

# SCIENTIFIC REPORTS



OPEN

## Sleep selectively stabilizes contextual aspects of negative memories

Roy Cox<sup>1,2,3</sup>, Marthe L. V. van Bronkhorst<sup>1</sup>, Mollie Bayda<sup>1</sup>, Herron Gomillion<sup>1</sup>, Eileen Cho<sup>1</sup>, Mittie E. Parr<sup>1</sup>, Olivia P. Manickas-Hill<sup>1</sup>, Anna C. Schapiro<sup>1,2</sup> & Robert Stickgold<sup>1,2</sup>

Sleep and emotion are both powerful modulators of the long-term stability of episodic memories, but precisely how these factors interact remains unresolved. We assessed changes in item recognition, contextual memory, and affective tone for negative and neutral memories across a 12 h interval containing sleep or wakefulness in 71 human volunteers. Our data indicate a sleep-dependent stabilization of negative contextual memories, in a way not seen for neutral memories, item recognition, or across wakefulness. Furthermore, retention of contextual memories was positively associated with the proportion of time spent in non-rapid eye movement sleep in a valence-independent manner. Finally, while affective responses to previously seen negative stimuli and to both old and new neutral stimuli decreased across an interval of sleep, effects for memorized items did not differ reliably between sleep and wake. These results add to our understanding of the complex interrelations among sleep, memory, and emotion.

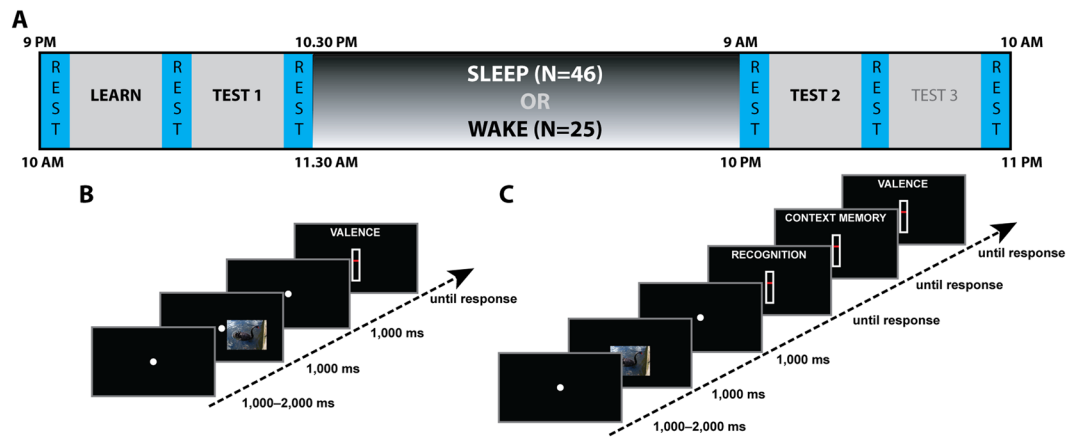
Episodic memories are rich, multisensory accounts of past events that guide our future behavior. Besides linking together various aspects of the external world into an integrated whole (e.g., objects, locations, temporal order), episodic memories also contain the emotional tone associated with an experience, resulting in qualitative and quantitative differences between emotional and neutral memories<sup>1</sup>. In particular, negative emotional memories are typically more durable than neutral memories, especially after longer retention intervals<sup>2–5</sup>.

Sleep plays a fundamental role in processing episodic memories, both stabilizing recent experiences and integrating them with stored knowledge<sup>6,7</sup>. These processes are thought to rely on the reinstatement of learning-related activity patterns in the neural structures originally encoding the memory, allowing hippocampal memory traces to be recoded in a more stable neocortical format<sup>8–10</sup>. Sleep-dependent memory consolidation effects are generally more reliable for associative and contextual memories compared to individual item recognition (e.g.<sup>11,12</sup>), consistent with stronger initial hippocampal involvement for memories with an associative component<sup>13</sup>.

Some evidence suggests that sleep and emotion interact, with enhanced retention of negative relative to neutral material following sleep<sup>14–17</sup>, although other studies have failed to find such an effect<sup>18–22</sup>. Furthermore, while several examinations into the contributions of specific sleep stages to emotional memory consolidation have emphasized the importance of rapid eye movement (REM) sleep<sup>17,23–25</sup>, others point to a role for non-rapid eye movement (NREM) sleep<sup>26–28</sup>, or find no clear stage dependence<sup>20,29–31</sup>. These inconsistencies may be related to various experimental design factors (e.g., sleep vs. wake; full night vs. nap, split-night; sleep-deprivation; targeted memory reactivation), as well as the memory paradigm employed. Specifically, the clearer effect of sleep on hippocampus-dependent associative memories compared to item recognition could result in greater emotional benefits for memories with an associative component, but again, the available evidence is contradictory<sup>25,32–34</sup>.

Another factor in sleep-related emotional processing concerns not the memories *per se*, but the modulation of their affective component. Whereas some physiological evidence suggests that sleep serves to decrease emotional reactivity to previously seen emotional items<sup>35,36</sup>, subjective ratings of emotionality have yielded discrepant findings. Some studies report no difference in emotional reactivity to emotional or neutral items, with similar decreases<sup>24</sup> or increases<sup>18</sup> across sleep versus wake, or comparable ratings after NREM and REM sleep for both emotional categories<sup>23</sup>. Other investigators have described a REM-dependent preservation of negative

<sup>1</sup>Department of Psychiatry, Beth Israel Deaconess Medical Center, Boston, MA, USA. <sup>2</sup>Department of Psychiatry, Harvard Medical School, Boston, MA, USA. <sup>3</sup>Department of Epileptology, University of Bonn, Bonn, Germany. Correspondence and requests for materials should be addressed to R.C. (email: [roycox.roycox@gmail.com](mailto:roycox.roycox@gmail.com))



**Figure 1.** Protocol overview. (A) Study timeline. (B) Trial structure at encoding (with lateralized stimulus presentation). (C) Trial structure at retrieval (with central stimulus presentation).

ratings relative to wake<sup>19</sup> or REM deprivation<sup>37</sup>, while yet others have reported enhanced emotional reactivity after REM-rich sleep<sup>32,38</sup>.

In sum, sleep's role in emotional processing remains unclear, both with regards to memory consolidation and affective processing (see<sup>39</sup> for a recent review). Here, we re-examine the role of sleep in emotional processing, focusing on both recognition and contextual memory, as well as subjective affective ratings. Participants encoded negative and neutral item-context associations in either the evening or morning, and were tested both immediately after encoding and after a 12 h interval containing sleep or wake. Polysomnography was recorded for the sleep group in order to investigate associations with sleep architecture. We hypothesized that sleep would preferentially stabilize the contextual aspects of negative memories. Given the conflicting findings reviewed above, we did not have strong predictions for recognition memory or affective ratings, nor did we have clear expectations on the involvement of specific sleep stages.

## Results

**Protocol.** Seventy-one healthy volunteers underwent a protocol consisting of memory encoding and both immediate and delayed retrieval, with retrieval tests organized around a ~12 h retention interval containing wakefulness or sleep (Fig. 1A). Participants in the wake group ( $N = 25$ ) performed encoding and immediate testing (Test 1) in the morning, and delayed testing at night (Test 2). This timing was reversed for participants in the sleep group ( $N = 46$ ). Furthermore, 62-channel EEG was recorded throughout the procedure in the sleep group to allow sleep staging as well as future examinations of the neural correlates of encoding, retrieval, and memory reactivation during sleep.

During encoding, participants maintained fixation in the center of the screen while a series of visual stimuli was briefly presented in either the left or right visual field (VF) (Fig. 1B). Importantly, all pictures shown on one side were of mild positive emotional valence and low arousal (“neutral valence”), while those presented on the opposite side were of strong negative valence and high arousal (“negative valence”); this manipulation served to facilitate future decoding of category-specific EEG signatures (see Discussion). Subjects rated each item's perceived valence immediately following its presentation on a continuous scale from very negative to very positive. Participants were not told that their memory would be tested later to render the encoding phase incidental. Assignment of emotional category to VF was counterbalanced across subjects. This additional factor SIDE (negative-left/neutral-right:  $N = 36$ ; negative-right/neutral-left:  $N = 35$ ) was included to ensure that any differences between negative and neutral categories could not be due to asymmetric emotional processing by the two hemispheres<sup>40–42</sup>.

During the immediate and delayed retrieval tests, subjects viewed previously seen (“old”) and novel (“new”) stimuli presented in the center of the screen (as opposed to lateralized; Fig. 1C). Different old and new items were used for immediate and delayed tests. Subjects then used sequentially presented continuous scales to rate each item's 1) old/new status (recognition memory), 2) VF of presentation during encoding (contextual memory), and 3) perceived valence. For items rated “new”, subjects rated VF as if they were actually “old” items (see Methods).

**Control Analyses.** We first assessed whether time of day affected vigilance and whether that, or the VF of negative item presentation, had any effect on valence ratings or sleep architecture.

**Vigilance.** We determined subjective vigilance in two ways both prior to encoding and prior to the delayed test. Sleepiness, as assessed by the Stanford Sleepiness Scale<sup>43</sup> (SSS), did not differ between the sleep and wake groups either before encoding (sleep:  $2.6 \pm 0.8$ ; wake:  $2.2 \pm 0.6$ ; Wilcoxon Rank Sum:  $Z = 1.7$ ,  $P = 0.09$ ) or after the 12 h interval (sleep:  $2.5 \pm 1.1$ ; wake:  $2.4 \pm 0.9$ ;  $Z = 0.2$ ,  $P = 0.81$ ). Similarly, self-reported ability to concentrate (continuous scale from 0–100) did not differ systematically pre-encoding (sleep:  $74.4 \pm 15.6$ ; wake:  $80.6 \pm 13.8$ ;  $t(69) = 1.7$ ,  $P = 0.10$ ) or before the delayed test (sleep:  $74.1 \pm 18.4$ ; wake:  $75.8 \pm 15.9$ ;  $t(69) = 0.4$ ,  $P = 0.70$ ), indicating no time of day effect on vigilance.

Category	SLEEP		WAKE		Group		Valence		Group*valence	
	negative	neutral	negative	neutral	F	P	F	P	F	P
encoding	-0.41 ± 0.12	0.28 ± 0.11	-0.39 ± 0.13	0.34 ± 0.16	3.4	0.07	<b>975.4</b>	<10 <sup>-41</sup>	0.4	0.51
Test 1-old	-0.50 ± 0.12	0.28 ± 0.11	-0.45 ± 0.13	0.30 ± 0.16	1.9	0.18	<b>1067.2</b>	<10 <sup>-42</sup>	0.4	0.53
Test 1-new	-0.46 ± 0.13	0.25 ± 0.13	-0.42 ± 0.13	0.27 ± 0.13	1.5	0.23	<b>1028.4</b>	<10 <sup>-42</sup>	0.2	0.64

**Table 1.** Valence ratings at encoding and immediate test (mean ± SD). F test degrees of freedom: (1,69). Significant effects indicated in bold.

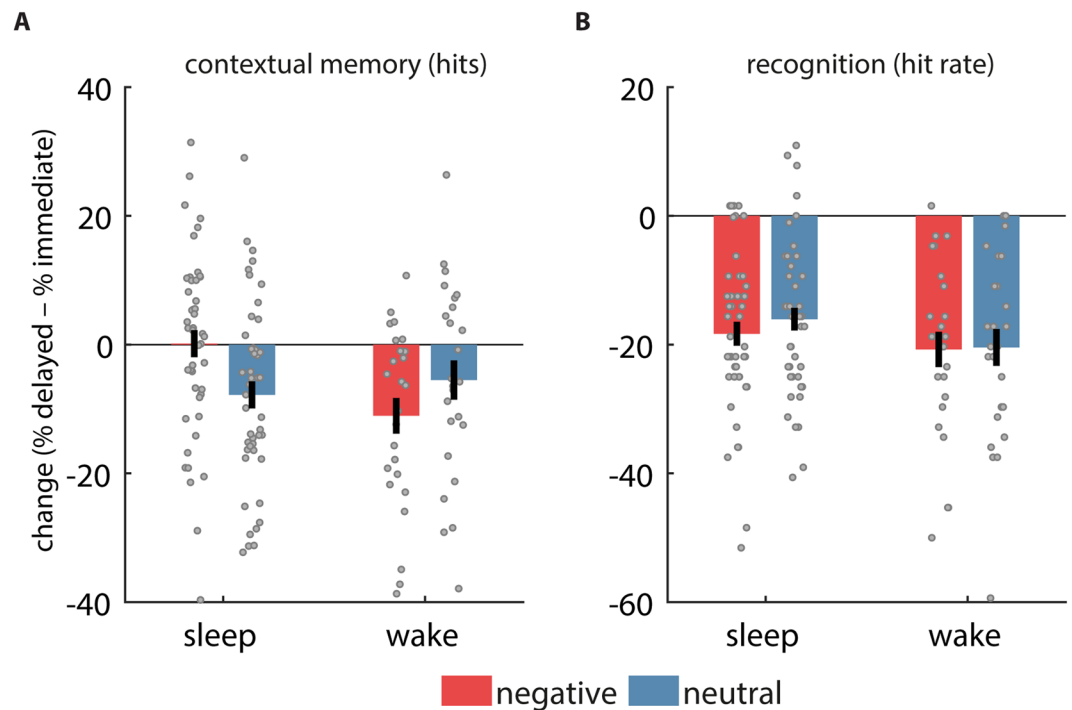
	Overall	Negative-left	Negative-right	Statistics		
	N = 46	N = 22	N = 24	t(44)	P	P <sub>adj</sub>
N1 (%)	3.9 ± 2.7	4.4 ± 2.8	3.4 ± 2.6	1.3	0.20	0.35
N2 (%)	51.1 ± 6.2	52.0 ± 6.1	50.3 ± 6.3	0.9	0.37	0.46
N3 (%)	24.5 ± 5.7	24.0 ± 5.8	25.1 ± 5.7	-0.7	0.51	0.59
REM (%)	20.5 ± 4.3	19.7 ± 4.2	21.3 ± 4.3	-1.3	0.22	0.35
N2-N3 (%)	75.6 ± 4.7	75.9 ± 4.0	75.4 ± 5.3	0.4	0.70	0.73
N1-N2-N3 (%)	79.5 ± 4.3	80.3 ± 4.2	78.7 ± 4.3	1.3	0.22	0.35
N1 (min)	18.0 ± 12.1	20.1 ± 12.1	16.1 ± 12.0	1.1	0.27	0.36
N2 (min)	243.0 ± 34.3	241.1 ± 34.5	244.7 ± 34.8	-0.3	0.73	0.73
N3 (min)	116.7 ± 28.4	111.0 ± 27.5	122.0 ± 28.7	-1.3	0.19	0.35
REM (min)	98.0 ± 23.6	91.9 ± 24.1	103.6 ± 22.2	-1.7	0.10	0.35
N2-N3 (min)	359.7 ± 34.9	352.1 ± 33.1	366.7 ± 35.7	-1.4	0.16	0.35
N1-N2-N3 (min)	377.7 ± 30.8	372.2 ± 31.4	382.8 ± 30.0	-1.2	0.25	0.36
WASO (min)	22.6 ± 24.6	27.6 ± 27.2	18.0 ± 21.6	1.3	0.19	0.35
total sleep (min)	475.7 ± 36.9	464.1 ± 39.7	486.4 ± 31.4	-2.1	<b>0.04</b>	0.21
sleep latency (min)	14.9 ± 16.7	20.6 ± 21.5	9.6 ± 8.0	2.4	<b>0.02</b>	0.18
sleep efficiency (%)	89.4 ± 5.6	87.2 ± 6.2	91.5 ± 4.1	-2.8	<b>0.008</b>	0.13

**Table 2.** Sleep architecture parameters (mean ± SD). WASO: wake after sleep onset; P<sub>adj</sub>: adjusted P value using FDR. Entries in bold indicate significance (P < 0.05).

**Encoding Valence.** During encoding, subjects rated the emotional valence of each picture on a scale from -1 (negative) to +1 (positive) immediately following its lateralized presentation. An ANOVA with between-subject factor GROUP (sleep/wake) and within-subject factor VALENCE (negative/neutral) on subjects' trial-averaged ratings indicated that negative items were rated significantly more negatively than neutral ones (main effect of VALENCE), with no effect of GROUP or GROUP\*VALENCE interaction (Table 1-encoding). Post-hoc tests indicated that the valence effect was present in both groups (sleep:  $t(45) = 30.4$ ,  $P < 10^{-30}$ ; wake:  $t(24) = 16.0$ ,  $P < 10^{-13}$ ), whereas subjective ratings did not differ systematically between the sleep and wake groups for either neutral ( $t(69) = 1.7$ ,  $P = 0.09$ ) or negative ( $t(69) = 0.8$ ,  $P = 0.41$ ) items, indicating no time of day effect on perceived valence. Moreover, single-subject independent t tests comparing negative and neutral items indicated that each of the 71 individuals rated negative items significantly more negatively than neutral items (all  $P_{adj} < 10^{-14}$ , where  $P_{adj}$  indicates the adjusted P value using the False Discovery Rate [FDR]<sup>44</sup>). We further added SIDE (negative-left/negative-right) as a between-subject factor, but found no significant main or interaction effects involving this factor (all  $F(1,67) < 2.5$ , all  $P > 0.12$ ), indicating that the VF where each emotional category was presented did not impact ratings. Thus, these results establish that subjects can accurately gauge the emotional valence of stimuli presented away from fixation, regardless of presentation side, and that these effects are similar in the morning and evening.

**Sleep Architecture.** Overall sleep architecture for the 46 overnight participants was in line with typical values for young, healthy subjects (Table 2). We compared sleep parameters as a function of SIDE to examine if the VF where negative vs. neutral items had been presented affected subsequent sleep. Time and percentages spent in each sleep stage did not differ between groups. We did observe differences significant at an uncorrected threshold in parameters related to sleep quality (total sleep time, sleep efficiency, sleep latency), which were driven by two outliers in the negative-left group with unusually long sleep latencies (76 and 81.5 min; all others < 36 min), one of whom also had the lowest observed sleep efficiency (72.6%). However, no comparison survived correction for multiple testing. Thus, macroscopic sleep structure did not vary depending on the hemisphere initially processing negative or neutral information prior to sleep.

**Contextual Memory.** We next analyzed subjects' memory for the VF in which each item had been presented during encoding (i.e., contextual memory). Continuous left/right ratings were dichotomized and relabeled "correct" and "incorrect" according to original presentation side. In keeping with previous approaches<sup>13,45</sup>, we focus below on contextual memory for correctly recognized items (i.e., hits).



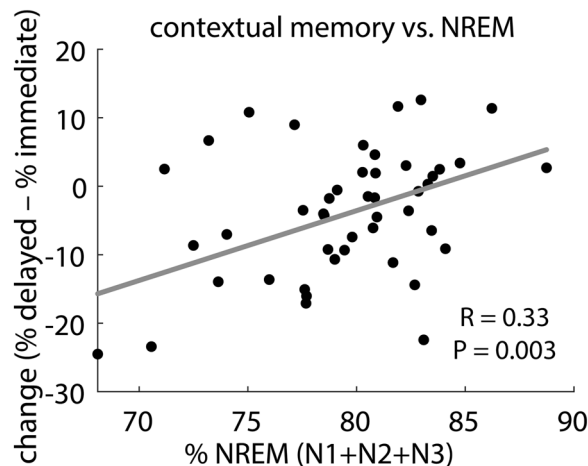
**Figure 2.** Change in memory across 12 h. **(A)** Contextual memory for hits was selectively preserved for negative items in the sleep group. **(B)** Recognition memory (hit rate) dropped similarly for each condition. Error bars reflect standard error of the mean with between-subject variability removed.

**Immediate Test.** At immediate testing, subjects correctly indicated the original presentation side significantly above chance in each condition (negative-sleep:  $73.9 \pm 14.0\%$ ; neutral-sleep:  $74.7 \pm 16.9\%$ ; negative-wake:  $75.6 \pm 16.5\%$ ; neutral-wake:  $78.3 \pm 15.5\%$ ; *t* tests vs. 50%: all  $P < 10^{-7}$ ). An ANOVA with between-subject factor GROUP and within-subject factor VALENCE did not reveal significant baseline differences (all  $F(1,69) < 0.8$ , all  $P > 0.37$ ; Table S3). Additional analyses including factor SIDE are presented in Supplementary Information.

**Change over 12 h.** Next, we evaluated the change in contextual memory across the 12 h interval (delayed – immediate). As predicted, we observed a selective preservation of contextual memory across 12 h for negative items in the sleep group (Fig. 2A). Whereas performance decreased across wake for negative items ( $-11.1 \pm 13.8\%$ ; one sample *t* test vs. zero:  $t(24) = 4.0$ ,  $P = 0.0005$ ) and neutral items ( $-5.5 \pm 15.2\%$ ;  $t(24) = 1.8$ ,  $P = 0.08$ ), and for neutral items across sleep ( $-7.8 \pm 14.2\%$ ;  $t(45) = 3.7$ ,  $P = 0.0005$ ), memory for negative items was unchanged across sleep ( $0.2 \pm 14.3\%$ ;  $t(45) = 0.1$ ,  $P = 0.94$ ). Across conditions, there was a significant GROUP\*VALENCE interaction ( $F(1,69) = 5.7$ ,  $P = 0.02$ ), a main effect of GROUP ( $F(1,69) = 4.2$ ,  $P = 0.04$ ), but no effect of VALENCE ( $F(1,69) = 0.2$ ,  $P = 0.67$ ). Post hoc tests indicated that the change in contextual memory for negative items in the sleep group was significantly different from that for neutral items in the same group (paired  $t(45) = 2.4$ ,  $P = 0.02$ ), and from that for negative items in the wake group (independent  $t(69) = 3.2$ ,  $P = 0.002$ ). The comparison with the neutral-wake condition was in the same direction as for the other two comparisons, but did not reach significance ( $t(69) = 1.6$ ,  $P = 0.12$ ). Within the wake group, forgetting of negative and neutral items did not differ significantly ( $t(24) = -1.5$ ,  $P = 0.26$ ). Further analyses indicated that these effects did not depend importantly on whether negative items had been presented to the left or right during encoding (Supplementary Information).

**Contextual Memory Control Analyses.** An exit questionnaire indicated that 73% of subjects (52/71) became explicitly aware of the mapping between VF and emotional category at some point during the protocol. These proportions were very similar for the sleep (74%, or 34/46) and wake (72%, or 18/25) groups ( $\chi^2 = 0.03$ ,  $P = 0.86$ ). A further breakdown of the participants who became aware of the VF-valence contingency indicated that 63% reached this insight prior to the 12 h interval (sleep: 62%, wake: 67%; encoding: 25% [sleep: 26%; wake: 22%], immediate test: 38% [sleep: 35%; wake: 44%]), while only 12% first realized this relation at delayed test (sleep: 15%; wake: 6%). The remaining 25% (sleep: 24%; wake: 28%) could not specify first rule awareness. Subject proportions were similarly distributed over these categories for the sleep and wake groups ( $\chi^2 = 1.28$ ,  $P = 0.73$ ). Thus, these findings indicate that when subjects attained explicit awareness of the VF-valence relation, they typically extracted this rule early in the protocol, with no indication that time of day and/or sleep vs. wakefulness affected the time course of reaching this insight.

Regardless of explicit insight, it is possible that, during retrieval, participants determined each item's original presentation side solely based on its perceived valence. To address this issue, we examined baseline performance at immediate test in two sets of analyses. Because the sleep and wake groups did not differ at immediate test with respect to either contextual memory or valence, groups were combined.



**Figure 3.** Overnight change in contextual memory for hits, pooled across emotional categories, was positively related to NREM sleep percentage. P value (uncorrected) and regression line from robust fit, R value from Spearman correlation.

First, we performed within-subject across-trial correlations, where we correlated items' valence ratings with their corresponding continuous "correct side/incorrect side" confidence ratings. This was done separately for negative and neutral, and old and new, items. We then extracted the slopes of these regression lines to examine the link between item valence and judgment of encoding side at the group level. Slopes differed significantly from zero for all four item categories (all  $t(70) > 5.4$ , all  $P < 10^{-6}$ ), indicating that more extreme valence ratings were associated with higher-confidence scores towards the correct side. We reasoned that if subjects exclusively relied on item valence to determine item placement, these relations should be similar for old and new items. However, slopes differed significantly between old and new items for both valence categories (negative-old:  $-0.42 \pm 0.32$ ; negative-new:  $-0.30 \pm 0.37$ ;  $t(70) = 3.8$ ,  $P = 0.0003$ ; neutral-old:  $0.33 \pm 0.31$ ; neutral-new:  $0.23 \pm 0.35$ ;  $t(70) = 2.6$ ,  $P = 0.01$ ), indicating that side judgments for old items went beyond the "baseline" association with valence seen for new items.

Second, if subjects merely relied on item valence, contextual memory for hits should not differ from performance for items with other recognition statuses (i.e., correct rejections, misses, and false alarms), as valence information is equally available in each of these cases. However, when we compared baseline VF ratings for hits to the other three categories, separately for negative and neutral items, we found significantly higher performance for hits for 5/6 comparisons (paired t tests, all  $P_{\text{adj}} < 0.00004$ ) except for hits vs. false alarms for negative items ( $P_{\text{adj}} = 0.70$ ). To follow up on the latter finding, we examined the change in performance on the contextual task for false alarms across 12 h. Unlike the differential retention of encoding side for previously presented negative and neutral items across sleep, no such valence-related difference was seen for proportions of false alarms assigned to the "correct" side (negative:  $-1.5 \pm 22.9\%$ ; neutral:  $-7.5 \pm 32.9\%$ ;  $t(43) = 1.0$ ,  $P = 0.34$ ), suggesting that the emotional advantage was tied to previously encoded items.

Combined, these findings suggest that while valence-based strategies contributed to judgments about encoding side, contextual memory for hits relied on additional memory-related processes. Moreover, the empirical finding that sleep, but not wake, selectively stabilizes contextual memories of negative, but not neutral, items cannot be explained by a pure valence-based account.

**Contextual Memory and Sleep Parameters.** Next, we examined whether overnight changes in contextual memory for hits were related to sleep parameters (all variables in Table 2, except WASO, sleep efficiency, and sleep latency). Separate analyses for negative, neutral, and pooled emotional categories, indicated a strong positive correlation between percentage of time spent in NREM (N1 + N2 + N3) sleep and contextual memory change across emotional categories (Fig. 3; Spearman  $R = 0.33$ , Robust regression  $P = 0.003$ ). A corresponding negative correlation was found for REM percentage. These correlations remained significant after correction for multiple comparisons ( $P_{\text{adj}} = 0.05$ ). We also observed a negative correlation with minutes spent in REM ( $R = -0.33$ ,  $P = 0.01$ ), although this relation did not survive correction for multiple comparisons ( $P_{\text{adj}} = 0.16$ ). Correlations with individual emotional categories were not significant for any sleep parameter (all  $P_{\text{adj}} > 0.40$ ), nor were correlations between proportion of time spent in NREM sleep and contextual memory change different for the two emotional categories (z test for correlated correlations:  $Z = 0.34$ ,  $P = 0.63$ ).

In sum, larger proportions of NREM (and smaller proportions of REM) sleep are related to better retention of contextual details regardless of emotional category, suggesting that NREM sleep promotes consolidation of episodic memories in a valence-independent fashion.

**Item Recognition.** Next, we assessed old/new recognition memory by dichotomizing subjects' continuous old/new response to each stimulus as "old" or "new", and then calculating standard metrics of hit rate (HR), and discriminability ( $d'$ ). Absolute performance at immediate and delayed test for these metrics, as well as for the false alarm rate (FAR), is reported in Table S1.

	SLEEP		WAKE		Group		Valence		Group*valence	
	Negative	Neutral	Negative	Neutral	F	P	F	P	F	P
HR	-18.3 ± 12.6	-16.1 ± 11.9	-20.8 ± 13.7	-20.4 ± 14.3	1.4	0.24	0.9	0.35	0.5	0.48
d'	-0.3 ± 0.5	-0.4 ± 0.6	-0.5 ± 0.4	-0.5 ± 0.5	1.9	0.18	0.3	0.61	0.2	0.63

**Table 3.** Change in recognition memory across 12 h (mean ± SD). F test degrees of freedom: (1,69).

*Change over 12 h.* We evaluated changes in recognition memory across the 12 h interval (delayed – immediate) as a function of GROUP and VALENCE (Fig. 2B). HR decreased significantly for both negative and neutral stimuli in both groups (one-sample t tests vs. zero: all  $P < 10^{-6}$ ), indicating robust forgetting. However, neither GROUP nor VALENCE affected the rate of forgetting (statistics in Table 3–HR). Adding SIDE as a between-subject factor did not reveal additional significant main or interaction effects involving this factor (all  $F(1,67) < 2.5$ , all  $P > 0.12$ ), indicating that forgetting was similar when negative items had originally been presented in the left or right VF. Similar to findings for HR, while subjects' ability to discriminate old from new items decreased significantly in each condition (one-sample t tests vs. zero: all  $P < 10^{-4}$ ), forgetting was not affected by GROUP or VALENCE (all  $F(1,69) < 1.9$ , all  $P > 0.17$ ; Table 3–d'). Again, adding SIDE did not yield significant effects (all  $F(1,67) < 2.7$ , all  $P > 0.10$ ). Changes in FAR are reported in Table S2.

For the sleep group we examined whether sleep parameters correlated with overnight memory changes. For each recognition metric (HR, FAR, d'), we performed separate analyses for negative, neutral, and pooled emotional categories. We found no significant correlations (FDR correction across 39 comparisons for each recognition metric: all  $P_{\text{adj}} > 0.58$ ).

In sum, while overall recognition memory deteriorated markedly across a 12 h interval, wake vs. sleep did not affect the overall rate of forgetting, nor did it impact retention of negative and neutral items differently, contrasting markedly with the effects reported for contextual memory.

**Emotional Valence.** Next, we examined whether and how sleep influences subjective ratings of item valence from immediate to delayed testing. We assessed valence separately for old and new pictures to determine whether changes in emotional reactivity are specific to memorized items.

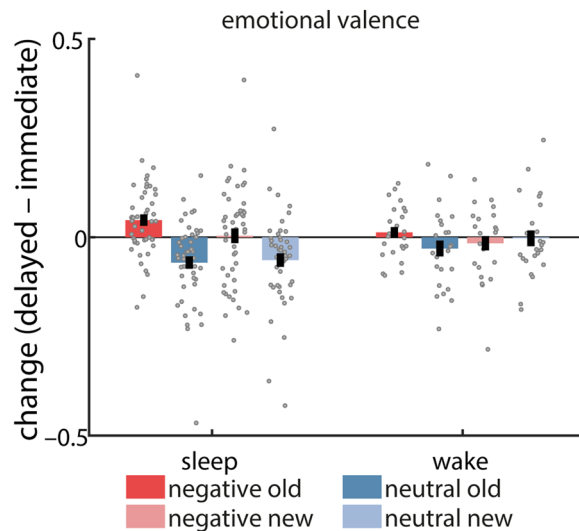
*Immediate Test.* We first sought to ensure that baseline valence ratings at immediate test differed reliably between negative and neutral items, and did so similarly for the sleep and wake groups. Separate ANOVAs for old and new items indicated highly significant main effects of VALENCE, with no effect of GROUP and no GROUP\*VALENCE interaction, indicating no effect of time of day on emotional ratings (Table 1, *Test 1–old* and *Test 1–new*). Post hoc paired t tests indicated that negative and neutral ratings differed for each combination of sleep/wake and old/new (all  $P < 10^{-13}$ ), and for every individual (all  $P_{\text{adj}} < 10^{-5}$ ). VF of presentation at encoding (SIDE) did not impact affective ratings at immediate test, for either old or new items (all  $F(1,67) < 1.7$ , all  $P > 0.20$ ).

Adding the factor OLD/NEW to the VALENCE\*GROUP ANOVA again revealed a significant effect of VALENCE ( $F(1,69) = 1126.6$ ,  $P < 10^{-43}$ ), as well as an OLD/NEW\*VALENCE interaction ( $F(1,69) = 28.0$ ,  $P < 10^{-3}$ ), but no effects involving GROUP (all other  $F(1,69) < 1.8$ ,  $P > 0.18$ ). Post hoc tests revealed that the OLD/NEW\*VALENCE interaction stemmed from more negative ratings to negative items for old compared to new items (sleep:  $t(45) = 3.3$ ,  $P = 0.002$ ; wake:  $t(24) = 2.3$ ,  $P = 0.03$ ; Table 1), and from old neutral items being rated more positively than new neutral items in both groups (sleep:  $t(45) = 3.2$ ,  $P = 0.002$ ; wake:  $t(24) = 2.3$ ,  $P = 0.03$ ; Table 1). These findings indicate that items previously encountered during encoding elicit a stronger emotional response at test than novel items, with this emotional potentiation effect similar for negative and neutral items and independent of time of day.

*Change over 12 h.* We next analyzed changes in valence ratings across the 12 h interval (delayed – immediate; Fig. 4). An ANOVA with factors OLD/NEW, GROUP, and VALENCE yielded significant OLD/NEW\*VALENCE ( $F(1,69) = 11.9$ ,  $P = 0.001$ ), GROUP\*VALENCE ( $F(1,69) = 4.5$ ,  $P = 0.04$ ), and VALENCE ( $F(1,69) = 8.8$ ,  $P = 0.004$ ) effects (all other  $F(1,69) < 1.3$ ,  $P > 0.26$ ). Post hoc tests indicated that whereas a period of wake did not result in significant changes for any condition (one sample t tests vs. zero, all  $P > 0.16$ ), an interval of sleep did.

Specifically, sleep led to less negative ratings of negative items for old ( $t(45) = 3.0$ ,  $P = 0.004$ ), but not new ( $t(45) = 0.2$ ,  $P = 0.84$ ) items, suggesting that sleep reduces the emotionality of previously memorized negative items. Moreover, sleep led to less positive ratings of neutral items (which, overall, had positive ratings at immediate test), but did so for both old ( $t(45) = 4.1$ ,  $P = 0.0002$ ) and new ( $t(45) = 3.4$ ,  $P = 0.002$ ) items, suggesting a depotentiation of emotional reactivity to mildly positive material regardless of item novelty. Indeed, while for negative items we observed a significant old/new difference in emotional change across sleep ( $t(45) = 3.3$ ,  $P = 0.002$ ), this effect was absent for neutral items across sleep ( $t(45) = 0.5$ ,  $P = 0.64$ ). No significant old/new differences were found for wake (negative:  $t(24) = 1.6$ ,  $P = 0.11$ ; neutral:  $t(24) = 1.8$ ,  $P = 0.09$ ), although, numerically, values followed the pattern seen in sleep. In line with this observation, emotional attenuation was stronger in sleep vs. wake only for the neutral new items ( $t(69) = 2.0$ ,  $P = 0.05$ ). In contrast, sleep and wake did not differ significantly for any of the other conditions (all  $t(69) < 1.5$ ,  $P > 0.16$ ), indicating that sleep's effect on emotional ratings was primarily the reduction of positive affect to previously unseen neutral items.

Correlational analyses with sleep parameters did not reveal associations with changes in affective ratings for any of the combinations of old/new and negative/neutral items (all  $P_{\text{adj}} > 0.94$ ). We also examined whether



**Figure 4.** Change (delayed – immediate) in affective ratings to centrally presented items. Error bars reflect standard error of the mean with between-subject variability removed.

valence changes were associated with the amount of forgetting of contextual memories. However, separate analyses for sleep/wake and negative/neutral revealed no significant associations, whether considering valence changes across old and new items, old items only, or correctly recognized old items (i.e., hits) only (all  $P_{\text{adj}} > 0.55$ ).

## Discussion

While the notion that sleep is an important factor in emotional memory processing is widely accepted, little is understood about precisely how these factors interact<sup>39</sup>. We found that sleep selectively stabilizes contextual aspects of negative memories, in a way not seen for neutral material, for either negative or neutral material across wake, or for recognition memory in any condition. Moreover, the proportion of time spent in NREM sleep predicted contextual memory retention in a valence-independent manner. In addition, an interval of sleep decreased affective responses to previously seen negative stimuli and to both old and new mildly positive stimuli, although effects for memorized items did not differ reliably between sleep and wake.

Our findings suggest a sleep-mediated benefit for emotional memory stabilization, but only for the spatial, contextual aspect of negative memories. The general observation of a sleep benefit for contextual (Fig. 2A) but not recognition memory (Fig. 2B; Table 3) fits with evidence that paradigms probing associative memory are generally more sensitive to sleep effects compared to recognition paradigms<sup>11,12</sup>. Studies investigating recognition have found stronger sleep-dependent effects for recollection-based recognition, which is thought to require stronger hippocampal involvement, than familiarity-based recognition<sup>46,47</sup>. Combined, these findings suggest that sleep's effect on declarative memory operates primarily on memories that rely on the binding machinery of the hippocampus<sup>13</sup>.

Considering the emotional dimension, the selective sleep benefit for negative contextual memory is consistent with sleep's negative vs. neutral advantages for associative memories of temporal order<sup>33</sup>, and spatial location<sup>32</sup>, although these split-night studies found this effect to depend on opposing sleep stages (NREM-rich early vs. REM-rich late sleep, respectively). More broadly, these observations fit with the notion that the sleeping brain preferentially processes information deemed of future relevance<sup>7</sup>. Specifically, there could be evolutionary advantages to remembering the spatiotemporal context in which a negative stimulus was originally encountered, allowing the organism to avoid such experiences in the future. The absence of a sleep benefit for neutral contextual information in our study could also be related to this prioritization process, in that preferential processing of emotional experiences could leave limited resources for the consolidation of neutral memories, although consolidation of neutral memories could still occur when only neutral items are learned (e.g.<sup>11,12,48,49</sup>).

Regarding the contributions of specific sleep stages to contextual memory consolidation, we found that greater proportions of NREM (and correspondingly, lower proportions of REM) sleep are associated with improved retention (Fig. 3). While these observations contrast with some evidence of REM sleep being primarily related to emotional memory processing<sup>16,17</sup>, they add to an accumulating body of findings emphasizing the importance of NREM and its electrophysiological markers in the consolidation of both neutral and emotional memories<sup>20,26,28,33</sup>. Indeed, the NREM-memory association was only observed when we pooled across emotional categories, not when we considered negative and neutral items separately. The absence of valence-specific relations may be due to the fact that sleep architectural parameters, as rather crude measures of brain state, do not have the required sensitivity to detect category-specific consolidation processes, which are likely reflected by subtler patterns of electrophysiological activity<sup>50–52</sup>. Indeed, our experiment was designed to allow future examinations of such category-specific activity patterns in the sleeping brain (see below).

Overall, studies investigating the link between sleep and emotional associative memory have been far from consistent in their findings. For example, a study employing an object-face paired associates approach found reduced forgetting of neutral-neutral pairs across a nap vs. wake, but no difference for negative-neutral pairs<sup>34</sup>,

whereas a picture-sound paradigm reported similar sleep benefits for emotional and neutral pairs<sup>18</sup>. Given that these studies focused on pairs of items, one possibility is that sleep's effects depend on whether the type of association tested relates to links with background contextual components (in which case sleep may favor the stabilization of links with negative memories), or links between individual items (in which case the emotional advantage may be suppressed or even reversed). Complicating this view, the aforementioned split-night study that found superior picture-location memory for negative vs. neutral items across late REM-rich sleep observed the opposite pattern (i.e., a benefit for neutral items) across early NREM-rich sleep<sup>32</sup>. Another split-night study observed a similar early-night benefit for neutral picture-color associations, but did not observe this finding for picture-location associations (for which no differential negative/neutral or early/late sleep effects were found)<sup>25</sup>. However, given that these studies focused on early/late night comparisons without direct contrasts to a wake control condition, it is unclear how specific the reported effects are to sleep. Lastly, a series of studies has reported that sleep enhances memory for central negative objects at the expense of their neutral background, with no such effect for neutral object-background pairings<sup>14,53,54</sup>, which has been taken to imply that sleep weakens rather than preserves negative item-context links. However, as this paradigm tests recognition memory for individual objects and backgrounds, as opposed to directly probing the associative link between them, it is presently unclear how these and our findings relate.

In the current study, we found no sleep/wake effect on recognition memory (Fig. 2B; Table 3), or any relationship between recognition memory and sleep architecture. As noted above, this is consistent with several studies reporting no effect of sleep on recognition memory<sup>11,12</sup> or familiarity-based recognition<sup>46,47</sup>. However, this contrasts with reports showing enhanced overall<sup>19,29</sup> or negative<sup>15,17,23–25</sup> item recognition after sleep, in some cases relating these effects to REM sleep<sup>17,23–25</sup>. Given the many differences in experimental design between these studies and ours, it remains to be seen under what conditions recognition metrics are sensitive to sleep-mediated consolidation processes.

A prominent theoretical account posits that sleep, and REM sleep in particular, serves to preserve the content of emotional memories while simultaneously removing their affective layer<sup>55</sup>. While the “emotional blunting” aspect of this model has received physiological support<sup>35,36</sup>, subjective ratings of arousal and/or valence have shown inconsistent sleep-related modulation of emotional tone<sup>18,19,23,24,32,37,38</sup>, as reviewed in the introduction. Likewise, our findings provide no clear support for a role of sleep in reducing the emotional tone of memories (Fig. 4). On the one hand, sleep reduced the perceived emotionality of both negative and mildly positive stimuli. For negative items, this effect was specific to previously seen items, suggesting a sleep-mediated emotional depotentiation of existing negative memories. In contrast, the decrease in emotional intensity for mildly positive pictures was similar for old and new items, suggesting that this overnight reduction in emotional reactivity was not specific to previously encoded materials. On the other hand, emotional attenuation was significantly stronger across sleep than wake only for new neutral items, but critically, not for old items of either negative or neutral valence. As such, the central tenet of the aforementioned theory – a sleep-specific reduction of emotional reactivity to memorized material – is not supported by our data. Moreover, changes in emotional ratings were unrelated to changes in contextual memory, or to aspects of sleep macrostructure, providing no direct support that sleep-mediated reductions in emotional reactivity accompany memory consolidation or occur in a particular sleep stage.

While the current paper reports behavioral evidence that sleep mediates the consolidation of contextual, negative memories, our study was designed to link these behavioral findings to neural evidence of category-specific memory reactivation. Specifically, we implemented the valence-VF manipulation at encoding to differentially “tag” the two cerebral hemispheres with emotional material, leading to the prediction that the hemisphere encoding negative items is more strongly activated during sleep. To achieve this goal of differential tagging, negative and neutral items were presented exclusively in one VF, as opposed to randomizing presentation side for both emotional categories (which would have been preferable from a purely behavioral standpoint). Consequent to this design choice, the link between VF and emotional valence was apparent to the majority of subjects. However, explicit awareness of this relation was similar for the sleep and wake groups, making it unlikely that our results are due to circadian or sleep/wake influences on rule extraction<sup>56,57</sup>. It is also worth noting that, due to several subjects becoming aware of the rule only after the 12 h interval, more subjects presumably had explicit insight after than before this interval. However, whereas a rule-based account would predict enhanced (or at least stabilized) performance across the interval, contextual performance instead dropped for three out of four conditions, most consistent with the forgetting of associative details. Additional control analyses at immediate test indicated that while item valence and side ratings were not independent, this relation differed for old and new items. Similarly, contextual performance was better for hits than for misses, correct rejections, and (neutral) false alarms, suggesting that the selective sleep benefit for negative contextual memories reflects a genuine memory effect.

That said, the present paradigm does not enable a pure separation of memory- and valence-based contributions to assessments of encoding side, which is a limitation of this study. In addition, we observed an unexpected influence of the factor SIDE (negative-left/negative-right) on contextual memory performance (Supplementary Information). (Note that both the direction of this effect and the observation that valence ratings are independent of VF do not support the notion of preferential right-hemisphere processing of negative information<sup>40–42</sup>). While we believe that a sleep-dependent stabilization of negative contextual memories is the most plausible explanation for our observations, these additional factors should be taken into consideration when interpreting our findings.

In conclusion, our data support the notion that sleep selectively stabilizes contextual aspects of negative episodic memories, possibly related to NREM sleep. These findings lay the foundation for more detailed investigations into the electrophysiological underpinnings of sleep-related reactivation processes for emotional and contextual information. Nonetheless, the determinants of emotional processing during sleep remain enigmatic, and continued effort will be needed to unravel the complex and important interrelations among sleep, memory, and emotion.



## Methods

**Participants.** Seventy-one young, healthy participants from the Boston area with normal or corrected-to-normal vision (24 male, 47 female; mean age  $\pm$  SD:  $21.3 \pm 3.1$ , range: 18–33) were included in this study. All told, 68/71 participants reported being right-handed (one reported being ambidextrous, handedness data for two subjects were lost), and none reported any history of neurological, psychiatric, or sleep disorders (except for one subject reporting a past anxiety disorder: exclusion did not affect pattern of results). Sleep logs indicated an average of  $\sim 8$  h of sleep across the three nights preceding the study ( $477 \pm 51$  min; range: 350–505). Participants were instructed to refrain from consuming recreational drugs or alcohol in the 24 h prior to the study, and to consume no more than one caffeinated beverage on the day of the study. They were instructed not to nap on the day of the first session (for the sleep group) or in between sessions (for the wake group). Post hoc evaluation of study questionnaires indicated that three subjects from the sleep group had not complied with this instruction. However, given that overnight sleep parameters appeared normal (total sleep time  $> 425$  min; sleep efficiency  $> 78\%$ ), they were included. Two additional participants were excluded because their valence ratings at encoding and/or testing failed to differentiate between negative and neutral items, suggesting poor compliance with instructions. Subjects were compensated monetarily for their participation. All subjects provided written informed consent, the experiment was carried out in accordance with relevant guidelines and regulations, and this study was approved by the institutional review board of Beth Israel Deaconess Medical Center.

**Protocol Overview.** The protocol (Fig. 1A) consisted of two sessions separated by  $\sim 12$  h containing either sleep or wakefulness (between-subject factor GROUP). The primary components of Session 1 were encoding followed by an immediate test (Test 1), whereas the primary component of Session 2 was a delayed test (Test 2; see below for detailed descriptions of each session). Participants in the sleep group ( $N = 46$ ) completed Session 1 in the evening, spent the night in the sleep lab, and completed Session 2 the next morning, whereas wake subjects ( $N = 25$ ) came to the lab for separate morning and evening visits. Group assignment was based on participants' availability. An additional between-subject factor SIDE defined the association between VF and emotional category (negative-left/neutral-right:  $N = 36$  [sleep: 22; wake: 14]; negative-right/neutral-left:  $N = 35$  [sleep: 24; wake: 11]). Males and females were equally distributed across GROUP and SIDE ( $\chi^2 = 0.79$ ,  $P = 0.94$ ).

Upon arriving at the lab, subjects in the sleep group were wired for high-density EEG recordings. Wake participants were wired with a simple electrooculography (EOG) montage in order to have similar setup procedures, experiences, and compliance with instructions in both groups. Session 1 typically started between 9–10 PM (sleep) or 9:15–10:15 AM (wake), and lasted  $98 \pm 13$  min (range: 80–158). Whereas wake subjects left the lab between sessions, sleep participants were given a 9 h sleep opportunity (lights out between 10:30 PM–12:15 AM). Immediately following waking, sleep participants were asked to provide details regarding dream content and sleep quality, and then had breakfast. Session 2 commenced between 8:30–9:30 AM (sleep) or 9–10 PM (wake), lasting  $89 \pm 16$  min (range: 60–149). The  $\sim 12$  h interval between the start of Sessions 1 and 2 differed slightly, but consistently, between groups (sleep:  $704 \pm 20$  min; wake:  $722 \pm 18$  min;  $t(69) = 3.9$ ,  $P = 0.0002$ ).

Of note, approximately two-thirds of subjects (46 vs. 25) were assigned to the sleep condition. Given that the primary objective of our EEG recordings was to allow future examinations of hemisphere-specific memory reactivation during sleep (see Discussion), the sleep group required a sufficient number of subjects in each SIDE condition to enable negative-left vs. negative-right comparisons of electrophysiological activity. In contrast, while the wake group was critical for testing the sleep-dependence of behavioral effects, EEG collection and comparisons based on SIDE were not deemed central for the wake group, allowing for smaller group size.

**Stimuli.** A total of 384 pictures (192 negative, 192 neutral) were selected from the Nencki Affective Picture System<sup>58</sup>, a collection of affective pictures providing extensive normative valence and arousal ratings, semantic category labels, and detailed descriptions of physical image properties. Normative valence and arousal scores, based on  $N = 204$  raters, ranged from 1 (very negative; relaxed) to 9 (very positive; aroused), with 5 being neutral. Picture selection was limited to items of  $1600 \times 1200$  pixels (landscape orientation) to ensure identical dimensions during presentation. Negative items were chosen that had normative ratings between 2–4.5 (valence) and 6–9 (arousal), while neutral items had ratings of 5–7 for valence and 2–5 for arousal, resulting in non-overlapping distributions for the two emotional categories. The number of negative and neutral items within each semantic category was matched (faces: 60 negative and 60 neutral; animals: 42 of each; objects: 42; people: 36, landscapes: 12). In addition, we selected negative and neutral items within each semantic category such that physical image properties were minimally different between emotional categories. As a result, selected negative and neutral items differed significantly in terms of valence ( $3.3 \pm 0.6$  vs.  $6.2 \pm 0.5$ ;  $t(382) = 50.0$ ,  $P < 10^{-168}$ ) and arousal ( $6.3 \pm 0.5$  vs.  $4.4 \pm 0.4$ ;  $t(382) = -44.6$ ,  $P < 10^{-152}$ ), but not in terms of luminance, contrast, complexity (as approximated by JPEG file size), or dimensions of CIELAB color space (all  $P > 0.24$ ).

**Task.** Task presentation was controlled using Psychtoolbox for Matlab (the Mathworks, Natick, MA). For sleep subjects, stimulus presentation markers were sent to the amplifier to allow synchronization of EEG activity to events of interest. Subjects were seated in front of a 27-inch computer display with their eyes approximately 60 cm away from the center of the screen. Subjects were instructed to keep their head in a steady position and to minimize eye blinking throughout the experiment.

For each subject, affective stimuli were randomly partitioned at runtime according to the following criteria. Two-thirds of the stimuli (256: 128 negative, 128 neutral) were reserved for encoding and to serve as “old” stimuli during retrieval tests, with separate subsets used for Test 1 and Test 2 (64 negative and 64 neutral per test). The remaining third (128: 64 negative, 64 neutral) served as foils during retrieval (“new”), again using separate subsets for Test 1 and Test 2 (32 negative and 32 neutral per test). Moreover, the relative proportion of items within each

semantic category was kept identical across encoding and retrieval tests, between negative and neutral items, and between old and new items.

**Session 1 – Pre-encoding.** Before EEG or EOG wiring, subjects indicated their level of vigilance on the Stanford Sleepiness Scale<sup>43</sup> (1: wide awake; 7: no longer fighting sleep) and the question “How would you describe your ability to concentrate right now?” on a visual analog scale (range: 0–100). Following wiring, Session 1 commenced with an extensive set of instructions regarding the encoding phase. Subjects practiced fixating, using the response slider, and reporting saccades (each described, as relevant, below). Subjects were not told their memory would be tested, nor were they explicitly informed of the relation between VF and stimulus emotionality.

**Session 1 – Encoding.** Encoding trials were presented in 8 blocks of 32 trials, with pauses between blocks. Subjects were encouraged to use pauses to move, stretch, and blink, in order to enable optimal compliance with task instructions during the subsequent block. Trial order was randomized with the constraint that the same emotional category could not be presented more than three times in a row. Each encoding trial began with a white fixation dot presented for an interval jittered uniformly between 1,000 and 2,000 ms (Fig. 1B). Next, with the fixation point still present, an affective stimulus was presented for 1,000 ms in either the left or the right VF (depending on emotional category and the subject’s VF-emotional category assignment). The lateralized stimulus was presented at a size of 21.8 × 16.4 cm, with the figure center displaced 13.9 cm laterally (13° of visual angle at 60 cm from the screen) and 4.2 cm downward (4°), resulting in a stimulus covering 20° × 15° of visual space. Another fixation display followed for 1,000 ms. Next, a screen appeared asking subjects to rate the perceived emotionality of the preceding image (“How EMOTIONAL would you say this picture is?”). Subjects indicated their response by placing a mouse-operated cursor on a vertically oriented continuous slider (visual analog scale) containing five equally spaced tick marks with associated labels (from bottom to top: very negative, somewhat negative, neutral/unsure, somewhat positive, very positive). Upon clicking, the slider position was recorded as a valence rating in the interval from –1 to +1, a blank display was presented for 500 ms, and the next trial commenced. No time restrictions were imposed for responses.

Subjects were instructed to keep their gaze on the fixation dot when present and to minimize eye blinks and other movements; they were free to move their eyes while the response screen was displayed. In addition, subjects were asked to click the mouse before onset of the response screen in the event that they made a saccade during the immediately preceding stimulus presentation. In that case, a confirmation (“You indicated you looked at the picture...”) was displayed for 2,000 ms, and the interrupted trial was repeated. On average, subjects reported making 12.6 ± 14.9 saccades (range: 0–65) during the 256 trials. Although it is unclear how reliably self-reported saccades reflect actual saccades, repeated instructions at the beginning of each block to report saccades, as well as saccades being reported by 65/71 subjects, likely helped maintain central fixation. To further dissuade subjects from making undesired movements, a warning (“Do not press any keyboard buttons unless you’re asked to!”) was presented for 2,000 ms upon detection of any key press, and the interrupted trial was repeated. Trials were repeated to maximize the number of trials available for future EEG analyses. However, behavioral findings were unaltered when excluding repeated trials.

Following the encoding phase, subjects used the slider (range from –1 to +1) to rate: (1) the visibility of laterally presented stimuli (ranging between “it was a complete blur” – “I saw everything perfectly”; sleep: 0.15 ± 0.34; wake: 0.22 ± 0.32), (2) difficulty fixating (“extremely easy” – “extremely difficult”; sleep: 0.05 ± 0.35; wake: 0.10 ± 0.41), (3) accuracy fixating (“I looked away all the time” – “I never looked away”; sleep: 0.61 ± 0.30; wake: 0.51 ± 0.34), and (4) accuracy reporting saccades (“I was too slow or forgot to click” – “I perfectly reported every time I looked”; sleep: 0.49 ± 0.43; wake: 0.58 ± 0.41). None of these parameters differed between the sleep and wake groups (all  $t(69) < 1.3$ , all  $P > 0.22$ ).

**Session 1 – Immediate Retrieval (Test 1).** Subjects were informed that a surprise memory test would be administered, and received instructions. Retrieval trials were presented in 6 blocks of 32 trials, with pauses between blocks. Trial order was randomized with the constraint that the same combination of emotional category and old/new status could not be presented more than three times in a row. Retrieval trials (Fig. 1C) consisted of a sequence of pre-stimulus (1,000–2,000 ms), stimulus (1,000 ms), and post-stimulus (1,000 ms) periods. Critically, stimulus presentation was now central (as opposed to lateralized), but still displaced downward with similar stimulus dimensions as during encoding. Following the post-stimulus fixation display, a first response screen asked subjects to indicate the preceding item’s OLD/NEW status (“How confident are you that this picture was shown during the LEARNING phase?”) using a response slider with equally spaced labels (certainly old, probably old, unsure, probably new, certainly new). A second response screen asked subjects for a continuous left/right judgment (“On what SIDE of the screen would you place this picture?”), with labels reading certainly left/probably left/unsure/probably right/certainly right. Subjects had been instructed to indicate where each picture “belongs”, and that their brain could still have processed the item’s left/right information even if they did not remember seeing it during encoding. Collecting left/right judgments regardless of old/new status and old/new response allowed us to compare these ratings as a function of recognition status. Finally, a third response screen asked for a valence rating as at encoding, after which a blank post-response display was shown for 500 ms, and the next trial began.

Following immediate retrieval, subjects were informed that a similar retrieval procedure would be administered during Session 2. This ensured that all subjects had similar test expectations, which are known to affect sleep-dependent memory consolidation<sup>59,60</sup>.

**Session 2.** Immediately upon waking, subjects in the sleep group provided details regarding dream experiences<sup>61</sup> and sleep quality. Then, for both groups, assessments of vigilance were administered, followed by the delayed test (Test 2) using trial timing identical to Test 1. Importantly, all items employed in Test 2 (both old and new)

were different from those in Test 1. In addition, a second delayed test (Test 3) was administered using the items previously used in Test 1 (Fig. 1A). However, contrary to the large drop in HR from Test 1 to Test 2 (Table 3), no significant decrease was observed from Test 1 to Test 3 in any condition (t-tests vs. zero: all  $P > 0.27$ ). This suggests a retrieval-related encoding effect such that at Test 3 subjects remembered the old/new judgment they had previously given at Test 1. Therefore, data involving Test 3 were not further analyzed. At the end of the protocol subjects completed an exit questionnaire probing their thoughts about the study's goal and their awareness of the VF-valence association.

**Resting states.** Before and after every encoding and retrieval phase we collected data from 5 min resting states (six in total; Fig. 1A). Subjects were instructed to close their eyes and rest quietly. After 5 min, a sound indicated they could open their eyes again. Resting states were collected to allow investigation of waking EEG signatures related to sleep-dependent memory consolidation, and were also included for wake subjects to minimize protocol differences between groups.

**Data Acquisition and Preprocessing.** EEG in the sleep group was collected using 62-channel EEG caps (EasyCap GmbH, Herrsching, Germany) with channel positions in accordance with the 10–20 system. Additionally, two single cup electrodes were placed on the mastoid processes, two lateral to each eye for EOG, two on the chin for electromyography (EMG), one on the forehead as an (online) reference, and one on the skin over the clavicle as a ground. For the wake group, only EOG, mastoids, reference, and ground electrodes were connected. An AURA-LTM64 amplifier and TWin software were used for data acquisition (Grass Technologies, Warwick, RI, USA). Impedances were kept below 25 k $\Omega$  and data were sampled at 400 Hz with hardware high-pass and low-pass filters at 0.1 Hz and 133 Hz, respectively. Impedances were checked repeatedly throughout the protocol.

Sleep staging was performed in TWin using 6 EEG channels (F3, F4, C3, C4, O1, O2) referenced to the contralateral mastoid (bandpass filtered 0.3–35 Hz), bipolar EOG (0.3–35 Hz) and bipolar EMG (10–100 Hz), on 30 s epochs according to standard criteria of the American Academy of Sleep Medicine<sup>62</sup>.

**Data Analysis and Statistics.** All analyses were performed using custom Matlab scripts. Valence ratings, ranging from  $-1$  to  $+1$ , were analyzed in continuous format. Continuous old/new confidence ratings were dichotomized as old ( $<0$ ) and new ( $>0$ ) to enable calculation of hits, false alarms, correct rejections, and misses, as well as  $d'$  (using Macmillan and Kaplan<sup>63</sup> adjustments as required). Similarly, continuous left/right ratings were converted into (continuous) incorrect side/correct side judgments and dichotomized as correct or incorrect. Percentage of items assigned to the correct side was calculated separately for each of the aforementioned recognition classes. Items with continuous recognition or side scores of exactly zero were excluded from dichotomization. We also performed memory analyses using continuous scores, but as the pattern of results was highly similar to that obtained using dichotomized scores, results based on continuous data are not reported.

Group comparisons were performed using conventional t tests (or Wilcoxon Rank Sum for non-normal distributions) and ANOVAs with between- and within-subject factors as required. Throughout the article, ANOVAs are typically first presented without factor SIDE, before reporting analyses including SIDE (or referring to these analyses in the Supplementary Information). Correlations were performed using robust regression procedures to minimize the influence of outliers. In addition, Pearson or Spearman correlations (as appropriate) were performed to obtain standard estimates of correlation strength (which are not provided by robust regression). Direct comparison of correlations was performed using the procedure outlined by Meng and colleagues<sup>64</sup>. Correction for multiple comparisons was done as appropriate using the False Discovery Rate procedure<sup>44</sup> with a threshold of 0.05.

## Data Availability

All data underlying the current report are available in Microsoft Excel and Matlab format (Supplementary Dataset 1).

## References

- Kensinger, E. A. Remembering the details: Effects of emotion. *Emot. Rev.* **1**, 99–113 (2009).
- Dolcos, F., LaBar, K. S. & Cabeza, R. Remembering one year later: Role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. *Proc. Natl. Acad. Sci.* **102**, 2626–2631 (2005).
- Bradley, M. M., Greenwald, M. K., Petry, M. C. & Lang, P. J. Remembering pictures: Pleasure and arousal in memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **18**, 379–390 (1992).
- Sharot, T. & Yonelinas, A. P. Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition* **106**, 538–547 (2008).
- Anderson, A. K., Yamaguchi, Y., Grabski, W. & Lacka, D. Emotional memories are not all created equal: Evidence for selective memory enhancement. *Learn. Mem.* **13**, 711–718 (2006).
- Rasch, B. & Born, J. About sleep's role in memory. *Physiol. Rev.* **93**, 681–766 (2013).
- Stickgold, R. & Walker, M. P. Sleep-dependent memory triage: evolving generalization through selective processing. *Nat. Neurosci.* **16**, 139–145 (2013).
- Frankland, P. W. & Bontempi, B. The organization of recent and remote memories. *Nat. Rev. Neurosci.* **6**, 119–130 (2005).
- Buzsáki, G., Buzsáki, G. & Buzsáki, G. Two-stage model of memory trace formation: A role for 'noisy' brain states. *Neuroscience* **31**, 551–570 (1989).
- Takashima, A. *et al.* Shift from Hippocampal to Neocortical Centered Retrieval Network with Consolidation. *J. Neurosci.* **29**, 10087–10093 (2009).
- Studte, S., Bridger, E. & Mecklinger, A. Nap sleep preserves associative but not item memory performance. *Neurobiol. Learn. Mem.* **120**, 84–93 (2015).
- van der Helm, E., Gujar, N., Nishida, M. & Walker, M. P. Sleep-Dependent Facilitation of Episodic Memory Details. *PLoS One* **6**, e27421 (2011).

13. Staresina, B. P., Fell, J., Do Lam, A. T. A., Axmacher, N. & Henson, R. N. Memory signals are temporally dissociated in and across human hippocampus and perirhinal cortex. *Nat. Neurosci.* **15**, 1167–1173 (2012).
14. Payne, J. D., Stickgold, R., Swanberg, K. & Kensinger, E. A. Sleep preferentially enhances memory for emotional components of scenes. *Psychol. Sci.* **19**, 781–788 (2008).
15. Hu, P., Stylos-Allan, M. & Walker, M. P. Sleep facilitates consolidation of emotional declarative memory. *Psychol. Sci.* **17**, 891–898 (2006).
16. Wagner, U., Gais, S. & Born, J. Emotional memory formation is enhanced across sleep intervals with high amounts of rapid eye movement sleep. *Learn. Mem.* **8**, 112–9 (2001).
17. Nishida, M., Pearsall, J., Buckner, R. L. & Walker, M. P. REM sleep, prefrontal theta, and the consolidation of human emotional memory. *Cereb. Cortex* **19**, 1158–1166 (2009).
18. Lehmann, M., Seifritz, E. & Rasch, B. Sleep benefits emotional and neutral associative memories equally. *Somnologie* **20**, 47–53 (2016).
19. Baran, B., Pace-Schott, E. F., Ericson, C. & Spencer, R. M. C. Processing of emotional reactivity and emotional memory over sleep. *J. Neurosci.* **32**, 1035–1042 (2012).
20. Cellini, N., Torre, J., Stegagno, L. & Sarlo, M. Sleep before and after learning promotes the consolidation of both neutral and emotional information regardless of REM presence. *Neurobiol. Learn. Mem.* **133**, 136–144 (2016).
21. Ackermann, S., Hartmann, F., Papassotiropoulos, A., de Quervain, D. J.-F. & Rasch, B. No Associations between Interindividual Differences in Sleep Parameters and Episodic Memory Consolidation. *Sleep* **38**, 951–9 (2015).
22. Lewis, P. A., Cairney, S., Manning, L. & Critchley, H. D. The impact of overnight consolidation upon memory for emotional and neutral encoding contexts. *Neuropsychologia* **49**, 2619–2629 (2011).
23. Groch, S., Wilhelm, I., Diekelmann, S. & Born, J. The role of REM sleep in the processing of emotional memories: Evidence from behavior and event-related potentials. *Neurobiol. Learn. Mem.* **99**, 1–9 (2013).
24. Wiesner, C. D. *et al.* The effect of selective REM-sleep deprivation on the consolidation and affective evaluation of emotional memories. *Neurobiol. Learn. Mem.* **122**, 131–141 (2015).
25. Groch, S., Zinke, K., Wilhelm, I. & Born, J. Dissociating the contributions of slow-wave sleep and rapid eye movement sleep to emotional item and source memory. *Neurobiol. Learn. Mem.* **122**, 122–130 (2015).
26. Lehmann, M., Schreiner, T., Seifritz, E. & Rasch, B. Emotional arousal modulates oscillatory correlates of targeted memory reactivation during NREM, but not REM sleep. *Sci. Rep.* **6**, 39229 (2016).
27. Hauner, K. K., Howard, J. D., Zelano, C. & Gottfried, J. A. Stimulus-specific enhancement of fear extinction during slow-wave sleep. *Nat. Neurosci.* **16**, 1553–1555 (2013).
28. Cairney, S. A., Durrant, S. J., Hulleman, J. & Lewis, P. A. Targeted Memory Reactivation During Slow Wave Sleep Facilitates Emotional Memory Consolidation. *Sleep* **37**, 701–707 (2014).
29. Morgenthaler, J. *et al.* Selective REM-sleep deprivation does not diminish emotional memory consolidation in young healthy subjects. *PLoS One* **9**, 1–7 (2014).
30. Ashton, J. E., Cairney, S. A. & Gaskell, M. G. No effect of targeted memory reactivation during slow-wave sleep on emotional recognition memory. *J. Sleep Res.* **27**, 129–137 (2018).
31. Cairney, S., Durrant, S. J., Power, R. & Lewis, P. a. Complementary Roles of Slow-Wave Sleep and Rapid Eye Movement Sleep in Emotional Memory Consolidation. *Cereb. Cortex* **25**, 1565–1575 (2015).
32. Sopp, M. R., Michael, T., Weeß, H.-G. & Mecklinger, A. Remembering specific features of emotional events across time: The role of REM sleep and prefrontal theta oscillations. *Cogn. Affect. Behav. Neurosci.* **17**, 1186–1209 (2017).
33. Groch, S. *et al.* Contribution of norepinephrine to emotional memory consolidation during sleep. *Psychoneuroendocrinology* **36**, 1342–1350 (2011).
34. Alger, S. E. & Payne, J. D. The differential effects of emotional salience on direct associative and relational memory during a nap. *Cogn. Affect. Behav. Neurosci.* **16**, 1150–1163 (2016).
35. Van Der Helm, E. *et al.* REM sleep depotentiates amygdala activity to previous emotional experiences. *Curr. Biol.* **21**, 2029–2032 (2011).
36. Cunningham, T. J. *et al.* Psychophysiological arousal at encoding leads to reduced reactivity but enhanced emotional memory following sleep. *Neurobiol. Learn. Mem.* **114**, 155–164 (2014).
37. Lara-Carrasco, J., Nielsen, T. A., Solomonova, E., Levrier, K. & Popova, A. Overnight emotional adaptation to negative stimuli is altered by REM sleep deprivation and is correlated with intervening dream emotions. *J. Sleep Res.* **18**, 178–187 (2009).
38. Wagner, U., Fischer, S. & Born, J. Changes in emotional responses to aversive pictures across periods rich in slow-wave sleep versus rapid eye movement sleep. *Psychosom. Med.* **64**, 627–634 (2002).
39. Tempesta, D., Succi, V., De Gennaro, L. & Ferrara, M. Sleep and emotional processing. *Sleep Med. Rev.* **40**, 183–195 (2018).
40. Sato, W. & Aoki, S. Right hemispheric dominance in processing of unconscious negative emotion. *Brain Cogn.* **62**, 261–266 (2006).
41. Pizzagalli, D., Regard, M. & Lehmann, D. Rapid emotional face processing in the human right and left brain hemispheres: An ERP study. *Neuroreport* **10**, 2691–2698 (1999).
42. Killgore, W. D. S. & Yurgelun-Todd, D. A. The right-hemisphere and valence hypotheses: Could they both be right (and sometimes left)? *Soc. Cogn. Affect. Neurosci.* **2**, 240–250 (2007).
43. Hoddes, E., Zarcone, V., Smythe, H., Phillips, R. & Dement, W. C. Quantification of Sleepiness: A New Approach. *Psychophysiology* **10**, 431–436 (1973).
44. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society* **57**, 289–300 (1995).
45. Waldhauser, G. T., Braun, V. & Hanslmayr, S. Episodic Memory Retrieval Functionally Relies on Very Rapid Reactivation of Sensory Information. *J. Neurosci.* **36**, 251–260 (2016).
46. Drosopoulos, S., Wagner, U. & Born, J. Sleep enhances explicit recollection in recognition memory. *Learn. Mem.* **12**, 44–51 (2005).
47. Daurat, A., Terrier, P., Foret, J. & Tiberge, M. Slow wave sleep and recollection in recognition memory. *Conscious. Cogn.* **16**, 445–455 (2007).
48. Talamini, L. M., Nieuwenhuis, I. L. C., Takashima, A. & Jensen, O. Sleep directly following learning benefits consolidation of spatial associative memory. *Learn. Mem.* **15**, 233–237 (2008).
49. Gais, S., Lucas, B. & Born, J. Sleep after learning aids memory recall. *Learn. Mem.* **13**, 259–62 (2006).
50. Schönauer, M. *et al.* Decoding material-specific memory reprocessing during sleep in humans. *Nat. Commun.* **8**, 15404 (2017).
51. Cairney, S. A., Guttesen, A. áV., El Marj, N. & Staresina, B. P. Memory Consolidation Is Linked to Spindle-Mediated Information Processing during Sleep. *Curr. Biol.* **28**, 948–954.e4 (2018).
52. Cox, R., Hofman, W. F., de Boer, M. & Talamini, L. M. Local sleep spindle modulations in relation to specific memory cues. *Neuroimage* **99**, 103–110 (2014).
53. Payne, J. D., Chambers, A. M. & Kensinger, E. A. Sleep promotes lasting changes in selective memory for emotional scenes. *Front. Integr. Neurosci.* **6**, 108 (2012).
54. Payne, J. D. *et al.* Napping and the Selective Consolidation of Negative Aspects of Scenes. *Emotion* **15**, 1–11 (2015).
55. Walker, M. P. & van der Helm, E. Overnight therapy? The role of sleep in emotional brain processing. *Psychol. Bull.* **135**, 731–748 (2009).

56. Batterink, L. J., Oudiette, D., Reber, P. J. & Paller, K. A. Sleep facilitates learning a new linguistic rule. *Neuropsychologia* **65**, 169–179 (2014).
57. Wagner, U., Gais, S., Haider, H., Verleger, R. & Born, J. Sleep inspires insight. *Nature* **427**, 352–355 (2004).
58. Marchewka, A., Żurawski, L., Jednoróg, K. & Grabowska, A. The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behav. Res. Methods* **46**, 596–610 (2014).
59. Wilhelm, I. *et al.* Sleep Selectively Enhances Memory Expected to Be of Future Relevance. *J. Neurosci.* **31**, 1563–1569 (2011).
60. van Dongen, E. V., Thielen, J. W., Takashima, A., Barth, M. & Fernández, G. Sleep supports selective retention of associative memories based on relevance for future utilization. *PLoS One* **7**, e43426 (2012).
61. Wamsley, E. J., Tucker, M., Payne, J. D., Benavides, J. A. & Stickgold, R. Dreaming of a learning task is associated with enhanced sleep-dependent memory consolidation. *Curr. Biol.* **20**, 850–855 (2010).
62. Iber, C., Ancoli-Israel, S., Chesson, A. & Quan, S. *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications. Sleep (Rochester)*, <https://doi.org/10.1002/ejoc.201200111> (American Academy of Sleep Medicine, 2007).
63. Macmillan, N. A. & Kaplan, H. L. Detection Theory Analysis of Group Data. Estimating Sensitivity From Average Hit and False-Alarm Rates. *Psychol. Bull.* **98**, 185–199 (1985).
64. Meng, X.-L., Rosenthal, R. & Rubin, D. B. Comparing correlated correlation coefficients. *Psychol. Bull.* **111**, 172–175 (1992).

## Acknowledgements

This work was supported by grants from The Netherlands Organization for Scientific Research (NWO) to RC (446-14-009); National Institutes of Health to ACS (F32-NS093901), and RS (MH048832, MH092638); The Harvard Clinical and Translational Science Center (TR001102); and Stanley Center for Psychiatric Research at Broad Institute.

## Author Contributions

Conceived and designed the experiment: R.C., M.L.V.B., A.C.S., R.S. Collected data: R.C., M.L.V.B., H.G. Preprocessed data: M.L.V.B., M.B., E.C., E.P., O.P.M. Analyzed data: R.C., M.L.V.B., A.C.S., R.S. Prepared figures: R.C. Wrote the manuscript: R.C. All authors contributed to interpretation of results and reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-35999-9>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018