Neurobiology of Learning and Memory 145 (2017) 181–189
http://dx.doi.org/10.1016/j.nlm.2017.10.009

Modulating influences of memory strength and sensitivity of the retrieval test on the detectability of the sleep consolidation effect.

Sarah F. Schoch[a], Maren J. Cordi[b,c], Björn Rasch[b,c]*

a. University of Zurich, CRPP Sleep and Health, Zurich, Switzerland
b. Zurich Center for Interdisciplinary Sleep Research (ZiS), University of Zurich, Zurich, Switzerland
c. University of Fribourg, Department of Psychology, Division of Biopsychology and Methods, Fribourg, Switzerland

*Corresponding author:

Björn Rasch, University of Fribourg, Department of Psychology, Division of Biopsychology and Methods, Rue P.A. de Faucigny 2, 1700 Fribourg, Phone: +41 26 300 7637, Fax: +41 26 300 9712, Email: Bjoern.rasch@unifr.ch

Running title: Detecting the sleep benefit

Abstract

Emotionality can increase recall probability of memories as emotional information is highly relevant for future adaptive behavior. It has been proposed that memory processes acting during sleep selectively promote the consolidation of emotional memories, so that neutral memories no longer profit from sleep consolidation after learning. This appears as a selective effect of sleep for emotional memories. However, other factors contribute to the appearance of a consolidation benefit and influence this interpretation. Here we show that the strength of the memory trace before sleep and the sensitivity of the retrieval test after sleep are critical factors contributing to the detection of the benefit of sleep on memory for emotional and neutral stimuli.

228 subjects learned emotional and neutral pictures and completed a free recall after a 12-hour retention interval of either sleep or wakefulness. We manipulated memory strength by including an immediate retrieval test before the retention interval in half of the participants. In addition, we varied the sensitivity of the retrieval test by including an interference learning task before retrieval testing in half of the participants. We show that a “selective” benefit of sleep for emotional memories only occurs in the condition with high memory strength. Furthermore, this “selective” benefit disappeared when we controlled for the memory strength before the retention interval and used a highly sensitive retrieval test. Our results indicate that although sleep benefits are more robust for emotional memories, neutral memories similarly profit from sleep after learning when more sensitive indicators are used. We conclude that whether sleep benefits on memory appear depends on several factors, including emotion, memory strength and sensitivity of the retrieval test.

Keywords: sleep-dependent memory consolidation, declarative memory, memory strength, emotion, sleep benefit, selectivity

1. Introduction

Emotion is important for the encoding and consolidation of memories. Numerous studies have shown that emotional content is remembered better than neutral content for words (e.g. Adelman and Estes, 2013; Kensinger and Corkin, 2003), stories (e.g. Cahill and McGaugh, 1995; Heuer and Reisberg, 1990) and pictures (e.g. Blake, Varnhagen, and Parent, 2001; Harris and Pashler, 2005). The strengthening effect of emotion on memory is linked to the activation of emotion-related brain regions including the amygdala, which modulates encoding and consolidation of memories in the hippocampus (e.g. Canli, Zhao, Brewer, Gabrieli, and Cahill, 2000; Phelps, 2004).

In addition to emotion, sleep also fosters memory consolidation after learning. Numerous studies have provided compelling evidence that sleep occurring shortly after learning results in an improvement of later memory retrieval performance as compared to a retention interval filled with wakefulness (see Rasch and Born, 2013 for a comprehensive review). Importantly, sleep's role in memory is considered active: According to the active system consolidation hypothesis, recently learned memories are spontaneously reactivated during sleep and thereby stabilized and integrated into neocortical memory networks for long-term storage. This system consolidation critically depends on several sleep-specific oscillatory brain signals, including hippocampal sharp-wave ripples, sleep spindles and slow-oscillations (Born and Wilhelm, 2012). Thus, according to this notion, sleep is not passively providing a shelter for temporary memory maintenance, but actively supports consolidation processes.

One important additional assumption in support of an active role of sleep for memory is that consolidation processes during sleep are selective (e.g. Rasch and Born, 2013; Stickgold and Walker, 2013). For example, it is assumed that sleep selectively consolidates memories that are relevant for the future, including emotional memories and memories associated with a reward. Regarding the selective consolidation of emotional memory content Hu, Stylos-Allan, and Walker (2006) showed that sleep after learning enhanced recognition performance only for emotional but not neutral images. Along similar lines, Payne and colleagues showed that a full night's sleep (Payne, Stickgold, Swanberg, and Kensinger, 2008) but also a brief nap (Payne, Kensinger, Wamsley, Spreng, Alger, Gibler, Schacter, and Stickgold, 2015) enhanced the recognition of central emotional objects in a scene, but not central neutral objects or the neutral backgrounds of that scene. Furthermore, Wagner, Gais, and Born (2001) reported that only emotional texts benefited from three hours of REM-rich sleep, but neutral texts did not.

Interestingly, in these studies the selective benefit of sleep for emotional memories is typically accompanied by a reduced or absent benefit of sleep for neutral memories. This is a puzzling finding, as numerous studies have consistently observed a sleep benefit for memory when using only neutral learning materials, including words, word-pairs and even senseless syllables (e.g. Jenkins and Dallenbach, 1924; Plihal and Born, 1997). Payne et al. (2008) have explained this paradox by arguing that when both emotional and neutral stimuli are presented together, a “trade-off” mechanism preferentially consolidates the emotional stimuli, as they are biologically more relevant. Bennion, Payne, and Kensinger (2015) suggest an early “emotional tagging” mechanism as a basis for this trade-off which selects memories for preferential consolidation. This could lead to a sleep-dependent alteration of brain activation as Payne and Kensinger (2011) report a strengthened connection between the amygdala, the hippocampus and the ventromedial prefrontal cortex during retrieval of emotional, but not neutral content after sleep compared to wakefulness.

However, not all studies find a prioritized consolidation of emotional items over neutral ones (Baran, Pace-Schott, Ericson, and Spencer, 2012; Lehmann, Seifritz, and Rasch, 2016; Lewis, Cairney, Manning, and Critchley, 2011). Baran et al. (2012) explain these divergent findings by highlighting the different experimental designs used across studies including variation of stimuli, range of emotional variance and presentation times. Bennion et al. (2015) discuss that besides the emotionality of the content, there are other important factors for the tagging mechanism such as stimulus novelty, state of the individual experiencing the event and neurochemical processes accompanying the emotional reaction.

Apart from methodological issues and tagging, an alternative explanation is that the benefit of sleep for memory is rather dependent on the strength of memory traces before sleep. Emotional memories are typically remembered better than neutral ones already during immediate recall attempts. Thus, sleep might simply require a certain threshold of memory strength during encoding to ensure that consolidation mechanisms during sleep can stabilize and integrate these new memories into long-term storage. One could argue that in some previous studies, emotional memories – which are more strongly encoded before sleep – pass the threshold for being consolidated during sleep, while weaker neutral memories do not. In contrast, when neutral material is repeatedly learned or studied until a certain criterion (e.g. 60%), a sleep benefit is observed also for solely neutral learning material (see e.g. Drosopoulos, Schulze, Fischer, and Born, 2007). Please note that the same study also reported that learning to a very high criterion (i.e., 90%) abolishes the beneficial effect of sleep on memory. This pattern of results suggests that the benefit of sleep on memory might be maximal at medium memory strengths, with lower or no effects for very weakly or very intensively encoded memories.

One possible way to increase memory strength in studies using emotional and neutral learning material is inserting an immediate retrieval right before the retention interval. Implementation of immediate recall of memories before the retention period should enhance the accessibility of these memories through a mechanism of effortful processing (Rowland, 2014), a phenomenon known as the testing effect (Sutterer and Awh, 2015). Thus, if the benefit of sleep relies on memory strength, we would expect that the neutral memories also benefit from sleep when their memory strength is enhanced due to immediate retrieval testing. This testing right after learning also enhances memory strength by changing the expectancy of the future relevance of this memory (Wilhelm, Diekelmann, Molzow, Ayoub, Molle, and Born, 2011), as subject's attention might be shifted towards the memory aspect of the task in contrast to the conditions without immediate recall. This offers the opportunity for an intentional reconsolidation after learning to the advantage of memory strength.

Another explanation for the divergent findings in the literature might be the sensitivity of the retrieval test. Several studies in this field used recognition measures, although very few studies have successfully observed sleep benefits on recognition using different kinds of learning material (see Diekelmann, Wilhelm, and Born, 2009). The authors of the review concluded that cued or free recall tests detect sleep benefits much more reliably than recognition tests (Diekelmann et al. 2009). While the reasons are still unclear, one might argue that sleep mostly facilitates retrieval access by reactivating memories during sleep which is less relevant for recognition, particularly with respect to familiarity judgements. Another possible reason is that recognition performance is typically very high, which might be less ideal to successfully detect sleep-benefits on memory. To increase the sensitivity of retrieval testing even further, an interference learning block can be inserted before the retrieval task. For example, Ellenbogen, Hulbert, Stickgold, Dinges, and Thompson-Schill (2006) showed that interference before retrieval testing resulted in a larger benefit of sleep for memory of neutral word-pairs. These findings were replicated even when controlling for circadian factors and using a more refined behavioral paradigm (Ellenbogen, Hulbert, Jiang, and Stickgold, 2009), as well as in both younger and older adults (Sonni and Spencer, 2015). However, some divergent findings have also been reported, where interference reduced or nullified the sleep benefit (Barsky, Tucker, and Stickgold, 2015; Deliens, Schmitz, Caudron, Mary, Leproult, and Peigneux, 2013), though the interference paradigm used in Deliens et al. (2013) was based on emotional interference rather than interference of another learning set. Thus, if the observation of the benefits of sleep on memory depends on the sensitivity of the retrieval testing, it might be possible to reveal sleep benefits on neutral memories even in the presence of emotional memories by inserting interference before retrieval testing.

Here we tested these two possibilities by systematically varying memory strength and the sensitivity of the retrieval testing. Participants viewed emotional and neutral pictures and freely recalled them after a 12-hour retention interval filled with either sleep or wakefulness (see Figure 1 for overview of the task design and experimental procedure). Half of the participants immediately recalled the pictures additionally before the interval, and the other half did not. Furthermore, half of the participants performed interference learning before free recall testing after the retention interval and the other half did not. We hypothesize that the appearance of the beneficial effect of sleep depends on both the memory strength and the sensitivity of the retrieval testing. Thus, we expect to observe a selective effect of sleep for emotional memories only when the memory strength of neutral memories is low (i.e. without immediate retrieval testing). In addition, we expect that the observed selectivity of sleep for emotional memories will disappear when the sensitivity of the retrieval test is increased (i.e. with interference learning before retrieval).

2. Materials and Methods

2.1 Participants

In total, 235 healthy subjects (158 female, mean age $\pm$ SD = 24.42 $\pm$ 4.11) participated in the experiment. Seven subjects were excluded from analysis due to non-compliance with either task instruction (n=1) or study protocol (n=3), technical difficulties (n=1) or being an outlier in the memory task (> 3 standard deviations of the overall mean; n = 2). This left a total of 228 subjects between 18 and 35 distributed across four experimental groups, each split into a sleep and a wake condition (see Figure 1). Participant characteristics for each experimental group are reported in Table 1. Group IV has previously been reported as the pilot study in Ackermann, Hartmann, Papassotiropoulos, de Quervain, and Rasch (2015). The experimental groups differed neither in their distribution of age, gender, sleep characteristics (sleep quality as indicated by subjective rating between 1 and 10, sleepiness, sleep disturbances) nor mood before the task (all $P \geq .09$). None of the subjects had a night shift or time zone shift of more than 6 hours within 6 weeks before participation in the experiment. Participants were instructed to abstain from caffeine and alcohol during the day(s) of the experiment. Subjects were not asked to follow a regular sleep schedule and adhered to their own sleep schedule in the night between the sessions. The study was approved by the ethics committee of the Department of Psychology, University of Zurich, and all subjects gave written informed consent prior to participating. Subjects received 15CHF per hour for participating in the experiment (between 30CHF and 45CHF in total).

Table 1. Participant characteristics.

Overview of the age, gender and sleep quality distribution between groups.

Group	Condition	N	Age M ± SD	Female	Subjective sleep quality between sessions
I	Sleep	30	24.03 ± 4.36	20	6.93 ± 1.87
	Wake	28	24.11 ± 5.21	20	-
II	Sleep	29	23.38 ± 3.98	19	6.93 ± 1.51
	Wake	28	24.43 ± 4.75	18	-
III	Sleep	30	24.5 ± 3.3	20	6.9 ± 2.33
	Wake	28	24.39 ± 3.25	20	-
IV	Sleep	28	24.96 ± 3.67	19	7.3 ± 2.3
	Wake	27	24.48 ± 3.47	18	-
P			.92	➢ .99	.87

Note. Table 1 presents mean age (M) ± standard deviation (SD), the number of female subjects per group and mean subjective sleep quality which had been rated on a 10 point scale with higher scores reflecting better sleep quality. *P values indicate group main effects.*

[Please insert Figure 1 here]

2.2 Picture Memory Task

The task was taken from Rasch, Spalek, Buholzer, Luechinger, Boesiger, Papassotiropoulos, and de Quervain (2009), and was originally designed to be used in an MRI scanner, while MRI was not used in this study. A total of 72 neutral, negative, and positive images from the International Affective Picture Set (IAPS, see supplementary for image numbers, Lang, Bradley, and Cuthbert, 2008) were presented in a quasirandomized order (24 per valence, see Figure 1A). This allowed for separate analyses of positive and negative items (see supplementary material). The participants were instructed to rate these pictures on valence and arousal, but were left naïve about the later retrieval of the images (incidental encoding). In each trial, a fixation cross was displayed for 500 ms, then the image was presented for 2.5 s before participants were asked to rate its emotional valence (1 = positive , 2 = neutral, 3 = negative) and arousal (1 = high, 2 = medium, 3 = low) each on a three-point-scale. Twenty-four images of geometrical figures on a scrambled background were interleaved (due to the MRI design of the task). Participants were asked to rate their form (wide, quadratic, high) and size (small, medium, large). Intertrial intervals varied between 9 and 12 seconds, including the time for the rating. Two additional neutral images were presented at the beginning and at the end, respectively, which were not analyzed to account for primacy and recency effects. The interference set used an additional 72 IAPS images with matched content (similar content but easily distinguishable) interleaved with 24 geometrical images, presented in a quasirandomized order, with the same valence distribution. The sets were the same for all the participants. The participants rated the valences of the pictures as significantly different ($F_{(2, 214)}$ = 3540.02, P < .001, η_p^2 = 0.94) on a 1-3 scale, with the negative images averaging at 2.79 (SEM: ± 0.01), the neutral images at 1.9 (± 0.01) and the positive images at 1.26 (± 0.01). There was also a trend effect of set, $F_{(1,104)}$ = 3.9, P = .05, η_p^2 = 0.04 showing that overall, the interference set was rated less negative.

The emotional pictures also significantly differed in arousal rates (on a scale of 1-3 with 1 as the most arousing), $F_{(1,214)}$ = 742.43, P < .001 (η_p^2 = 0.78). Negative images were rated most arousing 1.7 (± 0.03), and positive images were more arousing 2.2 (± 0.04) than neutral images, 2.62 (± 0.03), $t_{(215)}$ = 20.6, P < .001. Again there was also a significant effect of set, with the interference set being rated more arousing $F_{(1,106)}$ = 7.96, P = .006 (η_p^2 = 0.07).

For both immediate and delayed retrieval, participants were instructed to use keywords to describe as many images from picture memory task as possible within a maximum of 20 minutes. Participants that saw the interference image set recalled it immediately together with the delayed retrieval of the first set within a time limit of 25 minutes. Two independent raters decided for each described image whether it was present in the image set. When the two raters disagreed (agreement rate 80% - 89%), a third independent rater was consulted. At the end of the experiment, subjects were asked to indicate whether they had expected a later recall of the images during learning (0 = did not expect, 1

= expected). As the immediate retrieval test has not been announced before learning, there was no difference in expectancy between groups with and without immediate retrieval (*M* = 0.1 ± 0.03 and *M* = 0.06 ± 0.02 respectively) *P* = .25. We also tested if the mean of all groups was significantly different from 1 (expected recall) and found a significant difference $t_{(222)}$ = -50.28, *P* < .001. Therefore, we can assume that learning was incidental for most participants (overall only 7.9% participants said they expected a recall). However, including an immediate recall seemed to increase the expectation of another recall in session two as participants in group 4 compared to participants of group 3 showed greater expectation of recalling the images presented in session 1 in session 2 (*M* = 0.2 ± 0.06 and *M* = 0.05 ± 0.03) $t_{(105)}$ = 2.34, *P* = .02, d = 0.47). However, the mean of group 4 was still significantly different from 1 $t_{(48)}$ = -13.68 *P* < .001. Therefore, it seems that most participants (70.9 %) did not expect another recall of the images. Groups 1 and 2 had not answered this question. In the recognition task, 144 images were presented. Half of them were the same as in the first session; the others were unfamiliar (evenly distributed between the positive, negative and neutral images). Participants had to indicate whether the image was new, familiar or if they explicitly remembered the image. After a fixation cross was displayed for 500 ms the images were presented for 1 s.

2.3 Working Memory Task

The working memory task consisted of alternate blocks of 0- and 2-back tasks (versions of the n-back working memory task), and was originally adapted to be used in an MRI scanner. In the 0-back blocks participants were asked to indicate whether an X was presented by either pressing “b” or “n” on the keyboard. In the 2-back blocks subjects had to press “b” when the currently displayed letter matched the letter presented two items before and otherwise press “n”. This task requires maintenance and updating of two stimuli in the working memory. The task was used in all groups after the learning of the picture memory task and if applicable, before the retrieval and served as a distraction.

3. Experimental Design

Each participant was invited to two sessions with an interval of approximately 12 hours between sessions (Min: 10 h Max: 13.5 h, *M* = 11.96 h ± 0.47, see Figure 1, for an overview of the procedure). The sessions took place between 7 – 10 am and 7 – 10 pm. Half of the participants completed the first session in the morning (n= 111) and stayed awake between sessions and the other half started in the evening (n= 117) and slept between sessions. At the beginning, participants filled out a demographic questionnaire. Afterwards, all subjects completed the picture memory task and a working memory task. Participants were assigned to one of four groups (see Figure 1B), each split evenly into a wake and sleep condition. Only participants in groups II and IV completed an immediate retrieval of the learned content at the end of the first session. The second session began with questionnaires after which groups III and IV completed the picture memory task with different images (interference). After 10 minutes of working memory task all participants did a delayed retrieval of the image sets of the first session. Groups III and IV were also asked to recall the images of the Interference set. Groups I and II additionally performed a recognition task of the images they had learned in the first session. A recognition test was added to the design in group I and II due to the expected low performance levels in the free recall task after the retention interval. The session concluded with a post experiment questionnaire. Each session lasted between 1 and 1.5 hours.

4. Calculation

The data were analyzed in SPSS (Version 20.0.0.2) using repeated measure analyses of variance (rANOVAS). Emotion (emotion vs. neutral images) was used as within-subjects factor, interference (interference vs. no interference), condition (sleep vs. wake) and immediate recall (immediate recall vs. no immediate recall) were included as between-subjects factors. In some analysis additionally a between-subjects factor of expected recall was added. In case of significant main effects or interactions, follow-up analyses were performed with ANOVAs and t-tests. For the two nights with immediate retrieval a relative retention performance was calculated for the images remembered at session two with learning performance in session one set to 100%. To ensure that groups did not differ in covariates, several analyses were done using one-way ANOVAs. To examine the correlation of memory performance and sleep quality, we used spearman correlations. The level of significance was set to $P = .05$.

5. Results

5.1 Delayed free recall

As expected, emotional pictures were remembered better after the 12 hour retention interval compared to neutral ones ($M = 29.7 \pm 0.7\%$ of 48 emotional pictures vs. $M = 16.3 \pm 0.6\%$ of 24 neutral pictures correctly recalled, respectively $F_{(1,220)} = 430.7$, $P < .001$; $\eta_p^2 = 0.66$, see Table 1 for an overview and supplementary results for further effects independent of sleep vs. wake). Interestingly, we did not find a main effect of sleep as compared to wakefulness in our overall analysis of variance ($F_{(1,220)} = 1.21$, $P = .27$, $\eta_p^2 = 0.005$).

However, we did find a highly significant three-way interaction between the factors emotion (emotional / neutral), condition (sleep/wake) and immediate retrieval (yes/no) ($F_{(1,220)} = 7.18$, $P = .008$, $\eta_p^2 = 0.03$). In the following sections we will explore this three-way interaction in more detail.

Table 2. Mean Performances on the picture memory task.

Overview of the performance on the Picture Memory Task per group and condition for both the emotional and neutral images.

Group	Condition	N	Immediate Recall				Recall session 2			
			Emotional Images		Neutral Images		Emotional Images		Neutral Images	
			M (%)	SEM	M (%)	SEM	M (%)	SEM	M (%)	SEM
I	Sleep	30					31.32	± 1.62	15.69	± 1.42
	Wake	28					31.92	± 1.77	15.92	± 1.98
II	Sleep	29	45.19	± 1.87	25.43	± 1.62	44.9	± 1.99	23.13	± 1.89
	Wake	28	45.01	± 2.11	29.02	± 2.25	41.29	± 2.37	26.64	± 2.17
III	Sleep	30					14.79	± 1.88	7.92	± 1.46
	Wake	28					16.82	± 1.87	7.14	± 1.47
IV	Sleep	28	42.49	± 2.07	24.85	± 2.25	31.7	± 20.4	18.75	± 2.2
	Wake	27	40.82	± 2.16	24.38	± 2.08	24.23	± 1.99	14.81	± 1.89

Note: Standard error of the means (SEM) are reported.

To further explore the three-way interaction we first tested our hypothesis on the dependency of the selective emotional memory effect of sleep on memory strength. Therefore, we analyzed the groups with and without immediate retrieval separately (i.e., high vs. low memory strength, respectively). In contrast to our hypothesis, we observed a significant two-way interaction between the factors emotion and condition for participants with a high memory strength due to immediate retrieval before sleep ($F_{(1,110)}$ = 7.3, P = .008, η_p^2 = 0.06).

Post-hoc t-tests confirmed that in the groups with immediate retrieval, sleep only benefited memory for emotional images (sleep vs. wake: 38.41% ± 1.66% vs. 32.92% ± 1.92%, $t_{(110)}$ = 2.17, P = .03, d = 0.41) but not for neutral images *(P* = .94, see Figure 2A). This result pattern has been taken as evidence for a "selective" effect of sleep on emotional memories. However, in contrast with the existing literature, the selective benefit of sleep vanished when the pictures were encoded with a low memory strength: Participants without an immediate recall showed no sleep benefit (all $P \geq .57$), neither for emotional nor for neutral images (see Figure 2B). Thus, the "selectivity" of sleep for emotional memories only appeared at a certain memory strength. Interference learning before retrieval after sleep did not interact with the sleep benefit in this analysis (all interactions involving both factors conditions and interference $P > .17$). Rerunning this analysis with the additional factor emotional valence (negative vs. positive pictures) did not alter the general result pattern (all interactions with emotional valence $P > .22$, please see supplementary results for separate analyses for negative vs. neutral and positive vs. neutral images). Neither did adding the factor of expected recall (all interactions with expected recall $P \geq .13$).

[Please insert Figure 2 about here]

5.2 Controlling recall performance for memory strength

Results from the delayed free recall test suggested that the sleep benefit is selective for emotional pictures only when memory strength is high (i.e., with immediate retrieval). However, emotional arousal influences the strength of the memory trace already at encoding. Thus, already during immediate recall testing, emotional pictures are much better retrieved than neutral ones ($F_{(1,108)} = 277.78$, $P < .001$, $\eta_p^2 = 0.72$). Furthermore, immediate and delayed retrieval performance are highly correlated (emotion: r = .68; neutral: r = .79; both $P < .001$). Even more problematic, the sleep group that learned in the evening showed a slightly increased emotional memory already during immediate recall testing as compared to the wake group, although this potential circadian confound did not reach significance (interaction emotion * condition during immediate recall testing $F_{(1,110)} = 1.4$, $P = .24$, $\eta_p^2 = 0.01$). Note that independent of emotion, sleep and wake groups did not differ in their immediate recall performance ($P > .90$).

Due to the emotion-related difference in memory strength during the encoding phase and the potential confounding of circadian time, we re-analyzed our data including only groups that performed immediate retrieval before the retention interval (group II and IV, see Figure 1B). We defined the retention score as the relative change in retrieval performance across the retention interval, with immediate retrieval performance set to 100%. Importantly, relative retention scores did not correlate with immediate retrieval performance anymore (emotional: r = -.02. $P = .87$; neutral: r = .0, $P > .99$). In contrast to our previous analyses, emotion did not influence the retention score when controlling for memory strength at encoding, resulting in similar retention levels for emotional and neutral pictures (82.03 $\pm$ 1.75% vs. 80.4 $\pm$ 2.79%; $F_{(1,106)} < 1$, $P = .59$). Furthermore, the sleep benefit was no longer specific for emotional pictures, but sleep generally improved retention scores for both emotional and neutral pictures, resulting in a significant main effect of condition (sleep: $M = 86.66 \pm 2.47\%$; wake $M = 75.77 \pm 2.51\%$, $F_{(1,106)} = 9.57$, $P = .003$, $\eta_p^2 = 0.08$). The interaction between emotion and condition (sleep/wake) was not significant anymore ($F_{(1,106)}$ = 0.08, P = .78). Interestingly in this analysis, the interaction between condition and interference nearly gained significance ($F_{(1,106)}$ = 3.71, P = .057, η_p^2 = 0.03).

Post hoc tests revealed a larger sleep effect when interference learning was introduced before delayed retrieval (sleep M = 75.43 ± 3.18% vs. wake: M = 59.36 ± 3.5%; $t_{(53)}$ = 3.4, P = .001, d = 0.92), while the effect of sleep on memory for pictures was smaller without interference learning (sleep M = 97.55 ± 2.25% vs. wake: M = 91.84 ± 2.82%, $t_{(55)}$ = 1.59, P = .12, d = 0.43). Thus, interference learning before delayed retrieval testing more than doubled the effect size of sleep on memory. Exploratory pair-wise comparisons suggested that sleep benefited memory for emotional and neutral pictures robustly and equally when interference learning was performed before retrieval (emotional: $t_{(53)}$ = 2.73, P = .008, neutral: $t_{(52)}$ = 2.27, P = .027). The effect sizes of sleep on memory were highly comparable in this condition between emotional (d = 0.74) and neutral pictures (d = 0.62, see Figure 3A). In contrast, without interference, only the sleep benefit for emotional pictures reached significance ($t_{(55)}$ = 2.06, P = .045, d = 0.55, see Figure 3B), while no effect of sleep on memory for neutral pictures was observed in this condition (P > .99). However, please note that three-way interaction between emotion, interference and condition did not reach significance (3-way interaction P = .29), indicating that the effect of interference on the sleep benefit of memory was not significantly modulated by emotional arousal. We additionally calculated the ANOVA with the factor expected recall in the second session, however neither the main effect, nor any interaction with expected recall was significant (P ≥ .38). On average, participants that expected a recall performed slightly worse (M = 65.11 % ± 7.08% vs M = 67.51% ± 3.38%). In addition to free recall, we also tested recognition in these two groups. However, we only found a trend for a main effect of condition (sleep M = 83.4 ± 1.5 vs. wake M = 79.8 ± 1.5, P = .096), suggesting that recognition testing is less sensitive to capture sleep benefits on memories for pictures.

5.3 Correlation with subjective sleep quality

We calculated the correlation between the subjective sleep quality reported for the retention interval and the free recall memory performance. We found that subjective sleep quality correlated negatively with percentage of remembered emotional ($r_{s(114)}$ = -0.18, P = .049) and neutral pictures ($r_{s(114)}$ = -0.19, P = .042). The correlation with relative retention performance in groups II and IV with immediate recall was not significant (P > .29).

6. Discussion

Here we show that the previously reported benefit of sleep on emotional memory is highly sensitive to experimental manipulations of memory strength and sensitivity of the retrieval test. First, a selective sleep benefit for emotional pictures is only observed if the memory strength before the retention interval is sufficiently high (i.e., with immediate retrieval). Second, when controlling for differences in memory strength before the retention interval between emotional and neutral pictures, the advantage of emotional pictures disappears entirely, particularly when a more sensitive way of retrieval testing is used (i.e., challenging the memory trace with interference learning before retrieval). Thus, our results indicate that the effect of emotion on sleep dependent memory consolidation is only apparent under certain circumstances. To investigate factors that tag memories for consolidation, memory strength and retrieval test sensitivity have to be taken into account as they can influence detectability of effects. These results were all obtained using free recall measures. In the recognition task we did not find any significant effects of sleep. It seems that this retrieval test was not sensitive enough to uncover the sleep-related processes, as performance was very high across all participants.

Our result pattern first suggested that only emotional pictures with immediate recall had reached a sufficiently strong trace to benefit from sleep (immediate recall emotional pictures: 43.46 ± 1.0 % , neutral pictures: 26.02 ± 1.1 %). However, when controlling for differences in memory strength before the retention interval, the selectivity of sleep for emotional memories was no longer significant. Thus either both, emotion and immediate recall have contributed to detect the sleep benefit or memory strength acts as a moderator: As emotion influences memory strength, we consider it possible that not emotion per se influences consolidation, but that emotional events increase the chance of consolidation via their influence on memory strength. When memory performance level for emotional items without immediate recall was comparable to neutral items after immediate recall, both did not benefit from sleep. This could also explain why the sleep benefit seems absent for neutral items when emotional memories are interleaved. Memory strength for emotional content is stronger and increases the chances of sleep to benefit memory. Consequently, other factors that influence memory strength should show preferential consolidation as well. As literature shows, memory strength is influenced by several factors such as emotionality, intention and expectation. However, if memory strength is the deciding factor for memory consolidation during sleep, then whatever influences memory strength also alters the chance of detecting the sleep benefit. Besides mere effects on the detection of the benefit that are related to the way of testing, sleep might additionally differentially process weaker and stronger memory traces. Some studies have shown that a certain threshold has to be reached before consolidation in sleep can take place (Tucker and Fishbein, 2008). Conversely, Drosopoulos et al. (2007) have reported that also intense encoding (> 90% correct before sleep) abolishes the sleep benefit of memory, while reliable sleep effects are observed at a medium memory strength (e.g. 60% correct before the retention interval). This conclusion was also drawn by Creery, Oudiette, Antony, and Paller (2015) demonstrating limited cueing benefits for nearly perfectly remembered items. The sleep benefit might indeed depend on an inverted U-shaped curve of memory strength before sleep, with maximal benefits of sleep in a medium range and lower or even no sleep effect in the cases of low or high encoding intensity. Interestingly, while additional retrieval testing enhances the benefit of sleep in our study, Bauml, Holterman, and Abel (2014) reported in a series of studies that retrieval practice before sleep abolishes the sleep effect on memory. How can this apparent contradiction be concealed? Again, we would argue that the different outcome is due to the difference in initial memory strengths in both studies. In our study, retrieval practice increased memory strengths from 11-24% to 20-38 %. In the study of Bauml et al. (2014), retrieval practice increased initial memory strength from 75% to almost 90%. %. Thus, according to our model, retrieval practice can induce a sleep benefit or attenuate it depending on the initial location on the inverted U-shaped curve (e.g. low (11-38%) vs. high (75-90%)) describing the association between initial memory strength and the sleep benefit. Please note that some studies have reported small benefits of sleep on memory even with very high encoding levels (see e.g. Ellenbogen et al., 2006; Jenkins and Dallenbach, 1924), indicating that the detection of sleep-benefits on memory might still be possible at high initial encoding levels under certain circumstances (e.g. short word lists, low level of integration etc.). However, these examples might also indicate that immediate recall levels are possibly not the best estimator to determine initial memory strength, as these measures might be confounded by short-term memory and immediate retrieval processes in case sufficient initially not remembered items are close enough to the threshold to benefit from sleep.

What brain mechanisms could underlie the influence of initial memory strength on consolidation processes during sleep? First, involvement of the hippocampal system during encoding might be a necessary condition to trigger later sleep-dependent memory consolidation. For example in the context of motor memories, hippocampal activity as well as connectivity during encoding predict later consolidation during sleep (Albouy, King, Maquet, and Doyon, 2013; Albouy, Sterpenich, Vandewalle, Darsaud, Gais, Rauchs, Desseilles, Boly, Dang-Vu, Balteau, Degueldre, Phillips, Luxen, and Maquet, 2013). Similarly, for declarative memories, hippocampal activity at encoding predicts later sleep-consolidation in a directed forgetting paradigm (Rauchs, Feyers, Landeau, Bastin, Luxen, Maquet, and Collette, 2011). In addition, fast mapping of associative memories does not profit from sleep after learning, presumably because of direct cortical access during encoding with less hippocampal involvement (Himmer, Muller, Gais, and Schonauer, 2017).

Thus, one could argue that a certain level of encoding is needed to induce sufficient hippocampal activity and plasticity, which in turn is necessary to trigger later sleep-related memory consolidation. Furthermore, a certain degree of initial hippocampal-cortical connections during encoding might simultaneously be required to set the stage for successful systems consolidation during sleep. However, if memory traces are already strongly stored in cortical storage sites (for example by excessive training, several repetitions or overlearning), the benefit of sleep might be attenuated and only visible using highly sensitive retrieval procedures or longer time intervals. In sum, we would predict that benefits of sleep on memory are maximal when a) hippocampal and hippocampal-cortical involvement during encoding pass a certain threshold and b) strength of hippocampal-cortical and cortico-cortical memory traces is low. How can this interpretation be applied to the existing literature? Although in Payne et al. (2015), nap and wake groups did not differ in neutral memory performance, more SWS in the nap groups was still associated with better neutral object memory. The alternative explanation that neutral objects were just too weakly encoded to benefit from sleep to such an extent that it would withstand the between-group analysis cannot be ruled out by these findings. Indeed, in both studies (Payne et al., 2008; Payne et al., 2015) the only significant group effects were shown for emotional objects, which were remembered better than neutral objects as well as the backgrounds. Thus, possibly, the sleep effect only appeared for the negative object items as only they achieved a certain memory strength threshold that enabled preferential sleep-dependent consolidation.

Wagner et al. (2001) also found preferential consolidation of emotional texts compared to neutral content. However, they reported that subjects had rated the emotional texts as more comprehensible and that the narrative character of the emotional texts might have evoked differences in encoding levels. Additionally, the emotional texts were also rated as being easier, more arousing, more interesting and more important. This makes it hard to determine whether emotionality per se elicited the effect whether these components affected memory strength. In Hu et al. (2006) the pre-sleep learning level was not tested as recognition only followed after the retention interval. However, one might assume that memory strength of emotional pictures was higher as compared to neutral pictures already before the retention interval. Thus, differences in memory strength between emotional and neutral pictures might again explain the result pattern. Additionally, it is quite puzzling why the authors observed the sleep benefit only for “know” judgements and not for “remembered” judgements.

Thus, for all studies that found superior consolidation for emotional memory (Hu et al., 2006; Payne et al., 2015; Wagner et al., 2001) memory strength as a confounding variable cannot be excluded. Although our results of the current study support a role of initial memory strength, we cannot (and do not want to) exclude additional selective consolidation processes acting during sleep. For example, Bennion et al. (2015) provided the idea that information might be tagged before sleep rendering this marked content more likely to be consolidated during sleep. It is possible that relevance tagging and initial memory strength both act and interact with respect to later consolidation processes during sleep.

Besides memory strength, increasing sensitivity of retrieval test additionally enhanced the detectability of the sleep benefit for neutral items. It is very important to note that the benefits of sleep on memory (and also the mechanisms acting on the memory trace during sleep) do of course not depend on the choice of a specific retrieval procedure, as consolidation during sleep occurs before the retrieval test. However, the *statistical detection* of the sleep benefit strongly does. For example, adding interference before delayed recall in the current study did not change the sleep mechanism per se. But it revealed that sleep did strengthen memory for neutral items. This fact has strong implications for the interpretation of the results: For example, if the difference between a sleep and wake group is not significant, sleep might still have strengthened the underlying memory trace. In other words, the sleep effect might be masked by the specific retrieval procedure. Thus, the results of a retrieval task after sleep are not always indicative for mechanisms acting during sleep, but might be related to the sensitivity of a specific retrieval procedure.

Several limitations need to be considered regarding our study. First, our study used different experimental procedures and material than previous studies, which renders direct comparisons to other studies difficult. Our retrieval test differed for instance from Payne et al. (2009). While they used only recognition testing, we focused on free recall with and without interference using recognition only in a subset of groups. Free recall usually leads to lower memory performance compared to recognition and therefore recall is lower in our sample compared to other studies. The performance in the recognition task was rather high, compared to other studies, due to recognition being completed after the free recall.

Second, different from other studies, we included an immediate recall in half of our participants before the retention interval. While learning for all groups was incidental, this changed the expectation of participants about having to recall the images in the future. However, on average the expectancy was still low and we did not find a main effect or interaction with expected recall. It is therefore unclear, if the memory strengthening effect depends on a reconsolidation during the testing or on a change of expectations. It was however not the intent of the current study to identify the mechanism involved in the testing effect, which has been discussed in several other studies (see van den Broek, Takashima, Segers, Fernandez, and Verhoeven, 2013) but to demonstrate that the detection of the sleep benefit is influenced by memory strength. Thus, while immediate retrieval has enhanced memory as demonstrated by the data, the mechanism how this enhancement happened is unclear.

Third, while most studies have focused on negative and neutral items to test effects of emotion on memory, we included positive items as well. While it has been argued that sleep-dependent consolidation works differently on items of different affect (Jones, Schultz, Adams, Baran, and Spencer, 2016) and that using both valences at the same time can blur effects of emotionality, we could still demonstrate effects of emotion on sleep-dependent consolidation when not accounting for pre-sleep levels. Because we used both positive and negative images, we had a higher total of emotional than neutral images. However, as we always used percentages (of the total amount of images and of previously remembered images) we controlled for this. Furthermore we tested the two emotional valences (positive and negative) compared to neutral separately (see supplementary material) and found the effects where similar in both valences, however did not always reach significance in both.

Fourth, we did not measure any objective sleep data, but only asked for a subjective assessment of sleep quality between the sessions.

7. Conclusion

In sum, our results show that whether a sleep-dependent memory consolidation effect is detected in the data does not only depend on actual sleep consolidation processes. Several factors such as memory strength and sensitivity of the retrieval test can influence traceability of the sleep benefit. Neutral memories for instance also benefit from sleep, but the detection of these benefits depends on sufficient memory strength and sensitive retrieval testing procedures.

Given a certain memory strength and a sensitive retrieval measure, the sleep effect on emotional and neutral pictures results in the same effect size. These results add to the notion of a preferential consolidation of emotional memories during sleep although the preference for consolidation of emotional memory during sleep is not due to emotion alone, but also because emotion influences memory strength. Our results demonstrate that it is vital to determine factors that influence the appearance of the sleep benefit in order to avoid confounding variables and unclear conclusions. The appearance of a sleep benefit in behavioral measures alone is not sufficient to draw strong theoretical or mechanistic conclusions.

Acknowledgments

This work was supported by the Clinical Research Priority Program ‘Sleep and Health’ and of the University of Zurich and a grant from the Swiss National Science Foundation (SNF 100014_162388).

We thank Manuela Steinauer for assistance in data collection and Arlette Bär, Clemens Schröder, Fiona Hart-Hönig and Tobias Egli for help in analyzing the data.

References

Ackermann, S., Hartmann, F., Papassotiropoulos, A., de Quervain, D. J. F., & Rasch, B. (2015). No Associations between Interindividual Differences in Sleep Parameters and Episodic Memory Consolidation. *Sleep, 38*, 951-U263.

Adelman, J. S., & Estes, Z. (2013). Emotion and memory: A recognition advantage for positive and negative words independent of arousal. *Cognition, 129*, 530-535.

Baran, B., Pace-Schott, E. F., Ericson, C., & Spencer, R. M. (2012). Processing of emotional reactivity and emotional memory over sleep. *Journal of Neuroscience, 32*, 1035-1042.

Barsky, M. M., Tucker, M. A., & Stickgold, R. (2015). REM sleep enhancement of probabilistic classification learning is sensitive to subsequent interference. *Neurobiology of Learning and Memory, 122*, 63-68.

Bauml, K. H. T., Holterman, C., & Abel, M. (2014). Sleep Can Reduce the Testing Effect: It Enhances Recall of Restudied Items but Can Leave Recall of Retrieved Items Unaffected. *Journal of Experimental Psychology-Learning Memory and Cognition, 40*, 1568-1581.

Bennion, K. A., Payne, J. D., & Kensinger, E. A. (2015). Selective effects of sleep on emotional memory: What mechanisms are responsible? *Translational Issues in Psychological Science, 1*, 79.

Blake, T. M., Varnhagen, C. K., & Parent, M. B. (2001). Emotionally arousing pictures increase blood glucose levels and enhance recall. *Neurobiology of Learning and Memory, 75*, 262-273.

Born, J., & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological Research-Psychologische Forschung, 76*, 192-203.

Cahill, L., & McGaugh, J. L. (1995). A novel demonstration of enhanced memory associated with emotional arousal. *Consciousness and Cognition, 4*, 410-421.

Canli, T., Zhao, Z., Brewer, J., Gabrieli, J. D. E., & Cahill, L. (2000). Event-related activation in the human amygdala associates with later memory for individual emotional experience. *Journal of Neuroscience, 20*.

Creery, J. D., Oudiette, D., Antony, J. W., & Paller, K. A. (2015). Targeted Memory Reactivation during Sleep Depends on Prior Learning. *Sleep, 38*, 755-763.

Deliens, G., Schmitz, R., Caudron, I., Mary, A., Leproult, R., & Peigneux, P. (2013). Does Recall after Sleep-Dependent Memory Consolidation Reinstate Sensitivity to Retroactive Interference? *Plos One, 8*.

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

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.

Ellenbogen, J. M., Hulbert, J. C., Jiang, Y., & Stickgold, R. (2009). The Sleeping Brain's Influence on Verbal Memory: Boosting Resistance to Interference. *Plos One, 4*.

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.

Harris, C. R., & Pashler, H. (2005). Enhanced memory for negatively emotionally charged pictures without selective rumination. *Emotion, 5*, 191-199.

Heuer, F., & Reisberg, D. (1990). Vivid Memories of Emotional Events - the Accuracy of Remembered Minutiae. *Memory & Cognition, 18*, 496-506.

Himmer, L., Muller, E., Gais, S., & Schonauer, M. (2017). Sleep-mediated memory consolidation depends on the level of integration at encoding. *Neurobiology of Learning and Memory, 137*, 101-106.

Hu, P., Stylos-Allan, M., & Walker, M. P. (2006). Sleep facilitates consolidation of emotional declarative memory. *Psychological Science, 17*, 891-898.

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

Jones, B. J., Schultz, K. S., Adams, S., Baran, B., & Spencer, R. M. C. (2016). Emotional bias of sleep-dependent processing shifts from negative to positive with aging. *Neurobiology of Aging, 45*, 178-189.
Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition, 31*, 1169-1180.
Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. *Technical report*.
Lehmann, M., Seifritz, E., & Rasch, B. (2016). Sleep benefits emotional and neutral associative memories equally. *Somnologie, 20*, 47-53.
Lewis, P., Cairney, S., Manning, L., & Critchley, H. (2011). The impact of overnight consolidation upon memory for emotional and neutral encoding contexts. *Neuropsychologia, 49*, 2619-2629.
Payne, J., Stickgold, R., Swanberg, K., & Kensinger, E. (2008). Sleep preferentially enhances memory for emotional components of scenes. *International Journal of Psychology, 43*, 735-735.
Payne, J. D., & Kensinger, E. A. (2011). Sleep Leads to Changes in the Emotional Memory Trace: Evidence from fMRI. *Journal of Cognitive Neuroscience, 23*, 1285-1297.
Payne, J. D., Kensinger, E. A., Wamsley, E. J., Spreng, R. N., Alger, S. E., Gibler, K., Schacter, D. L., & Stickgold, R. (2015). Napping and the Selective Consolidation of Negative Aspects of Scenes. *Emotion, 15*, 176-186.
Phelps, E. A. (2004). Human emotion and memory: interactions of the amygdala and hippocampal complex. *Current Opinion in Neurobiology, 14*, 198-202.
Plihal, W., & Born, J. (1997). Effects of early and late nocturnal sleep on declarative and procedural memory. *Journal of Cognitive Neuroscience, 9*, 534-547.
Rasch, B., & Born, J. (2013). About Sleep's Role in Memory. *Physiological Reviews, 93*, 681-766.
Rasch, B., Spalek, K., Buholzer, S., Luechinger, R., Boesiger, P., Papassotiropoulos, A., & de Quervain, D. J. F. (2009). A genetic variation of the noradrenergic system is related to differential amygdala activation during encoding of emotional memories. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 19191-19196.
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. *Journal of Neuroscience, 31*, 2563-2568.
Rowland, C. A. (2014). The Effect of Testing Versus Restudy on Retention: A Meta-Analytic Review of the Testing Effect. *Psychological Bulletin, 140*, 1432-1463.
Sonni, A., & Spencer, R. M. C. (2015). Sleep protects memories from interference in older adults. *Neurobiology of Aging, 36*, 2272-2281.
Stickgold, R., & Walker, M. P. (2013). Sleep-dependent memory triage: evolving generalization through selective processing. *Nature Neuroscience, 16*, 139-145.
Sutterer, D. W., & Awh, E. (2015). Retrieval practice enhances the accessibility but not the quality of memory. *Psychonomic Bulletin & Review*, 1-11.
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.
van den Broek, G. S. E., Takashima, A., Segers, E., Fernandez, G., & Verhoeven, L. (2013). Neural correlates of testing effects in vocabulary learning. *Neuroimage, 78*, 94-102.
Wagner, U., Gais, S., & Born, J. (2001). Emotional memory formation is enhanced across sleep intervals with high amounts of rapid eye movement sleep. *Learning & Memory, 8*, 112-119.
Wilhelm, I., Diekelmann, S., Molzow, I., Ayoub, A., Molle, M., & Born, J. (2011). Sleep selectively enhances memory expected to be of future relevance. *Journal of Neuroscience, 31*, 1563-1569.

Figure 1

Figure 1. Task design (Adapted from Rasch et al., 2009) and experimental procedure. A) A schematic representation of one trial of the Picture Memory Task. After the fixation cross an image of the IAPS was presented for 2.5s and subsequently rated on valence and arousal. The intertrial interval including the ratings varied between 9 and 12 seconds. B) Participants were assigned to one of four groups and within this group to either the wake (white) or sleep condition (black).

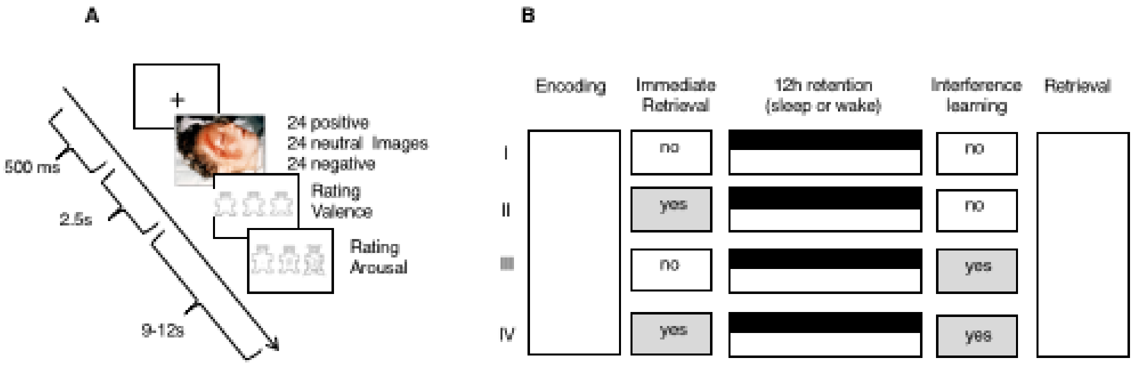

Figure 2.

Figure 2. Influence of condition (sleep vs. wake) on the amount of remembered images at delayed retrieval after the retention interval. The retrieval score is calculated relative to the total amount of pictures (48 emotional pictures, 24 neutral pictures). (A). When participants preformed an immediate retrieval test before the retention interval, sleep benefited memory significantly only for emotional, but not neutral pictures (interaction effect P = .008, see graph for post hoc pairwise comparisons). (B) Without immediate retrieval testing no sleep benefit on memory was observed, neither for emotional nor neutral pictures (all P > .57). **: P < .05. Means ± SEM are indicated.*

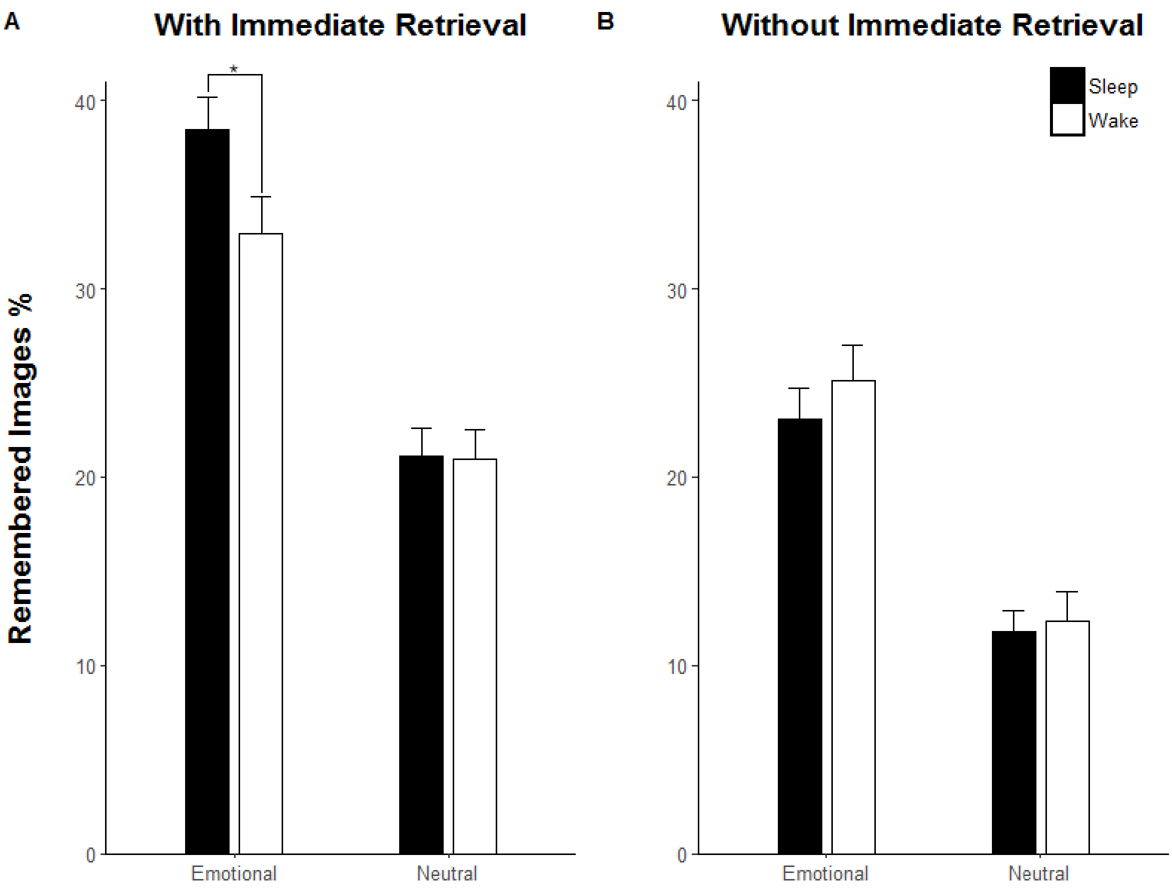

Figure 3.

Figure 3. Influence of interference learning before retrieval on retention of emotional and neutral pictures across sleep vs. wakefulness. Retention scores are calculated with learning performance before the retention interval set to 100%, thereby controlling for differences in memory performance at encoding. (A) With interference learning before delayed retrieval, both emotional and neutral pictures similarly benefited from sleep as compared to a retention interval filled with wakefulness. The interaction between condition (sleep/wake) and interference learning (yes/no) was significant (*P* = .05), whereas no significant three-way interaction with the factor emotional arousal emerged (*P* = .26). Thus, enhancing retrieval sensitivity by including interference did not lead to a difference between emotional and neutral pictures concerning the sleep benefit. (B) Without interference learning before delayed retrieval after the retention interval, the sleep benefit was only significant for emotional, but not neutral pictures. **: *P* < .01; *: *P* < .05. *Means ± SEM are indicated.*