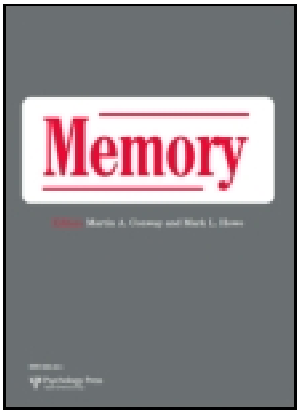

Memory
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pmem20

Retrieval-induced forgetting, delay, and sleep

Magdalena Abel [a] & Karl-Heinz T. Bäuml [a]
[a] Department of Experimental Psychology, Regensburg University, Regensburg, Germany

Version of record first published: 28 May 2012

MEMORY, 2012, 20 (5), 420–428

Psychology Press
Taylor & Francis Group

Retrieval-induced forgetting, delay, and sleep

Magdalena Abel and Karl-Heinz T. Bäuml

Department of Experimental Psychology, Regensburg University, Regensburg, Germany

Retrieval-induced forgetting (RIF) refers to the finding that retrieval of a subset of previously studied material can cause forgetting of related unpractised material. Prior work on the role of delay between practice and test for RIF reported mixed results. Whereas some studies found RIF to be a relatively transient phenomenon, others found RIF to persist over time. We examined whether the persistence of RIF can depend on whether wakefulness or sleep follows retrieval practice. We employed a variant of the retrieval-practice paradigm with short (20 minutes) and long (12 hours) delay conditions. In all conditions participants studied a perceptually categorised list followed by retrieval practice on some of the items from some of the categories. Participants studied and practised the material in the morning or the evening. RIF was present in the short-delay conditions, and it was present in the long-delay condition if the delay included nocturnal sleep; if the long delay was filled with diurnal wakefulness, RIF was absent. Results show that whether delay eliminates RIF or not can depend on whether sleep or wake follows retrieval practice. Connections of the results to recent findings about the role of delay and sleep for RIF are discussed.

Keywords: Retrieval-induced forgetting; Retrieval; Forgetting; Delay; Sleep.

Practising retrieval of a subset of formerly studied material can cause subsequent forgetting of related unpractised material. Such forgetting has repeatedly been demonstrated using the retrieval-practice paradigm (Anderson, Bjork, & Bjork, 1994). In this paradigm participants typically study a categorised list (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 (e.g., Fruit–Man__). Such retrieval practice creates three item types: practised items from practised categories (i.e., Fruit–Mango), unpractised items from practised categories (i.e., Fruit–Apple), and control items from unpractised categories that did not appear during the retrieval-practice phase (i.e., Sports–Tennis, Sports–Soccer). Memory for all three item types is ultimately tested. Compared to control items, memory performance for practised items is usually enhanced and memory performance for unpractised items is reduced. This reduction has been termed retrieval-induced forgetting (RIF) and 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 (for reviews, see Anderson, 2003; Bäuml, Pastötter, & Hanslmayr, 2010; for a non-inhibitory account of RIF, see Williams & Zacks, 2001).

Previous research has shown that RIF is a very general phenomenon that is found in a variety of settings (e.g., in eyewitness reports, Shaw, Bjork, & Handal, 1995; during the acquisition of a

Address correspondence to: Karl-Heinz T. Bäuml, Department of Experimental Psychology, Regensburg University, 93040 Regensburg, Germany. E-mail: karl-heinz.baeuml@psychologie.uni-regensburg.de

The project was conducted as part of Magdalena Abel's doctoral dissertation. We thank S. Kaltner, L. Kugler, and A. Schlichting for their help with data collection.

http://www.psypress.com/memory http://dx.doi.org/10.1080/09658211.2012.671832

foreign language, Levy, McVeigh, Marful, & Anderson, 2007; in conversations, Coman, Manier, & Hirst, 2009) and for a variety of stimulus materials (e.g., verbal material, Anderson et al., 1994; visual material, Ciranni & Shimamura, 1999; autobiographical material, Barnier, Hung, & Conway, 2004). However, past research has also identified a number of boundary conditions for RIF. For instance, negative moods (e.g., Bäuml & Kuhbandner, 2007), stress (e.g., Koessler, Engler, Riether, & Kissler, 2009), or divided attention (e.g., Roman, Soriano, Gomez-Ariza, & Bajo, 2009) during retrieval practice have been shown to abrogate the effect; no RIF has been observed when the encoded items were subject to directed forgetting (e.g., Bäuml & Samenieh, 2010; Storm, Bjork, & Bjork, 2007); and no RIF has been observed if practised and unpractised items were episodically or semantically highly integrated (e.g., Anderson & McCulloch, 1999; Bäuml & Hartinger, 2002).

Recently the question has been raised of whether the delay interval between retrieval practice and test might constitute another boundary condition for RIF, with RIF being present after short delays but not after long delays. Only a few studies have addressed the issue to date and the results are mixed. While numerous studies have shown that RIF is present after a delay of 20 minutes between practice and test (e.g., Anderson et al., 1994; Anderson & Spellman, 1995), the results of more recent studies suggest that RIF may gradually diminish as the time interval between retrieval practice and test increases. For instance, MacLeod and Macrae (2001) reported no RIF after a delay interval of 24 hours; while recall of practised and control items was reduced with delay, no such reduction was found for the unpractised items, which made the RIF effect disappear. Similarly, using low-integration study conditions, across two experiments Chan (2009) observed no RIF after a delay of 24 hours; again recall of the control items was reduced after the delay, whereas recall of the unpractised items was largely unaffected. Results for the practised items were less unequivocal, because 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 another recent study suggest that RIF can also be lasting. Using low-integration study conditions, Garcia-Bajos, Migueles, and Anderson (2009) found RIF still present after a delay of 1 week (for a similar result, see Tandoh & Naka, 2007; see also Chan, 2010, for the possible role of delay after 24 hours).[1]

To date, the inconsistency in results regarding the role of delay for RIF is still unresolved, and it remains unclear which factors might have been responsible for RIF surviving longer delay intervals in some studies but not in others. More recently, however, Racsmany, Conway, and Demeter (2010) reported results indicating that the persistence of RIF can depend on whether sleep or wake follows the retrieval-practice phase. In two experiments these authors found RIF to be absent if retrieval practice and test were separated by a 12-hour wake interval, but they found RIF to be present if retrieval practice and test were separated by a 12-hour sleep interval, suggesting that sleep after retrieval practice can be critical for the persistence of RIF. In their Experiment 2 the authors also provided (numerical) evidence that memory performance for control items, but not for unpractised items, may be reduced across a 12-hour wake interval, whereas memory performance for both item types may be largely unaffected across the 12-hour sleep interval, which suggests that sleep may affect RIF mainly through sleep-associated consolidation of the control items. However, Racsmany et al. did not directly address the issue of whether sleep and wake intervals affect the single item types differently, and therefore they also did not report corresponding statistical analyses.

Drawing firm conclusions from Racsmany et al.'s (2010) prior work regarding the role of sleep and wake delays for RIF and the single item types might be premature, for several reasons. First, regular RIF after a short delay was not examined (Experiment 1 of their study) or not present (Experiment 2 of their study), which makes it difficult to evaluate the effect of wake delay on RIF in this prior work. Second, no

[1] There are further studies in the literature reporting that RIF can persist with delay (Migueles & Garcia-Bajos, 2007; Saunders, Fernandes, & Kosnes, 2009; Storm, Bjork, Bjork, & Nestojko, 2006). However, these studies used a repeated-test design in which participants were given an initial test shortly after retrieval practice and then were tested on the same items 24 hours later. Unfortunately, because on the initial test more control items than unpractised items were recalled, and because recall of an item can enhance its recall on a subsequent test (e.g., Roediger & Karpicke, 2006), the RIF effect found in the delayed test may have been caused by the recall difference in the initial test, rather than by the lasting effects of inhibition.

appropriate control conditions of circadian effects were included in the study, which complicates evaluation of a possible sleep effect, because time of day may have affected the results (e.g., Ellenbogen, Hulbert, Stickgold, Dinges, & Thompson-Schill, 2006; Payne, Stickgold, Swanberg, & Kensinger, 2008). Third, the prior work did not analyse the impact of delay on the single item types, an analysis that could further elucidate when and why delay may be a boundary condition for RIF. The goal of the present study, therefore, was to run a fresh experiment that examines RIF for short and long delays, analyses effects of wake and sleep for both RIF and the single item types, and controls for circadian effects. To go beyond a "simple replication" of the prior work, 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), we used perceptual rather than semantic categories to examine how wake and sleep delays influence RIF and the single item types.

The experiment employed a variant of the retrieval-practice paradigm with two short and two long delay conditions. In all four delay conditions participants studied a perceptually categorised list followed by retrieval practice on some of the items from some of the categories (Spitzer & Bäuml, 2009; see Figure 1). Two groups of participants studied and practised the material in the morning (20-min wake, 12-h wake), the other two groups studied and practised the material in the evening (20-min sleep, 12-h sleep). The two short-delay groups (20-min wake, 20-min sleep) were tested after a relatively brief delay of 20 min, the two long-delay groups were tested after a 12-h interval of diurnal wakefulness (12-h wake) or nocturnal sleep (12-h sleep; see Figure 2a).

Following prior work indicating that RIF is relatively transient and can gradually diminish as the time interval between retrieval practice and test increases (e.g., Chan, 2009; MacLeod & Macrae, 2001), we expected to find RIF to be reduced, or even to be absent, after the 12-hour wake interval; in particular we expected the control items but not the unpractised items to show reduced recall after the delay interval. Following the prior work on the role of sleep for RIF (Racsmany et al., 2010), we expected RIF to be present after the 12-hour sleep interval; in particular we expected recall of both control items and unpractised items to be largely unaffected after the delay interval. Such results would indicate that sleep influences the role of delay for RIF, and that the persistence of RIF can depend on whether sleep or wake follows retrieval practice; in particular, the results would indicate that sleep affects RIF mainly through sleep-associated consolidation of the control items. Finally, the results would provide evidence that the role of wake and sleep interval is similar in purely episodic RIF tasks, as employed in the present study, and RIF tasks that employ a combination of episodic and semantic information (e.g., Anderson et al., 1994).

METHOD

Participants

A total of 96 undergraduates ($M = 23.5$; range 19–33 years) took part in the experiment

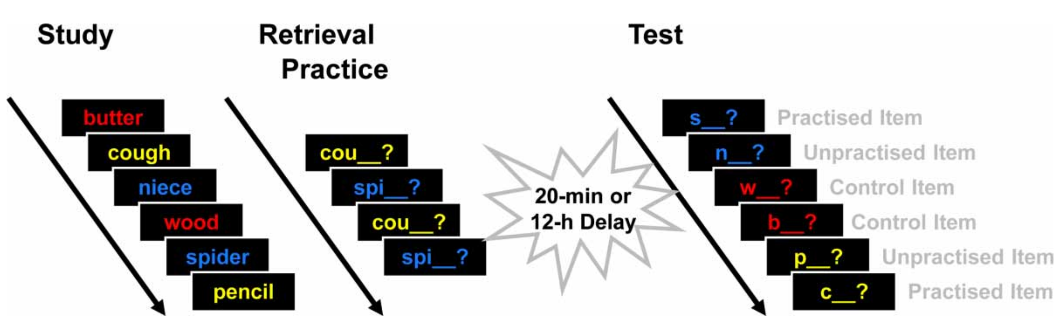

Figure 1. Participants studied a list of words that were categorised by three font colours (Study phase), and then practised retrieving half of the words from two of the categories using the items' font colour and unique word stem as retrieval cues (Retrieval Practice phase). Participants' memory for all initially studied words was tested either 20 minutes after retrieval practice or after a 12-hour retention interval that included regular sleep or wakefulness (Test phase).

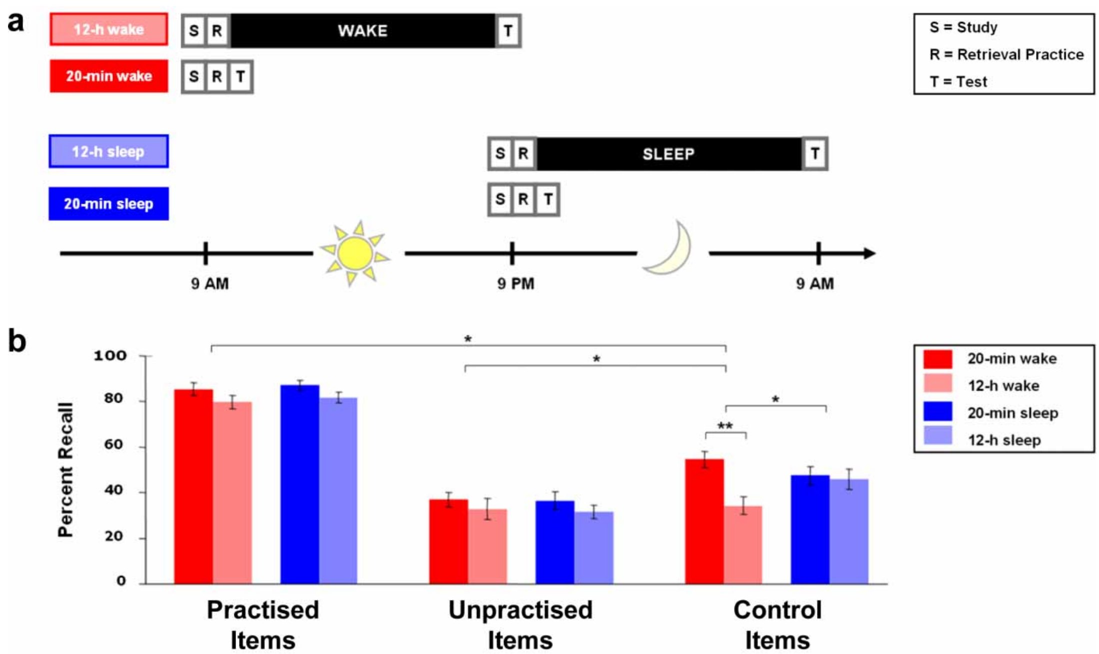

Figure 2. (a) Half of the participants started the experiment at 9 am (20-min wake and 12-h wake conditions), the other half at 9 pm (20-min sleep and 12-h sleep conditions). In addition, half of the participants completed the experiment immediately after 20 minutes (20-min wake and 20-min sleep), while the other half completed it after a delay of 12 hours (12-h wake and 12-h sleep). (b) Mean recall level is shown as a function of condition (20-min wake or 20-min sleep or 12-h wake or 12-h sleep) and item type (practised items from practised categories or unpractised items from practised categories or control items from unpractised categories). Error bars represent standard errors; * $p \leq .05$; ** $p \leq .01$.

voluntarily and in return for financial reimbursement. All participants completed a screening questionnaire and interview prior to selection (Ellenbogen et al., 2006). 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 having 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 a measure of speed of cognitive processing; Oswald & Roth, 1987), and habitual sleep duration, all *ps* $> .15$.

Materials

Item materials consisted of 24 semantically unrelated German nouns that were divided into three sets of eight items. Within sets, all items had unique initial letters. Each set was equally often assigned to the font colours red, blue, and yellow, and served equally often as practised and unpractised category. Each item was equally often used as practised and unpractised word (see also Spitzer & Bäuml, 2009; see Figure 1 for an illustration of the applied colour categories and the different item types). 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 $2 \times 2 \times 3$ mixed factorial design: CONDITION (wake, sleep) and DELAY (20-min delay, 12-h delay) were manipulated between participants, and ITEM TYPE (practised items, unpractised items, control items) was varied within participants. For each participant the experiment started at 9 am

(20-min wake, 12-h wake) or 9 pm (20-min sleep, 12-h sleep). Testing occurred either 20 minutes after the retrieval-practice phase (20-min wake and 20-min sleep) or after an additional 12-hour delay (12-h wake and 12-h sleep conditions; see Figure 2a for a depiction of the four experimental groups). 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 memorise the 24 items in their respective font colours. Items were presented for 5 seconds each and appeared centrally on a computer screen. They were shown individually, and in a pseudorandomised order with no two items of the same colour category following each other. Promptly following the first study cycle we conducted a second study cycle in exactly the same way.

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

Test phase. For the two control groups (20-min wake, 20-min sleep) testing occurred 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. At test, all originally studied words were tested; the words' font colours and unique first letters were provided as retrieval cues. The cues were presented successively in a blocked, randomised manner: The sequence of categories was randomly chosen, but all items of one colour category were tested successively. Within practised categories, practised and unpractised items were blocked and which block was tested first was determined randomly. 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 (see Figure 1 for an illustration of the different experimental phases).

RESULTS

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-practice phase

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, $F(1, 94) < 1.0$.

Test phase

A $2 \times 2 \times 3$ ANOVA with the factors of CONDITION (sleep, wake), DELAY (20-min delay, 12-h delay), and ITEM TYPE (practised, unpractised, 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 practised items than for control items, and of better recall for control items than for unpractised items (see also below); the main effect of DELAY reflects the decrease in recall with delay. No significant two-way interactions emerged, all ps > .20, but 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. Consistently, across the wake delay, reliable delay-induced forgetting was present for the control items, $t(46) = 3.94$, $p < .001$, but was absent for the practised and unpractised items, ts(46) < 1.41, ps > .15, whereas no delay-induced forgetting emerged for any of the three item types across the sleep delay, ts(46) < 1.65, ps > .10. Regarding the role of ITEM TYPE, planned comparisons revealed 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; in contrast, retrieval-induced enhancement—higher recall for the practised items than the control items—was present in all four groups, ts(23) > 7.45, ps < .001.

Circadian effects

When comparing the 20-minute wake and 20-minute sleep control groups, no circadian memory effects arose for any of the three item types, ts(46) < 1.35, ps > .15 (see also Figure 2b).

DISCUSSION

The goal of the present study was to examine RIF for short and long delays between practice and test, and to analyse effects of wake and sleep for both RIF and the single item types. We found that recall of unpractised items was reduced relative to control items after a 20-minute delay interval between retrieval practice and test, but that the two types of items no longer differed in recall after a delay interval of 12 hours, at least when the interval was filled with wakefulness; thus RIF was present after the short delay but was absent after the long delay. Consistently, analysis of the influence of wake delay on the single item types revealed that recall of the control items was reduced with delay, whereas recall of the unpractised items was unaffected. This pattern of results replicates Racsmany et al.'s (2010) prior finding of no RIF after a 12-hour wake interval, and it is consistent with the studies by MacLeod and Macrae (2001) and Chan (2009), who did not find RIF after intervals of 24 hours, and found delay-induced reductions in recall of the control items but not of the unpractised items (but see Garcia-Bajos et al., 2009, for other results).[2]

More importantly the present results show that the influence of delay on RIF can depend on whether the delay interval is filled with wakefulness or sleep. When sleep followed retrieval practice, RIF was maintained after a delay of 12 hours between retrieval practice and test, and memory performance for both the control and the unpractised items was unaffected, which differs from the results in the wake condition. Because the difference in results between wake and sleep interval is not attributable to circadian effects, the finding indicates that a 12-hour wake delay, but not a 12-hour sleep delay, can reduce or even eliminate RIF; whereas wake delay reduces recall of the control items but leaves recall of unpractised items unaffected, sleep delay leaves recall of both item types unaffected. The finding that RIF is present after a delay interval filled with sleep, but not after a delay interval filled with wakefulness, is consistent with the prior work by Racsmany et al. (2010). However, the present results go beyond this prior work by showing RIF for short delays, by controlling for circadian effects, and by analysing effects of sleep not only on RIF but also on the single item types. In addition, the results generalise the prior work by demonstrating effects of sleep in a purely episodic RIF task.

The present study demonstrates that the persistence of RIF can depend on whether wakefulness or sleep follows retrieval practice. Thus, contrary to previous suggestions (e.g., MacLeod & Macrae, 2001), delay per se does not seem to be a boundary condition for the detrimental effects of retrieval practice. Rather, whether RIF is present or not seems to depend on sleep-associated memory stabilisation of the control items. Prior work on the role of delay for RIF differed in whether RIF was present (Garcia-Bajos et al., 2009) or was absent (Chan, 2009; MacLeod & Macrae, 2001) after longer delays. However, this prior work did not control whether sleep or

[2] Intuitively one might expect to find reduced RIF to be accompanied by improved recall of the unpractised items. However, the results clearly show that it is the reduced recall of the control items that mediates the reduction in RIF, and that recall of the unpractised items remains relatively constant across the delay interval. This finding replicates results from several previous studies (e.g., Chan, 2009; MacLeod & Macrae, 2001), and raises the question of exactly why recall of unpractised items remains fairly constant with delay interval (see later).

wakefulness followed rather shortly upon retrieval practice. Because the amount of time passing before sleep onset can be crucial for the emergence of sleep-associated memory consolidation (e.g., Diekelmann & Born, 2010), this missing control might have contributed to the difference in prior results and the difference in conclusions arising.

This study examined effects of retrieval practice using a low-integration study condition, i.e., a condition in which practised and unpractised items show a low level of inter-item associations and retrieval practice typically induces RIF (see Anderson, 2003; Bäuml et al., 2010). In contrast, if practised and unpractised items are highly integrated, retrieval practice typically does not reduce recall of the unpractised items, and no RIF arises (e.g., Anderson & McCulloch, 1999; Bäuml & Hartinger, 2002). Interestingly, prior work indicates that delay may affect item types in low-integration and high-integration conditions similarly, with recall of control items being reduced and recall of unpractised items being stable across a wake delay. This leads to an elimination of RIF in low-integration conditions and to retrieval-induced facilitation—higher recall of unpractised than control items—in high-integration conditions after the delay (Chan, 2009; Chan, McDermott, & Roediger, 2006). If the present finding, that sleep stabilises the control items but leaves recall of unpractised items largely unaffected, generalised from low-integration to high-integration conditions, then the retrieval-induced facilitation observed in high-integration conditions after a wake delay should largely disappear after sleep. Together with the present finding such a result would indicate that, in general, RIF after sleep is similar to RIF after short (wake) delays.

The present results replicate prior RIF work by showing that retrieval practice can reduce, or even eliminate, delay-induced forgetting of the practised and related unpractised items (e.g., Chan, 2009; MacLeod & Macrae, 2001). The finding of reduced delay-induced forgetting of the practised items is consistent with numerous findings from the testing effect literature (e.g., Karpicke & Roediger, 2008; Roediger & Karpicke, 2006) and might indicate that retrieval practice creates particularly elaborated memory representations (Roediger & Butler, 2011) or particularly strong memory representations (Kornell, Bjork, & Garcia, 2011) for the practised items, so that these items may be less susceptible to delay-induced forgetting. Why retrieval practice reduces delay-induced forgetting of the unpractised items as well is less clear to date. Possibly the finding reflects a release from inhibition, so that with increasing delay the inhibitory effects on the unpractised items are reduced and the items show less delay-induced forgetting than the control items. However, other possibilities remain as well and further work is needed to examine the question in more detail.

Finally, our results may also offer suggestions for research on sleep-associated memory consolidation. Prior work on sleep-associated memory consolidation has shown that sleep does not influence all types of memories equally (e.g., Diekelmann & Born, 2010). For instance, sleep-related benefits have been found to depend on the future relevance of memories (e.g., Wilhelm et al., 2011), or their emotional tone (e.g., Payne et al., 2008). This study suggests that memories' sleep-associated consolidation might also depend on whether the studied material was previously retrieved or not. In fact, whereas we found the expected sleep-associated consolidation for the (non-retrieved) control items, practised items did not show sleep-associated consolidation, suggesting that prior retrieval might reduce the amount of sleep-associated consolidation.[3] The proposal might also account for the lack of sleep-associated consolidation for the related unpractised items, although only if the items' unconscious reactivation during retrieval practice (e.g., Kuhl, Dudukovic, Kahn, & Wagner, 2007; Staudigl, Hanslmayr, & Bäuml, 2010) was equivalent in effect to the practised items' conscious recall. The question might be addressed in future studies, possibly using EEG and/or fMRI recordings.

In sum, this study shows that a delay interval of several hours between retrieval practice and test per se does not reflect a boundary condition of RIF. Rather, whether delay eliminates RIF or not

[3] Many previous studies, examining sleep's impact on episodic memory and reporting evidence for sleep-associated memory consolidation, used a mixture of repeated presentation and repeated retrieval practice to train participants on the respective tasks prior to the critical delay manipulation (e.g., Plihal & Born, 1997, Wilhelm et al., 2011). Here, in contrast, we did not implement such study-test cycles, thus generating a less "contaminated" measure of the potential influence of prior retrieval practice. The results from the prior work, reporting consolidation effects after repeated study-retrieval cycles, and the results from the present study, reporting reduced consolidation effects after "pure" retrieval practice cycles, are thus not necessarily in conflict.

can depend on whether wake or sleep follows retrieval practice. While a wake interval can eliminate RIF after several hours because recall of the control items, but not of the unpractised items, is reduced, sleep following retrieval practice can preserve RIF. The reason for the persistence of RIF with sleep is that sleep consolidates the control items and thus leaves the RIF effect largely uninfluenced.

Manuscript received 16 November 2011
Manuscript accepted 22 February 2012
First published online 28 May 2012

REFERENCES

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

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

Barnier, A. J., Hung, L., & Conway, M. A. (2004). Retrieval-induced forgetting of emotional and unemotional autobiographical memories. *Cognition & Emotion*, *18*, 457–477.

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

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

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

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., McDermott, K. B., & Roediger, 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.

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.

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

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

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

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

Karpicke, J. D., & Roediger, H. L. III (2008). The critical importance of retrieval for learning. *Science*, *319*, 966–968.

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

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.

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.

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.

MacLeod, M. D., & Macrae, C. N. (2001). Gone but not forgotten: The transient nature of retrieval-induced forgetting. *Psychological Science*, *12*, 148–152.

Migueles, M., & Garcia-Bajos, E. (2007). Selective retrieval and induced forgetting in eyewitness memory. *Applied Cognitive Psychology*, *21*, 1157–1172.

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

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.

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.

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

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

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

Saunders, J., Fernandes, M., & Kosnes, L. (2009). Retrieval-induced forgetting and mental imagery. *Memory & Cognition*, *37*, 819–828.

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

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.

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.

Storm, B. C., Bjork, E. L., & Bjork, R. A. (2007). When intended remembering leads to unintended forgetting. *The Quarterly Journal of Experimental Psychology*, *60*, 909–915.

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.

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

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? *American Journal of Psychology*, *114*, 329–354.