this first list. In parallel to the results presented here, it was reported that memory performance was less impaired by interference in the testing condition compared to the restudy condition; however, the authors underlined that retroactive interference did impair performance in the testing condition as well, and concluded that testing provided only a relative protection from interference. As Halamish and Bjork (2011) set out to manipulate difficulty at test by introducing interference, their design did not include related, but unpracticed material. Based on the retrieval-practice paradigm, the current data extend the finding that directly retrieved material is less susceptible to interference by indicating that retrieval also protects related, unpracticed material from the detrimental effects of interference.

A point of critique frequently expressed when the retrieval-practice paradigm is applied is that observed effects do not necessarily have to be caused by retrieval, but could also be due to enhanced processing time. While practiced items are reprocessed during the retrieval practice phase, control items are not. This circumstance, it is argued, leaves room for the suspicion that a restudy instead of a retrieval practice phase could entail the same effects because it provides the same possibility to further process the contents. To be able to invalidate this objection, an additional data collection is necessary that clarifies, whether restudy instead of retrieval practice also protects from the detrimental effects of retroactive interference. The corresponding data will be presented in the following section.

6.2 Experiment 4b: Interference after Restudy

The design of Experiment 4b was similar to that of Experiment 4a. Again, subjects consecutively underwent two experimental blocks; yet, instead of a retrieval-practice phase, a partial restudy phase was placed after initial encoding. In parallel to Experiment 4a, retroactive interference was elicited prior to a final memory test in only one of the two blocks. Impairment due to interference was again assessed by comparing memory performance for restudied items as well as for items related and unrelated to these restudied items across the two experimental blocks.

By including this restudy manipulation it should additionally be probed whether retrieval or basic reprocessing in general caused the protection from interference that was observed in Experiment 4a. If mere reprocessing entailed the same protective effect as retrieval, restudied items (and possibly also related, but not restudied items) should show no or less impairment due to interference in comparison to control items. If, however, the protective effect was specific to retrieval, restudy should not provide any shelter from the detrimental effects of interference in this experiment. Interference effects should then be evident for all three item types.

Method

Participants

As in the previous experiment, 32 healthy and drug-free students participated either for partial course credit or a small financial reimbursement. Mean age was 22.5 years (range 19-33 years). All participants spoke German as their native language.

Material

Item material was selected in the same way as for Experiment 4a. The two item sets, each consisting of twelve exemplars of six semantic categories, were used across two experimental blocks that only differed in whether interference was induced before the final test or not.

Design

The experiment had a 3 x 2 design. The two factors of ITEM TYPE (relearned items, related items, control items) and INTERFERENCE (with interference, without interference) were both manipulated within subjects. Each participant successively underwent two highly similar experimental blocks. After initial study, a relearning phase created three different item types (restudied items, related items, control items). However, before a final test of all items was administered, subjects either studied additional items from the same semantic categories (experimental block with interference) or engaged in a distractor task for the same duration (experimental block without interference). The sequence of blocks with and without interference was balanced across subjects. Between the two blocks, subjects were given a short break of about 5 minutes.

Procedure

Study Phase. During each of the two experimental blocks, 36 items belonging to six different semantic categories were studied. Initially, participants were instructed to memorize the following items in relation to their categories. Items then appeared together with their respective categories centrally on a computer screen and were presented for 3 sec each. They were shown individually and in a pseudorandomized order with no two items of the same category following each other.

Restudy Phase. Immediately after study, participants were asked to restudy half of the words from four of the six semantic categories in an additional intensive study phase. Items together with their respective categories were presented for 8 sec each and on two consecutive relearning cycles. Subjects were told to focus on the presented items and to make use of this additional study opportunity. Thus, in parallel to the retrieval-practice paradigm, this partial restudy created three item types: Restudied items that were relearned themselves; related items that belonged to the same semantic categories as restudied items, but that were not reprocessed themselves; and unrelated control items from other categories that did not appear in the restudy phase at all.

Interference Phase. In the experimental block with interference, subjects encoded 36 additional items after the restudy phase that belonged to the same six semantic categories encoded during the initial study phase. Subjects were asked to additionally memorize as many of the following items as possible. Again, items and their respective categories were presented for 3 sec each and in a pseudorandomized order with the single restriction that no two items of the same category followed each other. Right before and after the interference manipulation, subjects had to count backwards in steps of three for 30 sec to control for working memory effects. All in all, the interference phase took three minutes. In the experimental block without interference, subjects had to solve simple arithmetic problems for three minutes instead, to rule out time as a confounding variable.

Test Phase. For each of the experimental blocks, testing occurred immediately after either the interference phase or the long distractor task had been absolved. The words' semantic categories and unique first letters were provided as retrieval cues. In parallel to Experiment 4a, the cues were presented successively in a blocked, randomized manner for 8 sec each centrally on a computer screen: The sequence of categories was randomly chosen, but all items of one category were tested successively. Within restudied categories, relearned and related items were blocked, and it was randomly chosen which block was tested first. Subjects had 8 sec to recall each item, and were asked to write down their respective answers. After 8 sec, the next retrieval cue appeared on the screen. In the experimental block with interference, all initially studied items were tested first; the additionally encoded items were tested second. After the final test phase, all subjects were debriefed and thanked for their participation.

Results

Interference Manipulation

Mean recall of the additionally encoded items in the experimental block with interference was 63.4 % ($SD = 15.0$).

Effects of Interference on the Different Item Types

A 3 x 2 ANOVA with the factors of ITEM TYPE (restudied, related, control items) and INTERFERENCE (with interference, without interference) revealed significant main effects of ITEM TYPE, $F(2, 62) = 56.26$, $MSE = 227.35$, $p < .001$, and INTERFERENCE, $F(1, 31) = 20.64$, $MSE = 204.40$, $p < .01$. The main effect of ITEM TYPE reflects the pattern of better recall for restudied items than for related as well as for control items (see below, for details); the main effect of INTERFERENCE reflects a decrease in recall in the experimental block with interference. Critically, no significant interaction between the two factors was found, $F(2, 62) < 1.0$, suggesting that INTERFERENCE did not affect the three item types differently (see Figure 6.3 for a plot of the results).

To further investigate the proposal that interference impaired recall irrespective of whether the items themselves or their categories had been restudied, two separate post-hoc 2 x 2 ANOVAs with the factors of INTERFERENCE (with interference, without interference) and ITEM TYPE (restudied, control items; or control items, related items) were calculated. For restudied and control items, significant main effects of ITEM TYPE, $F(1, 31) = 88.06$, $MSE = 251.41$, $p < .001$, and INTERFERENCE, $F(1, 31) = 19.79$, $MSE = 166.82$, $p < .001$, arose. However, no significant interaction between the two factors was found, $F(1, 31) < 1.0$. Enhancement due to restudy was present in both experimental blocks, $ts(31) > 6.95$, $ps < .001$. Memory performance for both item types was equally impaired by the interference manipulation, $ts(31) \geq 3.00$, $ps < .01$, with recall of restudied items decreasing from 91.9 % correct to 84.1 % correct, and recall of control items sinking from 68.0 % correct to 55.5 % correct.

For control and related items, a significant main effect of INTERFERENCE, $F(1, 31) = 12.76$, $MSE = 258.66$, $p = .001$, was found, while the main effect of ITEM TYPE, $F(1, 31) = 1.97$, $MSE = 282.29$, $p > .15$, as well as the interaction between the two factors, $F(1, 31) < 1.0$, did not reach significance. No forgetting due to restudy of related items was present in any of the two experimental blocks, $ts(31) \leq 1.40$, $ps \geq .17$. Again, recall of both item types was impaired by the interference manipulation, $ts(31) > 2.00$, $ps \leq .05$; recall of control items decreased from 68.0 % correct to 55.5 % correct, and recall of related items decreased from 69.8 % correct to 62.0 % correct.

Discussion

As outlined above, an objection concerning Experiment 4a could be raised. It could be argued, that the protection from interference observed in this data set was not caused by retrieval-associated reactivation and stabilization, but rather

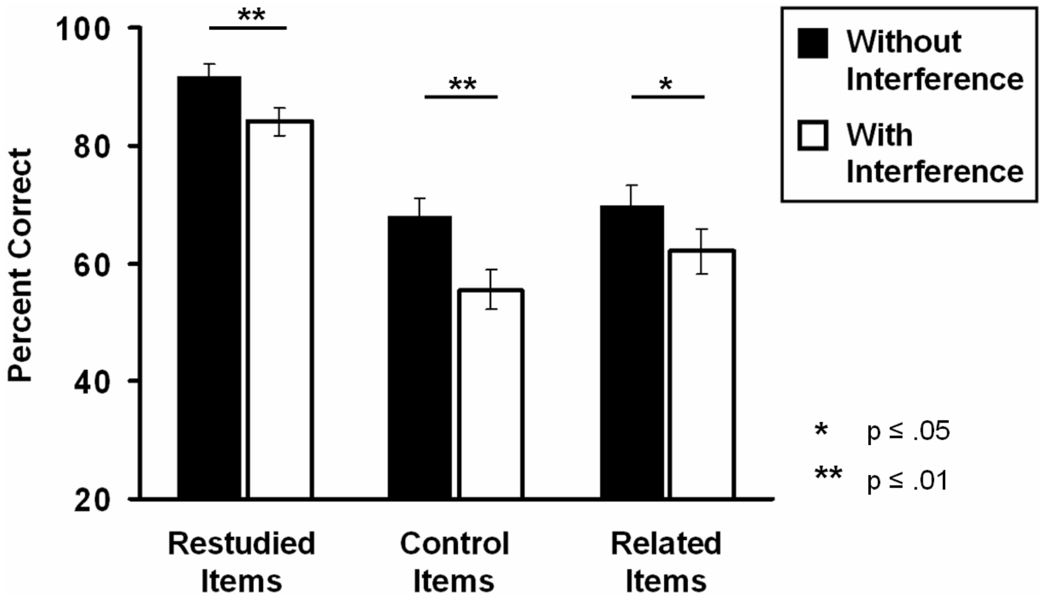

Figure 6.3: Results of *Experiment 4b*. Mean recall performance is shown for the three item types in the two experimental blocks. Error bars represent standard errors of the mean.

by mere reprocessing of the material. Yet, the results of Experiment 4b do not support this view. Memory reprocessing, when induced by additional restudy opportunities, did not yield any protection from retroactive interference. Memory performance for restudied as well as for related items was impaired by retroactive interference approximately to the same degree as for control items that did not experience any kind of reprocessing after initial study.

Therefore, as no evidence for a protective effect of restudy (and, thus, of mere reprocessing) was obtained, the results speak in favor of the hypothesis that retrieval stabilizes and protects memories. Of course, retrieval-induced reactivation can also be understood as a case of reprocessing; however, it appears to be a special case, because it seems to entail more powerful and pronounced effects than other types of reprocessing (e.g., restudy). This notion is well in line with findings on the testing effect (e.g., Roediger & Butler, 2011), showing that retrieval provides mnemonic benefits superior to those mediated by restudy. While most studies on the testing effect converge on the finding that retrieval is more beneficial than restudy after prolonged delays, the present results indicate that this benefit may already be detectable at an earlier stage. The results are also consistent with findings on retrieval-induced forgetting (e.g., Anderson et al., 2000; Staudigl et al., 2010), showing that retrieval provokes consequences quite different from those of restudy. In particular, the results are well in line with data provided by Halamish and Bjork (2011), who found less interference-induced impairment for a testing condition in comparison to a restudy condition.

6.3 Conclusions

The last two experiments showed that retrieval practice in particular, and not reprocessing in general, stabilizes memories and leaves them less prone to retroactive interference. Notably, this retrieval-associated benefit was not only evident for the retrieved material itself, but generalized to related, but nonretrieved material that was protected as well. All in all, this pattern resembles the data of Experiment 3, where it was shown that retrieval shields both the retrieved and related, but nonretrieved memories from time-dependent forgetting. Three different perspectives arise on the just presented findings and shall be described in the following.

One view arises when one considers that the retrieval-practice paradigm, which was used in these last three experiments, has previously brought forth a considerable amount of research (for reviews, see Anderson, 2003; Bäuml, 2008). The majority of these studies focused on the detrimental effects mediated by partial retrieval, thereby documenting the potentially negative side of retrieval practice (e.g., Macrae & MacLeod, 1999). It was demonstrated that such retrieval-induced forgetting can be of relevance in daily-life situations; for instance, Coman et al. (2009) showed that the social sharing of information in conversations can change the memories of both speakers and listeners by mediating forgetting of about the same information, thereby potentially impacting upon the contents of collective memory. Another example was brought forward by Dunn and Spellman (2003), who showed that the detrimental effects of retrieval practice may also be important in person perception; concentrating on stereotype-relevant traits can result in inhibition of specific individual aspects of a person's identity and vice versa, thereby leading to an altered perception of the respective person. All in all, most of the research applying the retrieval-practice paradigm has emphasized how partial retrieval can incidentally influence important aspects of our lives; not so much because the retrieved information is strengthened, but rather because related, nonretrieved information is weakened.

The data presented here suggest a different view, because a rather unexpected advantage seems to come with the detrimental impact of partial retrieval practice. Although the regular pattern of retrieval-induced enhancement and retrieval-induced forgetting was observed in the baseline conditions of Experiment 3 and Experiment 4a, both the directly retrieved and the related, but nonretrieved material were found to be shielded from further impairment due to time-dependent forgetting and interference in the experimental conditions. The standard perspective would be to regard the inhibited, unpracticed material as weakened; the findings presented here partly contradict this view by showing that such initially weakened material does not remain in a disrupted and fragile state, but is, as a consequence of partial retrieval practice, less susceptible to further detrimental influences than control material. As a consequence, the critical difference between control material and unpracticed material was abrogated after both a regular 12-h (wake) delay and after additional encoding of similar, interfering information. Ultimately, this finding suggests that the negative effects of retrieval practice may be at their greatest shortly after retrieval practice, but that they may very likely disappear with time and interfering experiences.

One situation that may be suitable to illustrate the importance of such a finding is the questioning of eyewitnesses after an accident or a criminal act has taken place. As previously shown by MacLeod (2002), the partial retrieval of specific information due to fragmentary or unbalanced questioning right after the witnessing of an event may result in forgetting of related facts that could be important for an appropriate assessment of the events, too. It was argued that incomplete questioning can entail consequences far worse for memory than if no questioning takes place at all. The results presented here suggest that this conclusion may be wrong. Even if, right after an incomplete questioning, specific information related to the actually retrieved material is less well remembered than if no information has been retrieved at all, it is very likely that this critical difference will diminish with time, because the material will show time-dependent forgetting when no retrieval has taken place, but should show no such forgetting when at least partial retrieval was encouraged. Thus, although effects of partial retrieval practice entail long-term consequences, the importance of retrieval-induced forgetting after longer delays may be overrated because the effect seems to diminish both with time and with the encoding of similar, interfering information. Nevertheless, the data presented here strengthen the proposal by Geiselman et al. (1984) who claimed that all information related to an eyewitness event should be retrieved as soon as possible. On the one hand this approach may avoid immediate retrieval-induced forgetting (e.g., Shaw et al., 1995), but more importantly its effectiveness may be based on the fact that the retrieved information itself is not only stabilized but also enhanced. From this perspective, the cognitive interview (e.g., Geiselman, Fisher, MacKinnon, & Holland, 1985), in which witnesses are encouraged to repeatedly recall all information related to an incident both from different perspectives and in various orders, seems to be a reasonable approach (at least as long as no misinformation is encountered; see Thomas et al., 2010).

A different view on the findings presented here arises when one takes the available literature on memory consolidation into consideration. As stated before, memory consolidation is assumed to become evident in less time-dependent forgetting and less susceptibility to retroactive interference (e.g., Wixted, 2004). Together, the present data provide evidence for the claim that memory retrieval may be related to memory consolidation, because they show that retrieval shields memories from both time-dependent forgetting and retroactive interference. The proposal is that active retrieval consolidates and stabilizes memories, thereby rendering them less susceptible to subsequent degradation and forgetting. Previously, similar evidence has been accumulated for sleep, reporting reduced normal forgetting and less susecptibility to interference after sleep compared to wakefulness (Ellenbogen et al., 2006a; Talamini et al., 2008). Thus, both sleep and retrieval seem to be associated with memory consolidation.

Another possible perspective on the present data is based on literature on the so-called testing effect (for a review, see Roediger & Butler, 2011). Retrieval from memory has previously been linked to less time-dependent forgetting and is, especially after longer delays, assumed to provide more distinct mnemonic benefits than mere restudy. The present results have several important implications for research on the testing effect. In comparison to previous research (e.g., Roediger & Karpicke, 2006b), here, the mnemonic advantage associated with retrieval became evident after a rather short delay interval of 12 hours as well as immediately after the induction of retroactive interference. As retrieval practice did only comprise two repetitions and, based on success rates, does not seem to have been difficult for the subjects, these findings are rather surprising. In contrast to previous conclusions (e.g., Pyc & Rawson, 2009) they suggest that also comparatively easy retrieval practice can entail benefits that are detectable after only a short delay and by means of probing the items' interference susceptibility. Moreover, the benefits of retrieval were shown to extend also to material related to the retrieved memories, which may offer further possibilities to investigate effects of testing. Ultimately, the results suggest that memory consolidation may lie at the heart of the testing effect. Of course this is just a proposal (for different accounts on the effects of retrieval, see for instance Kornell et al., 2011; Lasry et al., 2008; Pyc & Rawson, 2010); however, the proposal provides a plausible framework and testable hypotheses that may spark new research on the topic.

Although the data suggest an interesting parallel between sleep and retrieval, they do not indicate that effects of sleep and retrieval rely on the same mechanisms or entail the same consequences. For instance, sleep-associated reactivation is supposed to work more efficiently because the involved areas in the brain are not simultaneously occupied with perception (e.g., Born et al., 2006; Diekelmann, et al., 2009). Retrieval, on the other hand, seems to reactivate brain areas associated with encoding as well (e.g., Carr et al., 2011; Danker & Anderson, 2010; Eldrigde, Knowlton, Furmanski, Bookheimer, & Engel, 2000; Karlsson & Frank, 2009), but does so while perceptual input is still received. Another point of interest could be that sleep-associated reactivation is assumed to be related to a hippocampal-neocortical dialogue, which ensures the transfer of memories to sites of long-term storage (e.g., Born & Wilhelm, in press). That such a dialogue should also be initiated through retrieval-associated reactivation seems rather unlikely; not just because perceptual input needs to be continuously encoded during wakefulness, but also because specific sleep-related parameters have been identified in the past that may be involved in the coordination of the replay (e.g., Fogel & Smith, 2011). The reactivation of original memories during memory retrieval is only one proposal that may or may not be able to explain the mnemonic benefits related to retrieval (for further ideas, see Roediger & Butler, 2011). Future experiments should try to differentiate between the two types of reactivation (i.e., sleep-associated and retrieval-associated reactivation), and attempt to figure out what precisely it is that makes retrieval potent enough to leave its lasting traces irrespective of ongoing perceptual encoding.

Apparently, many questions remain unanswered or are initially raised by the present data. For instance, it remains unclear how exactly nonretrieved material is supposed to profit from retrieval of related material. If one assumes that the stabilization of directly retrieved memories relies on reactivation of the material, a similar line of argumentation could be followed for related, but nonretrieved memories: As there is evidence for the assumption that, during partial retrieval practice, related items interfere (e.g., Staudigl et al., 2010; Wimber, Rutschmann, Greenlee, & Bäuml, 2009) and are inhibited to enable the retrieval of target information (e.g., Anderson et al., 2000), it could be proposed that related material is reactivated as well during this chain of events. Following this reasoning, the reactivation and stabilization of related, but nonretrieved memories would only be an incidental by-product of selective retrieval. Although there is first evidence in favor of this idea (Kuhl, Rissman, Chun, & Wagner, 2011), it should be subject to closer investigation. For example it could be speculated that stabilization due to intentional and incidental reactivation should be linked to largely the same pattern of brain activity at later retrieval, further differentiating these two item types from control items. However, an alternative idea is that nonretrieved material is not incidentally reactivated during retrieval practice, but that subjects intentionally engage in covert retrieval of items related to the ones they are asked to retrieve.

Certainly, there are several conceivable proposals to explain the observed effects and several possible ways to investigate the potential role of retrieval-associated reactivation of practiced and related material proposed here. An interesting idea in this respect may also be to investigate whether the extent or frequency of reactivation after encoding is linked to the magnitude of retrieval-associated consolidation that can be observed. In the respective experiments presented here, two retrieval cycles were placed right after encoding; therefore it seems safe to say that this amount of retrieval-associated reactivation is sufficient to promote at least some detectable degree of consolidation. However, research on the testing effect indicates that both effort at retrieval and repetition of retrieval trials are supportive of the benefits of testing (e.g., Karpicke & Roediger, 2007; Pyc & Rawson, 2009). Thus, it might be interesting to manipulate retrieval difficulty or the number of retrieval cycles or both in future studies. Graded effects of retrieval-associated consolidation could then be observable either after prolonged retention intervals of weeks to months or after the induction of higher degrees of retroactive interference. The data presented in the last two chapters should in any case only be regarded as first evidence for the consolidating effect retrieval practice obviously can have. Whether this effect is subject to any kind of modulation or only emerges under specific conditions remains to be investigated in follow-up experiments.

Finally, it could be argued that the present findings are in conflict with recent reports on phenomena of reconsolidation after awake memory reactivation. Diekelmann, Büchel, Born, and Rasch (2011), for instance, examined whether the same kind of reactivation that had previously been reported to induce more effective consolidation if applied during sleep (Rasch et al., 2007), entailed the same or different effects if applied during wakefulness. Awake subjects were given reactivating odor cues 40 minutes after initial encoding and before additional, highly interfering learning. Interference after odor-related reactivation lead to more pronounced impairment in comparison to a control condition, in which no reactivation was provoked. Diekelmann et al. (2011) concluded that reactivation during wakefulness brings memories back into a labile state, in which they are again subject to modification and, thus, also to impairment. Although these results may at first glance seem to contradict the present results, they actually fit quite well within the proposed framework: As it was shown that retroactive interference specifically left retrieval practiced items unaffected, but not restudied items, the finding that subtle odor cues were not sufficient to promote reactivation sheltering from impairment extends the notion that retrieval is a special case of awake reactivation that, in comparison to less elaborate reprocessing, entails powerful mnemonic effects.

Part III

Summary

The data presented here suggest that both sleep and retrieval may consolidate memories. In the first part of this thesis, evidence for sleep-associated memory consolidation was collected by investigating sleep's effects on interference and list-method directed forgetting. Experiments 1a and 1b replicated and extended previous findings by Ekstrand (1967) and Drosopoulos et al. (2007). Applying paired-associate learning, it was shown that sleep counteracts both retroactive and proactive interference because memory performance for both double lists was found to be better after sleep compared to wake. As sleep-related benefits were found to be more distinct and also present for single lists under conditions of weak encoding, the results indicate that memory strength may indeed influence sleep-associated consolidation. Therefore, previous results are unlikely to have arisen due to strong encoding conditions, and sleep seems to counteract interference in a more general way.

Yet, sleep-associated memory consolidation was found to be preferentially present for double lists assumed to be reduced in accessibility (e.g., Tulving & Psotka, 1971). By preferentially stabilizing memories reduced in accessibility sleep did not only counteract time-dependent forgetting, but also forgetting experimentally induced by an interference manipulation. This finding suggests that it might indeed be advisable to go to sleep after one has, for instance, engaged in exam preparation or other learning scenarios during the day. When information on a complex topic is encoded, the single bits of information are usually related to at least one superordinate cue and may compete for recall when one tries to remember everything related to the cue. The present data on sleep's impact upon effects of interference do not only show that memory performance after sleep is better than after wake, they do more precisely show that sleep can abrogate interference and shape memory performance in such a way that information can be equally well remembered regardless of whether interfering learning occurred or not (this is at least the case for effects of retroactive interference, and most pronounced under conditions of weak encoding). Clearly, if this data set was the only one at hand, a reasonable conclusion would be to generally suggest to go to sleep after learning.

However, further evidence for sleep-associated memory consolidation was provided by investigating the effect of sleep and wake delays on list-method directed forgetting. The obtained pattern of results is interesting on more than one level. First, since previous studies only placed short distractor intervals of up to 5 minutes after encoding (e.g., Racsmany et al., 2008), the results show for the first time that list-method directed forgetting can persist over 12-h wake delays. Second, the results indicate that sleep may modulate the role of delay for directed forgetting by providing its benefits not only for memory contents one considers as relevant and wants to maintain, but to an even higher degree for information regarded as outdated, thereby abrogating the directed-forgetting effect. Third, as the forget cue is assumed to reduce the whole list's accessibility (e.g., Bäuml et al., 2010), sleep again was shown to abrogate an experimentally induced form of forgetting by preferentially strengthening memories that are harder to access. Notably, while Experiments 1a and 1b showed that sleep can abrogate incidental forgetting, Experiment 2 showed that it can also abrogate intentional forgetting.

Intriguingly, these data pose a whole different perspective on sleep's potential role in effective learning. The forgetting of outdated information has previously been related to processes of updating (e.g., Bäuml et al., 2008), as irrelevant information (here, list 1) is forgotten for the sake of highlighting more relevant and current information in memory (here, list 2). Although the presented data do only take list-1 performance into consideration and thus can not answer the question whether the finding of sleep reviving list-1 memories also impacted upon list-2 performance, they nevertheless indicate that the possibility exists. If so, the preferential consolidation of list-1 memories might hinder successful updating by affecting memory performance for the relevant information. In this scenario, advising students to generally go to sleep right after a day of exam preparations might be of little usefulness, at least if the day included the encoding of information that, in hindsight, turned out to be irrelevant and required processes of updating. However, Experiment 2 can clearly not resolve this issue; more data (preferentially to be collected in more applied settings, too) are needed to evaluate the exact impact the observed revival of outdated information may have on effective learning.

To sum up, the first part of this thesis documents effects of sleep-associated consolidation. All in all, across two experimental tasks and three experiments, memory performance was in general found to be better after 12 hours of nocturnal sleep than after 12 hours of diurnal wakefulness. In particular, sleep especially seemed to benefit memories that were experimentally reduced in accessibility, and, by doing so, worked against two forms of induced forgetting (i.e., forgetting due to retroactive interference, and list-method directed forgetting). Therefore, it is concluded that sleep does not only counteract normal (time-dependent) forgetting by stabilizing memories in general (e.g., Diekelmann & Born, 2010); in addition, sleep may also abrogate both incidental and voluntary forgetting by enabling the more extensive consolidation of memory traces reduced in accessibility.

In the second part of this thesis, first evidence for retrieval-associated memory consolidation was collected by applying the retrieval-practice paradigm (Anderson et al., 1994) and by investigating effects of retrieval practice on normal forgetting and susceptibility to interference. The data of Experiment 3 are again interesting on several different levels. First, the results replicate and extend previous findings by Racsmany et al. (2010), who showed that sleep can modulate the role of delay for retrieval-induced forgetting: While the effect was present at baseline, it was found to be abrogated after 12 hours of wakefulness and to be intact after 12 hours of sleep. By providing appropriate control groups and investigating the time-dependent forgetting of the single item types, it further was shown that this modulation relies on the fact that only control items benefit from sleep-associated consolidation, whereas practiced and unpracticed items are stable across both delays. Second, this finding indicates that prior retrieval may influence subsequent sleep-associated memory consolidation, because both practiced and related, unpracticed items did not additionally profit from any sleep-related benefits. Third, the results replicate prior work on the testing effect (e.g., Roediger & Butler, 2011) by documenting that retrieval stabilizes the practiced memories and protects them from time-dependent forgetting. This finding indicates that the mnemonic benefits associated with retrieval may, even under relatively simple retrieval conditions and with only two retrieval cycles, be detectable after rather short delay intervals (for previous results suggesting otherwise, see Pyc & Rawson, 2009, or Roediger & Karpicke, 2006b). The results further extend the prior work by showing that not only practiced items are protected from time-dependent forgetting, but that also unpracticed items may benefit from this protection (see also Chan, 2009, 2010). Ultimately, the results suggest that the effects of retrieval practice fulfill the first of two criteria claimed for effective memory consolidation (e.g., Wixted, 2004), as performance for both practiced and unpracticed material was found to remain stable over time.

Further evidence for retrieval-associated consolidation was obtained by probing whether retrieval also affects memories' interference susceptibility. In Experiment 4a, retroactive interference was elicited after partial retrieval practice and memory performance was compared to a baseline condition without interference. While the regular pattern of retrieval-induced enhancement and forgetting was present at baseline, retroactive interference increased retrieval-induced enhancement and abrogated retrieval-induced forgetting by selectively impairing memory performance for control items. Both practiced and unpracticed items were unaffected by retroactive interference. Experiment 4b further clarified that this finding was retrieval-specific and not due to more general reprocessing: Retroactive interference was elicited after partial restudy and memory performance was again compared to a baseline condition without interference. While retrieval practice in Experiment 4a had protected both the practiced and unpracticed items from interference, restudy in Experiment 4b did not entail such a protection; memory performance for all item types was equally impaired by retroactive interference. Therefore, in contrast to effects of restudy, effects of retrieval practice also fulfill the second criterion claimed for effective memory consolidation (e.g., Wixted, 2004), as memory performance for both practiced and unpracticed material was found to remain stable irrespective of interpolated learning.

These results further extend the prior work on the testing effect (e.g., Roediger & Butler, 2011) which showed that retrieval stabilizes memory performance for the retrieved material over time. Other previous reports indicate that the benefits of testing may also be evident in facilitated future encoding; it has been argued that such facilitation may rely on a reset of encoding (e.g., Pastötter et al., 2011) or a reduction of proactive interference (e.g., Szpunar et al., 2008) mediated by retrieval of previously encoded information. The results presented here indicate that retrieval also protects both practiced and related, unpracticed material from retroactive interference (for a related result, see Halamish & Bjork, 2011). Therefore, the benefits of testing may not only be evident in less time-dependent forgetting, but can already be observed rather shortly after retrieval (albeit by help of other paradigms and behavioral measures). Prior research moreover indicates that the previously described isolation against proactive interference may also be important for daily life, because the retrieval of initially encoded face-name pairs was shown to facilitate the encoding of new face-name pairs (Weinstein et al., 2011). The data at hand indeed suggest that retrieving face-name pairs should also protect these associations from subsequent encoding of other face-name pairs; if so, retrieval may counteract effects of interference in general and, as a consequence, constitute a valuable mnemonic aid that can be helpful in a variety of daily life situations in which interference is encountered (e.g., during the first sessions of a term, when university lecturers successively get to know the names and faces of the students taking part in their seminars).

To summarize, the second part of this thesis documents effects of retrieval-associated memory consolidation. Applying the retrieval-practice paradigm, it was shown across three experiments that specifically retrieval-related reprocessing protected memories from both time-dependent forgetting and retroactive interference. In particular, it was found that the benefits of retrieval were not restricted to the retrieved material itself, but spread to nonretrieved, yet related material which, as a consequence, was also stabilized and protected from time-dependent forgetting and retroactive interference.

All in all, this thesis links both sleep and retrieval to processes of memory consolidation. However, the data presented here do not speak to the question of what exact mechanisms lie at the bottom of the observed stabilization effects. As outlined in the first part, for sleep-associated memory consolidation a reactivation of memories during slow-wave sleep has been proposed to account for its entailed effects (e.g., Rasch et al., 2007; Rudoy et al., 2009). Although the reported results suggest that retrieval also stabilizes memories across time and protects them from retroactive interference, findings that are very similar to previous findings on sleep-associated memory consolidation (e.g., Ellenbogen et al., 2006a; Talamini et al., 2008), this parallel does not imply that sleep and retrieval also trigger similar processes. Indeed, neurocognitive studies showing that retrieval induces the awake reactivation of original memory traces (e.g., Carr et al., 2011; Kuhl et al., 2011) could be regarded as supportive of the idea that similar mechanisms are initiated by sleep and retrieval. Moreover, it could be argued that such a perspective could also explain why unpracticed material shows the same retrieval-associated benefits as practiced material, because previous work on retrieval-induced forgetting indicates that unpracticed items interfere during retrieval practice and, thus, are incidentally reactivated as well (e.g., Kuhl et al., 2007; Staudigl et al., 2010). However, the data at hand do not bear any features that support or contradict this view, and although the formal criteria linking retrieval to consolidation processes are fulfilled (e.g., Wixted, 2004) other accounts for the benefits related to testing have previously been proposed and seem plausible as well.

For instance, it was suggested that retrieval entails specific benefits by prompting the elaboration of memory traces (e.g., by creating new retrieval routes, Carpenter, 2009), by initiating better organization of the material (e.g., Zaromb & Roediger, 2010), or by establishing effective mediators between cue and target information (e.g., Pyc & Rawson, 2010). To assume that these and similar mechanisms could affect memory performance in the same way as is documented in the second part of this thesis is justifiable. Although, at first glance, all the proposed accounts only seem to be able to explain the benefits observed for practiced material, but to hardly offer a possibility to extend the respective reasoning to unpracticed material, it should be noted that a very simple assumption could do the trick: If one assumes that participants do not only engage in overt retrieval of the items they are cued to recall during the retrieval-practice phase, but additionally engage in partial covert retrieval of material also related to the superordinate cue (e.g., a category cue, as applied in the experiments described here), the same reasoning could be employed to explain the benefits for practiced and unpracticed material. As covert activity of subjects is difficult to control for, this assumption can not easily be discarded. Future research should try to figure out more precisely what it is that benefits memory during retrieval, and thereby clarify whether the mechanisms underlying retrieval-associated and sleep-associated memory consolidation are very much alike or totally different.

Nevertheless, further differences or parallels between sleep-associated and retrieval-associated consolidation may in general be interesting to investigate. Until now, both research areas on the mnemonic effects of sleep and retrieval existed completely independent of each other. Clearly established findings in one area could stimulate further research in the other area, thereby providing deeper insight into whether sleep and retrieval entail the same stabilizing effects mediated by roughly the same mechanisms or whether they differ in both underlying mechanisms and the more detailed outcomes. For instance, memory retrieval has been argued to lead to a reset of encoding (e.g., Pastötter et al., 2011), or to protect memories from proactive interference (e.g., Szpunar et al., 2008), while no such finding has been reported for sleep yet. The other way around, a role for sleep in the integration, abstraction, and schematization of memories has recently been discussed (e.g., Lewis & Durrant, 2011), whereas no connection between retrieval and such processes has been made.

In addition to the possibility of mutual scientific inspiration, a combination of the two research areas on sleep and retrieval may be of help in more applied settings as well. For example, sleep-dependent memory consolidation has been shown to be impaired in several psychiatric disorders (e.g., Genzel, Ali, Dresler, Steiger, & Tesfaye, 2011), findings that have been discussed to potentially underly the general memory impairments observed in corresponding and other clinical populations. Retrieval, in contrast, has been found to still provide mnemonic benefits for specific neurologically impaired patient groups (e.g., Sumowski, Chiaravalloti, & DeLuca, 2010a; Sumowski et al., 2010b). Therefore, the scenario of one consolidation mechanism 'helping out' when the other is for some reason deficient does not categorically seem far-fetched. Until now, however, it goes without saying that data in support of such an idea are still to be collected.

Pertaining to the results presented in this thesis, one data point is particularly suitable to discuss potential parallels or differences between sleep-associated and retrieval-associated memory consolidation. It could be argued, that the finding of more pronounced sleep-associated memory consolidation for memory contents that are harder to access (see also Drosopoulos et al., 2007) resembled the previously reported finding that benefits of retrieval turn out to be more distinct when retrieval is made more difficult (e.g., Pyc & Rawson, 2009). Underlying this reasoning is the assumption that sleep-associated reactivation grows 'more difficult' when memory contents are hard to access, and, in general, that effects of reactivation are boosted when reactivation is more complicated. The finding that difficulty at encoding increases sleep spindle density (Schmidt et al., 2006), a sleep parameter that has previously been linked to memory consolidation (e.g., Fogel & Smith, 2011), could be regarded as speaking in favor of this proposal. Yet, sleep-associated consolidation has also been found to be more effective when a task is already acquired to a high degree before sleep (e.g., Tucker & Fishbein, 2008). As a higher degree of task acquisition should facilitate the potential reactivation of contents, this finding contradicts the above reasoning. However, to my knowledge nothing is known about whether degree of task acquisition also impacts upon the effects of retrieval practice. Be that as it may, this paragraph illustrates one important point: All in all, both within and across the separate research areas little consistency has been reached so far. A more systematic and unified approach might be helpful to deepen our understanding of memory consolidation in general.

If future research further corroborates the involvement of retrieval in memory consolidation, it will be interesting to examine how best to optimize it. Transcranial brain stimulation during sleep has previously been shown to be able to boost memory consolidation (Marshall, Helgadottir, Mölle, & Born, 2006; Marshall et al., 2011), whereas it was only found to enhance encoding when applied during awake learning, without affecting memory consolidation when applied afterwards (Kirov, Weiss, Siebner, Born, & Marshall, 2009). Similarly, previous research indicates that retrieval may also offer possibilities to optimize its stabilizing impact. For instance, Finn and Roediger (2011) showed that inducing emotional arousal right after immediate retrieval can enhance the testing effect. While the authors embedded their results in a reconsolidation framework, the current data propose that they may also point to the possibility that retrieval-associated memory consolidation can be very easily further optimized, e.g., by inducing emotional arousal right after successful retrieval practice.

Arguably, whatever underlies retrieval-associated memory consolidation may be manipulated and optimized in a less complicated way than what underlies sleep-associated memory consolidation. For retrieval-associated consolidation, no special experimental apparatus is necessary to induce the manipulation and no specific time of day or sleep stage has to be awaited. Presumably, retrieval-associated consolidation can be triggered anywhere and anytime. Above all, however, memory retrieval during wakefulness has one major advantage to memory reactivation during sleep: Memory retrieval can be voluntarily induced by humans. Therefore, retrieval entails the possibility to decide which specific memories should be subject to the associated stabilization, thereby enabling its strategic use as a potent memory modifier. Such targeted and deliberate stabilization is - at least until now and for the bigger part of mankind - not possible during sleep.

References

Agarwal, P. K., Karpicke, J. D., Kang, S. H. K., Roediger III, H. L., & McDermott, K. B. (2008). Examining the testing effect with open- and closed-book tests. *Applied Cognitive Psychology*, *22*, 861-876.

Alberini, C. M., Milekic, M. H., & Tronel, S. (2006). Mechanisms of memory stabilization and de-stabilization. *Cellular and Molecular Life Sciences*, *63*, 999-1008.

Allen, G. A., Mahler, W. A., & Estes, W. K. (1969). Effects of recall tests on long-term retention of paired associates. *Journal of Verbal Learning and Verbal Behavior*, *8*, 463-470.

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanism of forgetting. *Journal of Memory and Language*, *49*, 415-445.

Anderson, M. C., Bjork, E. L. & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin & Review, 7*, 522-530.

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1063-1087.

Anderson, M. C. & McCulloch, K. C. (1999). Integration as a general boundary condition on retrieval-induced forgetting. *Journal* *of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 608-629.

Anderson, M. C. & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In: E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 237-313). San Diego, CA: Academic Press.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review*, *102*, 68-100.

Anderson, J. R. & Schooler, L. J. (2000). The adaptive nature of memory. In E. Tulving (Ed.), *Oxford Handbook of Memory* (pp. 557-570). Oxford: University Press.

Axmacher, N., Draguhn, A., Elger, C. E., & Fell, J. (2009). Memory processes during sleep: Beyond the standard consolidation theory. *Cellular and Molecular Life Sciences*, *66*, 2285-2297.

Baddeley, A. (1990). *Human memory: Theory and practice.* Needham Heights, MA: Allyn & Bacon.

Bäuml, K.-H. (2002). Semantic generation can cause episodic forgetting. *Psychological Science*, *13*, 356-360.

Bäuml, K.-H. (2008). Inhibitory processes. In: H. L. Roediger III (Ed.), *Cognitive psychology of memory.* Vol. 2 of Learning and memory - A comprehensive reference (pp. 195-220). Oxford: Elsevier.

Bäuml, K.-H. & Aslan, A. (2004). Part-list cuing as instructed retrieval inhibition. *Memory & Cognition*, *32*, 610-617.

Bäuml, K.-H., Hanslmayr, S., Pastötter, B., & Klimesch, W. (2008). Oscillatory correlates of intentional updating in episodic memory. *NeuroImage*, *41*, 596-604.

Bäuml, K.-H. & Hartinger, A. (2002). On the role of item similarity in retrieval-induced forgetting. *Memory*, *10*, 215-224.

Bäuml, K.-H. T., Pastötter, B., & Hanslmayr, S. (2010). Binding and inhibition in episodic memory - Cognitive, emotional, and neural processes. *Neuroscience and Biobehavioral Reviews*, *34*, 1047-1054.

Bäuml, K.-H. & Kuhbandner, C. (2007). Remembering can cause forgetting - but not in negative moods. *Psychological Science, 18*, 111-115.

Bäuml, K.-H. T. & Samenieh, A. (2010). The two faces of memory retrieval. *Psychological Science*, *21*, 793-795.

Barnes, J. M. & Underwood, B. J. (1959). “Fate” of first-list associations in transfer theory. *Journal of Experimental Psychology*, *58*, 97-105.

Benjamin, A. S. (2007). The effects of list-method directed forgetting on recognition memory. *Psychonomic Bulletin & Review*, *13*, 831-836.

Bjork, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior*, *9*, 255-268.

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In: H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309-330). Hillsdale, NJ: Erlbaum.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In: J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185-205). Cambridge, MA: MIT Press.

Bjork, R. A., LaBerge, D., & LeGrand, R. (1968). The modification of short-term memory through instructions to forget. *Psychonomic Science*, *10*, 55-56.

Born, J. (2010). Slow-wave sleep and the consolidation of long-term memory. *World Journal of Biological Psychiatry*, *11 (S1)*, 16-21.

Born, J., Rasch, B., & Gais, S. (2006). Sleep to remember. *The Neuroscientist*, *12*, 410-424.

Born, J. & Wilhelm, I. (in press). System consolidation of memory during sleep. *Psychological Research.*

Brickenkamp, R. (2002). *Test d2 - Aufmerksamkeits-Belastungs-Test [The d2 test of attention].* Göttingen: Hogrefe.

Brown, A. S. (1976). Spontaneous recovery in human learning. *Psychological Bulletin*, *83*, 321-338.

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *36*, 1118-1133.

Butler, A. C. & Roediger III, H. L. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology*, *19*, 514-527.

Butler, A. C. & Roediger III, H. L. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & Cognition*, *36*, 604-616.

Cai, D. J., Mednick, S. A., Harrison, E. M., Kanady, J. C., & Mednick, S. C. (2009). REM, not incubation, improves creativity by priming associative networks. *Proceedings of the National Academy of Sciences*, *106*, 10130-10134.

Camp, G., Pecher, D., & Schmidt, H. G. (2007). No retrieval-induced forgetting using item-specific independent cues: Evidence against a general inhibitory account. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *33*, 950-958.

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1563-1569.

Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1547-1552.

Carpenter, S. K. & Pashler, H. (2007). Testing beyond words: Using tests to enhance visuospatial map learning. *Psychonomic Bulletin & Review*, *14*, 474-478.

Carpenter, S. K., Pashler, H., Wixted, J. T., & Vul, E. (2008). The effects of tests on learning and forgetting. *Memory & Cognition*, *36*, 438-448.

Carr, M. F., Jadhav, S. P., & Frank, L. M. (2011). Hippocampal replay in the awake state: A potential substrate for memory consolidation and retrieval. *Nature Neuroscience*, *14*, 147-153.

Chan, J. C. K. (2009). When does retrieval induce forgetting and when does it induce facilitation? Implications for retrieval inhibition, testing effect, and text processing. *Journal of Memory and Language*, *61*, 153-170.

Chan, J. C. K. (2010). Long-term effects of testing on the recall of nontested materials. *Memory*, *18*, 49-57.

Chan, J. C. K. & Langley, M. M. (2011). Paradoxical effects of testing: Retrieval enhances both accurate recall and suggestibility in eyewitnesses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 248-255.

Chan, J. C. K., McDermott, K. B., & Roediger III, H. L. (2006). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General*, *135*, 553-571.

Ciranni, M. A. & Shimamura, A. P. (1999). Retrieval-induced forgetting in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1403-1414.

Cirelli, C. & Tononi, G. (2008). Is sleep essential? *PLoS Biology*, *6*, 1605-1611.

Coman, A., Manier, D., & Hirst, W. (2009). Forgetting the unforgettable through conversation: Socially shared retrieval-induced forgetting of September 11 memories. *Psychological Science*, *20*, 627-633.

Conway, M. A., Harries, K., Noyes, J., Racsmany, M., & Frankish, C. R. (2000). The disruption and dissilution of directed forgetting: Inhibitory control of memory. *Journal of Memory and Language*, *43*, 409-430.

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

Cuc, A., Koppel, J., & Hirst, W. (2007). Silence is not golden: A case for socially shared retrieval-induced forgetting. *Psychological Science*, *18*, 727-733.

Danker, J. F. & Anderson, J. R. (2010). The ghosts of brain states past: Remembering reactivates the brain regions engaged during encoding. *Psychological Bulletin*, *136*, 87-102.

Daurat, A., Terrier, P., Foret, J., & Tiberge, M. (2007). Slow wave sleep and recollection in recognition memory. *Conscious Cognition*, *16*, 445-455.

Debarnot, U., Castellani, E., Valenza, G., Sebastiani, L., & Guillot, A. (2011). Daytime naps improve motor imagery learning. *Cognitive, Affective, & Behavioral Neuroscience*, *11*, 541-550.

Debarnot, U., Creveaux, T., Collet, C., Gemignani, A., Massarelli, R., Doyon, J., & Guillot, A. (2009). Sleep-related improvements in motor learning following mental practice. *Brain and Cognition*, *69*, 398-405.

Dement, W. C. (2003). Knocking on Kleitmans door: The view from 50 years later. *Sleep Medicine Reviews*, *7*, 289-292.

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

Diekelmann, S., Büchel, C., Born, J., & Rasch, B. (2011). Labile or stable: Opposing consequences for memory when reactivated during waking and sleep. *Nature Neuroscience*, *14*, 381-386.

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

Drosopoulos, S., Harrer, D., & Born, J. (2011). Sleep and awareness about presence of regularity speed the transition from implicit to explicit knowledge. *Biological Psychology*, *86*, 168-173.

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*, 169-183.

Drosopoulos, S., Wagner, U., & Born, J. (2005). Sleep enhances explicit recollection in recognition memory. *Learning & Memory*, *12*, 44-51.

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annual Review of Psycholgy*, *55*, 51-86.

Dunn, E. W. & Spellman, B. A. (2003). Forgetting by remembering: Stereotype inhibition through rehearsal of alternative aspects of identity. *Journal of Experimental Social Psychology*, *39*, 420-433.

Durrant, S. J., Taylor, C., Cairney, S., & Lewis, P. A. (2011). Sleep-dependent consolidation of statistical learning. *Neuropsychologia*, *49*, 1322-1331.

Ekstrand, B. R. (1967). Effect of sleep on memory. *Journal of Experimental Psychology*, *75*, 64-72.

Ekstrand, B. R., Sullivan, M. J., Parker, D. F., & West, J. N. (1971). Spontaneous recovery and sleep. *Journal of Experimental Psychology*, *88*, 142-144.

Eldrigde, L. L., Knowlton, B. J., Furmanski, C. F., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: A selective role for the hippocampus during retrieval. *Nature Neuroscience*, *3*, 1149-1152.

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

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

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

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

Ferrara, M., Iaria, G., Tempesta, D., Curcio., G., Moroni, F., Marzano, C., De Gennaro, L., & Pacitti, C. (2008). Sleep to find your way: The role of sleep in the consolidation of memory for navigation in humans. *Hippocampus*, *18*, 844-851.

Ficca, G., Lombardo, P., Rossi, L., & Salzarulo, P. (2000). Morning recall of verbal material depends on prior sleep organization. *Behavioural Brain Research*, *112*, 159-163.

Finn, B. & Roediger III, H. L. (2011). Enhancing retention through reconsolidation: Negative emotional arousal following retrieval enhances later recall. *Psychological Science*, *22*, 781-786.

Fischer, S., Diekelmann, S., & Born, J. (2011). Sleep's role in the processing of unwanted memories. *Journal of Sleep Research*, *20*, 267-274.

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

Fogel, S. M. & Smith, C. T. (2011). The function of the sleep spindle: A physiological index of intelligence and a mechanism for sleep-dependent memory consolidation. *Neuroscience & Biobehavioral Reviews*, *35*, 1154-1165.

Foster, N. L. & Sahakyan, L. (2011). The role of forget-cue salience in list-method directed forgetting. *Memory*, *19*, 110-117.

Gais, S., Albouy, G., Boly, M., Dang-Vu, T. T., Darsaud, A., Desseilles, M., Rauchs, G., Schabus, M., Sterpenich, V., Vandewalle, G., Maquet, P., & Peigneux, P. (2007). Sleep transforms the cerebral trace of declarative memories. *Proceedings of the National Academy of Sciences*, *104*, 18778-18783.

Gais, S., Köster, S., Sprenger, A., Bethke, J., Heide, W., & Kimmig, H. (2008). Sleep is required for improving reaction times after training on a procedural visuo-motor task. *Neurobiology of Learning and Memory*, *90*, 610-615.

Gais, S., Lucas, B., & Born, J. (2006). Sleep after learning aids memory recall. *Learning & Memory*, *13*, 259-263.

Garcia-Bajos, E., Migueles, M., & Anderson, M. C. (2009). Script knowledge modulates retrieval-induced forgetting for eyewitness events. *Memory*, *17*, 92-103.

Gardiner, J. M., Craik, F. I. M., & Birtwistle, J. (1972). Retrieval cues and release from proactive inhibition. *Journal of Verbal Learning and Verbal Behavior*, *11*, 778-783.

Gardiner, J. M., Craik, F. I. M., & Bleasdale, F. A. (1973). Retrieval difficulty and subsequent recall. *Memory & Cognition*, *1*, 213-216.

Gates, A. I. (1917). Recitation as a factor in memorizing. *Archives of Psychology*, *40*, 1-104.

Geiselman, R. E., Bjork, R. A., & Fishman, D.L. (1983). Disrupted retrieval in directed forgetting: A link with posthypnotic amnesia. *Journal of Experimental Psychology: General*, *112*, 58-72.

Geiselman, R. E., Fisher, R. P., Firstenberg, I., Hutton, L. A., Sullivan, S., Avetissian, I., & Prosk, A. (1984). Enhancement of eyewitness memory: An empirical evaluation of the cognitive interview. *Journal of Police Science and Administration*, *12*, 74-80.

Geiselman, R. E., Fisher, R. P., MacKinnon, D. P., & Holland, H. L. (1985). Eyewitness memory enhancement in the police interview: Cognitive retrieval mnemonics versus hypnosis. *Journal of Applied Psychology*, *70*, 401-412.

Genzel, L., Ali, E., Dresler, M., Steiger, A., & Tesfaye, M. (2011). Sleep-dependent memory consolidation of a new task is inhibited in psychiatric patients. *Journal of Psychatric Research*, *45*, 555-560.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The Neural Bases of Cognitive Conflict and Control in Moral Judgment. *Neuron*, *44*, 389-400.

Groome, D. & Grant, N. (2005). Retrieval-induced forgetting is inversely related to everyday cognitive failures. *British Journal of Psychology*, *96*, 313-319.

Halamish, V. & Bjork, R. A. (2011). When does testing enhance retention? A distribution-based interpretation of retrieval as a memory modifier. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 801-812.

Hicks, J. L. & Starns, J. J. (2004). Retrieval-induced forgetting occurs in tests of item recognition. *Psychonomic Bulletin & Review*, *11*, 125-130.

Hinze, S. R. & Wiley, J. (2011). Testing the limits of testing effects using completion tests. *Memory*, *19*, 290-304.

Hobson, J. A., Pace-Schott, E. F., & Stickgold, R. (2000). Dreaming and the brain: Toward a cognitive neuroscience of conscious states. *Behavioral and Brain Sciences*, *23*, 793-842.

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

Hogan, R. M. & Kintsch, W. (1971). Differential effects of study and test trials on long-term recognition and recall. *Journal of Verbal Learning and Verbal Behavior*, *10*, 562-567.

Hupbach, A., Gomez, R., Hardt, O., & Nadel, L. (2007). Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learning & Memory*, *14*, 47-53.

Hupbach, A., Gomez, R., & Nadel, L. (2011). Episodic memory updating: The role of context familiarity. *Psychonomic Bulletin & Review*, *18*, 787-797.

Hupbach, A., Hardt, O., Gomez, R., & Nadel, L. (2008). The dynamics of memory: Context-dependent updating. *Learning & Memory*, *15*, 574-579.

Imeri, L. & Opp, M. R. (2009). How (and why) the immune system makes us sleep. *Nature Reviews Neuroscience*, *10*, 199-210.

Jacoby, L. L., Wahlheim, C. N., & Coane, J. H. (2010). Test-enhanced learning of natural concepts: Effects on recognition memory, classification, and metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1441-1451.

Jakab, E. & Raaijmakers, J. G. W. (2009). The role of item strength in retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 607-617.

Jenkins, J. G. & Dallenbach, K. M. (1924). Obliviscence during sleep and waking. *The American Journal of Psychology*, *25*, 605-612.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. A. (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, *59*, 193-224.

Kang, S. H. K., McDermott, K. B., & Roediger III, H. L. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, *19*, 528-558.

Karlsson, M. P. & Frank, L. M. (2009). Awake replay of remote experiences in the hippocampus. *Nature Neuroscience, 12*, 913-918.

Karpicke, J. D. & Bauernschmidt, A. (2011). Spaced retrieval: Absolute spacing enhances learning regardless of relative spacing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1250-1257.

Karpicke, J. D. & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, *331*, 772-775.

Karpicke, J. D., Butler, A. C., & Roediger III, H. L. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory*, *17*, 471-479.

Karpicke, J. D. & Roediger III, H. L. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language*, *57*, 151-162.

Karpicke, J. D. & Roediger III, H. L. (2010). Is expanding retrieval a superior method for learning text materials? *Memory & Cognition*, *38*, 116-124.

Keppel, G. (1968). Retroactive and proactive inhibition. In: T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory.* Englewood Cliffs, NJ: Prentice Hall.

Keppel, G. & Underwood, B J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, *1*, 153-161.

Kirov, R., Weiss, C., Siebner, H. R., Born, J., & Marshall, L. (2009). Slow oscillation electrical brain stimulation during waking promotes EEG theta activity and memory encoding. *Proceedings of the National Academy of Sciences*, *106*, 15460-15465.

Koessler, S., Engler, H., Riether, C., & Kissler, J. (2009). No retrieval-induced forgetting under stress. *Psychological Science, 20*, 1356-1363.

Kornell, N. & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, *14*, 219-224.

Kornell, N., Bjork, R. A., & Garcia, M. A. (2011). Why tests appear to prevent forgetting: A distribution-based bifurcation model. *Journal of Memory and Language*, *65*, 85-97.

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 989-998.

Kuhl, B. A., Dudukovic, N. M., Kahn, I., & Wagner, A. D. (2007). Decreased demands on cognitive control reveal the neural processing benefits of forgetting. *Nature Neuroscience*, *10*, 908-914.

Kuhl, B. A., Rissman, J., Chun, M. M., & Wagner, A. D. (2011). Fidelity of neural reactivation reveals competition between memories. *Proceedings of the National Academy of Sciences*, *108*, 5903-5908.

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.

Larsen, D. P., Butler, A. C., & Roediger III, H. L. (2008). Test-enhanced learning in medical education. *Medical Education*, *42*, 959-966.

Lasry, N., Levy, E., & Tremblay, J. (2008). Making memories, again. *Science*, *320*, 1720.

Levy, B. J. & Anderson, M. C. (2008). Individual differences in the suppression of unwanted memories: The executive deficit hypothesis. *Acta Psychologica*, *127*, 623-635.

Levy, B. J., McVeigh, N. D., Marful, A., & Anderson, M. C. (2007). Inhibiting your native language: The role of retrieval-induced forgetting during second language acquisition. *Psychological Science*, *18*, 29-34.

Lewis, P. A. & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Sciences*, *15*, 343-351.

MacLeod, C. M. (1998). Directed forgetting. In: J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 1-57). Mahwah, NJ: Erlbaum.

MacLeod, C. M. (1999). The item and list methods of directed forgetting: Test differences and the role of demand characteristics. *Psychonomic Bulletin & Review*, *6*, 123-129.

MacLeod, C. M. (2002). Retrieval-induced forgetting in eyewitness memory: Forgetting as a consequence of remembering. *Applied Cognitive Psychology*, *16*, 135-149.

MacLeod, M. D. & Macrae, C. N. (2001). Gone but not forgotten: The transient nature of retrieval-induced forgetting. *Psychological Science*, *12*, 148-52.

Macrae, C. N. & McLeod, M. D. (1999). On recollections lost: When practice makes imperfect. *Journal of Personality and Social Psychology*, *77*, 463-473.

Maquet, P. & Ruby, P. (2004). Insight and the sleep committee. *Nature*, *427*, 304-305.

Marsh, E. J., Roediger III, H. L., Bjork, R. A., & Bjork, E. L. (2007). The memorial consequences of multiple-choice testing. *Psychonomic Bulletin & Review*, *14*, 194-199.

Marshall, L., Helgadottir, H., Mölle, M., & Born, J. (2006). Boosting slow oscillations during sleep potentiates memory. *Nature*, *444*, 610-613.

Marshall, L., Kirov, R., Brade, J., Mölle, M., & Born, J. (2011). Transcranial electrical currents to probe EEG brain rhythms and memory consolidation during sleep in humans. *PLoS One*, *6*, e16905.

McDaniel, M. A., Agarwal, P. K., Huelser, B. J., McDermott, K. B., & Roediger III, H. L. (2011). Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology*, *103*, 399-414.

McDaniel, M. A. & Masson, M. E. J. (1985). Altering memory representations through retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 371-385.

McDaniel, M. A., Roediger III, H. L., & McDermott, K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review*, *14*, 200-206.

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

McGeoch, J. A. (1942). *The psychology of human learning.* New York: Longmans, Green.

McGeoch, J. A. & McDonald, W. T. (1931). Meaningful relation and retroactive inhibition. *American Journal of Psychology*, *43*, 579-588.

Mednick, S., Nakayama, K., & Stickgold, R. (2003). Sleep-dependent learning: A nap is as good as a night. *Nature Neuroscience*, *6*, 697-698.

Melton, A. W. & Irwin, J. M. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *The American Journal of Psychology*, *53*, 173-203.

Melton, A. W. & von Lackum, W. J. (1941). Retroactive and proactive inhibition in retention: Evidence for a two-factor theory of retroactive inhibition. *The American Journal of Psychology*, *54*, 157-173.

Moscovitch, M. & Nadel, L. (1998). Consolidation and the hippocampal complex revisited: In defense of the multiple-trace model. *Current Opinion in Neurobiology*, *8*, 297-300.

Moscovitch, M., Nadel, L., Winocur, G., Gilboa, A., & Rosenbaum, R. S. (2006). The cognitive neuroscience of remote episodic, semantic and spatial memory. *Current Opinion in Neurobiology*, *16*, 179-190.

Nadel, L. & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, *7*, 217-227.

Nader, K. & Einarsson, E. O. (2010). Memory reconsolidation: An update. *Annals of the New York Academy of Sciences*, *1191*, 27-41.

Nairne, J. S. & Pandeirada, J. N. S. (2008). Forgetting. In: H. L. Roediger III (Ed.), *Cognitive psychology of memory.* Vol. 2 of Learning and memory - A comprehensive reference (pp. 179-194). Oxford: Elsevier.

Nishida, M. & Walker, M. P. (2007). Daytime naps, motor memory consolidation and regionally specific sleep spindles. *PLoS One*, *4*, e341.

O'Neill, J., Pleydell-Bouverie, B., Dupret, D., & Csicsvari, J. (2010). Play it again: Reactivation of waking experience and memory. *Trends in Neuroscience*, *33*, 220-229.

Oswald, W. D. & Roth, E. (1987). *Der Zahlenverbindungstest (ZVT) [Connect-the-numbers Test]*. Göttingen: Hogrefe.

Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review*, *14*, 187-193.

Pastötter, B. & Bäuml, K.-H. (2007). The crucial role of postcue encoding in directed forgetting and context-dependent forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 977-982.

Pastötter, B. & Bäuml, K.-H. (2010). Amount of postcue encoding predicts amount of directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 54-65.

Pastötter, B., Schicker, S., Niedernhuber, J., & Bäuml, K.-H. T. (2011). Retrieval during learning facilitates subsequent memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 287-297.

Payne, J. D. (2011). Sleep on it!: Stabilizing and transforming memories during sleep. *Nature Neuroscience*, *14*, 272-274.

Payne, J. D., Ellenbogen, J. M., Walker, M. P., & Stickgold, R. (2008). The role of sleep in memory consolidation. In: H.L. Roediger III (Ed.), *Cognitive psychology of memory.* Vol. 2 of Learning and memory - A comprehensive reference (pp. 663-685). Oxford: Elsevier.

Payne, J. D. & Kensinger, E. A. (2010). Sleep's role in the consolidation of emotional episodic memories. *Current Directions in Psychological Science*, *19*, 290-295.

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

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

Poe, G. R., Walsh, C. M., & Bjorness, T. E. (2010). Cognitive neuroscience of sleep. *Progress in Brain Research*, *185*, 1-19.

Postman, L. & Stark, K. (1969). Role of response availability in transfer and interference. *Journal of Experimental Psychology*, *79*, 168-177.

Postman, L., Stark, K., & Fraser, J. (1968). Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior*, *7*, 672-694.

Pyc, M. A. & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language*, *60*, 437-447.

Pyc, M. A. & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science*, *330*, 335.

Racsmany, M., Conway, M. A., & Demeter, G. (2010). Consolidation of episodic memories during sleep: Long-term effects of retrieval practice. *Psychological Science*, *21*, 80-85.

Racsmany, M., Conway, M. A., Garab, E. A., & Nagymate, G. (2008). Memory awareness following episodic inhibition. *The Quarterly Journal of Experimental Psychology*, *61*, 525-534.

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

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

Rauchs, G,. Bertran, F., Guillery-Girard, B., Desgranges, B., Kerrouche, N., Denise, P., Foret, J., & Eustache, F. (2004). Consolidation of strictly episodic memories mainly requires rapid eye movement sleep. *Sleep*, *27*, 395-401.

Rauchs, G., Feyers, D., Landeau, B., Bastin, C., Luxen, A., Maquet, P., & Collette, F. (2011). Sleep contributes to the strengthening of some memories over others, depending on hippocampal activity at learning. *The Journal of Neuroscience*, *31*, 2563-2568.

Raven, J. C. (1999). *Standard Progressive Matrices.* Oxford: OPP.

Rechtschaffen, A. (1998). Current perspectives on the function of sleep. *Perspectives in Biology and Medicine*, *41*, 359-390.

Redondo, R. L., & Morris, R. G. (2011). Making memories last: The synaptic tagging and capture hypothesis. *Nature Reviews Neuroscience*, *12*, 17-30.

Richland, L. E., Kornell, N., & Kao, L. S. (2009). The pretesting effect: Do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology: Applied*, *15*, 243-257.

Roediger III, H. L. & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*, 20-27.

Roediger III, H. L. & Guynn, M. J. (1996). Retrieval processes. In: E. L. Bjork & R. A. Bjork (Eds.), *Memory.* Vol. 2 of Handbook of perception and cognition (pp. 197-236). San Diego: Academic Press.

Roediger III, H. L. & Karpicke, J. D. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*, 181-210.

Roediger III, H. L. & Karpicke, J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249-255.

Rohrer, D., Taylor, K., & Sholar, B. (2010). Tests enhance the transfer of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 233-239.

Roman, P., Soriano, M. F., Gomez-Ariza, C. J., & Bajo, M. T. (2009). Retrieval-induced forgetting and executive control. *Psychological Science*, *20*, 1053-1058.

Rudoy, J. D., Voss, J. L., Westerberg, C. E., & Paller, K. A. (2009). Strengthening individual memories by reactivating them during sleep. *Science*, *326*, 1079.

Sahakyan, L., & Delaney, P. F. (2003). Can encoding differences explain the benefits of directed forgetting in the list method paradigm? *Journal of Memory and Language*, *48*, 195-206.

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 1064-1072.

Saletin, J. M., Goldstein, A. N., & Walker, M. P. (2011). The role of sleep in directed forgetting and remembering of human memories. *Cerebral Cortex*, *21*, 2534-2541.

Sara, S. J. (2008). Reconsolidation: Historical Perspective and Theoretical Aspects. In: R. Menzel (Ed.), *Learning Theory and Behaviour.* Vol. 1 of Learning and memory - A comprehensive reference (pp. 461-475). Oxford: Elsevier.

Saunders, J. & MacLeod, M. D. (2006). Can inhibition resolve retrieval competition through the control of spreading activation? *Memory & Cognition*, *34*, 307-322.

Scheithe, K. & Bäuml, K.-H. (1995). Deutschsprachige Normen für Vertreter von 48 Kategorien. *Sprache & Kognition*, *14*, 39-43.

Schmidt, C., Peigneux, P., Muto, V., Schenkel, M., Knoblauch, V., Münch, M., de Quervain, D. J.-F., Wirz-Justice, A., & Cajochen, C. (2006). Encoding difficulty promotes postlearning changes in sleep spindle activity during napping. *The Journal of Neuroscience*, *26*, 8976-8982.

Schredl, M. & Erlacher, D. (2011). Frequency of lucid dreaming in a representative German sample. *Perceptual and Motor Skills*, *112*, 104-108.

Scullin, M. K. & McDaniel, M. A. (2010). Remembering to execute a goal: Sleep on it! *Psychological Science*, *21*, 1028-1035.

Sego, S. A., Golding, J. M., & Gottlob, L. R. (2006). Directed forgetting in older adults using the item and list method. *Aging, Neuropsychology, and Cognition*, *13*, 95-114.

Shaw, J. S., Bjork, R. A., & Handal, A. (1995). Retrieval-induced forgetting in an eyewitness-memory paradigm. *Psychonomic Bulletin & Review*, *2*, 249-253.

Siegel, J. M. (2005). Clues to the function of mammalian sleep. *Nature*, *437*, 1264-1271.

Siegel, J. M. (2009). Sleep viewed as a state of adaptive inactivity. *Nature Reviews Neuroscience*, *10*, 747-753.

Slamecka, N. J. & McElree, B. (1983). Normal forgetting of verbal lists as a function of their degree of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 384-397.

Spear, N.E. & Müller, C.W. (1984). Consolidation as a function of retrieval. In: H. Weingartner & E. S. Parker (Eds.), *Memory consolidation – Psychobiology of cognition* (pp. 111-147). Hillsdale, NJ: Erlbaum.

Spitzer, B. & Bäuml, K.-H. (2007). Retrieval-induced forgetting in item recognition: Evidence for a reduction in general memory strength.

Journal of Experimental Psychology: Learning, Memory, and Cognition, *33*, 863-875.

Spitzer, B. & Bäuml, K.-H. (2009). Retrieval-induced forgetting in a category recognition task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 286-291.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, *99*, 195-231.

Staudigl, T., Hanslmayr, S., & Bäuml, K.-H. T. (2010). Theta oscillations reflect the dynamics of interference in episodic memory retrieval. *The Journal of Neuroscience*, *30*, 11356-11362.

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

Stickgold, R., Scott, L., Rittenhouse, C., & Hobson, J. A. (1999). Sleep-induced changes in associative memory. *Journal of Cognitive Neuroscience*, *11*, 182-193.

Storm, B. C., Bjork, E. L., & Bjork, R. A. (2008). Accelerated relearning after retrieval-induced forgetting: The Benefit of being forgotten. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 230-236.

Storm, B. C., Bjork, E. L., Bjork, R. A., & Nestojko, J. F. (2006). Is retrieval success a necessary condition for retrieval-induced forgetting? *Psychonomic Bulletin & Review*, *13*, 1023-1027.

Storm, B. C. & Nestojko, J. F. (2010). Successful inhibition, unsuccessful retrieval: Manipulating time and success during retrieval practice. *Memory*, *18*, 99-114.

Sumowski, J. F., Chiaravalloti, N., & DeLuca, J. (2010a). Retrieval practice improves memory in multiple sclerosis: Clinical application of the testing effect. *Neuropsychology*, *24*, 267-272.

Sumowski, J. F., Wood, H. G., Chiaravalloti, N., Wylie, G. R., Lengenfelder, J., & DeLuca, J. (2010b). Retrieval practice: A simple strategy for improving memory after traumatic brain injury. *Journal of the International Neuropsychological Society*, *16*, 1147-1150.

Szpunar, K. K., McDermott, K. B., & Roediger III, H. L. (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1392-1399.

Talamini, L. M., Nieuwenhuis, I. L., Takashima, A., & Jensen, O. (2008). Sleep directly following learning benefits consolidation of spatial associative memory. *Learning & Memory*, *15*, 233-237.

Tandoh, K. & Naka, M. (2007). Durability of retrieval-induced forgetting. *Japanese Journal of Psychology, 78*, 310-315.

Thomas, A. K., Bulevich, J. B., & Chan, J. C. K. (2010). Testing promotes eyewitness accuracy with a warning: Implications for retrieval enhanced suggestibility. *Journal of Memory and Language*, *63*, 149-157.

Thompson, C. P., Wenger, S. K., & Bartling, C. A. (1978). How recall facilitates subsequent recall: A reappraisal. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 210-221.

Tononi, G. & Cirelli, C. (2006). Sleep function and synaptic homeostasis. *Sleep Medicine Reviews*, *10*, 49-62.

Tse, D., Takeuchi, T., Kakeyama, M., Kajii, Y., Okuno, H., Tohyama, C., Bito, H., & Morris, R. G. (2011). Schema-dependent gene activation and memory encoding in neocortex. *Science*, *333*, 891-895.

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*, 197-203.

Tucker, M. A., Tang, S. X., Uzoh, A., Morgan, A., & Stickgold, R. (2011). To sleep, to strive, or both: How best to optimize memory. *PLoS One*, *6*, e21737.

Tulving, E. (1967). The effects of presentation and recall of material in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, *6*, 175-184.

Tulving, E. & Psotka, J. (1971). Retroactive inhibition in free recall: Inaccessibility of information available in the memory store. *Journal of Experimental Psychology*, *87*, 1-8.

Tulving, E. & Watkins, M. J. (1974). On negative transfer: Effects of testing one list on the recall of another. *Journal of Verbal Learning and Verbal Behavior*, *13*, 181-193.

Underwood, B. J. (1948). Spontaneous recovery of verbal associations. *Journal of Experimental Psychology*, *38*, 429-439.

Underwood, B. J. (1957). Interference and forgetting. *Psychological Review*, *64*, 49-60.

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language*, *50*, 289-335.

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

Wagner, U., Hallschmid, M., Rasch, B., & Born, J. (2006). Brief sleep after learning keeps emotional memories alive for years. *Biological Psychiatry*, *60*, 788-790.

Walker, M. P. & Stickgold, R. (2010). Overnight alchemy: Sleep-dependent memory evolution. *Nature Reviews Neuroscience*, *11*, 218.

Walker, M. P. & van der Helm, E. (2009). Overnight therapy? The role of sleep in emotional brain processing. *Psychological Bulletin*, *135*, 731-748.

Weinstein, Y., McDermott, K. B., & Szpunar, K. K. (2011). Testing protects against proactive interference in face-name learning. *Psychonomic Bulletin & Review*, *18*, 518-523.

Wheeler, M. A. (1995). Improvement in recall over time without repeated testing: Spontaneous recovery revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 173-184.

Wheeler, M. A., Ewers, M., & Buonanno, J. F. (2003). Different rates of forgetting following study versus test trials. *Memory*, *11*, 571-580.

Wheeler, M. A., & Roediger III, H. L. (1992). Disparate effects of repeated testing: Reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychological Science*, *3*, 240-245.

Wickens, D. D. (1970). Encoding categories of words: An empirical approach to meaning. *Psychological Review*, *77*, 1-15.

Wilhelm, I., Diekelmann, S., Molzow, I., Ayoub, A., Mölle, M., & Born, J. (2011). Sleep selectively enhances memory expected to be of future relevance. *The Journal of Neuroscience*, *31*, 1563-1569.

Williams, C. C. & Zacks, R. T. (2001). Is retrieval-induced forgetting an inhibitory process? *The American Journal of Psychology*, *114*, 329-354.

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

Wimber, M., Bäuml, K.-H., Bergström, Z., Markopoulos, G., Heinze, H.-J., & Richardson-Klavehn, A. (2008). Neural markers of inhibition in human memory retrieval. *The Journal of Neuroscience*, *28*, 13419-13427.

Wimber, M., Rutschmann, R. M., Greenlee, M. W., & Bäuml, K.-H. (2009). Retrieval from episodic memory: Neural mechanisms of interference resolution. *Journal of Cognitive Neuroscience*, *21*, 538-549.

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

Wixted, J. T., Mickes, L., & Squire, L. R. (2010). Measuring recollection and familiarity in the medial temporal lobe. *Hippocampus*, *20*, 1195-1205.

Wixted, J. T. & Rohrer, D. (1993). Proactive interference and the dynamics of free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1024-1039.

Yonelinas, A. P., Aly, M., Wang, W.-C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, *20*, 1178-1194.

Yoo, S.-S., Gujar, N., Hu, P., Jolesz, F. A., & Walker, M. P. (2007). The human emotional brain without sleep - A prefrontal amygdala disconnect. *Current Biology*, *17*, R877-878.

Zaromb, F. M. & Roediger III, H. L. (2010). The testing effect in free recall is associated with enhanced organizational processes. *Memory & Cognition*, *38*, 995-1008.

Zellner, M. & Bäuml, K.-H. (2006). Inhibitory deficits in older adults: List-method directed forgetting revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 290-300.