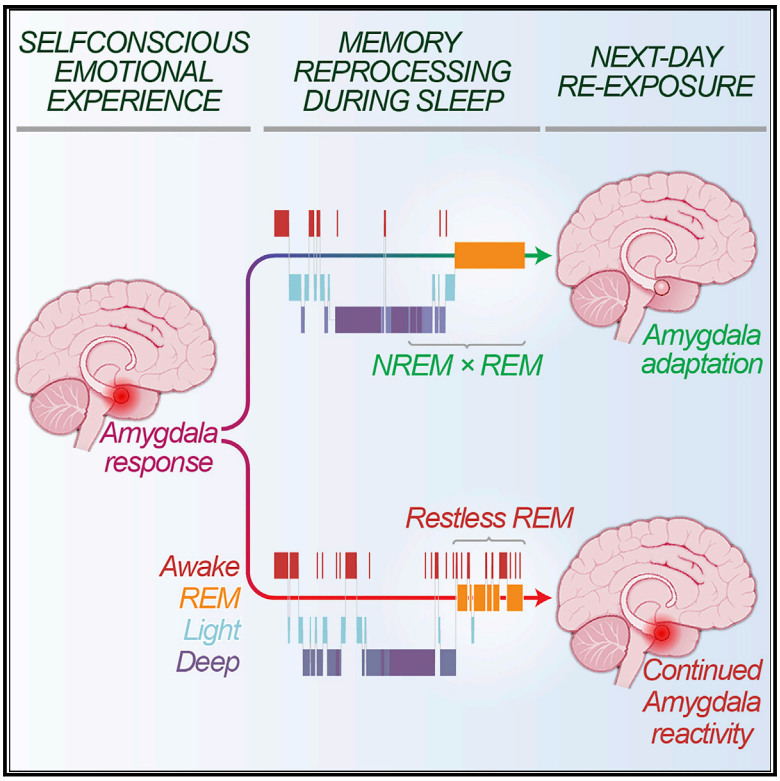
Article

Current Biology

Restless REM Sleep Impedes Overnight Amygdala Adaptation

Graphical Abstract



Highlights

- A novel self-conscious emotional experience elicits amygdala activation
- Overnight amygdala adaptation is proportional to the duration of sound REM sleep
- Preceding sleep spindles increase the benefit of sound REM sleep
- Overnight amygdala adaptation fails proportionally to the restlessness of REM sleep

Authors

Rick Wassing, Oti Lakbila-Kamal, Jennifer R. Ramautar, Diederick Stoffers, Frans Schalkwijk, Eus J.W. Van Someren

Correspondence

e.van.someren@nin.knaw.nl

In Brief

Sleep is considered to be good for about anything, but Wassing et al. reveal a maladaptive type of sleep: restless REM sleep impedes emotion processing in terms of amygdala reactivity. The findings provide a potential target for treatment of mental disorders characterized by restless REM sleep, including insomnia, depression, and anxiety disorders.

Wassing et al., 2019, Current Biology 29, 2351–2358
July 22, 2019
https://doi.org/10.1016/j.cub.2019.06.034

CellPress

Current Biology
Article

Restless REM Sleep Impedes Overnight Amygdala Adaptation

Rick Wassing,[1,2] Oti Lakbila-Kamal,[1] Jennifer R. Ramautar,[1] Diederick Stoffers,[3] Frans Schalkwijk,[4,7] and Eus J.W. Van Someren[1,5,6,7,8,*]

[1]Department of Sleep and Cognition, Netherlands Institute for Neuroscience, an institute of the Royal Netherlands Academy of Arts and Sciences, Meibergdreef 47, 1105 BA Amsterdam, the Netherlands
[2]Centre for Integrated Research and Understanding of Sleep (CIRUS), Woolcock Institute of Medical Research, The University of Sydney, 431 Glebe Point Road, Sydney, NSW 2037, Australia
[3]Spinoza Centre for Neuroimaging, Meibergdreef 75, 1105 BK Amsterdam, the Netherlands
[4]Department of Education, Program Group Forensic Child and Youth Care, University of Amsterdam, Nieuwe Achtergracht 127, 1018 WS Amsterdam, the Netherlands
[5]Department of Integrative Neurophysiology, Center for Neurogenomics and Cognitive Research, Amsterdam Neuroscience, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, the Netherlands
[6]Amsterdam UMC, Vrije Universiteit, Psychiatry, Amsterdam Neuroscience, Oldenaller 1, 1081 HJ Amsterdam, the Netherlands
[7]Senior author
[8]Lead Contact
*Correspondence: e.van.someren@nin.knaw.nl
https://doi.org/10.1016/j.cub.2019.06.034

SUMMARY

Animal studies show that insufficient silencing of the locus coeruleus (LC) during REM sleep impairs sleep-related brain plasticity. Restless REM sleep, a characteristic of several psychiatric disorders, likely reflects insufficient LC silencing. We investigated whether endogenous REM sleep interruptions interfere with overnight reorganization of limbic circuits in human volunteers with a wide range of insomnia severity, from no insomnia complaints to fulfilling community-sample criteria for insomnia disorder. We induced a self-conscious emotion during two functional MRI sessions and recorded sleep EEG in between. Amygdala reactivity decreased overnight in proportion to the total duration of consolidated REM sleep. Restless REM sleep, in contrast, impeded overnight amygdala adaptation. Using targeted memory reactivation with odors tagged to the self-conscious emotional stimulus, we could experimentally enhance both the favorable effect of consolidated REM sleep and the unfavorable effect of restless REM sleep. The findings reveal a maladaptive type of sleep, providing a target for interventions in mental disorders characterized by restless REM sleep.

INTRODUCTION

Several studies have shown that sleep aids the overnight resolution of emotional distress [1–4]. We here use the term "emotional distress" to refer to the combined unpleasant subjective experience and activation of the limbic circuit and autonomic nervous system, which can be elicited by exposure to emotional stimuli as well as by their recall or re-exposure. Sleep is thought to provide a time window for reactivation and reorganization of the neuronal circuits that were activated during the initial emotional experience [5]. For example, whereas the amygdala initially activates during an emotionally distressful experience, subsequent neuronal network reorganization that is facilitated by sleep results in amygdala inhibition if the experience is recalled or encountered again later [1, 6, 7]. While the reorganization of emotional memory circuits has been related to either rapid eye movement (REM) sleep or non-REM (NREM) sleep [8–10], their roles are best studied in an integrated way [8, 11]. Indeed, an animal model on the role of sleep in resolving emotional distress proposes an interplay of memory trace reactivation and synaptic plasticity during a time window that starts when slow waves subside and sleep spindle-rich sleep emerges, and lasts until the end of REM sleep [5]. The spindle-rich interval between slow-wave sleep and REM sleep is known as "transition to REM" (TTR) sleep in animal studies. In humans, it has its equivalent in an episode of predominantly stage N2 sleep that precedes REM sleep. This episode has been referred to as "ascending" sleep [12]. Spindles in TTR sleep may serve to spontaneously reactivate emotional memory traces [13], while subsequent REM sleep could support further memory transformation [5, 8]. However, no human study to date has investigated how TTR and REM sleep interact to promote the overnight reorganization of emotional memory circuits and dissolving of distress.

During wakefulness and NREM sleep, locus coeruleus (LC) activity maintains noradrenaline release at a level that promotes long-term potentiation and impedes depotentiation of synapses [14–16]. Uniquely before and during REM sleep, the LC is inhibited [17] while network activity is seen in limbic and paralimbic brain regions, including the amygdala and dorsal anterior cingulate cortex [18–20]. This time window of low noradrenaline release facilitates synaptic depotentiation [5, 17, 21]. In some mental disorders, electroencephalographic (EEG) recordings suggest insufficient inhibition of LC activity during sleep [22, 23]. Abundant arousals,

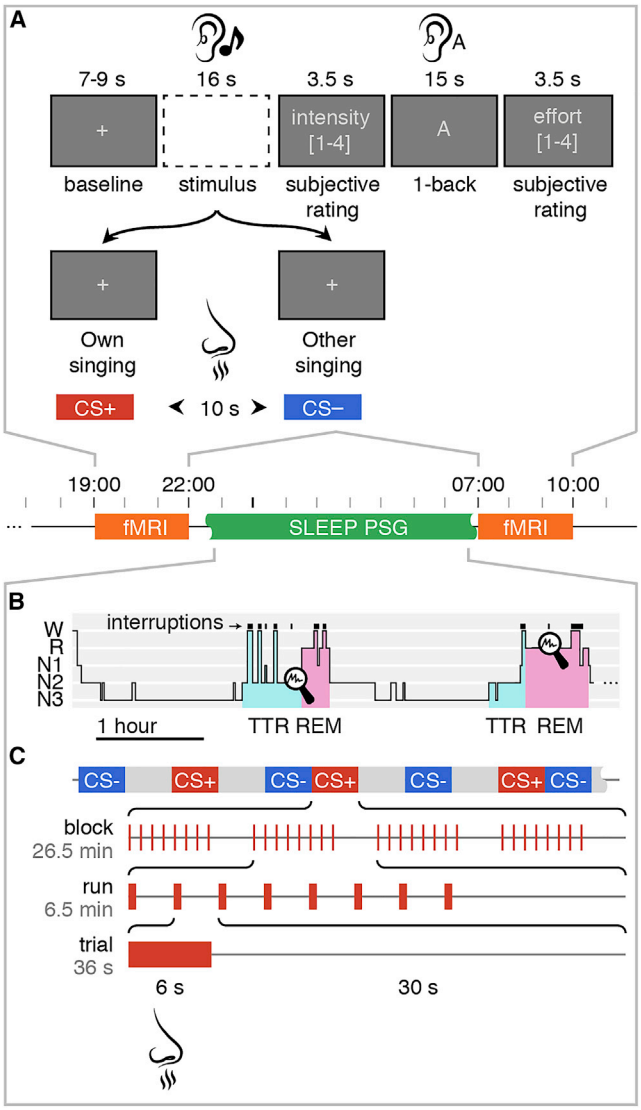


Figure 1. Procedure

After a habituation night (not shown), volunteers participated in two fMRI sessions (orange). Polysomnography (green) was assessed in between.

(A) fMRI paradigm. Trials started with a fixation-cross presented on the screen, after which their own out-of-tune singing or other in-tune singing stimulus was presented. For participants who successfully perceived and differentiated the odors (STAR Methods), the own-singing (US+) and other-singing stimuli (US–) were tagged with two olfactory cues (CS+, red; CS–, blue). Subjective emotional intensity was assessed with response options ranging from "none" (1) to "strong" (4). To prevent lingering of induced emotions, each trial closed with a 1-back task and a rating on the effort it took to perform that task.

(B) Polysomnography. Specific episodes can be distinguished in each sleep cycle. A REM episode (pink) covers the time between the first and last REM epoch. The transition to REM episode (TTR; turquoise) commences after the last two consecutive epochs of stage N3 sleep and lasts until the onset of a REM episode. The REM episode interruption density is the total number of cortical arousals and bouts of wakefulness or NREM sleep that interrupt REM episodes, divided by the total duration of REM episodes. Likewise, TTR episode interruptions are calculated as the density of cortical arousals and bouts of wakefulness or stage N1 NREM sleep that interrupt TTR episodes.

(C) Conditioned odors were presented overnight to induce targeted memory reactivation. To minimize interference and adaptation, the CS+ or CS– was presented in distributed blocks. PSG, polysomnography; fMRI, functional magnetic resonance imaging; CS, conditioned stimulus; REM, rapid eye movement; TTR, transition to REM.

suggesting that LC activity continues into TTR and REM sleep, have been observed as a result of early childhood adversity [24], in insomnia disorder (ID) [25], in people with post-traumatic stress disorder (PTSD) [26, 27], and also in animal models for PTSD [28]. It is well conceivable that persisting LC activity interferes with overnight emotion regulation by impeding REM sleep-related processes that utilize synaptic depotentiation. However, to date no human study experimentally addressed whether restless REM sleep impedes overnight adaptation of limbic circuit activity. The term "restless REM sleep" has been coined to refer to REM sleep with a high number of phasic events [3]. The denser the occurrence of arousals and stage transitions during an episode of REM sleep, the more restless it can be considered.

We here experimentally addressed the hypothesis that overnight reorganization of neuronal networks that include the limbic circuitry is best facilitated by TTR and REM sleep *only* if these sleep stages have the signature of low LC activity, just as was recently demonstrated for hippocampal network reorganization in rats [17]. While whole-night assessment of LC activity during REM sleep is currently not feasible in humans, cortical arousals and shifts to light sleep and wakefulness signal increased LC activity [22, 29–31]. Although we cannot strictly exclude other mechanisms underlying these proposed markers of LC activity, their sensitivity and specificity are corroborated by animal studies showing that stimulation of the LC causes immediate sleep-to-wake transitions, from both NREM and REM sleep, independently from other known arousal-promoting nuclei [30], and results in high-frequency EEG activity, i.e., EEG arousals [31]. Based on this theoretical framework, we expect interference with overnight limbic circuitry reorganization by abundant TTR and REM sleep interruptions that are indicative of persistent LC activity into sleep. Recent work showed that these interruptions are the hallmark of insomnia [25], and that the capacity of overnight dissolving of emotional distress decreases with increasing insomnia severity [3]. Therefore, to warrant sufficient variance with respect to individual differences in the duration and restlessness of TTR and REM sleep episodes, we included N = 29 participants covering a wide range of insomnia severity, i.e., from no insomnia complaints to fulfilling community-sample criteria for ID. We used functional magnetic resonance imaging (fMRI) to record the limbic response induced by the self-conscious emotional distress of listening to audio fragments of their own out-of-tune singing (Figure 1A; example: https://youtu.be/G3gWyua3grE). These stimuli have been validated before to elicit a self-conscious emotion [32, 33]. The exposure was repeated after a night of sleep to quantify reactivity adaptation across the night. EEG was recorded during sleep and scored according to standard criteria to obtain sleep stages and cortical arousals [34–36]. General linear models were used to assess whether an individual's overnight decrease in amygdala reactivity was (1) proportional to the total durations of TTR and REM episodes and (2) impeded by interruptions during these sleep episodes.

Table 1. Sleep Spindles, Total Duration, and Interruption Density of REM and TTR Episodes

	Mean (SD)	Range	r(ISI)[a]	p Value
Total duration of REM episodes (min)	99.6 (43.5)	[33.5–201.0]	–0.06	0.77
Total duration of TTR episodes (min)	37.8 (23.6)	[2.5–96.0]	0.35	0.06
REM episode interruption density (N/h)	15.2 (5.4)	[6.8–25.5]	0.69	3.3×10^{-5}
TTR episode interruption density (N/h)	14.8 (9.0)	[0.0–39.0]	0.13	0.51
Spindle count in TTR episodes (N)	112.2 (112.5)	[0.0–495.0]	0.20	0.30
Integrated spindle activity in TTR episodes (mV^2)	19.5 (21.5)	[0.0–100.7]	0.17	0.38

See also Tables S1 and S5. ISI, insomnia severity index; REM, rapid eye movement; TTR, transition to REM.
[a]Pearson correlation coefficient between the sleep variable and the total insomnia severity index

Moreover, in N = 13 participants that tested positively on their ability to perceive and differentiate odors, the initial distressful exposure was tagged with an olfactory cue to allow for subsequent targeted memory reactivation (TMR) during sleep [37, 38]. Employing differential conditioning, the shameful emotional experience (US+, own out-of-tune singing) was coupled to one odor (CS+) while a non-self-conscious control experience (US–, another in-tune singer) was coupled to another odor (CS–). Targeted memory reactivations were attempted by distributed re-exposure to the odors throughout sleep. For each individual, we assessed which part of the total TTR and REM episode duration coincided with CS+ re-exposure. We tested whether established positive and negative effects of REM-related sleep variables on overnight amygdala adaptation increased with the proportion of time the TTR and REM episodes coincided with CS+ re-exposure.

RESULTS

Manipulation Check: Subjective and BOLD Responses Support Induction of Self-Conscious Emotion

We followed the advice of Van Der Helm [1] to assess emotion as concisely as possible, to prevent the cognitive processing that would be required to distinguish and report different aspects of emotions. We followed their successful protocol of assessing, after each stimulus, only an intensity rating on a unipolar Likert-type scale ranging from "none" (1) to "strong" (4). Own-singing stimuli were rated more emotionally intense (mean (SD [range]) = 2.0 (0.7 [1–3.4]) than other-singing stimuli (1.5 (0.5 [1–2.4]); paired two-sample t test, $t(28) = 3.97$, $p = 0.0004$). At the end of each run, participants were asked to rate the intensity of 17 emotions (words) that they might have experienced during the run. The emotion words were "shame," "embarrassment," "fear," "fright," "anger," "upset," "sadness," "rage," "disgust," "aversion," "surprise," "interest," "pleasure," "excitement," "pride," "humiliation," and "guilt," each to be rated on a Likert-type scale from "none" (1) to "strong" (4). Analysis of these data confirms (1) that listening to karaoke fragments induced shame (one-sample t test, $t(28) = 12.35$, $p < 7.6^{-13}$) and embarrassment ($t(28) = 11.26$, $p < 6.6^{-12}$), (2) that the intensity ratings of shame and embarrassment were more intense than the ratings of eight other emotions (all $p < 0.05$), and (3) that none of the remaining seven emotions were rated more intense than shame (all $p > 0.92$) or embarrassment (all $p > 0.76$). Successful manipulation was also supported by significant BOLD responses to own-singing stimuli in limbic circuits, including the bilateral amygdalae, bilateral medial prefrontal cortex, and left posterior cingulate cortex (details shown in Table S4). Finally, applying olfactory stimulation in only part of the volunteers allowed us to evaluate possible confounding effects of odor presentation. Mixed-effects linear models indicated that TMR-exposed and non-TMR-exposed participants did not differ with respect to subjective emotional intensity ratings at the first session ($p = 0.21$), nor in their overnight change ($p = 0.25$). Similarly, there were no significant group differences in either the amygdala BOLD response during the first session ($p = 0.44$) or in its overnight change ($p = 0.25$).

Duration and Continuity of TTR and REM Predict Overnight Adaptation of Amygdala Reactivity

We evaluated whether individual differences in overnight amygdala reactivity changes were (1) proportional to the total duration of REM and TTR episodes and (2) moderated by the interruption density in these sleep episodes. For each individual, the overnight change in the bilateral amygdala BOLD response to own-singing stimuli was determined using a Brainnetome atlas mask [39].

Polysomnographically recorded sleep was staged according to standard procedures (Table S1) [40], and cortical arousals during sleep were indicated by transient high-frequency EEG activity (>16 Hz) lasting between 3 and 15 s [35]. The stages were used to calculate the total duration of four specific sleep episodes [12]: (1) the transition to deep sleep starts at sleep onset or after the final REM epoch of each sleep cycle and ends with the first two consecutive epochs of stage N3 sleep, (2) the deep sleep episode is the period comprised of mainly stage N3 NREM sleep, (3) the transition to REM (TTR) episode is the period following the last two consecutive epochs of stage N3 sleep until the onset of the REM episode, and (4) the REM episode is the period between the first and last REM epoch in each sleep cycle (Figure 1B). REM episode interruption density was calculated as total number of cortical arousals and bouts of wakefulness or NREM sleep that interrupted REM episodes, divided by the total duration of REM episodes. The TTR episode interruption density was likewise calculated as the total number of cortical arousals and bouts of wakefulness or stage-1 NREM sleep that interrupted TTR episodes, divided by the total duration of TTR sleep (Table 1) [25]. A general linear model evaluated whether the duration and interruption density of the REM and TTR episodes predicted the overnight change in amygdala reactivity in all participants who received TMR.

Table 2. Main and Interaction Effects of Sleep Variables on Overnight Change in Amygdala Reactivity

	β (SE)	*t* Statistic	p Value
Intercept			
Mean change in amygdala reactivity	–0.09 (0.02)	–4.07	0.001
Main Effects			
Total duration of REM episodes (h)[a]	–0.12 (0.04)	–2.84	0.01
Total duration of TTR episodes (h)[b]	–0.10 (0.07)	–1.43	0.17
REM episode interruption density (N/h)[c]	–0.008 (0.004)	–1.93	0.07
TTR episode interruption density (N/h)[d]	0.003 (0.003)	0.97	0.34
Interactions			
Total duration of TTR episodes × total duration of REM episodes	–0.33 (0.09)	–3.90	0.001
Total duration of REM episodes × REM interruption density	0.013 (0.006)	2.17	0.04
Total duration of TTR episodes × TTR interruption density	0.004 (0.005)	0.87	0.39

The overnight decrease in amygdala reactivity is proportional to the total duration of REM episodes. More time spent in the preceding TTR enhances the effect of the total duration of REM episodes, while more REM interruptions counteract it. Ancillary models, replacing the variable "total duration of TTR episodes" with the "number of spindles" or "integrated spindle activity," are shown in Data S1. In order to obtain a meaningful intercept (mean change in amygdala reactivity), all independent variables have been centered. REM, rapid eye movement; TTR, transition to REM.

[a]Period between the first and last REM epoch in each sleep cycle

[b]Period following the last two consecutive epochs of stage N3 sleep until the onset of the REM episode

[c]Total number of cortical arousals and bouts of wakefulness or NREM sleep that interrupted REM episodes, divided by the total duration of REM episodes

[d]Total number of cortical arousals and bouts of wakefulness or stage 1 NREM sleep that interrupted TTR episodes, divided by the total duration of TTR sleep

On average, amygdala reactivity decreased overnight (β = –0.09 (0.02), $t(21)$ = –4.07, p = 5.5 × 10^{-4}; Table 2). Individual differences in the decrease were proportional to the total duration of REM episodes (β = –0.12 (0.04), $t(21)$ = –2.84, p = 0.01; Figure 2B). Although individual differences in the decrease were not proportional to the total duration of TTR episodes themselves (β = –0.10 (0.07), $t(21)$ = –1.43, p = 0.17), a significant interaction with the total duration of REM episodes indicated that longer lasting TTR episodes boosted the effect of subsequent REM episode duration on the overnight decrease in amygdala reactivity (β = –0.33 (0.09), $t(21)$ = –3.90, p = 8.2 × 10^{-4}; Figure 2C).

A second moderation effect was indicated by an interaction of the total duration of REM episodes and interruption density of REM episodes (β = 0.013 (0.006), $t(21)$ = 2.17, p = 0.04; Figure 2C). The positive sign of this interaction indicates that with increasing interruption density, REM episodes become less supportive of the overnight decrease in amygdala reactivity. There was no significant interaction between the total duration of TTR episodes and interruption density of TTR episodes (β = 0.004 (0.005), $t(21)$ = 0.87, p = 0.39). There were no significant main effects of interruption density in either TTR or REM episodes ($0.07 \leq p \leq 0.34$).

In order to evaluate whether sleep spindles are an important factor in the contribution of TTR episode duration to the overnight decrease in amygdala reactivity, we used automated spindle detection [41] and calculated the total number and total integrated activity (STAR Methods) of spindles in TTR episodes. We evaluated two ancillary models, analog to the original model presented in Table 2, by replacing the variable "total duration of TTR episodes" with the "number of spindles" in one model and with "integrated spindle activity" in the other model. As to be expected, there were strong associations of individual differences in total duration of TTR episodes with both the total number and integrated activity of spindles during TTR (r = 0.84 and 0.82, respectively). Not surprisingly, therefore, the effect of REM sleep on overnight adaptation of amygdala reactivity to own-singing stimuli indeed increased with both the total number and integrated activity of spindles during TTR episodes (p = 0.003 and 0.006, respectively). The finding specifically indicates an interaction with REM sleep because neither the number nor the integrated activity of spindles themselves had a significant main effect (p = 0.44 and 0.48, respectively), as was the case for the original model with total duration of TTR episodes. The ancillary models that specified spindle-specific information about the TTR episodes did not explain more variance (R^2 = 0.47 and 0.44, respectively) than the original model that merely included total duration of TTR episodes (R^2 = 0.48). Details are presented in Data S1.

Ancillary analysis of a model that included the total durations of all four types of sleep episodes and their interactions with total REM episode durations verified specificity of REM episodes and their interaction with transition to REM episodes. Adding the other two types of sleep episodes and their interactions with REM episodes did not improve the model (F test, $F(4,17)$ = 0.38, p = 0.82), and none of the added effects were significant ($0.48 \leq p \leq 0.94$).

In summary, the findings indicate a stronger overnight decrease in amygdala reactivity with increasing duration of unperturbed REM sleep; this effect can be enhanced by longer preceding TTR episodes and counteracted if REM episodes have abundant interruptions, up to the point that the benefit of REM sleep is completely lost.

Overnight Adaptation of Amygdala Reactivity by Induced Memory Reactivation in REM Period Sleep

To experimentally support REM-related effects on overnight adaptation of amygdala reactivity, we promoted reactivation

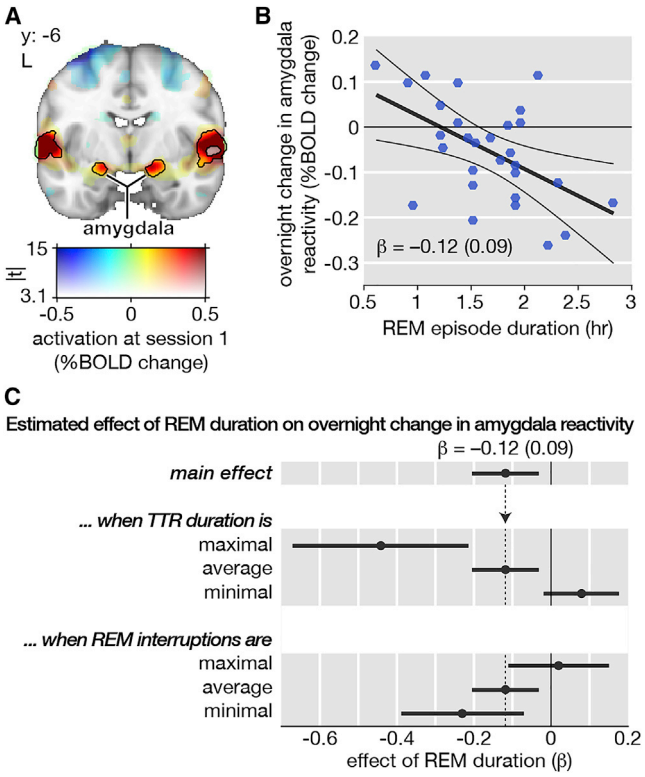


Figure 2. Overnight Amygdala Adaptation Is Proportional to the Duration of Sound REM Sleep but Fails Proportional to the Restlessness of REM Sleep

(A) Amygdala response with exposure to self-conscious own-singing stimuli. The magnitude of the BOLD response to one's out-of-tune solo singing is indicated with color hue, and the voxel's statistical significance ranges from transparent (t = 3.1) to opaque (t = 15). Areas delineated with a black outline indicate voxels with a significant BOLD response after whole-brain family-wise error correction (α = 0.05; bilateral amygdala and auditory cortex; see also Table S4).

(B) The overnight decrease in amygdala reactivity is proportional to the total duration of REM episodes. The thick line visualizes the effect of total duration of REM episodes on the overnight change in amygdala reactivity; the thin lines indicate the 95% confidence interval of the estimated effect.

(C) The β-coefficient and standard error of the main effect estimate of REM episode duration are shown for reference (top panel) and adjusted effect estimates of REM episode duration are shown for three levels of the moderating variable: the minimum, mean, and maximum values observed across all participants (middle and bottom panel). Middle: a significant interaction between total TTR and REM durations indicated that longer lasting TTR episodes boosted effects of subsequent REM episodes on the overnight decrease in amygdala reactivity. Bottom: a second significant moderation effect was indicated by an interaction of REM episode duration and interruption density. With increasing interruptions, REM episodes become less supportive of the overnight decrease in amygdala reactivity. BOLD, blood-oxygen-level-dependent signal; REM, rapid eye movement; TTR, transition to REM.

processes during sleep using TMR with differentially conditioned odors [37, 38]. During the first fMRI session, own-singing and other-singing stimuli (US+ and US–) were differentially tagged with two olfactory cues (CS+ and CS–). Throughout subsequent sleep, TMR was induced by distributed re-exposure to the CS+ or CS– odors, presented in blocks of 4 runs of 8 trials (Figure 1C). For each individual, we assessed the proportion of time that TTR episodes and REM episodes coincided with CS+ and CS– re-exposures. Naturally occurring individual differences in the distribution of sleep stages created a range of CS+ and CS– re-exposure proportions during TTR and REM episodes across participants (TTR, CS+ [0%–57.6%], CS– [0%–92.1%]; REM, CS+ [0%–44.6%], CS– [0%–54.6%]; for details, see Table S3). The proportions allowed us to investigate whether specifically CS+ TMR, but not CS– TMR, modified the effect of the total duration and interruption density of TTR and REM episodes on overnight adaptation in amygdala reactivity.

Table 3. Specifically CS+ TMR during REM Episodes Facilitates Overnight Amygdala Reactivity Adaptation

Main Effects	β Estimate (SE)	t Statistic	p Value
CS+ re-exposure proportion in REM episodes	–0.56 (0.23)	–2.47	0.04
CS– re-exposure proportion in REM episodes	–0.07 (0.18)	–0.39	0.71
CS+ re-exposure proportion in TTR episodes	0.11 (0.20)	0.53	0.61
CS– re-exposure proportion in TTR episodes	0.03 (0.13)	0.25	0.81

See Tables S2 and S3 for allocation of the four odor compounds for CS+ and CS–, and their re-exposure proportions during TTR and REM episodes. CS, conditioned stimulus; REM, rapid eye movement; TTR, transition to REM.

Regression analyses indicated that the overnight decrease in amygdala reactivity to own-singing stimuli became stronger in proportion to the time that REM episodes accommodated CS+ re-exposures (β = –0.56 (0.23), t(8) = –2.47, p = 0.04; Table 3). Specificity of CS+ during REM episodes was shown by the lack of effects of CS+ re-exposure proportions during TTR episodes (β = –0.07 (0.18), t(8) = –0.39, p = 0.71), or of CS– re-exposure proportions during either TTR episodes (β = 0.03 (0.13), t(8) = 0.25, p = 0.81) or REM episodes (β = 0.11 (0.20), t(8) = 0.53, p = 0.61; Table 3). Finally, a significant interaction effect indicated that the CS+ re-exposure proportion in REM episodes also enhanced the adverse effect of REM episode interruption density on the overnight decrease in amygdala reactivity (β = 0.06 (0.03), t(9) = 2.37, p = 0.04; Table 4).

In summary, the findings indicate that TMR during REM episodes using an odor that was conditioned to the own-singing stimuli during prior wakefulness added to the favorable effect of the total duration of REM episodes on the overnight decrease in amygdala reactivity, but also enhanced the adverse effect of the REM episode interruption density.

DISCUSSION

We addressed the hypothesis that TTR and REM sleep facilitate overnight amygdala adaptation, but *only* if these sleep stages are sufficiently consolidated, as indicated by relatively few interruptions. We found a stronger overnight decrease in amygdala reactivity with an increasing duration of unperturbed REM sleep. The effect of REM sleep on amygdala adaptation was enhanced if preceding episodes of TTR sleep were of longer duration, contained more spindles, or showed higher total integrated spindle activity. Notably, however, the effect was counteracted if REM

Table 4. CS+ TMR during REM Episodes Enhances the Adverse Effect of REM Interruption Density on Overnight Amygdala Reactivity Adaptation

	β Estimate (SE)	*t* Statistic	p Value
Main Effects			
CS+ re-exposure proportion in REM episodes	–0.58 (0.13)	–4.49	0.002
REM episode interruption density (N/h)	–0.01 (0.004)	–2.09	0.07
Interaction			
CS+ re-exposure × REM interruption density	0.06 (0.03)	2.37	0.04

CS, conditioned stimulus; REM, rapid eye movement; TTR, transition to REM.

episodes had abundant interruptions, up to the point that the benefit of REM sleep was completely lost.

The findings underscore the importance of an integrated approach to the functional role of sleep. Many previous studies have focused on the role of individual sleep variables (for example, the duration of stages or properties of sleep events like slow oscillations and spindles) in isolation. These studies indicated that NREM and REM sleep have complex and multifaceted roles in overnight neuronal network adaptations with relevance to emotion and cognition [9, 10]. It has, for example, been suggested that NREM sleep has an initial role in memory reactivation and consolidation, while subsequent REM sleep could support further memory transformation [5, 8]. The roles of sleep variables are therefore best investigated in an integrated way [8, 11].

Consequently, based on animal studies on the role of sleep in fear extinction and synaptic plasticity [5, 17, 28], our analyses integrated not only NREM and REM episodes but also the microstructure of interruptions and spindles within these sleep stages. Only because of this integrated approach were we able to reveal effects of sleep on overnight adaptive brain processes that could otherwise have cancelled out and gone unnoticed. First, while no main effects were found for individual differences in the total duration of the transition to REM episodes, longer durations boosted the effect of total REM episode duration on overnight amygdala adaptation. Second, while no main effects were found for individual differences in the density of interruptions in REM episodes, denser interruptions significantly interfered with the effect of total REM episode duration on overnight amygdala adaptation. Third, the results support the idea of spindles as an important factor in the contribution of prior TTR episodes to the role of REM sleep in regulating amygdala reactivity. The integrated approach to sleep stages and microstructure is a particular strength of our study.

Another strength is that these effects could consistently be enhanced by use of targeted memory reactivation. TMR with an odor conditioned to the own-singing stimuli offered during REM episodes added to the favorable effect of REM episodes on the overnight decrease in amygdala reactivity, but also enhanced the unfavorable effect of REM episode interruptions. These effect modifications were elicited only by the odor tagged to the own-singing stimuli: no effects were seen for another odor that was tagged to audio fragments of a professional singer.

A third strength of our study is that we included participants covering a range from no insomnia complaints to fulfilling community-sample criteria for ID. This approach provided sufficient individual differences in the duration and restlessness of TTR and REM sleep episodes. It was this variance that allowed us to uncover specificity and interactions of the duration and interruption density of episodes of REM sleep and transition to REM sleep.

Some limitations deserve mention. We propose that restless TTR and REM sleep are maladaptive across different types of distress and across disorders characterized by such restlessness during sleep, including disorders of affect and anxiety including PTSD. However, we demonstrated amygdala maladaptation only in a sample with a wide range of insomnia severity and only for the particular distress related to a shameful experience. Future studies are needed to evaluate whether a similar maladaptation can be demonstrated in disorders other than insomnia and for other self-conscious and basic emotions.

Another limitation is that while our theoretical framework proposes that continued LC activity during REM sleep hampers overnight restructuring of neuronal networks, our study did not directly assess LC activity. Instead, we assessed cortical arousals and shifts to light sleep and wakefulness as a proxy variable that signals increased LC activity [22, 29]. A recent animal study that employed subtle optogenetic stimulation of the LC during sleep showed that LC silencing during sleep is necessary for proper memory consolidation [17]. Future animal studies would be required to confirm that LC silencing during sleep supports adaptation of amygdala reactivity.

Another possible limitation is that the protocol did not allow us to directly infer whether the differential conditioning was successful. Odors have been used successfully as contextual cues; Hauner et al. performed a manipulation check during post-conditioning wakefulness and showed increased activations to the CS+ versus CS– in the limbic circuit (amygdala and hippocampus) and salience network (orbitofrontal cortex, insula, and anterior cingulate cortex) [42]. Our protocol did not allow for such a manipulation check after the conditioning procedure because presenting the olfactory cues would induce memory reactivation during waking. We specifically addressed effects of memory reactivation during sleep only, which would likely be confounded if preceded by wake reactivation. The finding that TMR during REM episodes indeed altered the overnight decrease in amygdala reactivity suggests that the pre-sleep differential conditioning was successful, in line with the findings of Hauner et al.

A final limitation is our relatively small sample size, inherent to the very demanding nature of the protocol. Model-predicted associations were evaluated in 29 participants, and ancillary support for causality by means of TMR was acquired in only 13 of them. Whereas replication is clearly desirable, the effects may be sufficiently strong to be demonstrated in samples of moderate size. Previous work showed that 18 subjects were sufficient to find an association between REM sleep and amygdala reactivity [1].

The findings significantly add to the proposed role of REM sleep in emotional adaptive processes involving changes in neuronal circuits including the amygdala [1]. Our findings are

also in line with an animal model of fear extinction, which proposes that the role of REM sleep is facilitated by the sleep spindle-rich transition period that precedes it [5, 28]. While this animal model also suggests that the drop in noradrenaline prior to and during REM sleep importantly modulates the restructuring of neuronal networks, we cannot provide direct support for this role because whole-night assessment of LC activity or central noradrenaline availability during REM sleep is currently not feasible in humans. Future studies employing pharmacological or optogenetic manipulations are needed to evaluate whether blocking or boosting noradrenaline during REM sleep facilitates or interferes with overnight amygdala adaptation, respectively.

In summary, we showed that REM sleep can support overnight regulation of amygdala reactivity. The effect increases with longer preceding episodes of transition to REM but is impeded by REM sleep interruptions. Chronically perturbed REM sleep has been observed as a result of early childhood adversity [24], in ID [25], and in people with PTSD [26, 27]. It is conceivable that chronic insufficiency of overnight adaptive processes in the amygdala could result in the daytime hyperarousal that is characteristic of these disorders [3]. Addressing overnight emotional memory processing deficits in these disorders is likely to provide clues to the mechanisms underlying hyperarousal, which have so far remained enigmatic.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Preparation
  - Experimental procedures
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Odor perception and differentiation tests
  - fMRI data processing
  - Polysomnography processing
  - BOLD responses to auditory stimuli
  - Main statistical analyses
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cub.2019.06.034.

ACKNOWLEDGMENTS

The authors would like to thank all the volunteers involved in this study, as well as the undergraduate students that helped with the acquisition of the data: C. Dekker, V. Hohn, M. Korpelshoek, K. Müller, L. Nijboer, C. Oosterwijk, S. Riemsma, M. Spaander, G. Thijsma, and N. Veerman. This work was supported by ZONMW Neuropsychoanalysis Fund grant 16.561.0001 of the Netherlands Organization of Scientific Research (NWO), by grant 253/2012 of the Bial Foundation, and by the European Research Council grants ERC-2014-AdG-671084 INSOMNIA and ERC-2016-PoC-737634 INSOMNIA BEAT IT.

AUTHOR CONTRIBUTIONS

Conceptualization, F.S. and E.J.W.V.S.; Methodology, R.W., F.S., and E.J.W.V.S.; Formal Analysis, R.W.; Investigation, R.W., O.L.-K., J.R.R., and D.S.; Data Curation, R.W. and O.L.-K.; Writing – Original Draft, R.W. and E.J.W.V.S.; Writing – Review & Editing, R.W., J.R.R., F.S., and E.J.W.V.S.; Supervision, J.R.R., D.S., F.S., and E.J.W.V.S.; Funding Acquisition, F.S. and E.J.W.V.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 18, 2019
Revised: May 14, 2019
Accepted: June 11, 2019
Published: July 11, 2019

REFERENCES

1. van der Helm, E., Yao, J., Dutt, S., Rao, V., Saletin, J.M., and Walker, M.P. (2011). REM sleep depotentiates amygdala activity to previous emotional experiences. Curr. Biol. *21*, 2029–2032.
2. Sterpenich, V., Albouy, G., Boly, M., Vandewalle, G., Darsaud, A., Balteau, E., Dang-Vu, T.T., Desseilles, M., D'Argembeau, A., Gais, S., et al. (2007). Sleep-related hippocampo-cortical interplay during emotional memory recollection. PLoS Biol. *5*, e282.
3. Wassing, R., Benjamins, J.S., Dekker, K., Moens, S., Spiegelhalder, K., Feige, B., Riemann, D., van der Sluis, S., Van Der Werf, Y.D., Talamini, L.M., et al. (2016). Slow dissolving of emotional distress contributes to hyperarousal. Proc. Natl. Acad. Sci. USA *113*, 2538–2543.
4. Pace-Schott, E.F., Shepherd, E., Spencer, R.M., Marcello, M., Tucker, M., Propper, R.E., and Stickgold, R. (2011). Napping promotes inter-session habituation to emotional stimuli. Neurobiol. Learn. Mem. *95*, 24–36.
5. Vanderheyden, W.M., Poe, G.R., and Liberzon, I. (2014). Trauma exposure and sleep: using a rodent model to understand sleep function in PTSD. Exp. Brain Res. *232*, 1575–1584.
6. Nieuwenhuis, I.L., and Takashima, A. (2011). The role of the ventromedial prefrontal cortex in memory consolidation. Behav. Brain Res. *218*, 325–334.
7. Takashima, A., Petersson, K.M., Rutters, F., Tendolkar, I., Jensen, O., Zwarts, M.J., McNaughton, B.L., and Fernández, G. (2006). Declarative memory consolidation in humans: a prospective functional magnetic resonance imaging study. Proc. Natl. Acad. Sci. USA *103*, 756–761.
8. Gais, S., Plihal, W., Wagner, U., and Born, J. (2000). Early sleep triggers memory for early visual discrimination skills. Nat. Neurosci. *3*, 1335–1339.
9. Wagner, U., Gais, S., and Born, J. (2001). Emotional memory formation is enhanced across sleep intervals with high amounts of rapid eye movement sleep. Learn. Mem. *8*, 112–119.
10. Walker, M.P., and Stickgold, R. (2006). Sleep, memory, and plasticity. Annu. Rev. Psychol. *57*, 139–166.
11. Stickgold, R., Whidbee, D., Schirmer, B., Patel, V., and Hobson, J.A. (2000). Visual discrimination task improvement: a multi-step process occurring during sleep. J. Cogn. Neurosci. *12*, 246–254.
12. Terzano, M.G., Parrino, L., Boselli, M., Smerieri, A., and Spaggiari, M.C. (2000). CAP components and EEG synchronization in the first 3 sleep cycles. Clin. Neurophysiol. *111*, 283–290.
13. Rasch, B., and Born, J. (2013). About sleep's role in memory. Physiol. Rev. *93*, 681–766.
14. Izumi, Y., Clifford, D.B., and Zorumski, C.F. (1992). Norepinephrine reverses N-methyl-D-aspartate-mediated inhibition of long-term potentiation in rat hippocampal slices. Neurosci. Lett. *142*, 163–166.
15. Thomas, M.J., Moody, T.D., Makhinson, M., and O'Dell, T.J. (1996). Activity-dependent beta-adrenergic modulation of low frequency stimulation induced LTP in the hippocampal CA1 region. Neuron *17*, 475–482.

16. Pace-Schott, E.F., and Hobson, J.A. (2002). The neurobiology of sleep: genetics, cellular physiology and subcortical networks. Nat. Rev. Neurosci. *3*, 591–605.

17. Swift, K.M., Gross, B.A., Frazer, M.A., Bauer, D.S., Clark, K.J.D., Vazey, E.M., Aston-Jones, G., Li, Y., Pickering, A.E., Sara, S.J., and Poe, G.R. (2018). Abnormal locus coeruleus sleep activity alters sleep signatures of memory consolidation and impairs place cell stability and spatial memory. Curr. Biol. *28*, 3599–3609.e4.

18. Nofzinger, E.A., Mintun, M.A., Wiseman, M., Kupfer, D.J., and Moore, R.Y. (1997). Forebrain activation in REM sleep: an FDG PET study. Brain Res. *770*, 192–201.

19. Braun, A.R., Balkin, T.J., Wesenten, N.J., Carson, R.E., Varga, M., Baldwin, P., Selbie, S., Belenky, G., and Herscovitch, P. (1997). Regional cerebral blood flow throughout the sleep-wake cycle. An H2(15)O PET study. Brain *120*, 1173–1197.

20. Maquet, P., Ruby, P., Maudoux, A., Albouy, G., Sterpenich, V., Dang-Vu, T., Desseilles, M., Boly, M., Perrin, F., Peigneux, P., and Laureys, S. (2005). Human cognition during REM sleep and the activity profile within frontal and parietal cortices: a reappraisal of functional neuroimaging data. Prog. Brain Res. *150*, 219–227.

21. Kemp, A., and Manahan-Vaughan, D. (2004). Hippocampal long-term depression and long-term potentiation encode different aspects of novelty acquisition. Proc. Natl. Acad. Sci. USA *101*, 8192–8197.

22. Aston-Jones, G., Gonzalez, M., and Doran, S. (2007). Role of the locus coeruleus-norepinephrine system in arousal and circadian regulation of the sleep-wake cycle. In Brain Norepinephrine: Neurobiology and Therapeutics, G.A. Ordway, M.A. Schwartz, and A. Frazer, eds. (Cambridge University Press), pp. 157–195.

23. Pace-Schott, E.F., Germain, A., and Milad, M.R. (2015). Sleep and REM sleep disturbance in the pathophysiology of PTSD: the role of extinction memory. Biol. Mood Anxiety Disord. *5*, 3.

24. Insana, S.P., Kolko, D.J., and Germain, A. (2012). Early-life trauma is associated with rapid eye movement sleep fragmentation among military veterans. Biol. Psychol. *89*, 570–579.

25. Riemann, D., Spiegelhalder, K., Nissen, C., Hirscher, V., Baglioni, C., and Feige, B. (2012). REM sleep instability–a new pathway for insomnia? Pharmacopsychiatry *45*, 167–176.

26. Mellman, T.A., Bustamante, V., Fins, A.I., Pigeon, W.R., and Nolan, B. (2002). REM sleep and the early development of posttraumatic stress disorder. Am. J. Psychiatry *159*, 1696–1701.

27. Germain, A., Buysse, D.J., and Nofzinger, E. (2008). Sleep-specific mechanisms underlying posttraumatic stress disorder: integrative review and neurobiological hypotheses. Sleep Med. Rev. *12*, 185–195.

28. Poe, G.R. (2017). Sleep is for forgetting. J. Neurosci. *37*, 464–473.

29. Aston-Jones, G., and Bloom, F.E. (1981). Activity of norepinephrine-containing locus coeruleus neurons in behaving rats anticipates fluctuations in the sleep-waking cycle. J. Neurosci. *1*, 876–886.

30. Carter, M.E., Yizhar, O., Chikahisa, S., Nguyen, H., Adamantidis, A., Nishino, S., Deisseroth, K., and de Lecea, L. (2010). Tuning arousal with optogenetic modulation of locus coeruleus neurons. Nat. Neurosci. *13*, 1526–1533.

31. Berridge, C.W., and Foote, S.L. (1991). Effects of locus coeruleus activation on electroencephalographic activity in neocortex and hippocampus. J. Neurosci. *11*, 3135–3145.

32. Wassing, R., Benjamins, J.S., Talamini, L.M., Schalkwijk, F., and Van Someren, E.J.W. (2019). Overnight worsening of emotional distress indicates maladaptive sleep in insomnia. Sleep *42*, zsy268.

33. Wassing, R., Schalkwijk, F., Lakbila-Kamal, O., Ramautar, J.R., Stoffers, D., Mutsaerts, H.J.M.M., Talamini, L.M., and Van Someren, E.J.W. (2019). Haunted by the past: old emotions remain salient in insomnia disorder. Brain *142*, 1783–1796.

34. Iber, C., Ancoli-Israel, S., Chesson, A., and Quan, S. (2007). The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology, and Technical Specification (American Academy of Sleep Medicine).

35. Bonnet, M.H., Carley, D., Carskadon, M., Easton, P., Guilleminault, C., Harper, R., Hayes, B., Hirshkowitz, M., Ktonas, P., and Keenan, S. (1992). EEG arousals: scoring rules and examples: a preliminary report from the Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. Sleep *15*, 173–184.

36. Berry, R.B., Budhiraja, R., Gottlieb, D.J., Gozal, D., Iber, C., Kapur, V.K., Marcus, C.L., Mehra, R., Parthasarathy, S., Quan, S.F., et al.; American Academy of Sleep Medicine; Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine (2012). Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. J. Clin. Sleep Med. *8*, 597–619.

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

38. Arzi, A., Shedlesky, L., Ben-Shaul, M., Nasser, K., Oksenberg, A., Hairston, I.S., and Sobel, N. (2012). Humans can learn new information during sleep. Nat. Neurosci. *15*, 1460–1465.

39. Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A.R., et al. (2016). The Human Brainnetome Atlas: a new brain atlas based on connectional architecture. Cereb. Cortex *26*, 3508–3526.

40. Berry, R.B., Brooks, R., Gamaldo, C.E., Harding, S.M., Lloyd, R.M., Marcus, C.L., and Vaughn, B.V. (2015). The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, Version 2.2 (American Academy of Sleep Medicine).

41. Lacourse, K., Delfrate, J., Beaudry, J., Peppard, P., and Warby, S.C. (2019). A sleep spindle detection algorithm that emulates human expert spindle scoring. J. Neurosci. Methods *316*, 3–11.

42. Hauner, K.K., Howard, J.D., Zelano, C., and Gottfried, J.A. (2013). Stimulus-specific enhancement of fear extinction during slow-wave sleep. Nat. Neurosci. *16*, 1553–1555.

43. Benjamins, J.S., Migliorati, F., Dekker, K., Wassing, R., Moens, S., Blanken, T.F., Te Lindert, B.H.W., Sjauw Mook, J., and Van Someren, E.J.W. (2017). Insomnia heterogeneity: characteristics to consider for data-driven multivariate subtyping. Sleep Med. Rev. *36*, 71–81.

44. Beckmann, C.F., Jenkinson, M., and Smith, S.M. (2003). General multilevel linear modeling for group analysis in FMRI. Neuroimage *20*, 1052–1063.

45. Woolrich, M.W., Behrens, T.E., Beckmann, C.F., Jenkinson, M., and Smith, S.M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. Neuroimage *21*, 1732–1747.

46. Woolrich, M.W., Ripley, B.D., Brady, M., and Smith, S.M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. Neuroimage *14*, 1370–1386.

47. Pruim, R.H.R., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J.K., and Beckmann, C.F. (2015). ICA-AROMA: a robust ICA-based strategy for removing motion artifacts from fMRI data. Neuroimage *112*, 267–277.

48. Todorov, A., and Engell, A.D. (2008). The role of the amygdala in implicit evaluation of emotionally neutral faces. Soc. Cogn. Affect. Neurosci. *3*, 303–312.

49. Somerville, L.H., Kim, H., Johnstone, T., Alexander, A.L., and Whalen, P.J. (2004). Human amygdala responses during presentation of happy and neutral faces: correlations with state anxiety. Biol. Psychiatry *55*, 897–903.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Compound fragrances: "strawberry," "gin," "tamarind," and "white tea"	International Flavors & Fragrances B.V., Liebergerweg 72-76, 1221 JT Hilversum, the Netherlands	N/A
Deposited Data		
Raw polysomnography and MRI data	This paper	Available upon request
Dataset containing calculated PSG-variables, amygdala BOLD responses and subjective emotional intensity ratings	This paper	http://doi.org/10.17026/dans-z3b-azw7
Software and Algorithms		
MATLAB R2016b, Statistics and Machine Learning Toolbox	The MathWorks, Natick, MA, USA	RRID: SCR_001622
E-Prime software	Psychology Software Tools, Sharpsburg, PA, USA	RRID: SCR_009567

LEAD CONTACT AND MATERIALS AVAILABILITY

This study did not generate new unique reagents. Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Eus J.W. Van Someren (e.van.someren@nin.knaw.nl).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

N = 29 participants (14 males and 15 females) were recruited by a newsletter emailed to volunteers of the Netherlands Sleep Registry [43]. Gender identity was not assessed and is therefore not reported. The inclusion criterion was an age between 18 and 70 years. Exclusion criteria were any diagnosed current or past neurological or psychiatric disorder, any current sleep disorder other than Insomnia Disorder (ID), chronic use of medication, the use of sleep medication during the prior 2 months, and any MRI contraindications. The Insomnia Severity Index ranged from 0 to 24 (mean (SD) = 10.7 (7.7)), indicative of inclusion of both good and poor sleepers. Indeed, a diagnostic interview upon inclusion indicated that N = 12 subjects fulfilled the diagnostic criteria for ID according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) and the International Classification of Sleep Disorders (ICSD-3). None had a diagnosis of any other disorder. Informed consent was obtained prior to study enrolment. The study was approved by the ethics review board of the University of Amsterdam, the Netherlands.

METHOD DETAILS

Preparation

One week prior to the experiment, subjects came to the lab for an intake interview, structural MRI-scans, karaoke-style audio recordings and tests for odor-perception and odor-discrimination. To obtain an audio recording of their singing, participants heard instrumentation and vocals of others over headphones while singing along with the lyrics presented in a Karaoke-style video. Their own voice was not presented over the headphones to impede pitch correction and thus promote out-of-tune singing. All participants underwent an odor-discrimination tests to evaluate whether they fulfilled the prerequisite for differential conditioning, i.e., capacity to differentiate the odors to be linked later to US+ and US–. Only those with sufficient odor discrimination capacity were exposed odors during fMRI and sleep for TMR (N = 13, 7 males, ISI mean (SD) = 11.1 (8.4), ISI range from 0 to 24, see below for details). Participants underwent an adaptation night in the lab, including polysomnography (PSG; Electrical Geodesic, Eugene, OR, United States of America) the night before the first fMRI session.

Technical description of the olfactometer

Participants were fitted with a polytetrafluoroethylene (Teflon) Y-shaped cannula that was placed directly below the nasal vestibules, connected to a custom build olfactometer (Figure S1 provides a technical description). A small and quiet air pump provided a constant airflow of 1.5 L/min, which was divided into a bypass airflow (1 L/min) and into solenoid valves (0.5 L/min) connected to four syringe-filter capsules containing different odors. The syringe filters were enclosed by one-way valves to ensure an airtight seal in the closed state, and a unidirectional airflow in the open state. The bypass airflow reconnected with the airflow from the syringe-filter capsules and ensured a constant clean unidirectional airflow in which the scented air from the odor-containing capsules could get "injected" into. We selected four compound solutions with discernible profiles normally used for industrial flavoring and

fragrances: ''strawberry,'' ''gin,'' ''tamarind,'' and ''white tea,'' of which two would be selected for each participant to be conditioned with the US+ and US– (International Flavors & Fragrances B.V., Liebergerweg 72-76, 1221 JT, Hilversum, the Netherlands). Capsules were loaded with 40 μL of compound solution, and next to the four odor-containing capsules, two capsules remained empty. With this approach, the olfactometer could provide an odor stimulation by switching from ''no-odor'' to one of the four odors, or provide a sham stimulation by switching from ''no-odor'' to another ''no-odor.'' The solenoid valves were operated by a program run on an Arduino UNO microprocessor or by serial connection with a computer running a valve switching script in E-Prime software (Psychology Software Tools, Sharpsburg, PA, USA).

Testing the ability to perceive odors

One week before the experiments, participants performed a 30-min odor perception test including 40 trials, with the initial instruction to breath normally and regularly. Furthermore, participants were instructed that with each trial, one out of four odors was presented for the duration of 3 s, and that in some trials no odor was presented. Each trial was preceded with the instruction to wait until their breathing cycle reached the end of exhalation, and to press spacebar to start the trial. In each trial one of the four odors or a sham stimulation was presented for 3 s (8 trials for each odor and 8 sham trials, in a random order). Directly after the stimulation, the participant was queried with the question ''the odor is pleasant,'' followed by ''the odor is intense'' and ''the odor is recognizable,'' with response options on a Likert-scale ranging from ''completely disagree'' (1) to ''completely agree'' (7). In case participants perceived an odor, they had to press a key between 1 and 7. In case no odor was perceived, participants had to press 0. The inter-trial interval varied between 30 and 120 s. The participant's overall ability to perceive odors was calculated as the percentage of correctly identified odor-trials.

Testing the ability differentiate odors

In addition to the ability to perceive the odors, a second prerequisite for differential conditioning is the ability to discern two odors. To this end, participants performed a 45-min odor differentiation test including 36 trials. Participants were instructed that they would be presented with three consecutive odor stimulations, of which two are the same, and that they had to indicate which odor was different. Each stimulation was preceded with the instruction to wait until the breathing cycle reached the end of exhalation, and the stimulation was initiated by pressing spacebar and lasted 3 s. Directly after the third stimulation, the question ''which odor was different'' appeared on screen with the answer options ''1,'' ''2,'' and ''3.'' The participant received feedback upon their response. We calculated the participant's overall ability to differentiate the odors as the overall percentage of correct trials, as well as the percentage of correct trials for each possible pair of odors.

Criteria for admission to odor conditioning

The criteria for participants to be admitted to odor conditioning and targeted memory reactivation, were a perception accuracy above 50% for at least two odors, and a differentiation accuracy score of at least 66% correct between that pair of odors (chance level = 33%).

Experimental procedures

MRI scans were made before (between 19:00 and 22:00 hr) and after (between 07:00 and 10:00 hr) PSG recorded sleep. Bedtimes were between 10:00 and 11:00 pm and rise times between 6:00 and 8:00 am, according to an individual's habitual sleep timing. During fMRI, the selfconscious emotion of shame was induced by confronting subjects with listening to fragments of their own often embarrassingly out-of-tune solo singing (example: https://youtu.be/G3gWyua3grE). Control stimuli consisted of the same fragments, sung in-tune by a professional singer. Participants with sufficient odor-discrimination capacity were moreover exposed to odors during initial audio fragment exposures, during sleep, and during post-sleep re-exposure to audio fragments (see below).

fMRI paradigm

Audio fragments were presented in a block design fMRI paradigm that consisted of two runs of five own-singing stimuli and five other-singing stimuli each, in counterbalanced order. Each trial started with a fixation-cross presented on the screen for 7 to 9 s, after which a recording was presented for 16 s. Selfconscious stimuli were audio fragments of the subject's own solo singing. Non-selfconscious stimuli were audio fragments of a professional singer. Subjects with sufficient odor-discrimination capacity underwent differential conditioning during stimulus exposure (see below, odor conditioning and targeted memory reactivation). Subjects were asked to rate their perceived emotional intensity on a unipolar Likert-type scale from ''none'' (1) to ''strong'' (4). To prevent possible lingering of induced emotions into subsequent trials, each trial then continued with an audio-visual 1-back task for 15 s to divert attention. A sequence of 9 letters was presented both on a screen and over headphones. Participants were instructed to compare the current letter with the preceding letter and respond with a button-press using their index-finger if the letter was different or with their middle finger if the letter was the same (0-3 targets). Finally, participants rated the effort it took them to perform the 1-back task on a unipolar Likert-type scale ranging from ''none'' (1) to ''strong'' (4).

BOLD responses were assessed using Echo Planar Imaging (EPI, N = 212 images per run, TR: 2.5 s, TE: 28 ms, 2.5 mm isotropic voxels, 43 slices, FoV: 240 by 240 mm) on a Philips Achieva 3T MRI scanner (Philips Healthcare Systems, Best, the Netherlands). A T1-weighted scan was used for anatomical registration (1 mm^3), and B0-fieldmaps were acquired to adjust the EPI images for magnetic-field distortions.

Polysomnography

On both the adaptation night and the night in between emotion inductions, polysomnography (PSG) recordings were obtained using a 256-channel HydroCel EEG net referenced to the Cz-electrode (Electrical Geodesic, Eugene, OR). We simultaneously assessed: EMG using Ag/AgCl electrodes placed on the submental area and on the anterior tibialis; ECG using Ag/AgCl electrodes placed

in accordance with the standard lead II configuration; and respiration, using respiratory belt transducers around the upper and lower chest. Electrode impedances were kept below 100 kΩ, which provides excellent signal quality due to the amplifier's high internal impedance. Signals were online band-pass filtered between 0.1-100 Hz and digitized at 1000 Hz.

Odor conditioning and targeted memory reactivation

Counterbalancing odors across participants, own-singing stimuli (US+) were coupled to one odor (CS+), and other-singing stimuli (US–) to the other (CS–) (Figure 1A). Odors were presented within a continuous airflow of 1.5 L/min. During subsequent nocturnal targeted memory reactivation (TMR; Figure 1C) the odors were presented in distributed blocks to minimize interference and adaptation. On average 10.9 blocks were presented throughout the night, and the mean (SD) time interval between blocks was 21.7 (23.0) minutes. In each block, either the CS+ or CS– was presented in 4 runs of 8 trials (Figure 1C). In each trial, the odor was presented for 6 s with an inter-trial interval of 30 s.

QUANTIFICATION AND STATISTICAL ANALYSIS

Odor perception and differentiation tests

For the participants that fulfilled the minimal odor perception and differentiation criteria to be admitted to odor conditioning and targeted memory reactivation, we assigned two out of the four odor compounds as the CS+ and CS– odor, aiming at distributed combinations of odors across participants (Table S2). A Chi-square test indicated that the difference between the expected and observed frequencies of odor allocation was not significant ($\chi;2(3) = 1.87$, $p = 0.60$). The perception accuracy for the allocated CS+ and CS– odors was very high (CS+: 95.5 ± 15.1%, CS–: 100.0 ± 0.0%), as was the differentiation accuracy between the odor pairs (92.4 ± 13.7%).

fMRI data processing

fMRI data were preprocessed with FMRIB's package FSL FEAT package version 5.0.10, including correction for subject-specific B0-field distortion maps [44–46]. In brief, EPI images were masked to strip non-brain tissue, spatially smoothed with a Gaussian kernel (5 mm FWHM) to reduce noise and remain sensitive to small brain responses, normalized to the grand mean intensity, and motion corrected with MCFLIRT. The ICA-AROMA algorithm and nuisance regression were applied to remove motion artifacts and noise [47], and subsequently the EPI-time series were high-pass filtered with a cut-off at 1/90 Hz. Transformation matrices were obtained by the combination of a boundary-based registration of the EPI scan to the anatomical scan with a linear registration of the anatomical scan to the standard-space image (MNI152-T1 image, 1 mm^3, FLIRT, 12 DOF).

Polysomnography processing

Polysomnographically recorded sleep was staged according to standard procedures (Table S1) [40], and cortical arousals during sleep were indicated by transient high-frequency EEG activity (> 16 Hz) lasting between 3 and 15 s [35]. The stages were used to calculate the total duration of four specific sleep episodes [12]: (1) the *transition to deep sleep* starts at sleep onset or after the final REM epoch of each sleep cycle and ends with the first two consecutive epochs of stage N3 sleep, (2) the deep sleep episode is the period comprised of mainly stage N3 NREM sleep, (3) the *transition to REM* (TTR) episode is the period following the last two consecutive epochs of stage N3 sleep until the onset of the REM episode, and (4) the *REM episode* is the period between the first and last REM epoch in each sleep cycle (Figure 1B). REM episode interruption density was calculated as total number of cortical arousals and bouts of wakefulness or NREM sleep that interrupted REM episodes, divided by the total duration of REM episodes. The TTR episode interruption density was likewise calculated as the total number of cortical arousals and bouts of wakefulness or stage-1 NREM sleep that interrupted TTR episodes, divided by the total duration of TTR episodes (Table 1) [25].

Finally, a recently developed spindle detection algorithm, that was validated to perform as well as manual identification of spindles by experts [41], was applied to identify spindles in the EEG recordings. The signal from electrode C3, re-referenced to mastoid 2 (M2), was bandpass filtered between 0.3 and 30 Hz and employed in the spindle detection algorithm. Spindles identified in epochs labeled as stage-2 or stage-3 NREM sleep were counted in TTR episodes. Next to 'total spindle count', 'total integrated spindle activity' was calculated by integrating the spectral power in the sigma-band across each sleep spindle and then summed over all identified spindles in TTR episodes.

BOLD responses to auditory stimuli

Subject-level fMRI data employed the FSL FEAT package version 5.0.10 to analyze the BOLD time-series with general linear models. The planned analysis was a within-subject model comparing the amygdala BOLD response to own-singing stimuli relative to baseline before and after sleep. To determine the nuisance regressors to be included, we evaluated the whole-brain BOLD response to the blocks of own-singing, of other-singing, of the 1-back task, and of subjective rating. This analysis showed no amygdala response to the 1-back task blocks and subjective rating blocks, and a significant amygdala response not only to one's own singing, but as well to someone else's singing. The unusual listening to someone else singing solo without accompanying music may have elicited an amygdala response due to novelty [48] or due to shame triggered by comparing one's own bad performance to that of the professional singer. Such undetermined amygdala activation to intended neutral stimuli within the context of an emotional experiment has been observed previously [49]. Given our detailed dissection of activation patterns, and the uncertainty about internal processes elicited by listening to someone else singing solo, the other-singing blocks were included as nuisance regressor in the final model.

In order not to lose statistical power in estimating nonsignificant effects, the final model integrated the volumes that were acquired during the 1-back task and blocks and subjective rating blocks among the other baseline volumes. Thus, the BOLD response to own-singing, and to other-singing stimuli, relative to baseline was modeled with two box-car regressors convolved with a double-gamma hemodynamic response function (HRF). To adjust for variation in timing of the actual HRF and slice-acquisition, the first-order derivatives of these HRF-regressors were added to the design-matrix. To control for motion artifacts, we added a confound-regressor for each time-sample where excessive motion was detected (RMS intensity difference: 75th percentile+1.5 × IQR). Two first-level β-coefficient contrasts were obtained for each subject and stimulus-type (own-singing and other-singing), which were used in whole-brain group-level general linear models, that included a covariate indicating group-membership for TMR, to estimate (1) the mean BOLD response at the first session and (2) the mean difference in BOLD response between the first and second session. Finally, for each individual, the overnight change in the bilateral amygdala BOLD response was extracted using a Brainnetome atlas mask and used in the main statistical analyses as the dependent variable.

Main statistical analyses

Group-level analyses employed general linear models (Statistics and Machine Learning Toolbox, MATLAB, The MathWorks, Natick, MA), regression coefficients were evaluated with two-tailed *t*-contrasts, and their statistical significance was considered at $\alpha = 0.05$.

The main general linear model (original model) evaluated whether the duration and interruption density of the REM and TTR episodes was predictive of the overnight change in amygdala reactivity to own-singing stimuli. Specifically, next to the four main effects, the model included three interaction effects of 'total duration of TTR episodes' × 'total duration of REM episodes', 'total duration of REM episodes' × 'REM interruption density', and 'total duration of TTR episodes' × 'TTR interruption density'. Two ancillary analyses were performed to evaluate the specificity of the found effects. First, to evaluate whether the other sleep episodes played a role in the overnight change in amygdala reactivity, two additional factors were added to the original model including their interaction with 'total duration of REM episodes', namely 'total duration of transition to deep-sleep episodes' and 'total duration of deep-sleep episodes'.

In order to evaluate whether spindles are an important factor in the contribution of TTR episodes in regulating amygdala reactivity to own-singing stimuli, we evaluated two models based on the original model by replacing the factor 'total duration of TTR episodes' with either 'total spindle count' or 'total integrated spindle activity'.

Finally, two general linear models evaluated the effects of our experimental manipulation of induced memory reactivation during sleep. First, two models evaluated whether the proportion of time that TTR episodes and REM episodes coincided with CS+ and CS– re-exposures facilitated the overnight change in amygdala reactivity to own-singing stimuli. Specifically, this model included four factors: CS+ re-exposure proportion in REM episodes and in TTR episodes, and CS– re-exposure proportion in REM episodes and in TTR episodes. Second, a model evaluated whether CS+ re-exposure proportion in REM episodes enhanced the adverse effect of REM interruptions. Specifically, the model included two factors and their interaction, namely 'CS+ re-exposure proportion in REM episodes', and 'REM episode interruption density'.

DATA AND CODE AVAILABILITY

The accession number for the data reported in this paper is http://doi.org/10.17026/dans-z3b-azw7.