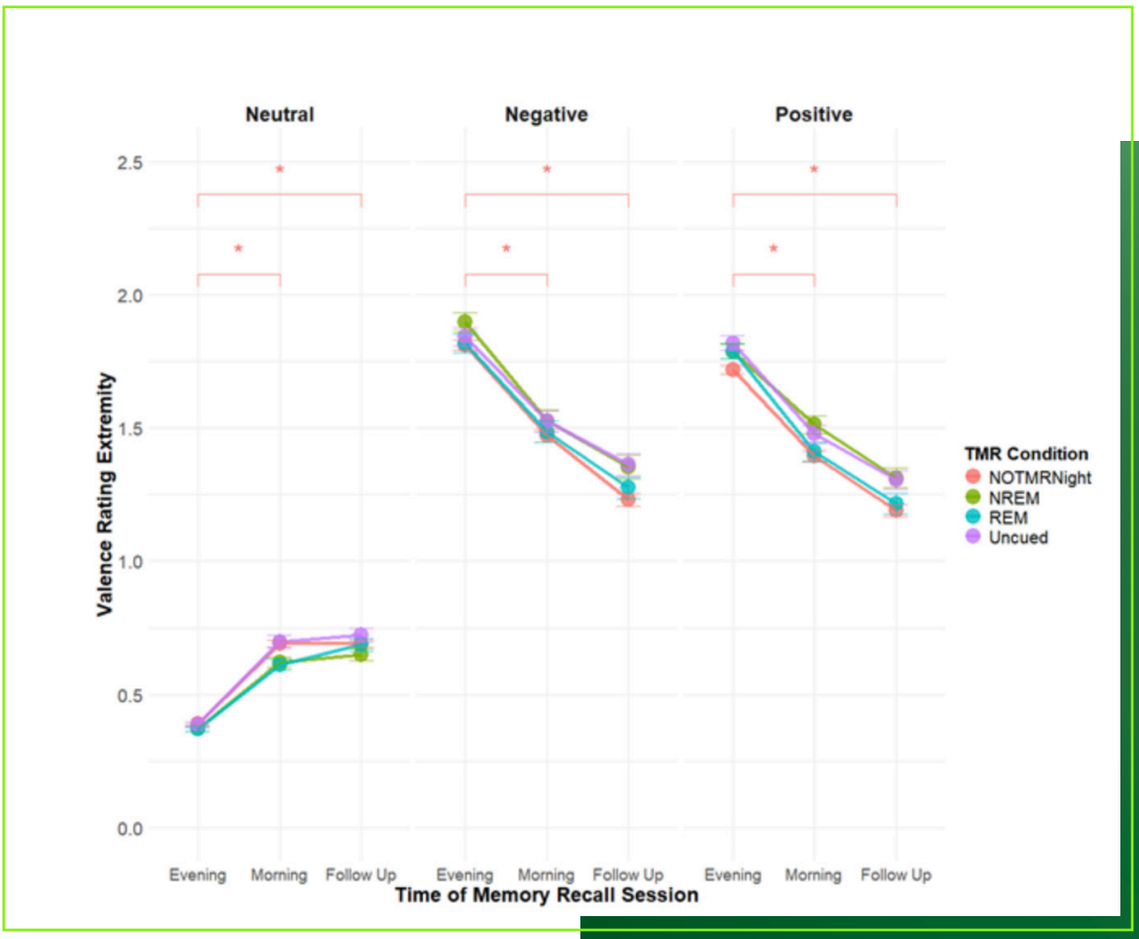
REM and NREM TMR Effects on Emotional Memory Processing During Sleep

Lisa Hazebroucq
Supervisor: Sarah Schoch

Donders Research Center for Cognitive Neuroscience

Nijmegen, the Netherlands

September 2024


Neutral
Negative
Positive
Valence Rating Extremity
2.5
2.0
1.5
1.0
0.5
0.0
*
*
*
*
*
*
Evening
Morning
Follow Up
Evening
Morning
Follow Up
Evening
Morning
Follow Up
Time of Memory Recall Session
TMR Condition
NOTMRNight
NREM
REM
Uncued

Abstract

Emotional memories are often better remembered, but their associated emotional tone tends to fade over time. The ”sleep to remember, sleep to forget” hypothesis suggests that sleep plays a crucial role in this phenomenon by selectively consolidating the memory while decreasing its associated emotional tone, a process potentially driven by different sleep stages. In this study, we aimed to investigate the effect of targeted memory reactivation during NREM and REM sleep on emotional memory processing. Specifically, we measured changes in arousal ratings and valence rating extremity for images initially rated as positive, neutral, or negative. Each image was paired with a neutral word. We compared the effects of reactivating images during NREM or REM sleep, or leaving them uncued, on a night when targeted memory reactivation was applied, to changes in arousal and valence ratings for images presented during a control night. We found that four days after TMR application, NREM cueing led to a decrease in arousal ratings of negative memories, while REM cueing distinctively reduced arousal ratings and dampened the decrease in valence rating extremity of negative images. These effects extended to uncued items and were primarily driven by forgotten images. Additionally, using support vector machine classification, we were able to detect reactivation of the arousal component of memories in NREM2 and consistently identified reactivation of negative valence compared to neutral in both NREM2 and REM sleep. To our knowledge, this is the first study demonstrating the feasibility of detecting reactivation of emotional components of memory during sleep, providing a novel approach to exploring the sleep stage-specific mechanisms that support emotional memory processing.

1 Introduction

1.1 Definition of Emotion

Emotions assume a central role in human decision-making processes (Ficca et al., 2000; Schwarz, 2000). They function as perceptual filters that accentuate salient elements (Niu et al., 2012) within a given situation, guiding individuals toward specific response patterns. In cognitive neuroscience, several theories have emerged to explain the nature and function of emotions. One prominent theory suggests that emotions are constructed through a process known as active inference (Seth and Friston, 2016), wherein physiological states are categorized as specific emotional experiences based on the context in which interoceptive stimuli occur and prior experiences (Barrett, 2017). This process enables rapid evaluation of situations (Dolan, 2002), facilitating quick decision-making (Finucane et al., 2000; Schwarz, 2000).

Within this perspective, emotions are often conceptualized along two primary dimensions: arousal, which ranges from low to high, and valence, which spans from negative to positive, with a neutral midpoint (Posner et al., 2005). Arousal is closely linked to autonomic nervous system activation, while valence refers to the cognitive appraisal of an experience. This cognitive appraisal has been proposed to depend on higher-order interpretations of the physiological state in a given context (Barrett, 2017).

1.2 Emotional Memory are Better Remembered while Their Emotional Tone Tend to Vanish with Time

In addition to their role in immediate decision-making, emotions—particularly their arousal component—are believed to enhance the identification and prioritization of significant events for memory consolidation, a concept known as ”emotional tagging” (Richter-Levin and Akirav, 2003). Indeed, numerous studies have demonstrated that emotionally arousing stimuli are more likely to be remembered over time than non-arousing ones, with this memory advantage becoming more pronounced as the retention interval lengthens. (Peace & Porter, 2004; Sharot and Phelps, 2004; Van Giezen et al., 2005; Dolcos et al., 2005; LaBar and Cabeza, 2006). This enhanced memory retention is linked to increased activation of the amygdala during the encoding of emotionally arousing memories (Richter-Levin and Akirav, 2003; Dolcos et al., 2004; Kensinger and Schacter, 2006; LaBar and Cabeza, 2006). Additionally, arousal triggers the activation of glucocorticoid and adrenergic systems, which work in concert with the basolateral amygdala to modulate memory formation by promoting plasticity in the hippocampus (LaBar and Cabeza, 2006; Roozendaal et al., 2009). While the ability to remember emotionally significant events is generally adaptive, the persistence of the associated emotional state can become pathological, as seen in post-traumatic stress disorder (Zoellner et al., 2001). Emotional memories are conceptualized as comprising two distinct components: the factual content of the event (episodic memory) and the emotional response it triggers (affective tone). Over time, the episodic memory typically persists, while its affective tone tends to fade (M. P. Walker, 2009). At the neuro-

physiological level, this phenomenon is linked to a dissociation between amygdala and hippocampal activity during emotional memory recollection (Dolcos et al., 2005).

1.3 Sleep to Forget and Sleep to Remember Hypothesis

Research has increasingly highlighted the essential role of sleep in the dual process of emotional memory consolidation (Bennion et al., 2015; Singh et al., 2022). The "sleep to forget and sleep to remember" hypothesis posits that sleep not only strengthens emotional memory consolidation but also reduces the affective charge associated with these memories (M. P. Walker, 2009; Figure 1). This hypothesis is underpinned by empirical findings demonstrating that sleep enhances the consolidation of salient memories (Wilhelm et al., 2011; Schoch et al., 2017; Alger et al., 2018), including emotional memory (P. Hu et al., 2006; Wagner et al., 2006; Wagner et al., 2007; Payne et al., 2008; Payne and Kensinger, 2010; Payne and Kensinger, 2011; Payne et al., 2012). Additionally, sleep is thought to facilitate a decrease in the emotional intensity of these memories, as evidenced by reduced amygdala reactivity during subsequent emotional recollection (Sterpenich et al., 2007), supporting the notion that sleep plays a critical role in dampening the emotional tone of memories (M. P. Walker and van Der Helm, 2009).

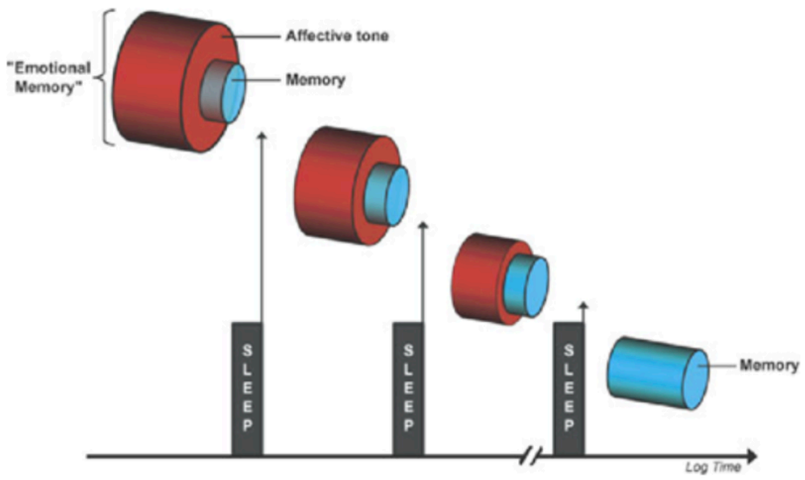


Figure 1: **Model of sleep-dependent emotional–memory processing (Walker 2009).** The "sleep to forget and sleep to remember" hypothesis suggests that, during sleep, emotional memories are strengthened while their emotional intensity gradually fades. Over time, the memory is retained, but the affective tone associated with it diminishes, leaving only the core memory without the initial emotional charge.

1.4 NREM Sleep and REM Sleep Are Hypothesized to Support Emotional Memory Consolidation and Affective Tone Reduction, Respectively

Sleep is not a homogeneous state, and different sleep stages are believed to play distinct roles in emotional memory processing. While Non-REM (NREM) sleep has been consistently associated with

memory consolidation (Rasch and Born, 2013), the specific role of REM sleep in emotional memory consolidation remains less well understood. In fact, the consolidation of ”simple” declarative memories, such as word pairs, appears to be independent of REM sleep (Smith, 2001; Ackermann and Rasch, 2014). Additionally, longer REM sleep duration has even been linked to decreased retention of low-value items (Oudiette et al., 2013). However, more complex emotional declarative memories seem to benefit from REM sleep. For example, REM sleep has been shown to enhance the retention of emotional prose (Wagner et al., 2001) and improve recognition of emotional images over neutral ones (Groch et al., 2013; Groch et al., 2015; Nishida et al., 2009), suggesting a unique role of REM sleep in processing emotional memories (Genzel et al., 2015).

The apparent inconsistency in REM sleep's role in declarative memory consolidation has been addressed by the Sleep Reinforcement and Refinement (SR2) hypothesis (MacDonald and Cote, 2021). This hypothesis proposes that NREM sleep primarily strengthens and stabilizes memories (reinforcement) by preserving the memory engram, thereby facilitating retrieval. In contrast, REM sleep is thought to refine memory by enhancing the accuracy of retrieval, selectively strengthening dominant memory traces (potentially marked by NREM reinforcement) and weakening less relevant or competing traces. This idea is supported by research showing that overnight retention of words is related to the integrity of NREM-REM sleep cycles (Ficca et al., 2000; Mazzoni et al., 1999), emphasizing the importance of a preserved sleep cycle for adaptive memory consolidation.

Additionally, REM sleep is thought to facilitate the gradual attenuation of the emotional intensity associated with emotional memories. This hypothesis is supported by evidence showing reduced amygdala reactivity to emotional experiences following REM sleep and the disruption of this effect when REM sleep quality is impaired (Gujar et al., 2011; Van Der Helm et al., 2011). Studies using night-half paradigms—where participants are selectively deprived of either the NREM-rich early part of the night or the REM-rich later part—further support the idea that emotional modulation of memories is particularly enhanced by REM-rich sleep during the latter half of the night. (Wagner et al., 2001; Wagner et al., 2002; Groch et al., 2013).

Thus, within the framework of the ”sleep to remember, sleep to forget” theory, NREM sleep is proposed to drive the specific consolidation of the factual content of emotional memories, an effect that may be further enhanced by REM sleep's selective strengthening of salient memories. At the same time, REM sleep is hypothesized to facilitate the gradual reduction of the affective tone initially associated with these memories, supporting emotional regulation over time.

At a functional level, during NREM sleep, hippocampal-cortical interactions are thought to promote memory consolidation through spontaneous hippocampal replay of prior waking experiences. This replay facilitates the selective strengthening and corticalization of reactivated memories (Paller et al., 2020). There is also evidence of replay during REM sleep, observed in both rodents (Louie and Wilson, 2001) and humans (Schönauer et al., 2017; M. E. A. Abdellahi, 2022). It has been hypothesized that during REM sleep, particularly during phasic REM, the replay of emotional memories sup-

ports their integration into existing associative networks (M. P. Walker and Stickgold, 2010; Simor et al., 2020) while concurrently reducing the arousal intensity initially associated with these memories (Wassing et al., 2019).

This theory is supported by functional evidence showing extensive limbic and paralimbic activation during REM sleep compared to wakefulness, especially heightened activity in the amygdala and hippocampus (Dang-Vu et al., 2010), potentially reflecting the reactivation of highly emotional memories. Additionally, REM sleep is marked by elevated theta power, which may facilitate enhanced communication between distributed neural networks (M. P. Walker, 2009). Coupled with the increase in cholinergic concentration during REM (Vazquez and Baghdoyan, 2001), these factors are thought to promote neural plasticity and, consequently, the consolidation of emotional memories (Power, 2004). As emotional memories become integrated within cortical networks, their emotional overlay gradually diminishes, potentially due to the reactivation of memories in a brain that is depleted of noradrenaline. Given that heightened noradrenaline levels are typically associated with stress (Pace-Schott and Hobson, 2002), this mechanism may facilitate the decoupling between the emotional memory and its corresponding physiological arousal state, allowing for emotional regulation over time.

1.5 TMR as a Tool to Study the Specific Sleep Stage Function in Emotional Memory Processing

One method employed to investigate the role of sleep in memory processing is Targeted Memory Reactivation (TMR, Rasch et al., 2007; X. Hu et al., 2020). In TMR, a sound or smell is paired with previously learned stimuli during wakefulness and reintroduced during subsequent sleep. This method biases the reactivations during sleep toward the learned memories and serves as an experimental tool to influence the endogenous replay of memories that naturally occurs during sleep (M. E. A. Abdellahi, 2022). TMR provides an opportunity to test the hypothesized functions of NREM and REM sleep by measuring behavioral changes induced by memory reactivation within specific sleep stages. Within the framework of the ”sleep to remember and sleep to forget” theory, TMR during NREM sleep is expected to enhance memory consolidation, thereby improving memory recall. Indeed, this memory-enhancing effect has been consistently observed when TMR is applied during NREM sleep (X. Hu et al., 2020), while this effect is less consistent and often absent when TMR is applied during REM sleep (Rasch et al., 2007; Sterpenich et al., 2007; Rihm & Rasch, 2015; X. Hu et al., 2020).Conversely, given the hypothesized role of REM sleep in progressively reducing the emotional intensity of memories, TMR during REM sleep is expected to decrease the subjective arousal associated with emotional memories. Some studies have provided support for this idea, showing that REM TMR can reduce arousal linked to conditioned stimuli and emotional declarative memories (Rihm and Rasch, 2015; Wassing et al., 2019; Hutchison et al., 2021). However, the results across studies have been inconsistent.

For example, Rihm and Rasch (2015) observed a decrease in affective tone for cues presented during both REM and NREM sleep, making it difficult to attribute the effect to a specific sleep stage. Additionally, because no uncued items were tested, it remains unclear whether the effect resulted from memory reactivation during both sleep stages or from the generalization of the effect from one stage to another. This is particularly relevant since REM sleep is thought to promote global memory reactivation and integration into associative networks. In fact, TMR during REM sleep has been associated with memory generalization (Sterpenich et al., 2014).

Similarly, Borghese et al. (2022) found no direct effect of TMR during REM sleep on arousal reduction, but they did observe a correlation between arousal decrease and the number of TMR stimulations, suggesting that TMR effectiveness may depend on multiple factors. For instance, TMR might not be effective at every stimulation, as demonstrated in a study by Abdellahi (2021) on procedural memory, where REM TMR did not lead to direct behavioral improvements. However, the number of detected memory reactivations, identified through machine learning classifiers, using multivariate pattern analysis (MVPA), was associated with enhanced procedural memory performance. This method of using classifiers to detect TMR-induced memory reactivation has been successfully applied in both NREM (Santamaria et al., 2024) and REM sleep (M. E. Abdellahi et al., 2021; M. E. Abdellahi et al., 2023). These findings suggest that while the behavioral effects of TMR may be subtle and difficult to detect in experimental settings, EEG-based detection of memory reactivation offers an additional and potentially more precise method for uncovering the neural mechanisms underlying the behavioral effects of TMR.

Although previous studies have provided results in line with the proposed functions of REM sleep in emotional regulation, the overall body of evidence supporting these hypotheses remains limited. In this study we aimed to investigate the distinct roles of REM and NREM sleep in emotional memory processing. Participants completed a memory task in which they associated neutral words with emotional images. We then assessed the effects of auditorily replaying the associated words (TMR) during REM and NREM sleep on their arousal and valence ratings of the cued images the next morning and four days later. These effects were compared to ratings for uncued images (i.e., images presented in the memory task during the TMR night but without auditory replay of the associated word while the participant was asleep) as well as to a control night in which no TMR was applied (No TMR night). To further explore the underlying neural mechanisms, we employed MVPA to identify whether TMR induced distinct EEG patterns related to the emotional components of reactivated memories. The implications of these findings are discussed in relation to the behavioral effects of REM and NREM TMR on changes in arousal and valence rating over time, offering insights into the differential roles of REM and NREM sleep in shaping emotional memory consolidation.

2 Methods

2.1 Inclusion Criteria and Recruitment

Sixty-four participants (45 female, 19 male, 0 non-binary), aged between 18 and 35 years (M = 23.59, SD = 2.97), were recruited for the study through the SONA database of the Donders Institute, as well as via social media and physical notice boards. Each participant's total involvement in the study spanned approximately four weeks. Selection criteria included good physical and mental health, as assessed by the General Health Questionnaire (developed in the lab, taking approximately 5 minutes), the Beck Depression Inventory (BDI, maximum score > 20 = excluded) (Beck et al., 1996), and the Beck Anxiety Inventory (BAI, maximum score > 15). Additionally, participants were only included if they had no history of or current sleep disorders, as assessed by the Pittsburgh Sleep Quality Index (PSQI, maximum score > 7) (Buysse et al., 1989) and the Munich Chronotype Questionnaire (MCTQ sleep) (Roenneberg, Wirz-Justice, and Merrow, 2003), and if their average sleep onset occurred between 21:00 and 1:00 during the week

Participants were excluded if they were taking any medication that might influence their sleep-wake cycle, consumed more than four cups of coffee per day, or had daily drug consumption. Pregnant or breastfeeding females were also excluded. Only participants who reported remembering their dreams at least twice a week, as measured by the MADRE questionnaire (Schredl et al., 2014), were selected to ensure sufficient dream recall during the study.

Finally, the study, including all associated questionnaires, was conducted in English. To maintain a consistent proficiency level in the language and to ensure that participants could recognize words for the word-picture association learning to a similar degree, we evaluated their English proficiency using the Boston Naming Test, with a minimum score requirement of 10/15 (Erdodi et al., 2017).

Participants expressing interest in joining the study underwent a 20-minute phone call orientation session, during which they received detailed information about the study procedures. Subsequently, an appointment for a 1.5-hour intake session was scheduled to assess their eligibility against the inclusion criteria by presenting them with the above-mentioned questionnaires in Castor EDC. Finally, a 20-minute structural T1 and T2 Magnetic Resonance Imaging (MRI) scan was performed on a Prisma or PrismaFit (3T).

During the intake session, participants were given a Fitbit Inspire 2 wristband to wear continuously throughout the study to monitor their sleep patterns. They were also instructed to complete a daily dream report via an email survey for the duration of the study. This process was designed to improve dream recall and ensure participants became accustomed to the dream reporting method used during the experimental nights.

2.2 Pre-Sleep Protocol and Monitoring Setup

On the days scheduled for lab-based sleep, participants were required to abstain from alcohol for 24 hours and from drug use for 48 hours prior to the study. Caffeine consumption was limited to no more than two cups of coffee in the morning, in line with their typical habits. Participants were also instructed to wake up before 08:00, a requirement verified through their sleep tracker data.

Upon arriving at the lab, participants prepared for bed and were then fitted with an EEG cap to monitor brain activity. EEG data were recorded using a 64-channel actiCAP and BrainAMP system by Brainproducts. Each electrode site was prepped with abrasive paste (Nuprep) and electrode gel (SuperVisc), with impedances checked to ensure they were below 20 kΩ. In addition to the EEG setup, two electrooculograms (EOG) tracked eye movements, three electromyogram (EMG) electrodes were placed on the chin to measure muscle activity, and two electrocardiogram (ECG) electrodes monitored heart rate. An optional electrogastrogram (EGG) with eight electrodes was also applied to track gastrointestinal activity, using BrainAMP ExG.

Data were recorded at a sampling rate of 500 Hz and referenced to the vertex. While the polysomnography setup was being applied, participants completed questionnaires tailored to each experimental night, focusing on their mood and sleep patterns.

2.3 Adaptation Night

To help participants adjust to sleeping in a lab environment while wearing an EEG cap and to improve sleep quality during the two experimental nights, an adaptation night was scheduled no more than seven days before the first experimental night. On the day of the adaptation night, participants arrived at the Donders EEG laboratory at 21:30. After the EEG setup was applied, participants completed a color-naming Stroop task and a trail-making test (TMT, 5 minutes) (Corrigan and Hinkeldey, 1987). Subsequently, at around 23:00 - 23:30, participants went to bed and were awakened at 7:00. Upon waking, they filled out a sleep quality questionnaire (SF-AR) (Görtelmeyer, 2011) and reported their dream recall. Participants who achieved at least 70% sleep efficiency during the designated sleep period and agreed to continue were included in the study.

2.4 Experimental Nights

The study employed a within-subjects design, conducted across two experimental nights. The two experimental sessions (with and without TMR) were scheduled at least 14 days apart, with random assignment and counterbalancing of the experimental night sessions as well as item categories used in the memory task (using random number generator samples in R) (R. S. Team, 2021). Participants were blinded to the condition. Participants arrived at the EEG lab at 19:30. After the EEG was applied, they completed the seven learning blocks of the memory task described below.

2.4.1 Memory Task

The memory task (Figure 2) involved 99 word-picture associations, with an additional pair to control for the primacy effect —the tendency to better remember items presented at the beginning of a list. Words were sourced from the Auditory English Lexicon (AELP) project (Goh, Yap, and Chee, 2020). They were selected based on the following criteria: two syllables, similar durations (636–805 ms), neutral valence, arousal levels between 4 and 6, recognition rates exceeding 88%, and no reference to image categories. Word-picture associations were randomized but remained consistent across all participants.

The images were divided into six categories (three per experimental night): mammals, vehicles, food, children, water, and buildings. Each category contained 11 positive, 11 negative, and 11 neutral images, sourced from databases such as NAPS (Marchewka et al., 2014), IAPS (Lang et al., 2005), NDPS (Merlhiot et al., 2018), DIRTI (Haberkamp et al., 2017), and Oasis (Kurdi et al., 2017). To complement the dataset, 55 Creative Commons-licensed images were sourced from platforms like Unsplash, Flickr, Pixahive, Wikipedia, Stocksnap, and Pxhere. The images underwent evaluation by 16 pilot participants to ensure suitability. Final selection was based on the following criteria: valence ratings (> 5.75 for positive, 4.25–5.75 for neutral, < 4.25 for negative), exclusion of adjacent or similar categories (e.g., no humans for children or other animals for mammals, flagged by at least three participants), and image quality (rated higher than 6 on a 0–9 scale). If more images met the criteria than required, selections were based on factors such as the lowest standard deviation for valence and arousal ratings, similarity to original database ratings, and uniqueness (e.g., avoiding repetition of the same mammal).

The memory task, conducted using Psychopy (Peirce et al., 2019), consisted of seven blocks. The first block involved rating neutral words presented auditorily. Participants rated the valence (1 = very negative, 4 = neutral, 7 = very positive) and arousal (1 = calm/low arousal, 7 = agitated/high arousal) of each word. Words were presented for 1000 ms, followed by the rating screen. In the second block, participants viewed images for 2000 ms on a computer screen and rated them for valence and arousal using the same scales. The third block initiated the learning phase. Participants first heard a word. After a 1000 ms pause, the corresponding image was shown for 1000 ms, and the word was repeated while the image remained on the screen for an additional 1500 ms. After half of the trials, participants could take a break of freely determined duration. Blocks 4 and 5 involved a combination of recall and learning in a different word-image order. Participants heard previously presented words and rated the associated picture for valence and arousal as quickly as possible using the same scales as earlier blocks. They were instructed to favor intuitive responses. After rating, the corresponding image was displayed for 1500 ms to reinforce the association. A break of freely determined duration was provided midway through the trials. After these five blocks, participants took a 10-minute break to complete questionnaires, including MDBF (Steyer et al., 1994) and SF-A/R (Görtelmeyer, 2011) for the previous night in Castor EDC. They then completed two memory recall blocks. In Block

6, participants were again presented with words auditorily and asked to rate the associated picture for valence and arousal, relying on intuition without considering previous ratings. They also rated their certainty of remembering the picture on a 1–7 scale (1 = completely uncertain, 7 = completely certain). Finally, in Block 7, participants heard a word and indicated whether they remembered the associated picture. If they did, they described it using 3–5 keywords via a computer keyboard.

Participants went to bed after completing the task (between 22:30 and 23:30). While in bed, resting-state EEG measurements were conducted, including 1.5 minutes with eyes open, 1.5 minutes with eyes closed, repeated twice.

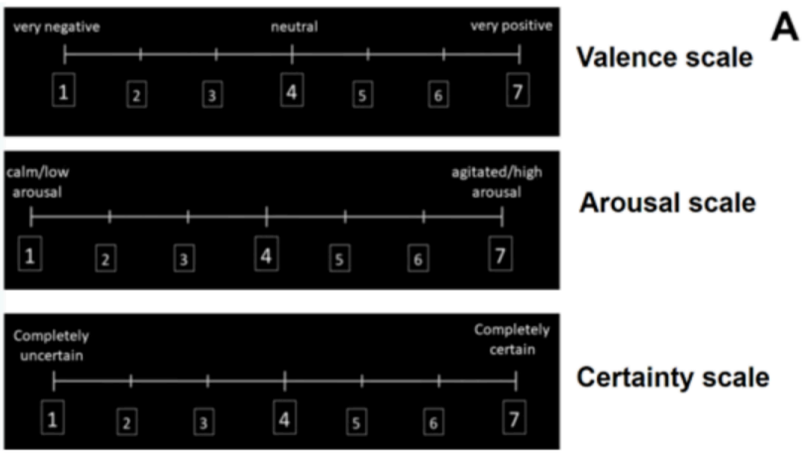


Figure 2: **Evening Memory Task**A) Example of the valence, arousal, and certainty scales presented during the memory task. B) Illustration of the seven blocks of the evening memory task. In Blocks 1 and 2, participants rated the valence and arousal of the auditorily presented words and visually presented images, respectively. In Block 3, the words and their associated images were presented together for the first time. Participants had to learn these word-image pairs, which remained constant throughout the session. In Blocks 4 and 5, participants were again presented with the words, followed by the associated images, providing an opportunity to learn the word-image associations. In Block 6, only the word was presented, and participants rated the arousal and valence of the associated image, as well as how certain they were about remembering it. In Block 7, participants described the image associated with each auditorily presented word using a.

2.4.2 Awakening Protocol Experimental Night

2.4.2.1 No TMR Night

Throughout the night, the investigators continuously monitored the participants' EEG activity. Participants were awakened up to eight times: four times during NREM sleep and four times during REM sleep, with awakenings occurring after 15 minutes into each sleep stage. The awakening procedure was as follows:

Using an intercom system, the researcher gently woke the participants. Upon waking, they were prompted to respond orally to a series of questions. Initially, participants were asked if they recalled any thoughts or imagery experienced during sleep and were encouraged to describe them aloud. Additional questions about the dream content followed. After completing the questions, participants were allowed to fall back asleep. Each awakening lasted approximately five minutes.

2.4.2.2 TMR Night

During this session, the investigators continuously monitored the EEG while participants were asleep. Once a stable sleep period of at least three minutes was observed in NREM or REM stages, researchers introduced auditory cues—specific words from word-image paired associations learned earlier. These cues were played through two loudspeakers placed near the participant's head for up to 12 minutes. For each sleep stage, words from a specific image category were randomly presented to each participant. These words were delivered at intervals ranging between 8,000 to 8,200 milliseconds, resulting in 99 to 104 word presentations depending on word length.

To determine the optimal volume, the audio started at 30 dB and was gradually increased until the researcher detected a K-complex or a significant change in EEG activity. Audio playback stopped if participants showed signs of arousal such as body movements, heightened respiration, mumbling or transitioned to a different sleep stage. After 15 minutes into one sleep stage, participants were then awakened, within 10 to 30 seconds after the last TMR cue, following the same awakening protocol as in the No TMR Night. The TMR and awakening were conducted four times in NREM sleep and four times in REM sleep.

2.4.3 Morning Routine

The sleep session concluded at 7:00 Participants first provided a free-form dream report, followed by a second report prompted by words visually presented to them, which were related to their dreams. These reports were additionally documented using the same questionnaire scales that were employed during the night awakenings. Afterward, participants completed a sleep experience questionnaire (adapted SF-A/R; Görtelmeyer, 2011), which inquired about their awareness of any words heard during the night and any spontaneous awakenings unrelated to the experiment.

Following this, participants repeated the recall tasks (Blocks 6 and 7) exactly as they had done the previous night. Next, they performed a localizer task in which they rated 67 new images from the task categories in three phases: valence, arousal, and prototypicality. Once the task was completed, electrodes were removed, and the study concluded between 8:00 and 8:30.

2.4.4 Follow-Up

Four days after each experimental session, participants underwent a follow-up memory recall assessment using the same recall blocks (Blocks 6 and 7). This follow-up was conducted online via Pavlovia, adapted from the original Psychopy experiment used in the lab. Participants received an email with a link to complete the memory recall session and were instructed to finish it within a single sitting, between noon and midnight on the 4th day following the experimental session.

2.5 Analysis

Baseline analysis focused on participants' ratings of images from Block 2 of the evening memory task. To evaluate changes in emotional ratings over time, further analysis was based on image ratings from Blocks 6 and 7, collected during evening, morning, and follow-up sessions. Additionally, EEG recordings taken during TMR application were analyzed to examine the neurophysiological effects of TMR. All statistical analyses were conducted in RStudio (RStudio Team, 2021).

2.5.1 Behavioral analysis

Memory accuracy of the images was evaluated by a single rater using a 5-point scale, based on the following criteria:

Rate 1: The description is entirely incorrect, referring to a completely different image.

Rate 2: The description is partially correct, capturing some elements of the image but also including incorrect details.

Rate 3: The description correctly identifies the general category (e.g., Animals, Children, Transportation/Vehicles, Buildings, Food, Water) but lacks specific details (e.g., "animal" without specifying which animal).

Rate 4: The description is mostly accurate but lacks sufficient detail to identify the exact image (e.g., "kitten" without more context).

Rate 5: The description is fully accurate, with enough detail to clearly identify the specific image.

In this thesis, images rated 4 or 5 were classified as "remembered," while those rated 3 or below were classified as "forgotten."

Behavioral analysis focused on arousal and valence ratings of images among 64 participants. On the TMR night, 61 participants completed the follow-up memory recall. Due to the need for data formatting and verification before analysis, and with priority given to TMR night data, 57 out of the 64 participants had usable data for the No TMR night, with 51 completing the follow-up recall.

First, we evaluated baseline differences in valence and arousal ratings of the images. Since participants were exposed to all images for the first time during Block 2 of the memory task (conducted in the evening), ratings from this block served as the baseline. Subsequent comparisons between TMR conditions and valence categories were made relative to these initial ratings.

Given that the images were selected to be approximately one-third positive, one-third negative, and one-third neutral on average, we accounted for individual differences in valence interpretation. The valence of each image was determined based on the participant's mean valence rating and half of the standard deviation of valence ratings across all images in Block 2. Images rated more than half the

standard deviation below or above the mean were classified as negative or positive respectively. Images that did not meet these criteria were assigned a neutral valence. The consistent distribution of images across valence categories for each TMR condition was verified using a Chi-square test of independence to ensure balanced representation.

Valence rating was recorded as valence rating extremity, defined as the absolute difference between a participant's mean valence rating across all images in Block 2 and the rating given to each specific image.

All analyses on arousal and valence ratings extremity were conducted using linear mixed models (Bates, 2014). Participants were included as random intercepts in the models to account for individual differences in baseline ratings. This approach controlled for variability in how participants perceived and rated stimuli, allowing the analysis to focus on the effects of valence categories and TMR conditions without being biased by personal rating tendencies. Valence and TMR condition variables were dummy-coded, with positive and negative valence compared against neutral valence, and REM, NREM, and uncued (during the TMR night) compared against the No TMR night.

Baseline differences in arousal rating and valence rating extremity between valence categories (Arousal rating Valence; Valence rating extremity Valence) and TMR conditions (Arousal rating TMR condition; Valence rating extremity TMR condition) were analyzed, with post hoc pairwise comparisons performed to evaluate differences between valence categories and TMR conditions.

We first assessed the main effect and interaction effect of time, TMR condition and images' initial valence, as rated by participants during Block 2 of the evening memory task on changes in arousal rating and valence rating extremity (Arousal rating Time * TMR condition * Valence; Valence rating extremity Time * TMR condition * Valence). For this and subsequent analyses, "Time" was dummy-coded, with morning and follow-up session ratings compared separately against evening ratings. Finally, we tested whether the effects of Time, TMR, and Valence depended on whether the participant remembered the image associated with the auditory cue in the morning (Arousal rating Time * TMR condition * Valence * Memory status; Valence rating extremity Time * TMR condition * Valence * Memory status). This analysis was motivated by the hypothesis that if the image associated with the word cue was remembered in the morning, it could influence the effectiveness of TMR. The memory status of the word-image pair was dummy-coded to compare the effects of remembering versus forgetting. Additionally, we conducted six post-hoc linear mixed model analyses, focusing separately on subsets of positive and negative images, without comparing them to neutral images (Arousal rating Time * TMR condition * Memory status; Valence rating extremity Time * TMR condition; Valence rating extremity Time * TMR condition * Memory status). This approach was based on the assumption that emotional memory processing operates through distinct mechanisms depending on the valence of the memory. Consequently, comparing changes in arousal ratings and valence rating extremity relative to neutral images might obscure the specific effects of TMR on positive and negative images.

2.5.2 EEG Analysis

2.5.2.1 EEG Data Conformity

To expedite EEG cleaning and ensure the accuracy of experimental trigger timings, all EEG recordings from each experimental night were appended in chronological order. Trigger timings were reviewed, and any missing triggers were added based on notes taken by the responsible researcher during the experimental sessions. These notes, recorded while completing the awakening questionnaires with participants, were documented in Castor Electronic Data Capture (Castor EDC, n.d.).

2.5.2.2 Sleep Quality Analysis

We first examined whether TMR influenced sleep structure. The sleep quality analysis included 44 participants on the TMR night and 37 participants on the No TMR night. Due to time constraints, not all participants from the behavioral analysis were included in the sleep scoring analysis. Our objective was to determine if TMR impacted sleep duration and the percentages of time spent in Wake, NREM1, NREM2, NREM3, and REM sleep. Sleep stages were scored using U-Sleep (Perslev et al., 2021) on the raw EEG recordings, based on 6 EEG and the two EOG channels. A Shapiro-Wilk test revealed that sleep duration and NREM1 percentage were not normally distributed on the No TMR night. To ensure consistency, we used the Wilcoxon test to assess differences across all sleep variables between the TMR and No TMR nights.

2.5.2.3 EEG Pre-processing

To clean the EEG data from the entire night, we used the High-density Sleep Cleaner, a semi-automatic artifact removal tool (Leach et al., 2023). This tool identified and removed artifacts from individual 10-second epochs by detecting outliers based on several Signal Quality Metrics (SQMs), including delta power, beta power, and the maximum squared deviation from the mean EEG signal. These SQMs were calculated from both robustly z-standardized and raw EEG data, with analyses performed on both the original and mean-referenced signals.

Out of the 64 EEG channels, TP9 and TP10 were excluded from further analysis due to suspected signal loss during the night, likely related to their positioning. To analyze the specific EEG correlates induced by TMR across different sleep stages, we selected continuous EEG segments labeled as REM and NREM2. This selection was guided by our objective to differentiate the effects of TMR on emotional memory processing in these stages. We chose to select only NREM2 and not NREM3 segments because the differing neurophysiological EEG correlates (Carskadon and Dement, 2005) and the potentially distinct functions of NREM3 could complicate further MVPA analysis. Additionally, we prioritized NREM2 as this stage spans a longer portion of sleep, providing a greater number of TMR trials for analysis.

For the continuous REM and NREM2 sleep segments, linear trends were removed using a first-order

polynomial applied across the 62 remaining channels. A band-pass filter between 1 and 45 Hz was applied to retain relevant EEG frequencies and reduce noise. The data were then re-referenced by averaging across all channels to enhance signal quality. Within these continuous REM and NREM2 segments, channels containing artifacts were excluded based on the High-density Sleep Cleaner results. For each 10-second epoch, channels with artifacts that had at least two neighboring artifact-free channels were interpolated using triangulation of neighboring channels. Epochs where interpolation was not possible were marked as NaN, and the data were segmented accordingly.

The cleaned segments were further processed using automated detection of artifactual Independent Component Analysis (ICA) components from the EEGLAB toolbox (Delorme and Makeig, 2004). Components identified as eye, muscle, heart, or line noise artifacts with 90% or greater accuracy were subsequently rejected from the analysis. Trials were segmented from −0.5 s to 3 s relative to the onset of the cue and undersampled to 200 Hz to expedite future multivariate pattern analysis (Ashton et al., 2022).

2.5.2.4 MVPA Analysis

We employed machine learning classifiers to determine whether the targeted reactivation of emotional memories during REM and NREM2 sleep could produce EEG patterns that differentiate between memories with varying levels of arousal and valence.

Our goal was to classify sleep EEG patterns from −0.5 to 3 seconds following auditory cues of remembered word-image associations across several emotional conditions: high arousal vs. low arousal, negative vs. positive, negative vs. neutral, and positive vs. neutral.

To ensure that the EEG trials used in the analysis corresponded to word-image pairs retained in memory throughout the night, classification of each EEG trial was based on the participant's memory recall conducted the following morning. Specifically, only word-image pairs that were remembered (score ≥ 4 in Block 7) were included.

For each word-image pair, classification into valence and arousal categories was based on the participant's ratings from Block 6 of the morning memory recall. As in the behavioral analysis, the valence of each image was determined using the participant's mean valence rating and half the standard deviation of valence rating across all images in Block 2. Images were categorized as positive, negative, or neutral according to the same mean and standard deviation criteria used in the behavioral analysis.

Similarly, arousal levels were determined using the participant's mean arousal rating and half the standard deviation of arousal rating across all images in Block 2. Images that were more than half the standard deviation above or lower than the mean arousal rating were defined as high or low arousal respectively. Images not meeting these criteria were classified as middle arousal and excluded from further analysis.

2.5.2.5 MVPA Analysis Across Participants

Due to the necessity of removing a significant number of trials where multiple channels contained artifacts—common in sleep EEG recordings—we first focused on classifying sleep EEG patterns across all participants. To remain consistent with the rest of the analysis for each classification analysis, only participants with trials available for both of the compared conditions were included. As a result, each analysis included between 11 and 14 participants, with 217 to 556 trials per class in each classification.

Single-trial data were classified over time using the MVPA-Light toolbox (Treder, 2020), with sleep ERP values from 62 EEG channels used as features. Classification was performed using a support vector machine (SVM) with a linear kernel. We employed a 5-fold cross-validation method with two repetitions. To avoid bias from imbalanced classes, surplus trials from the larger class were randomly excluded to ensure equal numbers of trials in both classes. Data within each fold were then z-scored to prevent bias, and we applied principal component analysis (PCA) to reduce dimensionality (n = 20). To increase the signal-to-noise ratio, trials were randomly averaged in groups of 4. Averaging by groups of 4 was chosen as a compromise to reduce noise while avoiding excessive smoothing, which could obscure subtle patterns important for classification (M. E. Abdellahi et al., 2021). Classifier performance was evaluated using the area under the curve (AUC) metric, which reflects the trade-off between true positive and false positive rates. To assess the significance of the classification, class labels were shuffled and the analysis was rerun 500 times. Cluster-based correction for multiple comparisons was applied (Maris and Oostenveld, 2007). Since memory reactivation of emotional tone might not be time-locked across trials and participants, we conducted a similar analysis using sliding time windows of 10 ms, 30 ms, 50 ms, and 70 ms, so that each time point represented the average of the preceding half and the following half of the time window.

Additionally, we conducted the same analysis but selected only trials with high variance. This approach was motivated by the likelihood that TMR might not be effective at every cueing, making high-variance trials more likely to reflect significant signal variations. Variance for each EEG trial was calculated across all channels and time points. Trials with variance exceeding the average variance across all trials for the two emotional categories being compared were selected for further analysis. We calculated the mean variance and standard deviation across both emotional conditions to ensure that if endogenous variance differences between emotional conditions were present, the MVPA would not be solely influenced by those differences. As a result, each classification had between 89 and 209 trials per class. The same classification method previously described was applied to single high-variance trial data. Additionally, the classification analysis was conducted on randomly averaged high-variance trials in groups of 4 to increase the signal-to-noise ratio.

Finally, we conducted the classification analysis on individual participants' trials using the same methodology as in the across-participant analysis. However, due to the smaller number of trials available per participant, we did not limit the analysis to high-variance trials and did not randomly average trials in groups of 4. Additionally, we increased the number of repetitions in the 5-fold cross-validation from

2 to 7. All other aspects of the analysis remained consistent.

3 Results

3.1 Sleep Quality

We first analyzed whether sleep quality was affected by TMR to ensure that differences between the TMR night and the No TMR night were not due to variations in sleep stage composition. To do this, we compared sleep duration and the percentage of time spent in Wake, NREM1, NREM2, NREM3, and REM between the two nights (Table 1).

The analysis revealed no significant differences in sleep duration or sleep structure when comparing the TMR night and the No TMR night (sleep duration: $W = 621$, $p = 0.07$; percentage in Wake: $W = 930$, $p = 0.28$; percentage in NREM1: $W = 883$, $p = 0.52$; percentage in NREM2: $W = 687$, $p = 0.23$; percentage in NREM3: $W = 755$, $p = 0.58$; percentage in REM: $W = 748$, $p = 0.54$). While a marginal trend was observed for sleep duration, overall, the sleep parameters were comparable across nights. These findings suggest that the differences in emotional processing observed between the nights are unlikely to be fully explained by variations in sleep stage composition and may instead reflect the effects of TMR. Additionally, although not statistically significant, sleep duration was descriptively longer on the TMR night compared to the No TMR night, indicating that TMR did not disrupt sleep by increasing awakenings.

Table 1: Sleep Statistics (Mean ± SD) and P-values by Night Condition.

	Sleep Duration	Wake%	N1%	N2%	N3%	REM%
NOTMR	443.93 ± 56	21.2 ± 12.27	6.69 ± 3.16	43.42 ± 8.15	14.36 ± 7.41	13.33 ± 5.82
TMR	461.45 ± 26.17	17.66 ± 9.24	6.16 ± 2.75	45.82 ± 7.87	15.52 ± 7.1	14.22 ± 5.13
Stat. Results	0.07	0.28	0.52	0.23	0.58	0.54

Note: Statistic description of sleep duration and sleep stages composition in function of the Night. Note that no significant difference was observed even though a trend for a shorter sleep duration in the No TMR Night appeared.

3.2 Behavioral Analysis of Emotional Processing

We proceeded to analyze whether the TMR conditions (No TMR night, NREM, REM, uncued during the TMR night) affected emotional processing over time. The analysis focused on arousal rating and valence rating extremity.

3.2.0.1 Baseline Differences We first examined whether arousal rating and valence rating extremity differed based on the initial valence of the images, while accounting for individual variations in valence interpretation.

A linear mixed model assessed the influence of valence (positive and negative) relative to neutral on arousal rating. Results indicated that negative images ($\beta = 1.41$, SE = 0.04, $t(12531.36) = 39.96$, $p < .001$) were rated significantly higher in arousal compared to neutral images, while positive images ($\beta = 0.05$, SE = 0.04, $t(12541.41) = 1.41$, $p = .16$) did not differ significantly from neutral. Furthermore, pairwise comparisons revealed that negative images were rated as significantly more arousing than positive images ($\beta = 1.36$, SE = 0.03, $Z = 39.84$, $p < .001$).

We then investigated differences in valence rating extremity as a function of valence at baseline. Pairwise comparisons indicated that negative images were rated with higher valence extremity than positive images ($\beta = 0.211$, SE = 0.02, $z = 16.67$, $p < .001$).

We also examined whether there were differences in arousal rating and valence rating extremity when participants first saw the images in Block 2 (evening) across the four TMR conditions. To ensure the proportionality of images per valence category and TMR condition was maintained, a Chi-square test of independence was conducted. There was no significant difference in the proportion of valence categories based on TMR condition ($\chi^2(6, N = 12,57) = 12.45$, $p = .053$).

linear mixed-effects model was used to assess statistical differences in baseline arousal rating. Arousal rating at baseline was significantly lower for images presented during the No TMR night compared to uncued images ($\beta = 0.20$, SE = 0.04, $t(12534) = 4.60$, $p < .001$) and images cued during REM ($\beta = 0.12$, SE = 0.04, $t(12534) = 2.73$, $p = .01$). However, arousal rating between images presented during the No TMR night and those cued during NREM did not differ significantly at baseline ($\beta = -0.004$, SE = 0.04, $t(12534) = -0.10$, $p = .92$). Further pairwise comparisons revealed that images cued during NREM were rated significantly less arousing than uncued images ($\beta = 0.20$, SE = 0.05, $z = 3.94$, $p < .001$). No other comparisons were significant ($p \geq 0.082$).

Finally, we analyzed baseline differences in valence rating extremity between the four TMR conditions using a linear mixed-effects model. Images presented during the No TMR night were rated significantly less extreme in valence than uncued images at baseline ($\beta = 0.07$, SE = 0.02, $t(12532.78) = 3.03$, $p = .002$). Additionally, pairwise comparisons showed that valence rating of images from the REM condition was significantly less extreme compared to uncued images ($\beta -0.09$, SE = 0.03, $t(6564.00) = -3.16$, $p = .002$). No other comparisons were significant ($p \geq 0.155$).

Baseline analysis indicated some differences in arousal rating and valence rating extremity between the TMR conditions. However, since the subsequent analysis focused on changes over time, these initial baseline differences are unlikely to significantly impact the overall findings.

3.3 Arousal change analysis

3.3.0.1 Arousal Rating as a Function of Time, TMR, and Valence Using linear mixed-effects model, we begin by examining how arousal rating changed across time and how this pattern depended upon the TMR condition and the initial valence of the image. During the No TMR night, arousal rating of neutral images did not significantly change from evening to morning ($\beta = 0.04$, SE = 0.04, $t(36.65) = 0.99$, $p = .32$), but significantly increased from evening to the follow-up session ($\beta = 0.12$, SE = 0.04, $t(36.65) = 3.15$, $p = .002$, Figure 2). In contrast, a significant negative interaction between time and negative valence indicated that arousal rating of negative images presented during the No TMR night significantly decreased compared to the change in arousal of the neutral images, both from evening to morning ($\beta = -0.28$, SE = 0.061, $t(36.65) = -4.52$, $p < .001$) and from evening to the follow-up session ($\beta = -0.39$, SE = 0.06, $t(36.65) = -6.19$, $p < .001$, Figure 2).

No significant interaction was found between time and positive valence ($p \geq .121$), suggesting that during the No TMR night, arousal rating over time for neutral and positive images did not differ. However, a significant positive three-way interaction between time, valence, and TMR was observed for positive images cued during REM sleep, specifically when comparing the change in arousal from evening to morning ($\beta = 0.25$, SE = 0.11, $t(36.65) = 2.23$, $p = .03$). This suggests that cueing in REM sleep led to a relative increase in arousal rating for positive images compared to the No TMR night. No other three-way interactions reached significance ($p \geq .06$, Figure 2), indicating that changes in arousal as a function of the images' valence in the NREM and uncued TMR conditions did not differ from those in the No TMR night.

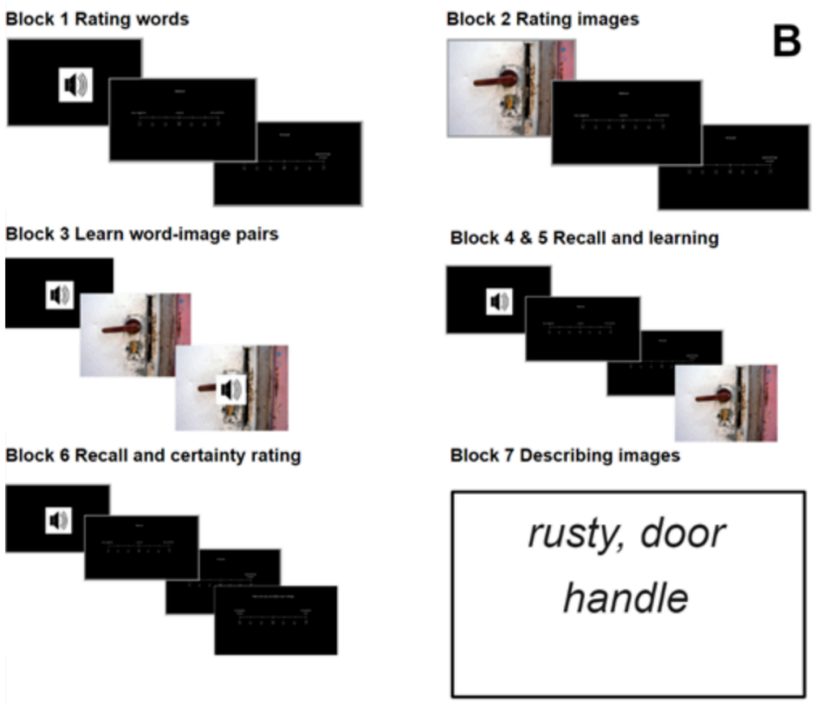


Figure 3: **Distribution of change in arousal rating as a function of valence and TMR condition** Differences in arousal rating between morning and evening and between follow up and evening are plotted as a function of the images' valence for each TMR condition. * Indicates significant comparisons with $p < 0.05$. Means and standard error of the mean (SEM) are shown. The colors of the comparison lines indicate which TMR conditions are compared to the No TMR night. Note that, other than for positive REM-cued images in the morning, there were no significant differences in changes in arousal ratings between TMR conditions.

3.3.0.2 Arousal rating as a function of time, TMR, Valence, and memory We then analyzed whether arousal rating changed as a function of time, TMR condition, valence, and memory of the word-image pair in the morning.

During the No TMR night, changes in arousal rating for neutral images were consistent with the analysis across memory conditions. From evening to morning ($\beta = 0.05$, SE = 0.05, $t(36.61) = 1.19$, $p = .24$) and from evening to the follow-up session ($\beta = 0.11$, SE = 0.05, $t(36.61) = 2.37$, $p = .02$), these changes did not depend on memory status (evening to morning: $\beta = -0.05$, SE = 0.08, $t(36.61) = -0.64$, $p = .52$; evening to follow-up: $\beta = 0.03$, SE = 0.08, $t(36.61) = 0.33$, $p = .74$, Figure 3). Cueing in NREM did not significantly affect arousal rating for forgotten neutral images from evening to morning compared to the No TMR night ($\beta = -0.15$, SE = 0.09, $t(36.61) = -1.73$, $p = .08$). However, the decrease in arousal from evening to follow-up approached significance ($\beta = -0.18$, SE = 0.09, $t(36.61) = -1.96$, $p = .05$), independently of memory ($p \geq .15$) (Figure 3). No other TMR condition had an interaction effect on changes in arousal rating for neutral images.

Similarly, to the analysis conducted across memory status, for positive images during the No TMR night, changes in arousal did not differ from neutral images, regardless of memory status ($p \geq .11$ for

time; $p \geq .07$ for memory).

In contrast, changes in arousal in the No TMR night for negative images depended on memory. Forgotten negative images showed a significant decrease in arousal compared to neutral images, both from evening to morning ($\beta = -0.41$, SE = 0.08, $t(36.61) = -4.95$, $p < .001$) and from evening to the follow-up session ($\beta = -0.20$, SE = 0.08, $t(36.61) = -2.31$, $p = .02$) (Figure 3). This effect was influenced by memory, with a significant three-way interaction between time, valence, and memory for remembered negative images. Specifically, the interaction was positive from evening to morning ($\beta = 0.27$, SE = 0.12, $t(36.61) = 2.21$, $p = .03$) and negative from evening to follow-up ($\beta = -0.39$, SE = 0.13, $t(36.61) = -3.14$, $p = .002$). This suggests that the decrease in arousal for remembered negative images over time was moderated in the morning but became more pronounced by the follow-up in comparison to forgotten negative images.

When accounting for memory status and analyzing the effect of time and valence on arousal ratings, multiple interactions with the TMR condition were observed. Notably, there was a significant positive three-way interaction between time, TMR, and valence for positive images cued in NREM, indicating that NREM cueing enhanced arousal for positive images over time compared to neutral images presented during the No TMR night, both from evening to morning ($\beta = 0.36$, SE = 0.15, $t(36.61) = 2.38$, $p = .02$) and from evening to follow-up ($\beta = 0.39$, SE = 0.16, $t(36.61) = 2.52$, $p = .01$). This effect depended on memory, with a significant negative four-way interaction for remembered positive images from evening to follow-up ($\beta = -0.51$, SE = 0.23, $t(36.61) = -2.22$, $p = .03$), suggesting that the increase in arousal observed at follow-up for positive images cued in NREM was dampened for remembered images.

For images cued in REM, there was an increase in arousal ratings for positive images compared to neutral images during the No TMR night, both from evening to morning ($\beta = 0.41$, SE = 0.15, $t(36.61) = 2.68$, $p = .007$) and from evening to follow-up ($\beta = 0.41$, SE = 0.16, $t(36.61) = 2.64$, $p = .008$), independent of memory ($p \geq .07$, Figure 3).

Uncued positive images during the TMR night also showed a significant increase in arousal from evening to morning compared to neutral images during the No TMR night ($\beta = 0.31$, SE = 0.15, $t(36.61) = 2.07$, $p = .04$), but this effect did not reach significance from evening to follow-up ($\beta = 0.26$, SE = 0.16, $t(36.61) = 1.68$, $p = .09$). This effect was memory-independent from evening to morning ($p = .23$) but dampened for remembered images from evening to follow-up, with a significant negative four-way interaction ($\beta = -0.48$, SE = 0.23, $t(36.61) = -2.08$, $p = .04$).

Two post-hoc linear mixed model analyses were conducted separately for positive and negative images, without comparing them to neutral images. The post-hoc analysis for positive images indicated that NREM cueing significantly increased arousal between evening and follow-up ($\beta = 0.25$, SE = 0.12, $t(11.21) = 2.06$, $p = .04$). However, contrary to the analysis relative to neutral images, NREM cueing did not affect arousal changes between evening and morning for negative images, regardless of

memory (p = .08, Figure 3). Regarding the results for REM and uncued TMR conditions, they were similar with those from the more complex model that included valence categories. Specifically, REM cueing significantly increased arousal from evening to morning (β = 0.39, SE = 0.12, t(11.21) = 3.28, p = .001) and from evening to follow-up (β = 0.35, SE = 0.12, t(11.21) = 2.85, p = .004). However, this effect was smaller for remembered positive images (β = -0.35, SE = 0.17, t(11.21) = -2.08, p = .04) (Figure 3). Similarly, uncued positive images during the TMR night showed an increase in arousal from evening to morning compared to the No TMR night (β = 0.38, SE = 0.12, t(11.21) = 3.29, p = .001), but the effect was smaller for remembered images (β = -0.33, SE = 0.16, t(11.21) = -2.08, p = .04). There were no significant differences in arousal ratings between evening and follow-up for uncued images.

When focusing on the effect of TMR on negative images, a late onset effect of TMR in the follow-up on arousal ratings was revealed (Figure 3). For negative images presented during the TMR night, changes in arousal ratings from evening to morning did not differ significantly from those presented during the No TMR night (NREM: β = 0.10, SE = 0.15, t(9.42) = 0.70, p = .48; REM: β = 0.24, SE = 0.14, t(9.42) = 1.67, p = .09; uncued: β = -0.02, SE = 0.15, t(9.42) = -0.12, p = .90). This pattern was observed regardless of memory retention (NREM: β = -0.23, SE = 0.20, t(9.42) = -1.14, p = .26; REM: β = -0.26, SE = 0.20, t(9.42) = -1.30, p = .19; uncued: β = 0.06, SE = 0.20, t(9.42) = 0.31, p = .76). However, arousal ratings for forgotten images presented during the TMR Night decreased more significantly from evening to follow-up, regardless of whether the images were cued in NREM (β = -0.32, SE = 0.15, t(9.42) = -2.12, p = .03), REM (β = -0.35, SE = 0.15, t(9.42) = -2.40, p = .02), or were uncued (β = -0.42, SE = 0.15, t(9.42) = -2.77, p = .006). This effect was independent of memory for images cued in NREM (β = 0.38, SE = 0.21, t(9.42) = 1.81, p = .07) and REM (β = 0.36, SE = 0.20, t(9.42) = 1.76, p = .08). For uncued images, however, the effect was dampened for remembered images (β = 0.47, SE = 0.20, t(9.42) = 2.29, p = .02).

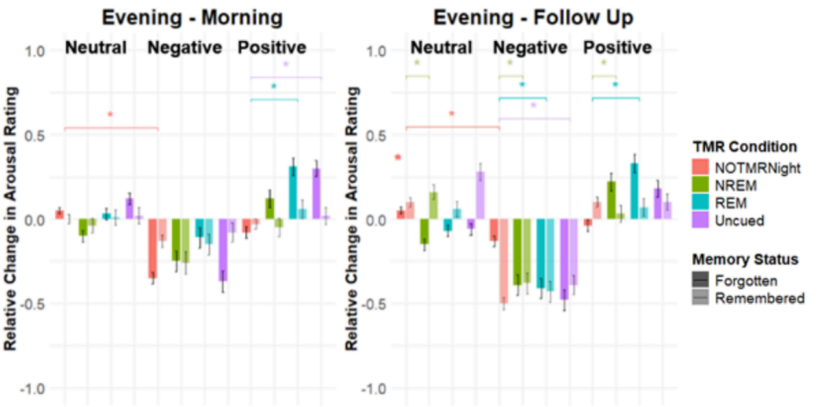


Figure 4: **Distribution of relative change in arousal rating between evening and morning and between evening and follow-up, analyzed by TMR condition, valence, and memory**Differences in arousal rating between morning and evening and between follow up and evening are plotted as a function of the images' valence for each TMR condition. Memory Statues indicate whether the image was remembered or not in the morning. * Indicates significant comparison, $p < .05$. Means and SEM are shown. The colors of the comparison lines indicate which TMR conditions are compared to the No TMR night.

3.4 Valence change analysis

3.4.0.1 Valence rating extremity as a function of Time, TMR and Valence In the next step, we proceeded to analyze changes in valence rating extremity as a function of time and how this pattern changed in function of the TMR condition and valence.

In the No TMR night, there was a significant increase in valence rating extremity for neutral images from evening to morning ($\beta = 0.30$, SE = 0.02, $t(36.65) = 15.79$, $p < .001$) and from evening to the follow-up ($\beta = 0.30$, SE = 0.02, $t(36.65) = 15.18$, $p < .001$, Figure 4). Conversely, a significant negative interaction between time and valence was observed for positive and negative images, indicating that valence extremity from evening to morning was decreased for positive and negative images compared to neutral images (positive: $\beta = -0.62$, SE = 0.03, $t(36.65) = -20.45$, $p < .001$; negative: $\beta = -0.64$, SE = 0.03, $t(36.65) = -20.32$, $p < .001$). This effect persisted from evening to the follow-up (positive: $\beta = -0.83$, SE = 0.03, $t(36.65) = -26.26$, $p < .001$; negative: $\beta = -0.88$, SE = 0.03, $t(36.65) = -27.24$, $p < .001$).

In addition, we tested the effect of time and TMR condition on valence rating extremity for the subset of positive and negative images to examine the direct impact of time and TMR condition on these specific valence categories (using the ratings of images from each valence category in the evening as the reference level). Valence rating extremity of positive and negative images was significantly lower in the morning compared to the evening (positive: $\beta = -0.32$, SE = 0.03, $t(11.22) = -12.33$, $p < .001$; negative: $\beta = -0.34$, SE = 0.03, $t(9.43) = -11.63$, $p < .001$) and in the follow-up compared to the evening (positive: $\beta = -0.54$, SE = 0.03, $t(11.22) = -19.99$, $p < .001$; negative: $\beta = -0.59$, SE = 0.03,

t(9.44) = -19.69, p < .001, respectively, Figure 4). No interaction between time, TMR, and valence was significant.

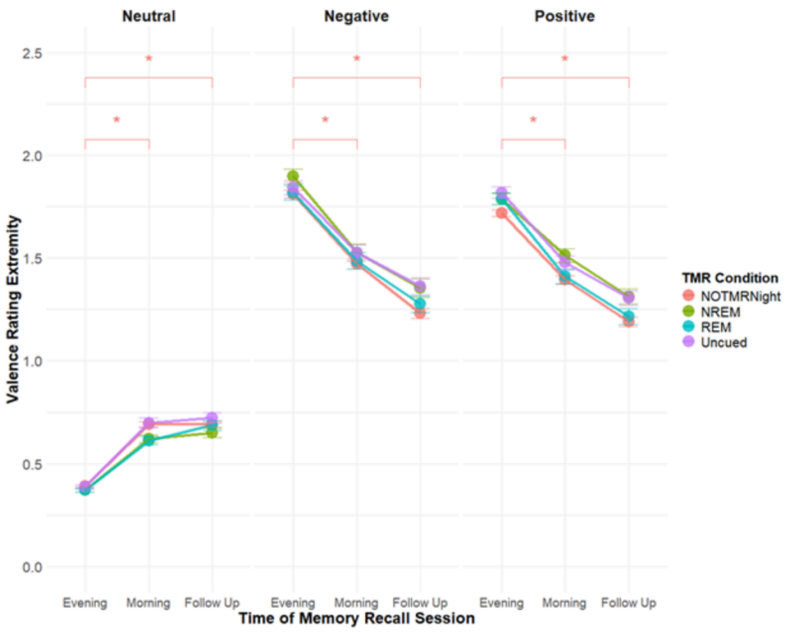


Figure 5: **Distribution of valence rating extremity as a function of time of the memory recall session for each TMR condition and valence** Valence rating extremities are plotted as a function of time of the memory recall session for each TMR condition and valence. * Indicates significant comparison, p < .05. Means and SEM are shown. The color of the comparison bar indicates which TMR condition is being considered. Note that no interaction effect with the TMR condition was significant.

3.4.0.2 Valence ratings extremity as a function of time, TMR, Valence and memory

Lastly, we aimed to see if the patterns of change in valence rating extremity over time, as a function of valence and TMR condition, differed between remembered and forgotten items (Figure 5).

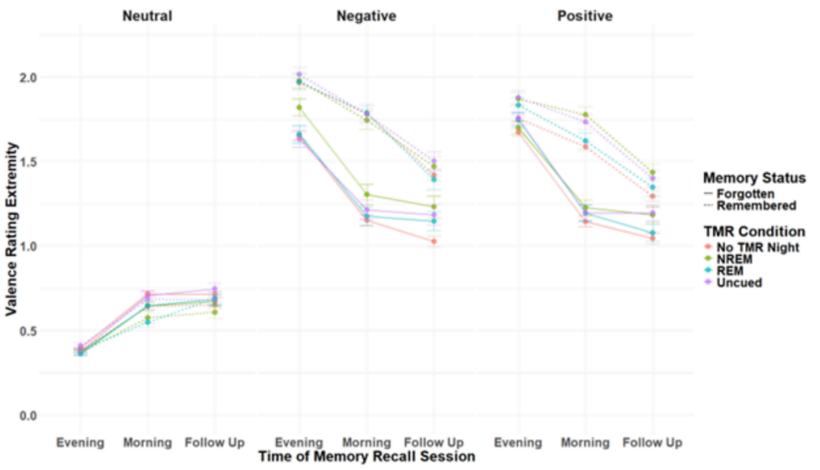


Figure 6: **Distribution of valence rating extremity as a function of time of the memory recall session for each TMR condition, valence, and memory status in the morning** Valence rating extremities are plotted as a function of time of the memory recall session for each TMR condition and valence. Results are plotted separately in function of Memory Status - whether or not the participant remembered the image in the morning. Means and SEM are shown. Note that comparison effects of time as a function of valence and images' memory statues are not plotted.

In the No TMR night, the extremity of valence ratings significantly decreased between evening and morning, as well as between evening and follow-up, for positive (β = -0.85, SE = 0.04, t(36.61) = -20.22, $p < .001$; β = -0.96, SE = 0.04, t(36.61) = -21.99, $p < .001$, respectively) and negative images (β = -0.82, SE = 0.04, t(36.61) = -19.72, $p < .001$; β = -0.95, SE = 0.04, t(36.61) = -22.17, $p < .001$, respectively) compared to neutral images. This effect was dampened for the images remembered in the morning, as a significant positive three-way interaction between valence, time, and memory was observed for positive and negative images when comparing the change in valence rating extremity both from evening to morning and from evening to follow-up ($p \leq .027$).

Additionally, the mixed-effects model revealed a significant three-way interaction between time, valence, and TMR condition on valence rating extremity at follow-up. This effect was observed for forgotten negative images cued during REM sleep (β = 0.17, SE = 0.08, t(36.61) = 2.07, p = .038) as well as for uncued forgotten negative images presented during the TMR Night (β = 0.17, SE = 0.09, t(36.61) = 1.97, p = .049). For both REM and uncued images, this effect was independent of memory (REM: β = -0.23, SE = 0.12, t(36.61) = -1.91, p = .056; uncued: β = -0.12, SE = 0.12, t(36.61) = -1.00, p = .318). These results were confirmed by the post-hoc analysis on the subset of negative images. Indeed, the decrease in valence rating extremity for forgotten negative images from evening to follow-up (β = -0.64, SE = 0.04, t(9.43) = -15.40, $p < .001$) was significantly smaller when the images had been cued during REM sleep (β = 0.17, SE = 0.08, t(9.43) = 2.19, p = .029) and when they were uncued (β = 0.21, SE = 0.08, t(9.42) = 2.60, p = .009). This effect was also independent of memory (REM: β = -0.21, SE = 0.11, t(9.42) = -1.88, p = .061; uncued: β = -0.17, SE = 0.11, t(9.42) = -1.55, p = .121, Figure 6). No additional interaction effect with TMR was found on the valence rating extremity of

positive images relative to neutral images, and when the analysis was conducted on the positive image subset ($p \geq .0502$).

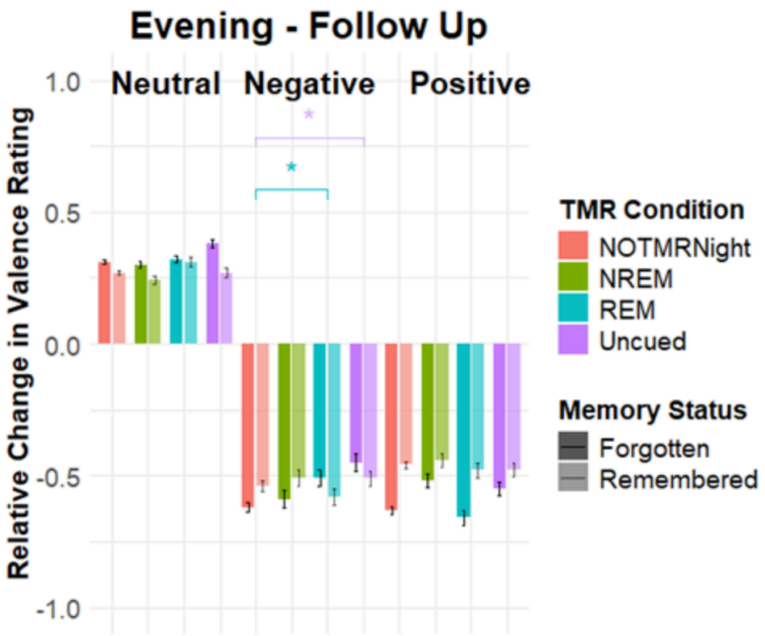


Figure 7: **Distribution of relative change in valence rating extremity between evening and follow-up, analyzed by TMR condition valence and memory** Differences in valence rating extremity between follow up and evening are plotted as a function of the images' valence for each TMR condition. Memory Statues indicate whether the image was remembered or not in the morning.* Indicates significant comparison, $p < .05$. Means and standard error of the mean (SEM) are shown. The colors of the comparison lines indicate which TMR conditions are compared to the No TMR night. Note that the plotted effects were independent of memory ($p \geq .061$).

3.5 MVPA Analysis

After analyzing the impact of TMR on arousal ratings and valence extremity over time, we aimed to investigate whether the emotional tone of remembered images cued during NREM2 sleep and REM sleep could elicit distinct EEG patterns using multivariate pattern analysis. We compared high arousal versus low arousal, negative versus positive, negative versus neutral, and positive versus neutral images.

3.5.1 Classification Across Participants

We first conducted classification analysis by randomly averaging every trial in groups of 4 to increase the signal-to-noise ratio. This analysis was conducted using no averaging time window, as well as 10 ms, 30 ms, 50 ms, and 70 ms averaging time windows.

In NREM2, the comparison between high arousal and low-arousal images revealed a significantly higher AUC compared to chance when using a 10 ms time window ($n = 24/80$, $p = .04$), a 30 ms time window ($n = 26/82$, $p = .048$), and a 50 ms time window ($n = 14/67$, $p = .01$), around 0.55 seconds after cueing (Figure 7 A).

In addition, in REM, the comparison between negative and neutral images revealed a significantly higher AUC compared to chance when using a 30 ms time window (n= 19/94 p= 0.02) (Figure 9 A) around 0.3 seconds after cueing.

No other class comparisons in NREM2 (negative vs. positive, negative vs. neutral, and positive vs. neutral), and in REM (high arousal vs. low arousal, negative vs. positive, and positive vs. neutral) yielded significant classification results compared to chance (AUC = 0.5), indicating that the SVM classifier could not distinguish these emotional tones from one another beyond chance levels.

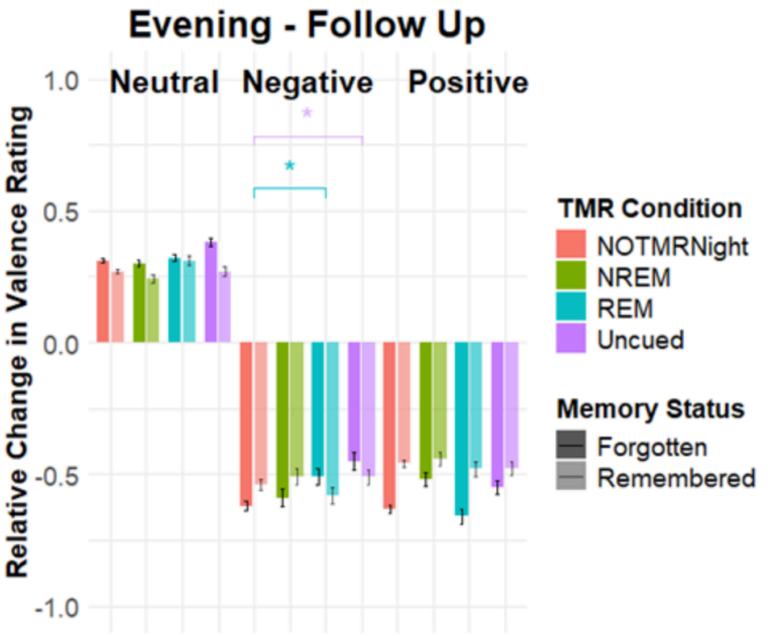


Figure 8: **AUC distribution across time of High Arousal vs Low arousal cued images in NREM2** AA) Classification across participants using a 10 ms averaging time window, with trials randomly averaged in groups of 4. One cluster was significant around 0.55 seconds after TMR onset (n = 24/80, p = .04). B) Classification across participants of high-variance trials. Two clusters out of 78 were significant, one around 0.1 seconds after TMR onset (n = 12, p = .01) and another around 2.1 seconds after TMR onset (n = 61, p = .03). C) Classification across participants of high-variance trials with a 70 ms averaging time window, with trials randomly averaged in groups of 4. Two clusters out of 61 were significant, one around 0.1 seconds after TMR onset (n = 9, p = .01) and another around 1.4 seconds after TMR onset (n = 36, p = .02). Bold lines indicate significant clusters.

We then conducted the same classification analysis using only high-variance trials, as they are more likely to reflect significant signal variations.

In NREM 2, the comparison between high arousal and low-arousal images revealed a significantly higher AUC compared to chance. When the analysis was conducted without an averaging time window, significant results were found for two clusters out of 78: one around 0.1 seconds after TMR onset (n = 12/78, p = .01) and another around 2.1 seconds after TMR onset (n = 61/78, p = .03) (Figure 7 B). Significant clusters were also identified when the analysis was conducted with a 10 ms

averaging time window (n = 53/69, p = .01) around 2.1 seconds, with a 30 ms (n = 6/55, p = .048), 50 ms (n = 13/76, p = .03), and 70 ms (n = 10/59, p = .03) averaging time window around 0.1 seconds after TMR onset.

In addition, using this method, the comparison between negative and neutral images cued in NREM2 revealed a significantly higher AUC compared to chance. When the analysis was conducted with an averaging time window of 30 ms (n = 58/91, p = .02, Figure 8 A) and 50 ms (n = 51/97, p = .02), significant results were observed around 1.4 seconds after cueing.

No other class comparisons in NREM2 (negative vs. positive and positive vs. neutral) and none of the class comparisons in REM yielded significant classification results compared to chance (AUC = 0.5) when using high-variance trials.

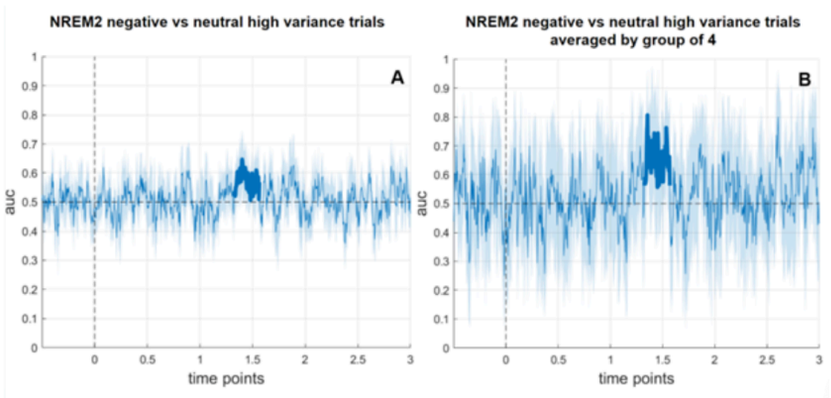


Figure 9: **AUC distribution across time of negative vs neutral cued images in NREM2** A) Classification across participants of high-variance trials with a 30 ms averaging time window. One cluster was significant around 1.4 seconds after TMR onset (n = 58/91, p = .02). B) Classification across participants of high-variance trials, averaged in groups of 4 with a 30 ms averaging time window. One cluster out of 86 was significant around 1.4 seconds after TMR onset (n = 50, p < .001). Bold lines indicate significant clusters.

Finally, we conducted the same classification analysis by randomly averaging high-variance trials in groups of 4 to increase the signal-to-noise ratio.

The comparison between high arousal and low arousal images cued in NREM2 revealed a significantly higher AUC compared to chance. Specifically, significant results were found without an averaging time window (n = 52/93, p = .008) around 1.4 seconds after cueing, with a 10 ms time window (n = 14/78, p = .024) around 0.1 seconds after cueing, with a 50 ms time window (n = 48/85, p = .002) around 1.4 seconds after cueing, and with a 70 ms averaging time window (n = 9/61, p = .014) around 0.1 seconds and (n = 36/61, p = .022) around 1.4 seconds (Figure 7 C).

In addition, using this method, the comparison between negative and neutral images cued in NREM2 revealed a significantly higher AUC compared to chance when using a 30 ms time window (n = 50/86,

p < .001), a 50 ms time window (n = 34/67, p = .046), and a 70 ms time window (n = 46/82, p = .034) around 1.4 seconds after cueing (Figure 8 B).

In addition, using this method, the comparison between negative and neutral images cued in REM revealed a significantly higher AUC than chance, with significant results for a 10 ms time window (n = 21/98, p = .04), a 30 ms time window (n = 19/101, p = .04), and a 70 ms time window (n = 18/100, p = .018) around 0.3 seconds after cueing (Figure 9 B).

Further, the comparison between positive and neutral images cued in REM also revealed a significantly higher AUC compared to chance when using a 30 ms time window (n = 110/114, p = .048) and (n = 114/114, p = .02, Figure 10)

No other comparisons in NREM2 (negative vs. positive and positive vs. neutral) or REM (high arousal vs. low arousal and negative vs. positive) elicited significant classification results compared to chance (AUC = 0.5).

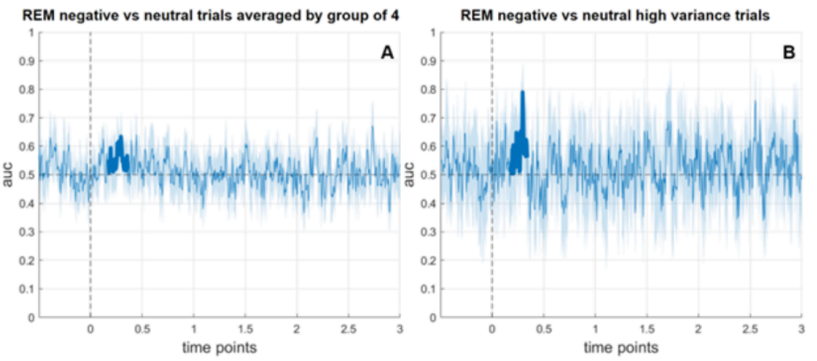


Figure 10: **AUC distribution across time of negative vs neutral cued images in REM classification across participants** A) Classification across participants using a 30 ms averaging time window, with trials randomly averaged in groups of 4. One cluster was significant around 0.3 seconds after TMR onset (n = 19/94, p = .018). B) Classification across participants of high-variance trials with a 10 ms averaging time window, averaged in groups of 4. One cluster out of 98 was significant around 0.3 seconds after TMR onset (n = 21, p = .04). Bold lines indicate significant clusters.

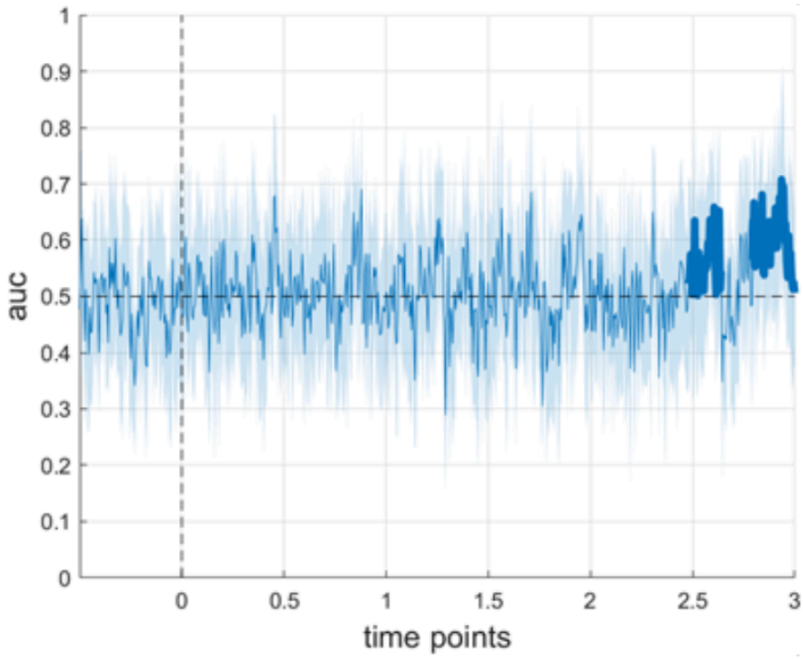


Figure 11: **AUC distribution across time of positive vs. neutral cued images in REM classification across participants for high variance trials with 30 ms averaging time window averaged by group of 4** 2 clusters out of 114 were significant (n=110, p=0.048) around 2.5 sec after TMR onset (n=114, p=0.02) around 2.7 sec after TMR onset. Bold lines indicate significant clusters.

In brief, depending on the method classification between positive and negative cued images was possible in both REM and NREM while classification between high and low arousal cued images was only achieved in NREM. Finally, classification between positive and neutral images was only possible in REM.

3.5.2 Classification for Each Participant

We then analyzed whether the emotional tone of remembered images cued during sleep could elicit distinct EEG patterns using support vector machine (SVM) learning, conducted separately for each participant. The class comparisons remained the same: high arousal vs. low arousal, negative vs. positive, negative vs. neutral, and positive vs. neutral remembered images. Participants with fewer than 10 trials per category were excluded from the analysis. Due to the low number of trials per participant, the reliability of these results is limited.

In brief, in NREM2 sleep, classification was possible for a subset of participants across all comparisons: high vs. low arousal, negative vs. neutral, negative vs. positive, and positive vs neutral. In REM sleep, classification was successful for high vs. low arousal, negative vs. neutral, and negative vs. positive comparisons, but not for the positive vs. neutral comparison.

4 Discussion

This study aimed to investigate the effect of TMR during NREM and REM sleep on emotional memory processing. Specifically, we measured changes in arousal rating and valence rating extremity for images initially rated as positive, neutral, or negative. These Images were paired with a neutral word and reactivated during either NREM or REM sleep, or left uncued, compared to a control night with no TMR application.

4.1 TMR Effect on Emotional Memory Processing Depended on Valence and Was Driven by Forgotten Images

The findings suggest that TMR had specific effects on emotional memory processing depending on both the valence of the image and whether the word-image pair was remembered in the morning following the TMR application. Indeed, no significant effect of TMR was observed when the analysis was conducted across valence categories, and when the analysis was conducted across memory status, only a significant effect of REM cueing on increased arousal ratings for positive images in the morning emerged. However, when the analysis accounted for potential interactions between valence and memory status, additional effects of TMR became evident. In addition to the effect of REM cueing, which led to increased arousal ratings for positive images in the morning, uncued images from the TMR night also exhibited a similar rise in arousal. However, this increase was less pronounced for uncued images that were remembered. Four days later, the heightened arousal for REM-cued positive images remained significant, though it was reduced for remembered images. Interestingly, the arousal effect on uncued images faded by the follow-up, while for NREM-cued positive images, the increase in arousal became significant at this later time point.

For negative images, the effects of TMR on both arousal and valence rating extremity were similarly driven by forgotten images. When considering all negative images across memory status, no significant effect of the TMR condition on arousal or valence rating extremity was apparent. However, when memory status was taken into account, a delayed TMR effect on emotional ratings became evident during the follow-up. Specifically, arousal ratings for all negative images presented during the TMR night (REM, NREM, and uncued) significantly decreased by the follow-up. Even though this effect was statistically independent of memory status for the images cued during REM and NREM sleep, the effect was descriptively damped for the remembered images which may explain why no overall effect was detected when memory status was not taken into account. Additionally, for uncued images, the decrease in arousal was significantly dampened for those that were remembered.

While the subjective arousal of negative images appeared to decrease more for the images presented during the TMR night, the valence rating extremity of these negative images was more preserved (i.e., decreased less compared to the No TMR night) at follow-up for those cued during REM sleep and for uncued images presented during the TMR night. Similarly, although this effect was statistically

independent of memory status, the direction of the effect differed for remembered images, which again obscured the overall effect when analyzed across memory statues.

The observation that TMR effects on arousal and valence rating extremity are mainly driven by forgotten items raises the possibility that these changes reflect random fluctuations rather than genuine TMR effects. For example, on the TMR night, positive images that later showed increased arousal were initially rated as less arousing than neutral ones, whereas the opposite pattern occurred on the No TMR night—suggesting a general drift toward the scale's midpoint. A similar argument could be made for the decrease in arousal ratings of forgotten negative images at follow-up. Initially rated as more arousing, these images, once forgotten, may have been rated more randomly, drifting toward the midpoint and creating the appearance of a decrease. Furthermore, the higher baseline arousal ratings of images later cued in REM and left uncued—compared to those presented during the No TMR night—could also contribute to this apparent reduction, as ratings naturally regress toward the mean. However, this drift was also observed for negative images cued in NREM, despite no baseline differences with the No TMR night, suggesting that additional factors may be at play. Furthermore, it is important to note that for valence rating extremity, the general decrease observed for negative images was dampened for forgotten images cued in REM sleep and for uncued images. If these ratings were becoming more random, a larger decrease in valence rating extremity would be expected, rather than the dampening effect that is observed. Therefore, while caution is warranted when interpreting these results, it remains reasonable to hypothesize that the observed effects may indeed be attributable to TMR rather than random fluctuations in the data.

The observation that TMR appeared to be more effective on the emotional ratings of forgotten items should be considered in light of previous research on TMR and memory. For instance, Creery et al. (Creery et al., 2015) reported that NREM TMR enhanced memory recall accuracy for items that were better remembered before sleep, suggesting that TMR may be more effective for the consolidation of memories that are strongly encoded before sleep. Although our study did not directly assess evening memory accuracy, we relied on morning recall, which likely reflects similar trends, implying that images remembered in the morning were also well remembered in the evening. Consequently, our results suggest that the emotional processing triggered by covert reactivation may be more effective for labile or weaker memories, with these effects becoming more noticeable after some delay.

Within the framework of Walker's (M. P. Walker, 2009) "sleep to remember, sleep to forget" theory, sleep's role in emotional memory may vary depending on the strength and stability of the initial encoding. Indeed, these findings suggest that sleep may preferentially consolidate memories that are strongly encoded, while the processing of the emotional tone associated with memories may occur more effectively or more quickly for weakly encoded, labile memories, as evidenced in our study within a four-day period.

4.2 REM TMR Effect Transferred to Uncued Items

Consistent with prior research (Rihm and Rasch, 2015), we found that both REM and NREM cueing led to a decrease in arousal ratings four days later for negative images, compared to the NoTMR night. Interestingly, this effect extended to uncued items presented during the TMR night, suggesting a possible transfer of TMR's influence.

Transfer of TMR effect to uncued items has previously been reported, for instance Oudiette et al. (Oudiette et al., 2013), demonstrated that TMR during slow-wave sleep (SWS) could rescue memory for low-reward value items, which were otherwise forgotten more often than high-reward items, and that this effect extended to uncued items. By assigning reward value, participants likely imbued these items with emotional significance, and the authors suggested that reactivating low-value items might trigger generalized reactivation within the broader category of low-value items.

Our study suggests a similar mechanism: the observed decrease in arousal ratings for forgotten negative images and the increase in arousal ratings for positive images at follow-up—across REM, NREM, and uncued conditions—as well as the dampened decrease in valence extremity for REM-cued and uncued items, raises the possibility of generalized reactivation occurring within valence categories.

The observed decrease in arousal for forgotten negative images and the increase in arousal for forgotten positive images across all conditions during the TMR night complicate the interpretation of which sleep stage may have driven these effects. Given the mixed findings in previous research on TMR and emotional memory processing, disentangling the specific contributions of NREM and REM sleep remains challenging. Indeed, while NREM TMR has been linked to enhanced fear extinction (Hauner et al., 2013; He et al., 2015), other studies report no effect on arousal or traumatic-like memories (Pereira et al., 2022; Gvozdanovic et al., 2023). Some research suggests both NREM and REM cueing reduce arousal (Rihm, 2015), though REM-specific effects have also been observed (Hutchison et al., 2021; Borghese et al., 2022).

Nevertheless, the results of this study suggest that the emotional processing effects of TMR might have primarily been driven by REM sleep. Specifically, the effects of REM cueing on valence rating extremity generalized to uncued images, a pattern not observed with NREM cueing. Additionally, an increase in arousal ratings of positive images was observed in the morning for images cued during REM sleep, and this effect also extended to uncued items. However, this generalization to uncued items was not seen when NREM cueing influenced arousal at the follow-up. These findings indicate that REM TMR effects on emotional memory processing may more readily transfer to uncued items compared to NREM TMR effects. This aligns with the proposed distinct roles of REM and NREM sleep in memory, particularly concerning generalization and discrimination (Witkowski et al., 2020). NREM sleep, especially SWS, is believed to support memory discrimination and pattern separation, processes linked to NREM-specific oscillations (Hanert et al., 2017). Conversely, REM sleep is thought to promote memory generalization, facilitating the extraction of broader patterns or

regularities from experiences, sometimes at the expense of detailed memory content (Sterpenich et al., 2014). The hypothesized integrative function of REM sleep appears to exert a broader influence on emotional processing, as evidenced by the greater transfer of REM TMR effects to uncued memories compared to NREM TMR. We propose that this generalization may result from the global reactivation of related memories during REM sleep, potentially supported by theta/gamma oscillations enabling long-range communication (M. P. Walker, 2009), low noradrenaline (Pace-Schott and Hobson, 2002), and high cholinergic activity (Vazquez and Baghdoyan, 2001), which together may promote integrative memory processing (Power, 2004; Almeida-Filho et al., 2018). In contrast, NREM TMR may foster more selective consolidation (Paller et al., 2020) potentially reinforced by higher noradrenaline levels that strengthen specific memory traces.

4.3 REM and NREM TMR Effect on Emotional Rating of Negative Images

4.3.1 NREM TMR-Related Decrease in Arousal for Negative Images May Reflect Enhanced Consolidation

In fact, the observed decreases in arousal ratings for images cued in REM and NREM sleep may result from distinct mechanisms, with each sleep stage independently producing similar effects. Given the different brain states and functions of NREM and REM sleep, separate neurophysiological processes could underlie these effects on arousal ratings.

According to the ”sleep to remember, sleep to forget” hypothesis (M. P. Walker, 2009), NREM sleep is primarily hypothesized to support memory consolidation. Within the framework of system-level consolidation theory, memory reactivation during slow-wave sleep (SWS) is believed to play a key role in the corticalization of initially hippocampus-dependent memories (Rasch and Born, 2013; Paller et al., 2020). TMR during NREM sleep may facilitate this consolidation process, potentially making memories less aversive as they become more familiar, with emotional intensity diminishing over time and only the core, or ”gist,” of the memory being retained. Indeed, SWS TMR has been shown to promote fear extinction, an effect linked to reduced hippocampal activity (Hauner et al., 2013) and altered amygdala responses (Carbone and Diekelmann, 2024). Consistent with this hypothesis, neutral images cued during NREM showed a near-significant drop in arousal at follow-up ($p = .05$), unlike REM and uncued items. If this decrease in arousal reflects enhanced consolidation, similar effects would be expected for both neutral and negative memories. This partially aligns with Rihm and Rasch (Rihm and Rasch, 2015), who found reduced arousal for both negative and neutral images after NREM cueing. However, unlike our findings, they also reported similar effects following REM cueing.

4.3.2 REM TMR Reduced Arousal of Negative Images While Preserving Valence Extremity

Conversely, within the ”sleep to remember, sleep to forget” hypothesis, REM sleep is hypothesized to play a key role in processing the emotional tone of memories. Our results suggest that REM sleep

may have a differential role in processing the autonomic arousal response and the higher-order cognitive appraisal of the valence of negative memories, affecting these aspects of the emotional memory in contrasting ways. Specifically, REM cueing was associated with a more pronounced decrease in arousal for forgotten negative images, while the decrease in valence rating extremity was lessened. This suggests that REM sleep may differentially influence the emotional components of memory, potentially diminishing physiological arousal while preserving some degree of cognitive appraisal of the emotional content.

Walker (2009) proposed that REM sleep reduces the emotional tone of memories by reactivating them during periods of low noradrenaline, allowing emotional content to be recalled without its original intensity. This may decouple the memory from its emotional charge while preserving the core content. Supporting this hypothesis, it is noteworthy that in this study, while classification across participants of high vs. low arousal and negative vs. neutral memories during NREM sleep was possible, only the classification of negative vs. neutral memories was consistently achievable in REM sleep (note however that even if high vs. low arousal classification during REM sleep did not generalize across participants, it reached statistical significance in 3 out of 9 participants). Given that baseline arousal ratings of positive and neutral images did not significantly differ, it is most likely that the high-arousal memories primarily involved negative images. These results suggest that during NREM sleep, TMR of negative images reactivates the memory's emotional valence and arousal components, whereas during REM sleep, TMR may selectively reactivate the valence without the associated arousal, potentially contributing to the fading of arousal over time. However, these findings should be interpreted with caution. Indeed, the classification was based on a relatively small sample of 12 participants, and to address classification bias from uneven trial numbers, under sampling was used. Additionally, given that TMR-induced REM reactivations may be compressed in time (M. E. Abdellahi et al., 2021), the lack of significant results could be due to variations in the timing and duration of reactivation across participants. Despite using multiple sliding averaging time windows in the analysis, this intra-participant variability might not have been adequately compensated for, potentially influencing the classification outcomes.

On the other hand, REM cueing appeared to dampen the decrease in valence rating extremity of forgotten negative images at follow-up, suggesting that the cognitive appraisal of these images was more preserved. Given the accumulating evidence that REM sleep preferentially supports the consolidation of negative memories (Wagner et al., 2001; Wiesner et al., 2015; Groch et al., 2013) or the negative emotional component of memories (Payne et al., 2008; Payne et al., 2012), we propose that REM sleep may facilitate the preservation of negative emotional valence, particularly when the specific memory content is no longer consciously accessible. This covert retention of emotional valence—without heightened arousal—may support adaptive decision-making by subtly guiding responses to future situations without the cognitive load of recalling the full memory.

4.4 TMR Increased Arousal Rating of Positive Images

If our hypothesis that REM cueing supports a certain level of preservation of the cognitive appraisal of memory is correct, one would also expect to see this effect for positive valence. Previous studies have reported results consistent with this hypothesis. For example, Rihm and Rasch (Rihm and Rasch, 2015) found that both REM and NREM cueing enhanced positive valence ratings for conditioned and unconditioned stimuli, with stronger effects observed for REM cueing. However, in contrast to their findings, we observed an increase in arousal ratings for positive images rather than changes in valence rating extremity. The typically lower arousal of positive images in laboratory settings, which contrasts with real-life experiences, may explain why positive emotional content was less engaging and could potentially account for this inconsistency.

The heightened arousal ratings for positive images suggest that sleep may support Fading Affect Bias (FAB)— a cognitive phenomenon in which the emotional intensity of negative autobiographical memories fades more rapidly than that of positive memories (Landau and Gunter, 2009). FAB supports emotional regulation and well-being (W. R. Walker et al., 2003; M. P. Walker, 2009). Given that sleep, particularly REM sleep, disturbances (Harrington et al., 2018; Pesonen et al., 2019) are associated with mood disorders, and the fact that Fading Affect Bias can emerge within 12 hours (Gibbons et al., 2011), sleep may play a critical role in this bias. Our findings suggest that REM sleep may particularly contribute to Fading Affect Bias, as positive images cued during REM sleep exhibited this bias in the morning, while this effect only emerged four days later for images cued during NREM sleep. Additionally, the detection of positive vs. neutral memories in REM sleep may support the hypothesis that REM sleep contributes to Fading Affect Bias by being more sensitive to the reactivation of positive emotional content. However, this finding was derived from a single classification method and was not observed at the participant level, significantly limiting the reliability of this result.

However, the presented results of TMR on positive images should be interpreted cautiously due to initial differences in arousal ratings between positive images on the two nights (higher than neutral in the No TMR night vs. lower than neutral in the TMR night), which represent a potential bias. This suggests that the observed effects could be partially attributable to these baseline disparities. It is also possible that TMR facilitated a ”catch-up” effect, wherein arousal ratings of positive and neutral images were normalized, but this hypothesis was not tested in this study.

4.5 Temporal dynamic of valence and arousal classification in NREM2

In this study, we used a SVM algorithm to demonstrate that TMR of emotional memories induces distinct neurophysiological responses, which are detectable at the scalp level through specific ERP patterns. Previous studies have shown that memory reactivation triggered by TMR occurs during both REM (M. E. Abdellahi et al., 2021; M. E. Abdellahi et al., 2023) and NREM sleep (Santamaria et al., 2024). However, to our knowledge, this is the first study to demonstrate that TMR effectively reactivates the emotional content of a memory, which is evident in EEG correlates.

While our methodology does not provide detailed information about the specific EEG patterns responsible for classification, it does offer valuable insight into the reactivation of emotional content and its timing. In NREM2 sleep, we were able to consistently differentiate high-arousal from low-arousal memories, as well as negative from neutral memories. However, the timing at which EEG patterns differentiating high and low arousal became significantly distinct varied depending on the methodological approach.

Previous research has demonstrated that aperiodic EEG activity is critical for MVPA classification of arousal in waking states (Borah et al., 2024). It is possible that when only high-arousal trials are selected, the aperiodic differences between trials are normalized (Gerster et al., 2022), whereas when all trials are included, the aperiodic component may contribute to classification alongside other EEG signal components.

Moreover, when trials were randomly averaged in groups of four, an additional time point (1. 4 sec) at which classification accuracy for high vs. low arousal trials became significant, which corresponds to the timing at which the distinction between negative and neutral memories also became significant in NREM2 sleep. Since negative but not positive images were significantly rated as more arousing than neutral, this suggests that the time-locked classification of arousal and valence might be driven by the same memory being reactivated.

However, the arousal classification in NREM2 that was not time-locked with negative vs. neutral classification is more challenging to interpret. Given that negative emotions tend to trigger fast, automatic processing while positive emotions involve slower, more reflective processing (Thiruchselvam et al., 2011), it is possible that the classification time points are influenced by these differing emotional processing dynamics. This interpretation is however limited by the absence of significant classification of positive vs. neutral or positive vs. negative images. Alternatively, it is possible that the classification was driven by a subset of participants who exhibited strong reactivation for negative and high-arousal trials. However, this interpretation is limited by the fact that classification between negative and neutral memories was still possible without averaging and by using only high-variance trials. When trials are not averaged, classification may reflect a broader participant pool, which could account for the multiple time points of classification observed. The multiple time points of classification may also suggest repeated reactivation of the arousal component. Previous research in REM sleep (M. E. Abdellahi et al., 2021) and NREM sleep (Schreiner et al., 2018) has shown that TMR can lead to recurrent reactivations of memory components, which may explain the multiple arousal classification time points observed in our NREM2 analysis.

4.6 Limitations and future directions

One notable limitation of our study is that participants were awakened eight times during the night to report their dreams. Although group-level analyses showed no significant differences in sleep quality between the two experimental nights, this remains a potential confound, as previous research has

demonstrated that TMR effects on emotional processing (Wassing et al., 2019) and memory consolidation (Göldi and Rasch, 2019; Whitmore and Paller, 2023) can be affected by sleep quality. Therefore, future studies investigating the effects of TMR should aim to conduct experiments without awakening participants to assess the impact of TMR on undisturbed sleep. Furthermore, the duration of REM sleep has been found to modulate TMR effects, particularly concerning anxiety (Borghese et al., 2022). Incorporating sleep stage duration into the statistical model could provide more nuanced insights into the relationship between sleep architecture and TMR effects.

Previous studies have shown that TMR alone does not directly influence emotional processing or procedural learning; instead, these effects are driven by the number of TMR stimulations (Borghese et al., 2022) and, in the case of procedural learning, the number of detected memory reactivations (M. E. A. Abdellahi, 2022). Controlling for these factors—as well as the number of correctly cued memories within the appropriate sleep stage—could enhance the interpretability of our results. Given that many participants did not have enough trials for reliable multivariate pattern analysis, future analyses could focus on a subset of participants with sufficient trials. This would allow for a more precise investigation into which aspects of emotional memory are reactivated by TMR and how these reactivations relate to behavioral outcomes, providing deeper insights into the mechanisms underlying sleep-related emotional memory processing. For instance, this method offers an opportunity to further investigate Walker's (2009) hypothesis, which posits that the attenuation of emotional memory during REM sleep arises from a dissociation between the memory itself and its emotional tone. Wassing et al. (Wassing et al., 2019), previously demonstrated that the effect of REM TMR on amygdala adaptation to emotional memory depends on REM sleep quality. Our methodology presents a valuable approach to further test the "sleep to forget" hypothesis. Indeed, future studies could specifically explore whether the simultaneous reactivation of both arousal and negative valence during REM sleep leads to increased or unchanged future arousal ratings compared to instances where only valence is reactivated, which may result in a decrease in arousal ratings.

Additionally, understanding the temporal dynamics of arousal and valence components of emotional memory reactivation during sleep, and how these dynamics vary with the valence of the memory, could further enhance our understanding of the mechanisms underlying emotional memory processing during sleep. In this study, using SVM classification we were able to significantly detect reactivation of negative vs. neutral memories induced by TMR but not positive vs. neutral, likely due to the more engaging nature of the negative images compared to the positive images used. To address this limitation, future research should use a paradigm that induces more engaging positive emotions to investigate whether positive valence can be reliably detected during sleep and whether arousal reactivation occurs concurrently with valence reactivation or independently, without being time-locked to specific valence categories. A potential approach to explore the temporal dynamics of emotional memory reactivation in greater detail would be to use SVM classification to differentiate between high and low arousal trials within the same valence category. This method could help clarify whether the reactivation of arousal and valence follows similar or divergent time courses, depending on the valence

of the emotion. Such insights could provide a deeper understanding of how emotional memories are processed during sleep and the distinct roles of arousal and valence reactivation in this process.

A key limitation in interpreting the classification results concerning the behavioral analysis is that, at the behavioral level, the effects of TMR on emotional processing appear primarily driven by forgotten word-image pairs. In contrast, the multivariate pattern analysis (MVPA) classification was performed on remembered word-image pairs. To strengthen our interpretation, it is crucial to test whether similar classifications can be performed on forgotten items. Additionally, the absence of a TMR effect on remembered images in the follow-up might have been partly influenced by the fact that remembered images were described in the morning, while forgotten ones were not; this recall may have induced emotional processing that masked the TMR effect. For instance, in the No TMR night, arousal ratings of forgotten negative images decreased by the morning but returned to baseline levels four days later. In contrast, for remembered negative images, the pattern was reversed: these images showed a smaller decrease in arousal ratings in the morning compared to forgotten images but exhibited a greater decrease at follow-up. This greater reduction in arousal for remembered negative images at follow-up may be partly due to emotional desensitization resulting from repeated recall (Adler, 2022). Repeated recall can facilitate emotional fading, an effect that would not occur for images forgotten by the morning. Therefore, to accurately measure delayed TMR effects on emotional memory processing, future studies should avoid memory recall before this effect is tested. Instead, the strength of the memory before sleep could be used to assess word-image pair retention rather than relying on morning recall.

Moreover, NREM and REM sleep were treated as homogeneous stages in this study, potentially overlooking important neurophysiological distinctions. NREM sleep includes distinct sub-stages, such as NREM2 and SWS, and although some studies report no significant differences between SWS and NREM2 TMR effects (Carbone et al., 2023; Wick and Rasch, 2023), these sub-stages may still differentially influence emotional memory processes. Similarly, REM sleep can be divided into tonic and phasic states, each with unique neurophysiological characteristics and potentially distinct functions (Simor et al., 2020). Future research should examine TMR effects across these specific sleep stages to gain a more comprehensive understanding of their respective roles in emotional memory processing.

It is, however, important to keep in mind that attributing a singular role in processing the affective components of memory to any specific sleep stage might oversimplify the complex interactions at play (Rauchs et al., 2005). For example, recall of word pairs is impaired following fragmented sleep that disrupts the sleep cycle, while recall remains intact when awakenings preserve the integrity of the sleep cycle (Ficca et al., 2000). More directly related to emotional memory, the effect of SWS TMR on amygdala reactivity to emotional memory depends on the subsequent time spent in REM sleep (Pereira et al., 2022). These findings highlight the importance of cyclic interactions between NREM and REM sleep stages in emotional memory processing rather than isolated effects of specific sleep stages. Hence, future studies should analyze TMR effects in one sleep stage in relation to the other

sleep stages characteristics.

Lastly, it has been hypothesized that the timing of TMR relative to the phase of ongoing brain oscillations may lead to varying behavioral outcomes (Talamini and Juan, 2020; Carbone and Diekelmann, 2024). Our study did not account for this factor, and it is possible that controlling for the phase of brain oscillations during TMR delivery could produce different, potentially more pronounced, or even opposite effects depending on the timing of reactivation. Future research should explore this temporal aspect to better understand how brain oscillations interact with TMR and influence emotional memory processing during sleep.

5 Conclusion

In this study, we aimed to investigate the effect of TMR during NREM and REM sleep on emotional memory processing. We found that the TMR effect on emotional memory processing was primarily driven by memories that were not consciously accessible following TMR application. NREM cueing induced a delayed decrease in arousal ratings of negative images, an effect we hypothesized to be due to increased memory consolidation, resulting in greater familiarity with the memory, this hypothesis being further supported by the same near-significant effect of NREM cueing on neutral memories.

Additionally, we found that REM TMR had a differential delayed effect on the processing of negative memories by enhancing the decrease in arousal associated with these memories while dampening the reduction in valence rating extremity. This suggests that REM sleep may play an adaptive role in processing negative memories, allowing individuals to retain emotionally relevant information without the associated increase in arousal. This effect also extended to uncued items, leading us to hypothesize that REM cueing induces a general reactivation of emotional memories within the same valence category. This could be supported by the characteristic neurophysiological underpinnings of REM sleep, including long-range neuronal communication facilitated by high power theta oscillations, low noradrenaline levels, and high cholinergic activity.

Furthermore, we found that sleep, particularly REM sleep, may support the Fading Affect Bias, as suggested by the delayed increase in arousal of positive images presented during the TMR night—an effect already visible in the morning for REM cued memories. However, we emphasize that this effect should be interpreted cautiously due to baseline disparities in arousal ratings between TMR conditions.

Using SVM classification, we were able to distinguish between emotional memory components induced by TMR, providing insights into the nature of the components being reactivated depending on the sleep stage, as well as the temporal dynamics of these reactivations. We suggest that future research employing this method could further elucidate the mechanisms through which emotional memories are processed during sleep and how these processes relate to changes in affective responses.

Our findings align with Walker's (2009) "sleep to remember, sleep to forget" hypothesis, suggesting

that NREM sleep may facilitate the consolidation and attenuation of arousal associated with negative memories, enhancing familiarity, while REM sleep plays a distinct role by selectively reducing arousal and preserving emotionally relevant information, allowing for adaptive processing without the burden of recalling the full emotional intensity. Additionally, our findings demonstrate that the classification of emotional content induced by TMR is feasible during sleep, providing a powerful tool to further investigate the sleep stage-specific mechanisms that support emotional memory processing.

References

Abdellahi, M. E., Koopman, A. C., Treder, M. S., & Lewis, P. A. (2021). Targeted memory reactivation in human rem sleep elicits recurrent, temporally compressed reactivation. *bioRxiv*, 2021–12.

Abdellahi, M. E., Koopman, A. C., Treder, M. S., & Lewis, P. A. (2023). Targeted memory reactivation in human rem sleep elicits detectable reactivation. *Elife*, *12*, e84324.

Abdellahi, M. E. A. (2022). *Detecting neural replay in sleep with eeg classifiers* [Doctoral dissertation]. Cardiff University.

Ackermann, S., & Rasch, B. (2014). Differential effects of non-rem and rem sleep on memory consolidation? *Current Neurology and Neuroscience Reports*, *14*, 1–10.

Alger, S. E., Kensinger, E. A., & Payne, J. D. (2018). Preferential consolidation of emotionally salient information during a nap is preserved in middle age. *Neurobiology of Aging*, *68*, 34–47.

Almeida-Filho, D. G., Queiroz, C. M., & Ribeiro, S. (2018). Memory corticalization triggered by rem sleep: Mechanisms of cellular and systems consolidation. *Cellular and Molecular Life Sciences*, *75*, 3715–3740.

Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*(1), 1–23.

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

Borah, R. M., Pathak, A., & Banerjee, A. (2024). Emotion arousal but not valence is strongly represented in aperiodic eeg activity stemming from thalamocortical interactions. *bioRxiv*, 2024–03.

Borghese, F., Henckaerts, P., Guy, F., Perez Mayo, C., Delplanque, S., Schwartz, S., & Perogamvros, L. (2022). Targeted memory reactivation during rem sleep in patients with social anxiety disorder. *Frontiers in Psychiatry*, *13*, 904704.

Carbone, J., Bibian, C., Born, J., Forcato, C., & Diekelmann, S. (2023). Targeted memory reactivation is not more effective during slow wave sleep than sleep stage 2. *bioRxiv*, 2023–10.

Carbone, J., & Diekelmann, S. (2024). An update on recent advances in targeted memory reactivation during sleep. *npj Science of Learning*, *9*(1), 31.

Carskadon, M. A., & Dement, W. C. (2005). Normal human sleep: An overview. In *Principles and practice of sleep medicine* (pp. 13–23, Vol. 4).

Castor EDC. (n.d.). Electronic data capture solution [Accessed: 2025-11-03].

Corrigan, J. D., & Hinkeldey, N. S. (1987). Relationships between parts a and b of the trail making test. *Journal of Clinical Psychology*, *43*(4), 402–409.

Creery, J. D., Oudiette, D., Antony, J. W., & Paller, K. A. (2015). Targeted memory reactivation during sleep depends on prior learning. *Sleep*, *38*(5), 755–763.

Dang-Vu, T. T., Schabus, M., Desseilles, M., Sterpenich, V., Bonjean, M., & Maquet, P. (2010). Functional neuroimaging insights into the physiology of human sleep. *Sleep*, *33*(12), 1589–1603.

Delorme, A., & Makeig, S. (2004). Eeglab: An open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Dolcos, F., LaBar, K. S., & Cabeza, R. (2004). Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron*, *42*, 855–863.

Dolcos, F., LaBar, K. S., & Cabeza, R. (2005). Remembering one year later: Role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. *Proceedings of the National Academy of Sciences*, *102*(7), 2626–2631.

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

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, *13*(1), 1–17.

Genzel, L., Spoormaker, V. I., Konrad, B. N., & Dresler, M. (2015). The role of rapid eye movement sleep for amygdala-related memory processing. *Neurobiology of Learning and Memory*, *122*, 110–121.

Gerster, M., Waterstraat, G., Litvak, V., Lehnertz, K., Schnitzler, A., Florin, E., & Nikulin, V. (2022). Separating neural oscillations from aperiodic 1/f activity: Challenges and recommendations. *Neuroinformatics*, *20*(4), 991–1012.

Gibbons, J. A., Lee, S. A., & Walker, W. R. (2011). The fading affect bias begins within 12 hours and persists for 3 months. *Applied Cognitive Psychology*, *25*(4), 663–672.

Göldi, M., & Rasch, B. (2019). Effects of targeted memory reactivation during sleep at home depend on sleep disturbances and habituation. *NPJ Science of Learning*, *4*(1), 5.

Görtelmeyer, R. (2011). *Sf-a/r und sf-b/r: Schlaffragebogen a und b*.

Groch, S., Wilhelm, I., Diekelmann, S., & Born, J. (2013). The role of rem sleep in the processing of emotional memories: Evidence from behavior and event-related potentials. *Neurobiology of Learning and Memory*, *99*, 1–9.

Groch, S., Zinke, K., Wilhelm, I., & Born, J. (2015). Dissociating the contributions of slow-wave sleep and rapid eye movement sleep to emotional item and source memory. *Neurobiology of Learning and Memory*, *122*, 122–130.

Gujar, N., McDonald, S. A., Nishida, M., & Walker, M. P. (2011). A role for rem sleep in recalibrating the sensitivity of the human brain to specific emotions. *Cerebral Cortex*, *21*(1), 115–123.

Gvozdanovic, G., Schoch, S., Stämpfli, P., Seifritz, E., & Rasch, B. (2023). Neural correlates of sleep-induced benefits on traumatic memory processing. *Human Brain Mapping*, *44*(9), 3506–3518.

Haberkamp, A., Glombiewski, J. A., Schmidt, F., & Barke, A. (2017). The disgust-related-images (dirti) database: Validation of a novel standardized set of disgust pictures. *Behaviour Research and Therapy*, *89*, 86–94.

Hanert, A., Weber, F. D., Pedersen, A., Born, J., & Bartsch, T. (2017). Sleep in humans stabilizes pattern separation performance. *Journal of Neuroscience*, *37*(50), 12238–12246.

Harrington, M. O., Johnson, J. M., Croom, H. E., Pennington, K., & Durrant, S. J. (2018). The influence of rem sleep and sws on emotional memory consolidation in participants reporting depressive symptoms. *Cortex*, *99*, 281–295.

Hauner, K. K., Howard, J. D., Zelano, C., & Gottfried, J. A. (2013). Stimulus-specific enhancement of fear extinction during slow-wave sleep. *Nature Neuroscience*, *16*(11), 1553–1555.

He, J., Sun, H. Q., Li, S. X., Zhang, W. H., Shi, J., Ai, S. Z., & Lu, L. (2015). Effect of conditioned stimulus exposure during slow wave sleep on fear memory extinction in humans. *Sleep*, *38*(3), 423–431.

Hu, P., Stylos-Allan, M., & Walker, M. P. (2006). Sleep facilitates consolidation of emotional declarative memory. *Psychological Science*, *17*(10), 891–898.

Hu, X., Cheng, L. Y., Chiu, M. H., & Paller, K. A. (2020). Promoting memory consolidation during sleep: A meta-analysis of targeted memory reactivation. *Psychological Bulletin*, *146*(3), 218.

Hutchison, I. C., Pezzoli, S., Tsimpanouli, M. E., Abdellahi, M. E., Pobric, G., Hulleman, J., & Lewis, P. A. (2021). Targeted memory reactivation in rem but not sws selectively reduces arousal responses. *Communications Biology*, *4*(1), 404.

Kensinger, E. A., & Schacter, D. L. (2006). Processing emotional pictures and words: Effects of valence and arousal. *Cognitive, Affective, & Behavioral Neuroscience*, *6*, 110–126.

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the open affective standardized image set (oasis). *Behavior Research Methods*, *49*, 457–470.

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience*, *7*(1), 54–64.

Landau, J. D., & Gunter, B. C. (2009). Don't worry; you really will get over it: Methodological investigations of the fading affect bias. *The American Journal of Psychology*, *122*(2), 209–217.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (iaps): Affective ratings of pictures and instruction manual*. NIMH, Center for the Study of Emotion & Attention. Gainesville, FL.

Leach, S., Sousouri, G., & Huber, R. (2023). High-density-sleepcleaner: An open-source, semi-automatic artifact removal routine tailored to high-density sleep eeg. *Journal of Neuroscience Methods*, *391*, 109849.

Louie, K., & Wilson, M. A. (2001). Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron*, *29*(1), 145–156.

MacDonald, K. J., & Cote, K. A. (2021). Contributions of post-learning rem and nrem sleep to memory retrieval. *Sleep Medicine Reviews*, *59*, 101453.

Marchewka, A., Żurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The nencki affective picture system (naps): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods*, *46*, 596–610.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of eeg- and meg-data. *Journal of Neuroscience Methods*, *164*(1), 177–190.

Mazzoni, G., Gori, S., Formicola, G., Gneri, C., Massetani, R., Murri, L., & Salzarulo, P. (1999). Word recall correlates with sleep cycles in elderly subjects. *Journal of Sleep Research*, *8*(3), 185–188.

Merlhiot, G., Mermillod, M., Le Pennec, J. L., & Mondillon, L. (2018). Introduction and validation of the natural disasters picture system (ndps). *PLoS One*, *13*(8), e0201942.

Nishida, M., Pearsall, J., Buckner, R. L., & Walker, M. P. (2009). Rem sleep, prefrontal theta, and the consolidation of human emotional memory. *Cerebral Cortex*, *19*(5), 1158–1166.

Oudiette, D., Antony, J. W., Creery, J. D., & Paller, K. A. (2013). The role of memory reactivation during wakefulness and sleep in determining which memories endure. *Journal of Neuroscience*, *33*(15), 6672–6678.

Pace-Schott, E. F., & Hobson, J. A. (2002). The neurobiology of sleep: Genetics, cellular physiology and subcortical networks. *Nature Reviews Neuroscience*, *3*(8), 591–605.

Paller, K. A., Mayes, A. R., Antony, J. W., & Norman, K. A. (2020). Replay-based consolidation governs enduring memory storage. In *The cognitive neurosciences* (pp. 265–276).

Payne, J. D., Chambers, A. M., & Kensinger, E. A. (2012). Sleep promotes lasting changes in selective memory for emotional scenes. *Frontiers in Integrative Neuroscience*, *6*, 108.

Payne, J. D., & Kensinger, E. A. (2010). Sleep's role in the consolidation of emotional episodic memories. *Current Directions in Psychological Science*, *19*(5), 290–295.

Payne, J. D., & Kensinger, E. A. (2011). Sleep leads to changes in the emotional memory trace: Evidence from fmri. *Journal of Cognitive Neuroscience*, *23*(6), 1285–1297.

Payne, J. D., Stickgold, R., Swanberg, K., & Kensinger, E. A. (2008). Sleep preferentially enhances memory for emotional components of scenes. *Psychological Science*, *19*(8), 781–788.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., & Lindeløv, J. K. (2019). Psychopy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*, 195–203.

Pereira, S. I. R., Tsimpanouli, M. E., Hutchison, I., Schneider, J., Anderson, I. M., McFarquhar, M., & Lewis, P. A. (2022). Cueing emotional memories during slow wave sleep modulates next-day activity in the orbitofrontal cortex and the amygdala. *NeuroImage*, *253*, 119120.

Perslev, M., Darkner, S., Kempfner, L., Nikolic, M., Jennum, P. J., & Igel, C. (2021). U-sleep: Resilient high-frequency sleep staging. *NPJ Digital Medicine*, *4*(1), 72.

Pesonen, A. K., Gradisar, M., Kuula, L., Short, M., Merikanto, I., Tark, R., & Lahti, J. (2019). Rem sleep fragmentation associated with depressive symptoms and genetic risk for depression in a community-based sample of adolescents. *Journal of Affective Disorders*, *245*, 757–763.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(3), 715–734.

Power, A. E. (2004). Muscarinic cholinergic contribution to memory consolidation: With attention to involvement of the basolateral amygdala. *Current Medicinal Chemistry*, *11*(8), 987–996.

Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological Reviews*.

Rasch, B., Büchel, C., Gais, S., & Born, J. (2007). Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science*, *315*(5817), 1426–1429.

Rauchs, G., Desgranges, B., Foret, J., & Eustache, F. (2005). The relationships between memory systems and sleep stages. *Journal of Sleep Research*, *14*(2), 123–140.

Richter-Levin, G., & Akirav, I. (2003). Emotional tagging of memory formation—in the search for neural mechanisms. *Brain Research Reviews*, *43*(3), 247–256.

Rihm, J. S. (2015). *Sleep-dependent declarative memory consolidation and its application to emotional and extinction memory processing* [Doctoral dissertation, University of Zurich].

Rihm, J. S., & Rasch, B. (2015). Replay of conditioned stimuli during late rem and stage n2 sleep influences affective tone rather than emotional memory strength. *Neurobiology of Learning and Memory*, *122*, 142–151.

Roozendaal, B., McEwen, B. S., & Chattarji, S. (2009). Stress, memory and the amygdala. *Nature Reviews Neuroscience*, *10*(6), 423–433.

Santamaria, L., Kashif, I., McGinley, N., & Lewis, P. A. (2024). Memory reactivation in slow wave sleep enhances relational learning in humans. *Communications Biology*, *7*(1), 288.

Schoch, S. F., Cordi, M. J., & Rasch, B. (2017). Modulating influences of memory strength and sensitivity of the retrieval test on the detectability of the sleep consolidation effect. *Neurobiology of Learning and Memory*, *145*, 181–189.

Schönauer, M., Alizadeh, S., Jamalabadi, H., Abraham, A., Pawlizki, A., & Gais, S. (2017). Decoding material-specific memory reprocessing during sleep in humans. *Nature Communications*, *8*(1), 15404.

Schreiner, T., Doeller, C. F., Jensen, O., Rasch, B., & Staudigl, T. (2018). Theta phase-coordinated memory reactivation reoccurs in a slow-oscillatory rhythm during nrem sleep. *Cell Reports*, *25*(2), 296–301.

Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition Emotion*, *14*(4), 433–440.

Seth, A. K., & Friston, K. J. (2016). Active interoceptive inference and the emotional brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708), 20160007.

Sharot, T., & Phelps, E. A. (2004). How arousal modulates memory: Disentangling the effects of attention and retention. *Cognitive, Affective, Behavioral Neuroscience*, *4*, 294–306.

Simor, P., van der Wijk, G., Nobili, L., & Peigneux, P. (2020). The microstructure of rem sleep: Why phasic and tonic? *Sleep Medicine Reviews*, *52*, 101305.

Singh, D., Norman, K. A., & Schapiro, A. C. (2022). A model of autonomous interactions between hippocampus and neocortex driving sleep-dependent memory consolidation. *Proceedings of the National Academy of Sciences*, *119*(44), e2123432119.

Smith, C. (2001). Sleep states and memory processes in humans: Procedural versus declarative memory systems. *Sleep Medicine Reviews*, *5*(6), 491–506.

Sterpenich, V., Albouy, G., Boly, M., Vandewalle, G., Darsaud, A., Balteau, E., & Maquet, P. (2007). Sleep-related hippocampo-cortical interplay during emotional memory recollection. *PLoS Biology*, *5*(11), e282.

Sterpenich, V., Schmidt, C., Albouy, G., Matarazzo, L., Vanhaudenhuyse, A., Boveroux, P., & Maquet, P. (2014). Memory reactivation during rapid eye movement sleep promotes its generalization and integration in cortical stores. *Sleep*, *37*(6), 1061–1075.

Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1994). Testtheoretische analysen des mehrdimensionalen befindlichkeitsfragebogen (mdbf). *Diagnostica*.

Talamini, L. M., & Juan, E. (2020). Sleep as a window to treat affective disorders. *Current Opinion in Behavioral Sciences*, *33*, 99–108.

Thiruchselvam, R., Blechert, J., Sheppes, G., Rydstrom, A., & Gross, J. J. (2011). The temporal dynamics of emotion regulation: An eeg study of distraction and reappraisal. *Biological Psychology*, *87*(1), 84–92.

Treder, M. S. (2020). Mvpa-light: A classification and regression toolbox for multi-dimensional data. *Frontiers in Neuroscience*, *14*, 289.

Van Der Helm, E., Yao, J., Dutt, S., Rao, V., Saletin, J. M., & Walker, M. P. (2011). Rem sleep depotentiates amygdala activity to previous emotional experiences. *Current Biology*, *21*(23), 2029–2032.

Van Giezen, A. E., Arensman, E., Spinhoven, P., & Wolters, G. (2005). Consistency of memory for emotionally arousing events: A review of prospective and experimental studies. *Clinical Psychology Review*, *25*(7), 935–953.

Vazquez, J., & Baghdoyan, H. A. (2001). Basal forebrain acetylcholine release during rem sleep is significantly greater than during waking. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *280*(2), R598–R601.

Wagner, U., Fischer, S., & Born, J. (2002). Changes in emotional responses to aversive pictures across periods rich in slow-wave sleep versus rapid eye movement sleep. *Psychosomatic Medicine*, *64*(4), 627–634.

Wagner, U., Gais, S., & Born, J. (2001). Emotional memory formation is enhanced across sleep intervals with high amounts of rapid eye movement sleep. *Learning Memory*, *8*(2), 112–119.

Wagner, U., Hallschmid, M., Rasch, B., & Born, J. (2006). Brief sleep after learning keeps emotional memories alive for years. *Biological Psychiatry*, *60*(7), 788–790.

Wagner, U., Kashyap, N., Diekelmann, S., & Born, J. (2007). The impact of post-learning sleep vs. wakefulness on recognition memory for faces with different facial expressions. *Neurobiology of Learning and Memory*, *87*(4), 679–687.

Walker, M. P. (2009). The role of sleep in cognition and emotion. *Annals of the New York Academy of Sciences*, *1156*(1), 168–197.

Walker, M. P., & Stickgold, R. (2010). Overnight alchemy: Sleep-dependent memory evolution. *Nature Reviews Neuroscience*, *11*(3), 218–218.

Walker, M. P., & van Der Helm, E. (2009). Overnight therapy? the role of sleep in emotional brain processing. *Psychological Bulletin*, *135*(5), 731.

Walker, W. R., Skowronski, J. J., & Thompson, C. P. (2003). Life is pleasant—and memory helps to keep it that way! *Review of General Psychology*, *7*(2), 203–210.

Wassing, R., Lakbila-Kamal, O., Ramautar, J. R., Stoffers, D., Schalkwijk, F., & Van Someren, E. J. (2019). Restless rem sleep impedes overnight amygdala adaptation. *Current Biology*, *29*(14), 2351–2358.

Whitmore, N. W., & Paller, K. A. (2023). Sleep disruption by memory cues selectively weakens reactivated memories. *Learning Memory*, *30*(3), 63–69.

Wiesner, C. D., Pulst, J., Krause, F., Elsner, M., Baving, L., Pedersen, A., & Göder, R. (2015). The effect of selective rem-sleep deprivation on the consolidation and affective evaluation of emotional memories. *Neurobiology of Learning and Memory*, *122*, 131–141.

Wilhelm, I., Diekelmann, S., Molzow, I., Ayoub, A., Mölle, M., & Born, J. (2011). Sleep selectively enhances memory expected to be of future relevance. *Journal of Neuroscience*, *31*(5), 1563–1569.

Witkowski, S., Schechtman, E., & Paller, K. A. (2020). Examining sleep's role in memory generalization and specificity through the lens of targeted memory reactivation. *Current Opinion in Behavioral Sciences*, *33*, 86–91.

Zoellner, L. A., Sacks, M. B., & Foa, E. B. (2001). Stability of emotions for traumatic memories in acute and chronic ptsd. *Behaviour Research and Therapy*, *39*(6), 697–711.