The relationship of memory consolidation with task incorporations into dreams – A registered report

Sarah F. Schoch[1,2]*, Somayeh Ataei [1,3], Leila Salvesen[1,4], Michael Schredl[5], Jennifer Windt[6,7], Giulio Bernardi[4], Björn Rasch[8], Nikolai Axmacher[3], and Martin Dresler[1]*

[1] Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands
[2] Center of Competence Sleep & Health Zurich, University of Zurich, Switzerland
[3] Ruhr University Bochum, Germany
[4] IMT School for Advanced Studies Lucca, Italy
[5] Central Institute of Mental Health, Mannheim, Germany
[6] Department of Philosophy, Monash University, Australia
[7] Monash Centre for Consciousness and Contemplative Studies, Monash University, Australia
[8] Cognitive Biopsychology and Methods, University of Fribourg, Switzerland

* Corresponding authors: Sarah F. Schoch (sarah.schoch@donders.ru.nl); Martin Dresler (martin.dresler@donders.ru.nl).

Abstract

Sleep is crucial for memory consolidation, but whether dreams play an essential role in memory consolidation is still unknown. This research will examine if incorporating a memory task into a dream benefits memory strength in a sleep-stage-dependent fashion. We will investigate spontaneous and experimentally induced incorporations using targeted memory reactivations. Ninety-two participants will be invited to spend two nights in the sleep laboratory, where they will learn a memory task before dream reports are collected. Memory performance will be measured before and after sleep as well as four days later.

Introduction

Memory is essential to humans throughout their lifespan, and sleep plays a crucial role in memory processing[for a review see 1]. It has been proposed that sleep provides an optimal brain state for memory consolidation[2]. However, it is unclear if the subjective experience during sleep, i.e., dreaming, plays a role in sleep-dependent memory consolidation processes.

Several studies have shown that dreams incorporate recent waking-life experiences[3,4]. In fact, the content of dreams can be influenced by having participants learn a specific task before sleep[5–7]. But whether this incorporation of a task into dreams is beneficial for memory consolidation remains inconclusive. A review article summarizing 12 published studies investigating the association between task incorporation into dreams and memory task performance has shown inconsistent results[8]. Seven studies have demonstrated at least a partial association between incorporating the memory task into the dream and subsequent memory performance[9–13]. Two early studies found that incorporating an explicit verbal memory task (story recall and language learning) into dreams is associated with better memory[11,13]. However, this effect was not found in another study that used meaningless sentences as stimuli[14]. For visuospatial tasks, Wamsley et al. showed an effect of incorporating a Maze task into dreams on memory performance both in a nap and overnight paradigms[10,12], but not in two other overnight studies[15,16]. A multisensory visuospatial task benefitted from the incorporation of both the task and the experimental setting[17]. For procedural tasks, an effect of dream incorporations was found for a virtual reality flying task[18], but not for a mirror tracing[19], balancing[20], or video game task[21].

There are several possible reasons why the findings so far have been discordant. One potential explanation is that the studies used memory tasks relying on different memory systems. Hippocampus-based declarative memory tasks have been more consistently shown to benefit from sleep than procedural memory tasks[22–25]. Therefore they might be more likely to benefit from incorporation into dreams. Further, the previous studies have several limitations, including the small sample sizes, with six studies relying on fewer than 20 subjects[11–14,20,26]. Often, very few participants incorporated the task into dreams (< 10%)[10,12,15,19,26], further reducing the sample size for testing possible associations. Therefore, many studies may have been underpowered to find associations, even if they existed.

A final possible explanation for the inconsistent results could be the different sleep stages during which the dream reports were collected. Some studies collected dream reports without distinguishing between the sleep stages in their analysis, while others focused only on either rapid eye movement sleep (REM) or non-REM sleep (NREM). Humans report dreams when awoken from all sleep stages[27]. However, dream reports are more frequent, longer, more emotional, and vivid upon awakenings from REM sleep[28]. The different sleep stages are also associated with a markedly different neurobiological background[29]. Therefore, it has been hypothesized that the different sleep stages are critical for different aspects of memory consolidation. Specifically, it has been proposed that during NREM sleep, there is a tight coupling between the hippocampus and neocortex, which leads to a high-fidelity replay of recent memories. In contrast, in REM sleep, memories are integrated with more remote memories and lower-fidelity replay, aiming to protect old knowledge from interference[30]. This could explain why many studies only find an association between NREM sleep with declarative memory strength the next morning[22,31,32].

The active systems consolidation hypothesis[33] proposes that sleep plays an active role in memory consolidation through spontaneous (i.e., not externally triggered) and repeated neural reactivations (i.e., activations of the same neurons in the same or reversed sequence), which have been measured in rodents[34–36] and suggested in humans[37–41]. The hypothesis suggests that reactivations in the hippocampus trigger associated reactivations in cortical areas orchestrated by slow waves and spindle-ripple events[42,43], both hallmarks of NREM sleep. Evidence for memory reactivations during REM sleep is more debated, potentially due to more remote memories being reactivated or the reactivations being of lower fidelity (and combining recent and remote memories). Studies in humans have shown that these reactivations can also be induced by presenting cues (e.g., sounds, odors) previously associated with the memory trace during sleep, so-called targeted memory reactivations (TMR)[44]. In rats, it has been shown that these cues induced neural reactivations related to the specific associated memory[45]. The evidence for memory-strengthening effects comes mainly from reactivating in NREM but not REM sleep[46–48], including a meta-analysis, which only found a significant effect for TMR in NREM sleep[49].

When dividing the studies included in the above-mentioned review by sleep stage the dream reports were collected from (ignoring studies that mixed dream reports from different sleep stages), we find weaker evidence for the association of REM dreams with memory performance (only 1 out of 5 studies with positive findings), while NREM dreams potentially show an association, but have been studied less (2/2 studies). Therefore, it seems plausible that dreams are biased by memory consolidation processes during sleep and thus reflect the specific consolidation processes happening during each sleep stage.

In summary, NREM and REM sleep appear to have complementary roles in memory consolidation during sleep. However, it is currently unclear whether dreams represent a functionless epiphenomenon of sleep-dependent memory processing or whether they play a direct role in sleep-dependent memory consolidation - and if so, whether that role differs for NREM and REM conscious experiences. In this study, we will use a declarative memory task (word-picture association task), which has previously been shown to be affected by sleep-dependent memory consolidation processes and suitable for TMR[46] and has a high incorporation rate into dreams[9], therefore overcoming many of the limitations of previous studies. We will use a serial awakening paradigm in NREM and REM sleep to systematically disentangle the effects of task incorporations on the different sleep stages. Furthermore, we will also address the sample size issue by collecting dream reports from 92 participants. Lastly, we will employ a two-step approach: spontaneous incorporations (correlational approach) and auditory TMR (experimental approach). Using TMR enables us to manipulate memory processes during sleep, therefore probing if we can experimentally modify dream content by inducing memory replay events. While a recent study has found that TMR did not affect the incorporation of a motor task into dreams[50], the study used only a single short reactivation period without collecting a dream report immediately afterward. During the sleep onset period, dream content has been successfully biased by using auditory stimulation[51], and during lucid dreams, participants were able to reply to questions presented aurally (among others)[52]. Using TMR also enables us to manipulate NREM and REM sleep independently. Considering that the function of dreams has long been a topic of interest and continues to be debated [53,54], this study will provide a large empirical dataset to understand two potential functions of dreaming: memory and emotional processing.

In this study, we will test the following hypotheses in a sample of 92 participants:

- Hypothesis 1a) Incorporations of the picture categories of the memory task into NREM dreams, but not REM dreams, are associated with improved performance on the memory task the next morning and 4-days later.

- Hypothesis 2) TMR leads to the subsequent incorporation of the associated image categories into dreams during both NREM and REM sleep stages.

Methods

Ethics information

The research was approved by the CMO Regio Arnhem-Nijmegen (NL75927.091.20). All participants will give written consent after the procedures have been fully explained. Participants will be paid 250 € for full participation.

Design

Procedure

Exact details on the procedure can be found in the supplemental methods. Here, we provide a brief overview of the study design.

Data will be collected in a within-subjects design across an intake session, adaptation night, and two experimental nights. After volunteers have signed up for the study, they will be invited to a short intake session. Volunteers fill out the informed consent and complete the screening questionnaires (see Supplementary Table 1 and Figure 2). If a volunteer is eligible to participate, they will receive a structural T1 and T2 magnetic resonance imaging (MRI) scan. Then the adaptation night and experimental sessions are scheduled. The participant gets a sleep tracker (Fitbit Inspire 2) and instructions on a sleep and dream diary. Participants will start wearing the tracker and digitally fill out the diaries for one week before the first experimental session.

For the adaptation night, participants will be invited to the Donders electroencephalography (EEG) laboratory at 21:30. The adaptation night enables participants to get used to the sleep laboratory environment and sleep while wearing the EEG cap. During the adaptation night, participants will complete a Stroop task and answer several questionnaires, including sleep and mood questionnaires. Participants will sleep while EEG, electrooculography (EOG), electromyography (EMG), electrocardiography (ECG), and electrogastrography (EGG, opt-in) are recorded. Participants will be provided with a sleep opportunity from 23:00 to 07:00. In the morning, they will fill out a questionnaire about their sleep quality and be asked to recall their dreams.

The two experimental sessions, separated by at least 14 days, will be counterbalanced between the participants with random assignment (see Figure 1). Both the order and the images used in the task will be randomized among all participants. Participants will be blinded to the experimental session. However, experimenters cannot be blinded. Both experimental sessions will start at 19:30 and end at approximately 8:30. Participants will fill out several questionnaires during the application of the electrodes. Again an EEG, EOG, EMG, ECG, and EGG are recorded. Participants will complete a memory task (word-picture association learning task) similar to the one used in a previous study[9] with three learning blocks and two recall blocks separated by a 10-minute break. Words are presented on two speakers 100 cm from the head on each side. In experimental session A, participants will be woken up a maximum of four times from NREM and four times from REM sleep, at least 15 minutes after the first start of the respective NREM/REM sleep stage. A free dream report for the last minute of sleep will be elicited during each awakening, followed by ratings on several scales. Then dream reports for previous parts of the dreams or previous dreams are collected and rated. Participants will have been trained to collect such dream reports concerning the minute preceding awakening during the week before each experimental session. In experimental session B, the awakenings are preceded by auditory cueing of the words used in the memory task (TMR). The words will be presented for 5 - 15 minutes before each awakening, and the awakening takes place 10 - 30 seconds after the last audio cue. The words associated with different image categories will be used as cues in NREM and REM sleep (with one remaining uncued category). The sleep opportunity will end at 7 am. After giving a detailed dream report, they will rate their sleep. Then they will complete another recall of the memory task. Lastly, they will do a localizer task.

Four days after each experimental session, there will be a follow-up on the memory recall performance using the same recall blocks.

Memory Task

To measure memory performance, we will use an adapted version of the word-picture association task we have used previously[9]. The task consists of 99 word-picture associations of neutral words with positive and neutral pictures, which are now extended with negative pictures. The pictures are related to 6 categories (3 different categories for each experimental night): mammals, vehicles, food, children, water, and buildings. Each category has 11 positive, 11 negative, and 11 neutral pictures. At the beginning of the task, one image unrelated to the categories will be presented at the very beginning (primacy effect). The pictures are taken from the NAPS, IAPS, NDPS, DIRTI, and Oasis databases which contain large sets of images rated on emotional valence and arousal[55–59]. Still, the images had to be supplemented with 55 images because not enough were available to fit our criteria (see supplemental info). The words are taken from the auditory English Lexicon (AELP) project[60]. The chosen words have two syllables as well as a similar length (636 – 805 ms), neutral valence and arousal (between 4 - 6), and be well known (> 88% recognition). Furthermore, words are selected not to contain any reference to the image categories. The association between words and pictures was done randomly but will be consistent across participants.

The memory task has six blocks: two rating blocks, two learning blocks (the second is repeated once), and two recall blocks. The recall blocks contain a valence/arousal recall and a cued recall, where participants hear the word and describe the associated picture with keywords.

Sleep Recording

EEG will be recorded with 64 channels cap (actiCAP original) and the BrainAMP by Brainproducts. Each electrode location will be prepared using an abrasive paste (Nuprep) and electrode paste (Abralyt). Impedances will be checked to be below 20 kΩ. Additionally, two electrodes will be used to measure EOG, ECG, and three electrodes for chin EMG (using BrainAMP ExG, impedance level below 10 kΩ) and an 8 channel EGG (participants can opt-out of the EGG if they are unable to sleep with it, impedance level below 25 kΩ). Data will be recorded with a 500 Hz sampling frequency and referenced to the vertex.

Targeted Memory Reactivation

The words from the word-picture association task will be used. The words associated with different image categories are used as cues in either NREM or REM sleep (with one category used as an uncued control). Words will be presented for maximally 15 minutes before each awakening after 3 minutes of stable sleep (NREM2/NREM3 or REM) has been reached. Words are presented starting from 30dB SPL via two loudspeakers situated 230 cm from the head of the subject. Sound levels will be increased until a K-complex (NREM), or arousal (REM) is elicited in each sleep stage and then kept at that sound level (NREM) or one below (REM) or to the maximum of 65dB SPL.

Sampling plan

Participants

Ninety-two healthy male and female volunteers aged 18-35 will be recruited from the general area around Nijmegen, Gelderland, Netherlands. The inclusion criteria to participate in the study are to be physically and mentally healthy, have a dream recall frequency of more than once a week, have high English language proficiency, and can sleep in the sleep laboratory. Exclusion criteria are history of or current sleep disorder, current physical or mental illness, intake of medication that influences sleep/wake cycle and/or memory consolidation, frequent coffee consumption (> 4 cups/day), skin disease at intended electrode sites, chronotype incompatible with the study time window, inability to sleep during adaptation night, contraindications for MRI (including pregnancy/nursing), irregular sleep pattern leading up to experimental sessions. Supplementary Table 1 reports the exact criteria for each inclusion/exclusion and the corresponding measurement used. Data will be excluded from single experimental nights if sleep duration is too short (<= 3 hours). The specific awakening is excluded if less than 85% of auditory cues are presented in the correct sleep stage or less than 5 minutes of auditory cueing can occur. Any participant replacements, dropouts, and exclusions will be reported.

Sample Size Calculation

We conducted a power analysis using simulations[61] based on the results of our previous study[9]. Simulations were done in RStudio[62] and using the packages *tidyverse*[63], *lme4*[64], *lmerTest*[65], *fitdistrplus*[66], *broom.mixed*[67], *faux*[68]. For hypothesis 1, we simulated datasets containing 10 – 120 participants (across 1000 repetitions) based on estimates from the data of our previous study (n = 22). 95% power was reached with 90 participants (suppl Fig 2a). Using a sensitivity analysis with 92 participants and 1000 repetitions while varying the beta for the interaction of interest (NREM incorporation * time) from 3.0 to 6.0 (in 1.0 steps), we estimate that b ≥ 5 will be detected with 95% power and b ≥ 3.9 with 80% power (b = 5.14 estimated from the previous study, suppl Fig 2b). The same sensitivity analysis was done for the model controlling incorporation for chance level, estimating 95% power for b ≥ 2.4 and 80% power for b ≥ 1.8 (b = 7.12 estimated from the previous study, range tested 0 – 6.0, suppl Fig 2c). For hypothesis 2, we simulated datasets based on data from our previous study on incorporating the task into the dreams (comparison task from before sleep and the one 10 weeks before/after). We estimate that the effect size of TMR will be similar (based on similar effect sizes reported for TMR on memory performance compared to general sleep effects). For 92 participants (1000 repetitions), we showed that the sensitivity of our analyses was 95% for $b \geq 0.4$ and 80% for $b \geq 0.3$ (0.45 estimated from the previous study).

Analysis Plan

EEG Data will be analyzed in MATLAB[69] using SpiSOP/Sleeptrip[70] and Fieldtrip[71]. Behavioral data will be analyzed using R and R Studio[72].

Sleep Scoring

EEG data will be imported into MATLAB. Data will be filtered (0.5 – 50 Hz bandpass Butterworth filter) and downsampled to 128 Hz. Data will then be re-referenced to Mastoids (F3/F4, C3/C4, O1/O2), and sleep will be scored in 30-second epochs using an automatic sleep scoring algorithm and one blind rater based on the AASM criteria[73]. A second rater will go over epochs where there is a disagreement between the algorithm and human scoring. Next, we will check if all the awakenings were in the correct sleep stage (preceding 60 s). If not, data for that awakening will be excluded. Then we will check that the reactivations were within the correct sleep stage. If < 85% of reactivations previous to an awakening are in the correct sleep stage, the awakening will be excluded from the analysis. We will calculate descriptive information on the sleep stages of the adaptation night and experimental nights (mean +/- sd).

Memory Task

We will average the performance score across all images. Two raters will rate the image description from the cued recall if the image description fits with the associated image. If the two raters disagree, they will discuss the disagreement and come to a final score. If the correct image is remembered, 1 will be assigned, otherwise, 0. We will then calculate a percentage of how many images were correctly remembered (0 – 100).

Dream Reports

Dream reports are recorded and later transcribed. The reports from the nighttime awakenings will be used to calculate the incorporation scores. Irrelevant information will be removed (e.g., “I dreamed that…”). Dreams will then be shuffled into a random order. The dreams will be rated by two independent raters blinded to condition and experimental night. Both raters will be trained beforehand. The raters will rate all dreams according to a prespecified manual on the incorporation of any of the image categories as well as of the laboratory and experimental setting and unusual auditory experiences. Furthermore, they will rate how realistic/bizarre the dreams were and the arousal and valence of the dreams, and the length of each dream. The ratings from the two raters will be compared to see if an acceptable agreement is reached (kappa > 0.6 for each category). If kappa is below that, the dreams will have to be re-rated. For the disagreements, a third trained blinded rater will decide on the final rating. Incorporations are analyzed as % of incorporated categories within each dream report across all dream reports from a specific sleep stage for hypothesis 1 and separately for each awakening per specific category for hypothesis 2.

Statistical Analysis

All statistical analyses will be performed in R Studio[62]. Analyses will be performed using the *lme4*[64] and *lmerTest*[65] packages for the multilevel models. Additionally, the packages *ggplot2*, *ggpubr*, *cowplot*, *RColorBrewer*, *plotly*, *sjPlot*, *dplyr*, *magrittr*, *tidyr*, *reshape*, *kableExtra* will be used for data handling and plotting[74–83]. First, we will examine outliers in each variable. Outliers will be inspected but not removed unless there is a reason to believe they are due to measurement error (e.g., the wrong task presented, audio not working, etc.). Our primary analyses are in a Null Hypothesis Significance Testing (NHST) framework) but are extended with a Bayesian Framework in the case of non-significant results.

Control Analyses

In the first step, we will run two control analyses to determine if our task was incorporated into dreams and if the TMR benefits memory performance.

To check if the task was successfully incorporated into dreams, we will run the following multilevel model with random intercepts:

Incorporation_Dreams ~ Sleep_stage + **Task** + (1 | SubjectID)

Incorporation_Dreams (numeric) will reflect the incorporation of all the task categories for each awakening separately across the task categories seen in this experimental night (% of 3 categories) and the categories seen in the other experimental night (% of 3 categories).

Sleep_stage (sum coded categorical) will reflect the sleep stage of the awakening (NREM = -0.5, REM = 0.5).

Task (sum coded categorial) will reflect if the incorporation is the task seen in this experimental session or the other one (other session = -0.5, this session = 0.5).

SubjectID (categorical) refers to the participant ID to model individual intercepts.

If *Task* shows a significant effect, we will interpret this as evidence that the task was incorporated into dreams beyond the level of random incorporations.

To control if the TMR worked, we will run the following multilevel model with random intercepts per participant

Correct_response_category ~ **TMR** + sleep_stage + (1 | SubjectID)

Correct_response_category (numeric) will be the memory performance per category (0 – 33 items).

TMR (dummy coded categorical) will reflect if TMR was performed for this category (no = 1, yes = 0).

Sleep_stage (sum coded categorical) will refer to the sleep stage the TMR was performed in (none = 0.5, NREM = -0.25, REM = -0.25).

SubjectID refers to the participant ID to model individual intercepts.

If *TMR* shows a significant effect, we will interpret this as evidence that TMR significantly improved memory performance. Furthermore, we can look at the effect of *Sleep_stage* to examine if this was evident for both NREM and REM sleep.

We will analyze our two hypotheses regardless of the control analyses, however, if either control analysis fails to show an effect, then the interpretation of the results will be limited.

Hypothesis 1

To analyze H1, we will run two models, one including the raw incorporation rates of the task categories into dreams and one with adjusted incorporation rates by the baseline level estimate from the incorporation in the other night.

The primary multilevel model with random intercept per participant is the following:

Correct_response ~ Timepoint + Night (spontaneous/TMR) + NREM_Dream_Incorporations + REM_Dream_Incorporations + **NREM_Dream_Incorporations:Timepoint** + **REM_Dream_Incorporations:Timepoint** + (1 | SubjectID/Night)

Correct response (numeric) reflects the number of correctly remembered images (0 – 99).
Timepoint (dummy coded categorical) reflects the timepoint of recall (Evening = 0, Morning = 1, Follow up = 1).

Night (sum coded categorical) reflects which experimental night (Sponteanous = -0.5, TMR = 0.5).

NREM_Dream_Incoporations (numeric) reflects the incorporation percentage of the task seen in the experimental night across all reported NREM dreams.

REM_Dream_Incoporations (numeric) reflects the incorporation percentage of the task seen in the experimental night across all reported REM dreams.

NREM_Dream_Incorporations:Timepoint (interaction) Interaction effect to quantify changes between baseline (evening) and morning/follow-up dependent on incorporations into NREM dreams.

REM_Dream_Incorporations:Timepoint (interaction) Interaction effect to quantify changes between baseline (evening) and morning/follow-up dependent on incorporations into REM dreams.

SubjectID refers to the participant ID to model individual intercepts.

The secondary multilevel model will be the same except that the incorporations are conceptualized differently. Incorporations will be difference scores between incorporation in the experimental night when the image category was presented compared to 'incorporation' (spontaneous appearance) in the other night.

Correct_response ~ Timepoint + Night + NREM_inc_cor + REM_inc_cor + **NREM_inc_cor:Timepoint + REM_inc_cor:Timepoint** + (1 | SubjectID/Night)

NREM_inc_cor (numerical) reflects incorporation into NREM dreams in the night the image category was presented minus incorporations in the other night.

REM_inc_cor (numerical) reflects incorporation into REM dreams in the night the image category was presented minus incorporations in the other night.

NREM_inc_cor:Timepoint (Interaction) Interaction effect to quantify changes between baseline (evening) and morning/follow-up dependent on incorporations into NREM dreams (baseline-adjusted).

REM_inc_cor:Timepoint (Interaction) Interaction effect to quantify changes between baseline (evening) and morning/follow-up dependent on incorporations into REM dreams (baseline-adjusted).

If the interaction *NREM_Dream_Incorporations:Timepoint* is significant in either model, we will interpret this as evidence for H1 that NREM dream incorporations are significantly associated with memory performance after sleep. If the interaction *REM_Dream_Incorporations:Timepoint* is significant in either model, we will interpret this as evidence against H1 that REM dream incorporations are not significantly associated with memory performance after sleep. If the interaction is only significant in the secondary but not primary model this means that baseline adjustment for dream incorporations is necessary to detect association with memory performance.

Hypothesis 2

For hypothesis 2, we will run the following generalized multilevel model (binomial distribution) using random intercepts:

Incorporation_Dreams ~ **Cued_Topic** + Sleep_stage + (1 | SubjectID)

Incorporation_Dreams (numeric) will reflect the incorporation of the task category (separately) for each awakening individually across the task categories seen in this experimental night (% of 3 categories)

Cued_topic (dummy coded categorical) will reflect if the topic was cued prior to the awakening or not (yes = 0, no = 1)

Sleep_stage (sum coded categorical) will reflect the sleep stage from which the awakening occurred (NREM = -0.5, REM = 0.5)

If *Cued_Topic* is significant, we will interpret this as evidence for H2, meaning that TMR significantly influences dream content. Furthermore, if *Sleep_stage* is significant, we will interpret this as evidence that this effect depends on the sleep stage (i.e., it works better in one of the sleep stages).

If the initial NHST results in a p-value above our 0.5 alpha threshold for the specified fixed effects, we plan to explore further the extent to which our data provides evidence against/for our hypotheses by using Bayesian methods, specifically Bayes factors BF01 to quantify how much more likely the null hypothesis is relative to the alternative hypothesis. We will use the *bmrs*[84] and *BayesFactor package*[85] to implement the Bayesian analyses. We will use a balanced null comparison to test for the presence/absence of the fixed effect[86]. We will follow the guidelines proposed by [87] and consider the evidence to be: inconclusive/null if BF01 = 1; weak in favor of H0 if 1 < BF01 < 3; moderate in favor of H0 if 3 < BF01 < 10; strong in favor of H0 if 10 < BF01 < 30; weak in favor of H1 if 1/3 < BF01 < 1; moderate in favor of H1 if 1/10 < BF01 < 1/3; strong in favor of H1 if 1/30 < BF01 < 1/10.

To ensure the robustness of the results, models will be additionally analyzed with outliers (> 3 SD for each specific measure) removed at the cell level. While interpretations will be based on the models with outliers included, these additional analyses will be used to interpret if the effects are robust or dependent on a few participants with extreme values.

Data availability

All data used in this manuscript will be available on the Donders Data Repository and the DREAM database for the Stage 2 review.

Code availability

Code will be made available on the Donders Data Repository and OSF for the Stage 2 review and will be made public upon acceptance.

Results

Do **not** include a **Results** section.

Discussion

Do **not** include a **Discussion** section.

References

1. Rasch, B. & Born, J. About sleep's role in memory. *Physiological Reviews* **93**, 681–766 (2013).
2. Marshall, L., Mölle, M., Hallschmid, M. & Born, J. Transcranial Direct Current Stimulation during Sleep Improves Declarative Memory. *J. Neurosci.* **24**, 9985–9992 (2004).
3. Schredl, M. & Hofmann, F. Continuity between waking activities and dream activities. *Consciousness and Cognition* **12**, 298–308 (2003).
4. Fosse, M. J., Fosse, R., Hobson, J. A. & Stickgold, R. J. Dreaming and episodic memory: a functional dissociation? *Journal of cognitive neuroscience* **15**, 1–9 (2003).
5. Wamsley, E. J., Perry, K., Djonlagic, I., Reaven, L. B. & Stickgold, R. Cognitive replay of visuomotor learning at sleep onset: temporal dynamics and relationship to task performance. *Sleep* **33**, 59–68 (2010).
6. Stickgold, R., Malia, a, Maguire, D., Roddenberry, D. & O'Connor, M. Replaying the game: hypnagogic images in normals and amnesics. *Science (New York, N.Y.)* **290**, 350–353 (2000).
7. Oudiette, D. *et al.* Evidence for the Re-Enactment of a Recently Learned Behavior during Sleepwalking. *PLOS ONE* **6**, e18056 (2011).
8. Plailly, J., Villalba, M., Vallat, R., Nicolas, A. & Ruby, P. Incorporation of fragmented visuo-olfactory episodic memory into dreams and its association with memory performance. *Scientific Reports* (2019) doi:10.1038/s41598-019-51497-y.
9. Schoch, S. F., Cordi, M. J., Schredl, M. & Rasch, B. The effect of dream report collection and dream incorporation on memory consolidation during sleep. *Journal of Sleep Research* **28**, e12754 (2019).

10. Wamsley, E. J., Tucker, M., Payne, J. D., Benavides, J. A. & Stickgold, R. Dreaming of a Learning Task Is Associated with Enhanced Sleep-Dependent Memory Consolidation. *Current Biology* **20**, 850–855 (2010).
11. de Koninck, J., Christ, G., Hébert, G. & Rinfret, N. Language learning efficiency, dreams and REM sleep. *Psychiatric Journal of the University of Ottawa* **15**, 91–92 (1990).
12. Wamsley, E. J. & Stickgold, R. Dreaming of a learning task is associated with enhanced memory consolidation: Replication in an overnight sleep study. *Journal of Sleep Research* **28**, e12749 (2019).
13. Fiss, H., Kremer, E. & Litchman, J. The mnemonic function of dreaming. *Sleep Res* **6**, 122 (1977).
14. Cipolli, C., Fagioli, I., Mazzetti, M. & Tuozzi, G. Incorporation of presleep stimuli into dream contents: evidence for a consolidation effect on declarative knowledge during REM sleep? *Journal of Sleep Research* **13**, 317–326 (2004).
15. Wamsley, E. J., Hamilton, K., Graveline, Y., Manceor, S. & Parr, E. Test Expectation Enhances Memory Consolidation across Both Sleep and Wake. *PLOS ONE* **11**, e0165141 (2016).
16. Stamm, A. W. *et al.* Negative reinforcement impairs overnight memory consolidation. *Learn. Mem.* **21**, 591–596 (2014).
17. Plailly, J., Villalba, M., Vallat, R., Nicolas, A. & Ruby, P. Incorporation of fragmented visuo-olfactory episodic memory into dreams and its association with memory performance. *Scientific Reports* **9**, 15687 (2019).
18. Picard-Deland, C., Aumont, T., Samson-Richer, A., Paquette, T. & Nielsen, T. Whole-body procedural learning benefits from targeted memory reactivation in REM sleep and task-related dreaming. *Neurobiology of Learning and Memory* **183**, 107460 (2021).

19. Schredl, M. & Erlacher, D. Is sleep-dependent memory consolidation of a visuo-motor task related to dream content? (2010) doi:10.11588/IJODR.2010.1.486.
20. Nefjodov, I., Winkler, A. & Erlacher, D. Balancing in dreams: Effects of playing games on the Wii balance board on dream content. (2016) doi:10.7892/BORIS.85628.
21. Pantoja, A. L. *et al.* Assessment of the adaptive value of dreams. (2009).
22. Tucker, M. A. *et al.* A daytime nap containing solely non-REM sleep enhances declarative but not procedural memory. *Neurobiology of Learning and Memory* **86**, 241–247 (2006).
23. Barham, M. P., Enticott, P. G., Conduit, R. & Lum, J. A. Transcranial electrical stimulation during sleep enhances declarative (but not procedural) memory consolidation: evidence from a meta-analysis. *Neuroscience & Biobehavioral Reviews* **63**, 65–77 (2016).
24. Prehn-Kristensen, A. *et al.* Sleep in children enhances preferentially emotional declarative but not procedural memories. *Journal of experimental child psychology* **104**, 132–139 (2009).
25. Wilhelm, I., Diekelmann, S. & Born, J. Sleep in children improves memory performance on declarative but not procedural tasks. *Learning & memory* **15**, 373–377 (2008).
26. Koninck, J. D., Prévost, F. & Lortie-Lussier, M. Vertical inversion of the visual field and REM sleep mentation. *Journal of Sleep Research* **5**, 16–20 (1996).
27. Foulkes, W. D. Dream reports from different stages of sleep. *The Journal of Abnormal and Social Psychology* **65**, 14–25 (1962).
28. Moorcroft, W. H. & Belcher, P. Understanding sleep and dreaming. *Understanding Sleep and Dreaming* 1–339 (2005) doi:10.1007/0-387-28698-5.
29. McCarley, R. W. Neurobiology of REM and NREM sleep. *Sleep Medicine* **8**, 302–330 (2007).

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

31. Holz, J. *et al.* EEG sigma and slow-wave activity during NREM sleep correlate with overnight declarative and procedural memory consolidation. *Journal of Sleep Research* (2012) doi:10.1111/j.1365-2869.2012.01017.x.

32. Mander, B. a *et al.* Prefrontal atrophy, disrupted NREM slow waves and impaired hippocampal-dependent memory in aging. *Nature neuroscience* **16**, 357–64 (2013).

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

34. Hirase, H., Leinekugel, X., Czurkó, A., Csicsvari, J. & Buzsáki, G. Firing rates of hippocampal neurons are preserved during subsequent sleep episodes and modified by novel awake experience. *Proc Natl Acad Sci U S A* **98**, 9386–9390 (2001).

35. Pavlides, C. & Winson, J. Influences of hippocampal place cell firing in the awake state on the activity of these cells during subsequent sleep episodes. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* **9**, 2907–2918 (1989).

36. Wilson, M. A. & McNaughton, B. L. Reactivation of Hippocampal Ensemble Memories During Sleep. *Science* **265**, 676–679 (1994).

37. Peigneux, P. *et al.* Are Spatial Memories Strengthened in the Human Hippocampus during Slow Wave Sleep? *Neuron* **44**, 535–545 (2004).

38. Deuker, L. *et al.* Memory Consolidation by Replay of Stimulus-Specific Neural Activity. *J. Neurosci.* **33**, 19373–19383 (2013).

39. Jiang, X. *et al.* Replay of large-scale spatio-temporal patterns from waking during subsequent NREM sleep in human cortex. *Scientific reports* **7**, 1–17 (2017).

40. Zhang, H., Fell, J. & Axmacher, N. Electrophysiological mechanisms of human memory consolidation. *Nat Commun* **9**, 4103 (2018).
41. Rubin, D. B. *et al.* Learned Motor Patterns Are Replayed in Human Motor Cortex during Sleep. *J. Neurosci.* **42**, 5007–5020 (2022).
42. Staresina, B. P. *et al.* Hierarchical nesting of slow oscillations, spindles and ripples in the human hippocampus during sleep. *Nature Neuroscience* **18**, 1679–1686 (2015).
43. Schreiner, T., Petzka, M., Staudigl, T. & Staresina, B. P. Endogenous memory reactivation during sleep in humans is clocked by slow oscillation-spindle complexes. *bioRxiv* 2020.09.16.299545 (2020) doi:10.1101/2020.09.16.299545.
44. Rasch, B., Büchel, C., Gais, S. & Born, J. Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science* **315**, 1426–1429 (2007).
45. Bendor, D. & Wilson, M. A. Biasing the content of hippocampal replay during sleep. *Nature Neuroscience* (2012) doi:10.1038/nn.3203.
46. Lehmann, M., Schreiner, T., Seifritz, E. & Rasch, B. Emotional arousal modulates oscillatory correlates of targeted memory reactivation during NREM, but not REM sleep. *Scientific Reports* (2016) doi:10.1038/srep39229.
47. Cousins, J. N., El-Deredy, W., Parkes, L. M., Hennies, N. & Lewis, P. A. Cued memory reactivation during slow-wave sleep promotes explicit knowledge of a motor sequence. *Journal of Neuroscience* (2014) doi:10.1523/JNEUROSCI.1011-14.2014.
48. Cordi, M. J., Diekelmann, S., Born, J. & Rasch, B. No effect of odor-induced memory reactivation during REM sleep on declarative memory stability. *Frontiers in Systems Neuroscience* (2014) doi:10.3389/fnsys.2014.00157.
49. Hu, X., Cheng, L. Y., Chiu, M. H. & Paller, K. A. Promoting memory consolidation during sleep: A meta-analysis of targeted memory reactivation. *Psychological bulletin* **146**, 218 (2020).

50. Picard-Deland, C., Samson-Richer, A., Paquette, T. & Nielsen, T. Immediate and delayed effects of targeted memory reactivation during sleep on procedural learning and dream content. *Sleep Medicine* **64**, S300–S301 (2019).

51. Haar Horowitz, A., Cunningham, T. J., Maes, P. & Stickgold, R. Dormio: A targeted dream incubation device. *Consciousness and Cognition* **83**, 102938 (2020).

52. Konkoly, K. *et al.* Real-Time Dialogue between Experimenters and Dreamers During rem Sleep. (2021).

53. Scarpelli, S., Bartolacci, C., D'Atri, A., Gorgoni, M. & De Gennaro, L. The Functional Role of Dreaming in Emotional Processes. *Front. Psychol.* **0**, (2019).

54. Schredl, M. Functions of Dreaming. in *Researching Dreams: The Fundamentals* (ed. Schredl, M.) 175–181 (Springer International Publishing, 2018). doi:10.1007/978-3-319-95453-0_9.

55. Lang, P. J., Bradley, M. M. & Cuthbert, B. N. International affective picture system (IAPS): Affective ratings and instruction manual. *The Center for Research in Psychophysiology, University of Florida* (2005).

56. Marchewka, A., Żurawski, Ł., Jednoróg, K. & Grabowska, A. The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behav Res* **46**, 596–610 (2014).

57. Kurdi, B., Lozano, S. & Banaji, M. R. Introducing the Open Affective Standardized Image Set (OASIS). *Behav Res Methods* **49**, 457–470 (2017).

58. Merlhiot, G., Mermillod, M., Pennec, J.-L. L. & Mondillon, L. Introduction and validation of the Natural Disasters Picture System (NDPS). *PLOS ONE* **13**, e0201942 (2018).

59. Haberkamp, A., Glombiewski, J. A., Schmidt, F. & Barke, A. The DIsgust-RelaTed-Images (DIRTI) database: Validation of a novel standardized set of disgust pictures. *Behaviour Research and Therapy* **89**, 86–94 (2017).
60. Goh, W. D., Yap, M. J. & Chee, Q. W. The Auditory English Lexicon Project: A multi-talker, multi-region psycholinguistic database of 10,170 spoken words and nonwords. *Behavior Research Methods* 1–30 (2020).
61. DeBruine, L. M. & Barr, D. J. Understanding Mixed-Effects Models Through Data Simulation. *Advances in Methods and Practices in Psychological Science* **4**, 2515245920965119 (2021).
62. R Core Team, Rf. R: A language and environment for statistical computing. (2013).
63. Wickham, H. *et al.* Welcome to the Tidyverse. *Journal of open source software* **4**, 1686 (2019).
64. Bates, D., Sarkar, D., Bates, M. D. & Matrix, L. The lme4 package. (2007).
65. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. lmerTest package: tests in linear mixed effects models. *Journal of statistical software* **82**, 1–26 (2017).
66. Delignette-Muller, M. L. & Dutang, C. fitdistrplus: An R package for fitting distributions. *Journal of statistical software* **64**, 1–34 (2015).
67. Bolker, B. & Robinson, D. broom. mixed: Tidying Methods for Mixed Models. R package version 0.2. 6. (2020).
68. DeBruine, L. Faux: Simulation for factorial designs. *R package version 0.0* **1**, (2020).
69. Toolbox, S. M. Matlab. *Mathworks Inc* (1993).
70. Weber, F. SleepTrip as a Tool to Create Analysis Pipelines for Investigations of Polysomnography in Clinical Populations. in *International Journal of Psychophysiology* vol. 168 S65–S65 (Elsevier RADARWEG 29, 1043 NX AMSTERDAM, NETHERLANDS, 2021).

71. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience* **2011**, e156869 (2010).
72. Team, Rs. RStudio: integrated development for R. *RStudio, Inc., Boston, MA URL http://www. rstudio. com* **42**, 14 (2015).
73. Iber, C., Ancoli-Israel, S., Chesson, A. L. & Quan, S. F. *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*. vol. 1 (American academy of sleep medicine Westchester, IL, 2007).
74. Wickham, H. ggplot2: elegant graphics for data analysis. (2016).
75. Kassambara, A. ggpubr:"ggplot2" based publication ready plots. *R package version 0.1* **7**, (2018).
76. Wilke, C. O. cowplot: Streamlined Plot Theme and Plot Annotations for "ggplot2". R package version 0.9. 4. (2019).
77. Neuwirth, E. RColorBrewer: ColorBrewer palettes. R package version 1.1-2. (2014).
78. Sievert, C. plotly for R. (2018).
79. Lüdecke, D. sjPlot: Data visualization for statistics in social science. (2018).
80. Wickham, H., Francois, R., Henry, L. & Müller, K. dplyr: A grammar of data manipulation. (2015).
81. Bache, S. M. & Wickham, H. magrittr: a forward-pipe operator for R. (2014).
82. Wickham, H. Reshaping data with the reshape package. *Journal of statistical software* **21**, 1–20 (2007).
83. Zhu, H. KableExtra: Construct complex table with'kable'and pipe syntax. (2018).
84. Bürkner, P.-C. brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software* **80**, 1–28 (2017).

85. Morey, R. D., Rouder, J. N., Jamil, T. & Morey, M. R. D. Package ‘bayesfactor’. *URLh http://cran/r-projectorg/web/packages/BayesFactor/BayesFactor pdf i (accessed 1006 15)* (2015).
86. Linde, M. & van Ravenzwaaij, D. Bayes Factor Model Comparisons Across Parameter Values for Mixed Models. *Comput Brain Behav* (2021) doi:10.1007/s42113-021-00117-y.
87. van Doorn, J. *et al.* The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychon Bull Rev* **28**, 813–826 (2021).
88. Erdodi, L. A., Jongsma, K. A. & Issa, M. The 15-item version of the Boston Naming Test as an index of English proficiency. *The Clinical Neuropsychologist* **31**, 168–178 (2017).
89. Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R. & Kupfer, D. J. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* **28**, 193–213 (1989).
90. Beck, A. T., Steer, R. A. & Brown, G. Beck depression inventory–II. *Psychological Assessment* (1996).
91. Steer, R. A. & Beck, A. T. Beck Anxiety Inventory. (1997).
92. Schredl, M., Berres, S., Klingauf, A., Schellhaas, S. & Göritz, A. S. The Mannheim Dream Questionnaire (MADRE): Retest reliability, age and gender effects. *International Journal of Dream Research* **7**, 141–147 (2014).
93. Roenneberg, T., Wirz-Justice, A. & Merrow, M. Life between clocks: daily temporal patterns of human chronotypes. *Journal of biological rhythms* **18**, 80–90 (2003).
94. Siclari, F., LaRocque, J., Postle, B. & Tononi, G. Assessing sleep consciousness within subjects using a serial awakening paradigm. *Frontiers in Psychology* **4**, 542 (2013).
95. Görtelmeyer, R. SF-A/R und SF-B/R: Schlaffragebogen A und B. (2011).

96. Steyer, R., Schwenkmezger, P., Notz, P. & Eid, M. Testtheoretische Analysen des Mehrdimensionalen Befindlichkeitsfragebogen (MDBF). [Theoretical analysis of a multidimensional mood questionnaire (MDBF).]. *Diagnostica* **40**, 320–328 (1994).

97. Giambra, L. M. The influence of aging on spontaneous shifts of attention from external stimuli to the contents of consciousness. *Experimental Gerontology* **28**, 485–492 (1993).

98. Carver, C. S. You want to measure coping but your protocol' too long: Consider the brief cope. *Int. J. Behav. Med.* **4**, 92 (1997).

99. Webster, D. M. & Kruglanski, A. W. Individual differences in need for cognitive closure. *Journal of personality and social psychology* **67**, 1049 (1994).

100. Walach, H., Buchheld, N., Buttenmüller, V., Kleinknecht, N. & Schmidt, S. Measuring mindfulness—the Freiburg Mindfulness Inventory (FMI). *Personality and Individual Differences* **40**, 1543–1555 (2006).

101. Corrigan, J. D. & Hinkeldey, N. S. Relationships between parts A and B of the Trail Making Test. *J Clin Psychol* **43**, 402–409 (1987).

102. Esfahani, M. J., Daraie, A. H., Weber, F. D. & Dresler, M. Dreamento: an open-source dream engineering toolbox for sleep EEG wearables. Preprint at https://doi.org/10.48550/arXiv.2207.03977 (2022).

103. Schredl, M., Rieger, J. & Göritz, A. S. Measuring lucid dreaming skills: A new questionnaire (LUSK). *International Journal of Dream Research* **11**, 54–61 (2018).

104. Marks, D. F. Visual imagery differences in the recall of pictures. *British journal of Psychology* **64**, 17–24 (1973).

105. Rosenberg, M. *Society and the adolescent self-image*. (Princeton university press, 2015).

106. Carver, C. & White, T. Behavioral Inhibition, Behavioral Activation, and Affective Responses to Impending Reward and Punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology* **67**, 319–333 (1994).

Acknowledgments

This study was funded by the Swiss National Science Foundation (P2ZHP1_195248 to S.F.S), the Cogito foundation (20_123-S to S.F.S., M.D., and J.W.), the DreamScience Foundation, and International Association for the Study of Dreams (IASD) (to S.F.S. and M.D.), and the Netherlands Organization for Scientific Research (NWO 016.Vidi.185.142 to M.D.) the European Research Council (ERC-2020-STG #948891 to G.B.), the International Brain Research Organization (IBRO Early Career Award 2019 to G.B.), the German Academic Exchange Service (DAAD, Research Grants - Doctoral Programmes in Germany, 2021/22 to S.A.), and Australian Research Council Discovery Early Career Researcher Award (DE170101254 to J.W.) and National Health and Medical Research Council Ideas Grant (APP2002454 to J.W.) and the Boehringer Ingelheim Fonds (BIF) travel grant (to L.S.). Further, we received support from the COS for this registered report. The funders have/had no role in study design, data collection, analysis, decision to publish, or manuscript preparation.

Author contributions

B.R.: Conceptualization, Methodology, Software, and Writing - review & editing. **G.B.:** Conceptualization, Methodology, Supervision, and Writing - review & editing. **J.W.:** Conceptualization, Funding acquisition, Methodology, and Writing - review & editing. **L.S.:** Conceptualization, Investigation, Methodology, Software, and Writing - review & editing. **M.D.:** Conceptualization, Funding acquisition, Resources, Supervision, and Writing - review & editing. **M.S.:** Conceptualization, Methodology, Supervision, and Writing - review & editing. **N.A.:** Conceptualization, Supervision, and Writing - review & editing. **S.A.:** Data curation, Investigation, and Writing - review & editing. **S.F.S.:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft, and Writing - review & editing.

Competing interests

The authors declare no competing interests.

Figures

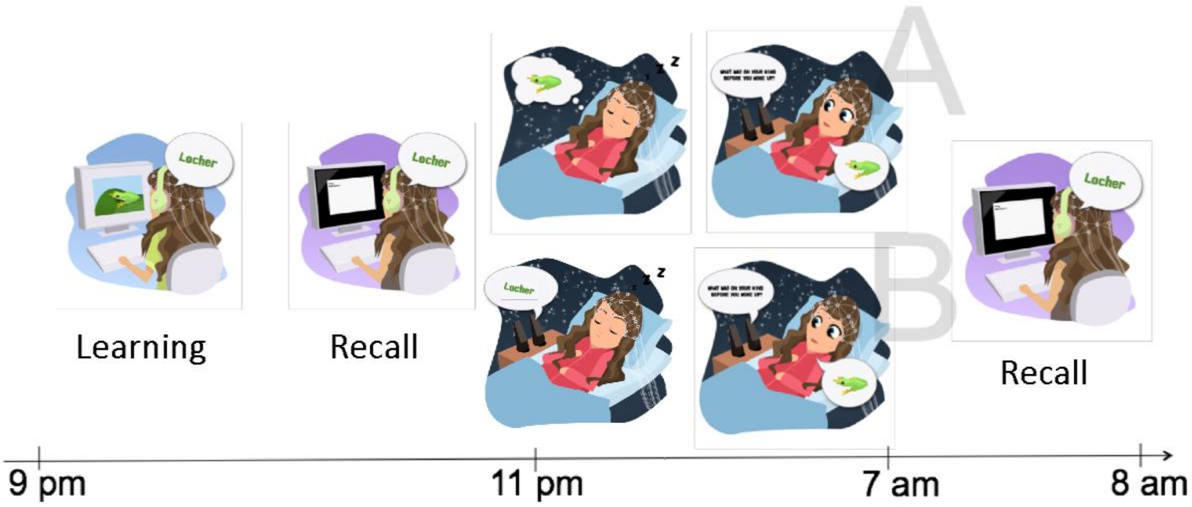


Fig 1. The procedure of the two experimental nights. On both nights, participants will learn a task with a recall session before and after sleep, and dream reports will be collected from NREM and REM sleep. In night B, targeted memory reactivation will be applied for approximately 15 minutes prior to awakenings.

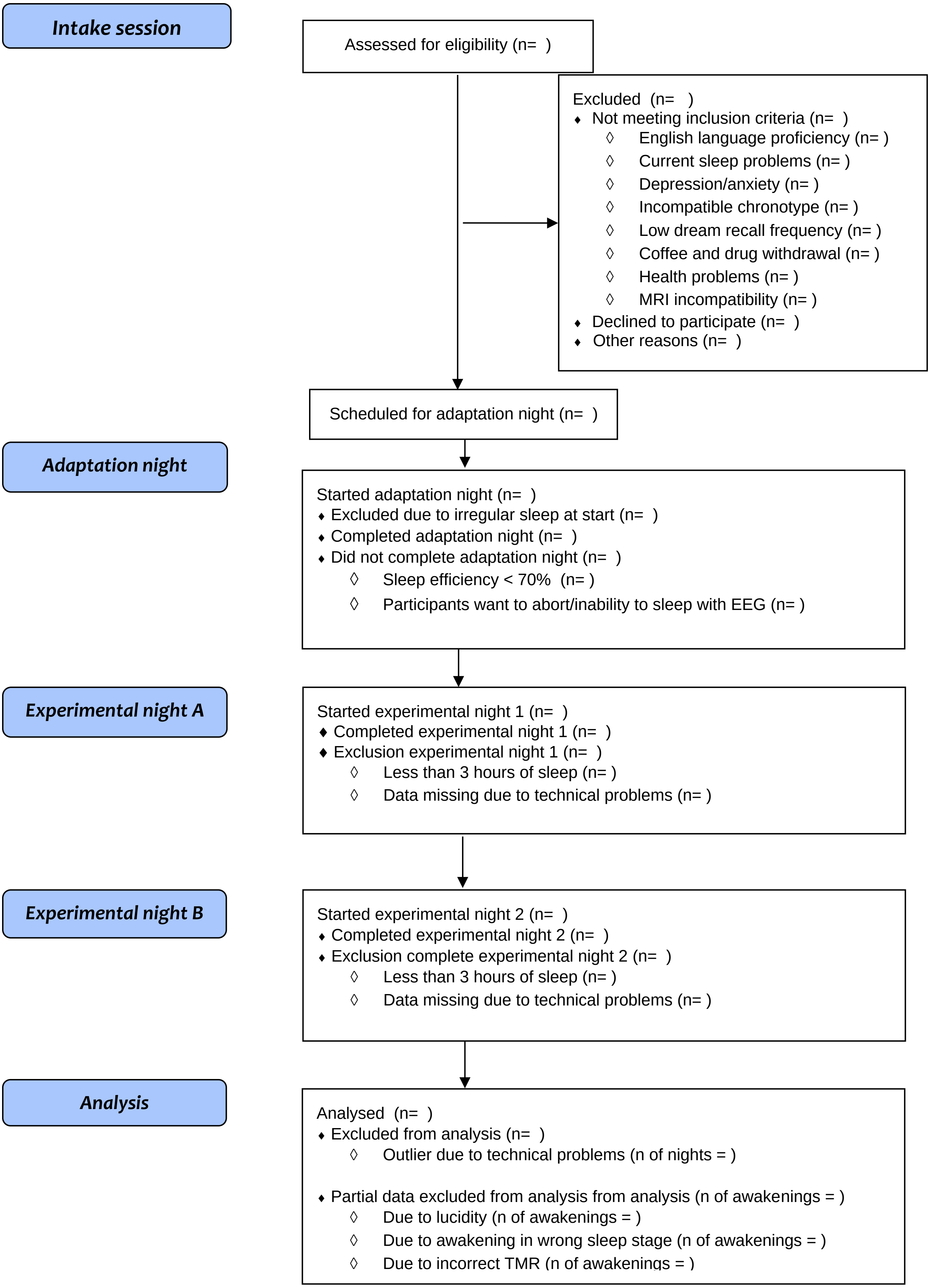

Intake session
Assessed for eligibility (n=)
Excluded (n=)
Not meeting inclusion criteria (n=)
English language proficiency (n=)
Current sleep problems (n=)
Depression/anxiety (n=)
Incompatible chronotype (n=)
Low dream recall frequency (n=)
Coffee and drug withdrawal (n=)
Health problems (n=)
MRI incompatibility (n=)
Declined to participate (n=)
Other reasons (n=)
Scheduled for adaptation night (n=)
Adaptation night
Started adaptation night (n=)
Excluded due to irregular sleep at start (n=)
Completed adaptation night (n=)
Did not complete adaptation night (n=)
Sleep efficiency < 70% (n=)
Participants want to abort/inability to sleep with EEG (n=)
Experimental night A
Started experimental night 1 (n=)
Completed experimental night 1 (n=)
Exclusion experimental night 1 (n=)
Less than 3 hours of sleep (n=)
Data missing due to technical problems (n=)
Experimental night B
Started experimental night 2 (n=)
Completed experimental night 2 (n=)
Exclusion complete experimental night 2 (n=)
Less than 3 hours of sleep (n=)
Data missing due to technical problems (n=)
Analysis
Analysed (n=)
Excluded from analysis (n=)
Outlier due to technical problems (n of nights =)
Partial data excluded from analysis from analysis (n of awakenings =)
Due to lucidity (n of awakenings =)
Due to awakening in wrong sleep stage (n of awakenings =)
Due to incorrect TMR (n of awakenings =)

Figure 2. **CONSORT Style diagram of inclusion and exclusion across the different steps of the study.**

Table 1. Design Table

Question	Hypothesis	Sampling plan (e.g., power analysis)	Analysis Plan	Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis	Interpretation given to different outcomes	Theory that could be shown wrong by the outcomes
Control analyses	Images of the task learned prior to sleep are incorporated more often into dream content	NHST N = 92	Incorporation_Dreams ~ **Task** + Sleep_stage + (1 | SubjectID)	Sample size determined by H1/H2	Task P < 0.05 Task is significantly more often incorporated as expected from random incorporations P > 0.05 Task incorporation could be random	If task incorporation is random and not above chance, this would greatly limit the interpretation of the study.
Control analyses	TMR was successful in improving memory performance	NHST N = 92	Correct_response_category ~ **TMR** + Sleep_stage + (1 | SubjectID)	Sample size determined by H1/H2	TMR P < 0.05 We see an effect of TMR on memory performance P > 0.05 no effect of TMR on memory performance	If TMR does not show an effect on memory performance, this will limit the interpretation of hypothesis 2.
Are task incorporations into dreams associated with the memory strength of the task (measured as memory performance) in a sleep-stage-	H1) Incorporations of the picture categories of the memory task during NREM dreams are associated with improved performance on the memory task	NHST N = 92 based on simulations from data from the previous study	*Primary multilevel model* Correct_response ~ Timepoint + Night (spontaneous/TMR) + NREM_Dream_Incorporations + REM_Dream_Incorporations + **NREM_Dream_Incorporations:Timepoint** + **REM_Dream_Incorporation**	Simulation of 1000 datasets based on estimates from the previous study, with 92 participants, we have 95% power to detect effect sizes similar to the previous study	NREM_Dream_Incorporations_Experimental_Night:Timepoint in either model P < 0.05 Support for H1 P > 0.05 (in both models) Follow up Bayes analysis 1<BF<10 = unclear evidence 10<BF<30 = strong evidence for H0	Task incorporation into NREM sleep is not significantly associated with memory strength.

dependent fashion?	the next morning and 4-days later.		**s:Timepoint** + (1 | SubjectID/Night) *Secondary Multilevel model correcting for baseline incorporation of each category (frequency in the other night)* Correct_response ~ Timepoint + Night + NREM_inc_cor + REM_inc_cor + **NREM_inc_cor:Timepoint + REM_inc_cor:Timepoint** + (1 | SubjectID/Night)		BF>30 = very strong evidence for H0 If either model shows a significant effect this is support for H1, however, interpretation is different. If the secondary model is significant but not the primary this means that only when adjusting for the baseline effects of task in dreams can a significant effect be detected.	
Does TMR influence dream content?	H2: TMR leads to subsequent incorporation of the associated image categories into dreams during NREM and REM sleep stages.	NHST N = 92 based on simulations from the previous study	Multilevel generalized model (binomial distribution) Incorporation_Dreams ~ **Cued_Topic** + Sleep_stage + (1 | SubjectID)	Simulation of 1000 datasets based on estimates of task incorporation vs. random incorporation into dreams from the previous study, with 92 participants, we have 95% power to detect effect sizes in the range that the memory task had an influence on incorporation, as TMR data is not directly available. However, based on the literature, TMR effects should be similar in effect size.	Cued_Topic P < 0.05 Support for H2 P > 0.05: Follow up Bayes analysis 1<BF<10 = unclear evidence 10<BF<30 = strong evidence for H0 BF>30 = very strong evidence for H0 Sleep_stage P < 0.05 Support that this is sleep stage-dependent P > 0.05 no support that this is sleep stage-dependent	TMR does not significantly influence dream content; therefore, dreaming does not directly reflect memory consolidation processes.

Supplementary information

Supplemental Methods

Design

We will collect data in a within-subjects design across an intake session, adaptation night, and two experimental nights. The study, including all questionnaires, will be conducted in English. This registered report will not analyze several measures collected within the study.

Recruitment: Volunteers will be recruited via the SONA database of the Donders Institute, social media, and physical notice boards. After participants have signed up for the study, a telephone call will explain the details of the study, and the study information will be provided by email. Participants will then be invited to a short intake session (1 hour). A brief recap of the study procedure will be given during this session. Participants will also be informed that they will be excluded from participation in case they (i) do not fit one of the inclusion criteria, (ii) fit any exclusion criteria, or (iii) when no data of sufficient quality can be acquired due to any unforeseen reasons. This explicit declaration is followed by the opportunity for the participant to ask any remaining questions. Once all questions are answered, the participants will sign the informed consent agreement (5 minutes). Then they will fill out all questionnaires and tasks used to screen eligibility for the study. The questionnaires will be presented digitally using Castor EDC. The questionnaires include the Boston Naming test (15-item form, 5 minutes)[88], the Pittsburgh Sleep Quality Index (PSQI, 5 minutes)[89], the Beck Depression Inventory (BDI, 5 minutes)[90], the Beck Anxiety Inventory (BAI, 3 minutes)[91], a General Health Questionnaire (lab developed on Project OSF, 5 minutes), a question on dream recall frequency (taken from MADRE,1 minute)[92], the Munich Chronotype Questionnaire (MCTQ, 5 minutes)[93], an MRI screening questionnaire (developed by the Donders Institute, 5 minutes), and a questionnaire on the frequency of dream categories (lab developed on Project OSF, 10 minutes). The questionnaires are then checked for exclusion criteria (see Supplementary Table 1). If a participant meets one of the exclusion criteria, they will be excluded from participation and paid (6 €), and a replacement participant will be recruited. If all criteria are fulfilled, the participants will do a structural T1 and T2 Magnetic Resonance Imaging (MRI) scan on a Prisma or PrismaFit (3T) (20 minutes). The MRI data will not be analyzed as part of the registered report. Then the three nights in the sleep laboratory (adaptation and both experimental nights) are scheduled. The participants will start collecting sleep data using a sleep tracker (Fitbit Inspire 2) and a sleep diary, as well as a dream diary (project OSF) for one week before the first experimental session. The dream diary is based on the dream protocol used in the laboratory so that participants are familiarized with the questions[94]. Both are presented digitally and can be completed on a computer or phone. The sleep and dream tracking procedure is explained in detail, and participants can ask questions (10 minutes). Participants will be reminded on their phones to fill out their questionnaires each morning.

Adaptation night: The adaptation night is scheduled as closely as possible to the first experimental night (the night before the first experimental night, maximally seven nights before) and at least 6 days after the intake session. Participants will be invited to the Donders EEG laboratory at 21:30. They will be asked to refrain from any alcohol/drug intake during the study day, caffeine intake after lunch (maximum of 2 coffees in the morning according to their usual intake), and get up at or before 08:00 (checked with participant report and sleep tracker). The participants will get a short description to read of the adaptation night and make themselves ready for bed. Then we will apply the EEG cap and EOG, EMG, ECG, and EGG (EGG is opt-in for participants) electrodes. During this time, the participants will fill out the following questionnaires: a check on alcohol/drug/caffeine intake (2 minutes, on project OSF), the “Schlaffragebogen A” (sleep questionnaire A, lab translated from German, SF-A/R, 10 minutes)[95] about the previous night and the “Mehrdimensionaler Befindlichkeitsfragebogen” (multidimensional mood questionnaire, lab translated from German, MDBF, 3 minutes)[96], a lab-developed dream memory questionnaire (30 minutes on project OSF), and the daydreaming frequency scale (DDFS, 5 minutes)[97]. They will complete a color-naming Stroop task across one practice and five experimental blocks (24 congruent, 12 incongruent trials, 10 minutes). At 23:00, participants will go to bed and be able to sleep until 07:00. An investigator will always be present in the experimenter room, and participants are instructed to call out if they need anything (e.g., go to the toilet). If participants cannot fall asleep (either after 1.5 hours or when participants request it), we will first remove the EGG. If they still cannot sleep (after 3 h or when they request it), we will remove all electrodes and discontinue the study (they will have the option to sleep in the laboratory or go home). At 07:00, the sleep opportunity will end. They will fill out a questionnaire about their sleep quality (SF-AR) and recall their dreams. Then the EEG and other electrodes will be removed, and participants can shower and get dressed. Afterward, we will confirm that they want to continue the study and are eligible based on sleep efficiency. At around 7:40, the adaptation night will be done.

Experimental Sessions

The two experimental sessions will be counterbalanced between the participants with random assignment (random number generator (sample in R) will be used for each participant) and additional counterbalancing of the memory task categories. Participants are blinded to the condition. The two experimental conditions are scheduled at least 14 days apart. Participants are instructed to abstain from alcohol and drugs on experimental days and to get up before 08:00. No caffeine intake is allowed after lunch, with a maximum of two coffees in the morning. Alcohol and caffeine intake is checked with a questionnaire.

Furthermore, sleep tracker data will be checked to confirm that no sleep nights have been skipped in the previous week. A stool sample is collected by the participant with a kit (OMNIgene•GUT | OM-200) on the day of the experimental session (not analyzed within this registered report, opt-in by participants). The experimental sessions will start at 19:30. The participants will get written instructions explaining the experimental session. Afterward, they will get ready for bed. Then polysomnography will be applied.

Session A: Awakenings

During the EEG application, the participants can ask questions about the awakening protocol (the same questions as those used at home). For the remaining time during EEG application, the participant will fill out the following questionnaires: the alcohol/coffee check (2 minutes), the Mannheim Dream Questionnaire (MADRE, 10 minutes)[92], the Brief-COPE questionnaire (10 minutes)[98], the MDBF[96] (3 minutes), the need for closure scale (15 minutes)[99], and the Freiburg Mindfulness Inventory (FMI, 5 minutes)[100]. Additionally, they will complete the trail-making test (TMT, 5 minutes)[101]. Afterward, the participants will undergo the learning blocks of the memory task. Between the learning blocks and the recall, there will be a 10 minutes break during which the participants will fill out the MDBF[96] again and the SF-A/R[95] for the previous night. Recall happens in 2 blocks which take approximately 40 minutes.

At 23:00, participants will go to bed. When the participant is lying in bed, we will do a resting-state EEG measurement (1.5 min eyes open, 1.5 min eyes closed, 1.5 min eyes open, 1.5 min eyes closed). The investigators will monitor the EEG while the participant is asleep visually, aided by information provided by the dreamento toolbox[102]. The participants will be woken up to 8 times during the night following an awakening protocol (on project OSF) – four times from NREM and four times from REM sleep (at least 15 minutes into each sleep stage). For NREM sleep, N2 will be used as the start of the sleep stage, however, the awakening can be done in any NREM (N1, N2, or N3) sleep stage. The preceding 1 minute of each awakening should not contain any wake or the opposite sleep stage (i.e., REM for a NREM awakening and NREM for a REM awakening). After each awakening, the participants will be prompted to report their dreams orally and rate them on several scales. After this, participants can go back to sleep. The sleep opportunity ends at 7 am. They will fill out a dream report, where they will report dreams not previously reported as well as dreams reported in the night. If they forget some of the dreams, we will give them a related one-word prompt to each dream to trigger the memory. Afterward, they fill out a questionnaire about their sleep (adapted SF-A/R, the question on dream recall removed, an additional question regarding "Did you hear any words presented last night?" (Yes/No), and a question about spontaneous, non-experimenter awakenings). After this, both recall rounds of the memory task will be repeated exactly as during the night before. Then the participants will complete a localizer task in which they rate 67 new images corresponding to the task categories three times (first-round valence, second round arousal, third round prototypicality). Then electrodes will be removed, and participants can shower. The study will be finished around 8:30 am.

Session B: TMR + Awakenings

During the EEG application, the participants will again read the protocol used for the awakenings to ensure that the participant understands all the questions. The participant can ask questions if they do not understand them. For the remaining time during EEG application, the participant will fill out the following questionnaires: alcohol check (2 minutes), the Lucid Dreaming Skills Questionnaire (LUSK, 5 minutes)[103], the Vividness of Visual Imagery Questionnaire (VVIQ, 10 minutes)[104], the Rosenberg self-esteem scale (5 minutes)[105] and the behavioral inhibition/activation scale (BIS/BAS, 10 minutes)[106]. Afterward, the participants will undergo the learning blocks of the memory task. The task will be the same as in session A but using different image categories. At 23:00, participants will go to bed. When the participant is lying in bed, we will do a resting-state EEG measurement (1.5 min eyes open, 1.5 min eyes closed, 1.5 min eyes open, 1.5 min eyes closed). The investigators will monitor the EEG while the participant is asleep visually, aided by information provided by the dreamento toolbox[102].

After at least 3 minutes of stable NREM (N2 or N3) and REM sleep, experimenters will play audio cues for 5 to 15 minutes using two loudspeakers placed at 230 cm from the participants' heads (position kept consistent across participants). Words associated with one specific image category will be used for cueing in each sleep stage (randomly chosen for each participant). Words from the category will be presented randomly every 8,000 to 8,200 ms. Cueing will start at 30dB SPL and increase in 5 dB steps until the participant shows a K-complex (NREM) or arousal (REM). Audio will then be played at the level (NREM) or one step below the level (REM) for the remainder of the sleep cycle. Audio levels will be determined for each cycle as thresholds vary throughout the night. Audio cues will be stopped if participants show a sign of arousal or change into a different sleep stage. The participants will be awoken between 10 – 30s after the last TMR at least 15 minutes into each sleep stage. The protocol for the awakenings is identical to session A. After this, participants can go back to sleep. In the morning, the sleep opportunity ends at 7 am. They will fill out a dream report, where they will report dreams not previously reported as well as dreams reported in the night. If they forget some of the dreams, we will give them a related one-word prompt to each dream to trigger the memory. Afterward, they will fill out a questionnaire about their sleep. After this, both recall rounds will be repeated exactly as during the evening before. Then the participants will complete another localizer task corresponding to the image categories presented in this session. Then electrodes will be removed, and participants can shower. The study is finished around 8:30 am.

Follow-up Memory Recall

Four days after each experimental session, there will be a follow-up on the memory recall performance using the same recall blocks. This recall will be presented online using Pavlovia (based on the psychopy experiment used in the laboratory). Participants will have to complete the follow-up in a single session within a 12-h timeframe.

Memory Task

To measure memory performance, we will use an adapted version of the word-picture association task we have used previously[9]. The task consists of 99 word-picture associations of neutral words and positive and neutral pictures, which we have extended with negative pictures. The pictures are related to 6 categories (3 per experimental night): mammals, vehicles, food, children, water, and buildings. Each category has 11 positive, 11 negative, and 11 neutral pictures. At the beginning of the task, one image unrelated to the categories is presented (primacy effect). The pictures are taken from the NAPS (90), IAPS (15), NDPS (10), DIRTI (7), and Oasis (21) databases which contain large sets of images that have been rated on emotional valence and arousal[55–59]. Still, the images had to be supplemented with 55 open Creative Commons license images (from Unsplash, Flickr, Pixahive, Wikipedia, Stocksnap, pxhere) because not enough images were available to fit our criteria (see project OSF for a complete list).

All potentially fitting images from the databases and the additional images were rated by 16 pilot participants to ensure adequacy for the task. The final images were selected using the following criteria: appropriate valence rating (> 5.75 (on a 1 - 9 scale) for positive, 4.25 to 5.75 for neutral, and < 4.25 for negative), the appearance of none of the other five categories as well as no adjacent categories (e.g., adult humans for children category, or other animals for mammal category, flagged by >= 3 participants) and image quality (rated higher than 6 on a 0 - 9 scale). If more images than needed fitted the criteria, the images were selected for the lowest standard deviation on the valence and arousal rating, the most similar rating to the original database, and the highest discriminability (e.g., not two images of the same mammal).

The words are taken from the auditory English Lexicon (AELP) project[60]. The words are chosen to have two syllables as well as a similar length (636 – 805 ms), neutral valence and arousal (between 4 - 6), and be well known (> 88% recognition). Furthermore, words were selected not to contain any reference to the image categories. The association between word and picture was done randomly but will be kept consistent across participants.

The memory task has six blocks: two rating blocks, two learning blocks (the second done twice), and two recall blocks. In the first block, the participants will hear all the neutral words and rate them for valence and arousal. In the second block, the participants will see all pictures and rate them for valence and arousal. During the first learning block, they will see the picture and hear the associated word. The second learning block will be completed twice, where the participants will hear the word and then indicate the expected valence (negative/neutral/positive) and arousal (negative/neutral/positive). Then they will see the picture presented to enable another learning possibility. After a 10 minutes break, there will be two recall blocks. First, the participants will hear the words and indicate the associated picture valence, arousal, and certainty. Then there will be a cued recall. The participants will hear the word and describe the associated picture with 3-5 keywords. The task is implemented using Psychopy.

Sleep Recording

EEG will be recorded with 64 channels cap (actiCAP original) and the BrainAMP by Brainproducts. Each electrode location will be prepared using an abrasive paste (Nuprep) and electrode paste (Abralyt). Impedances will be checked to be below 20 kΩ. Additionally, two electrodes will be used to measure EOG, ECG, and three electrodes for chin EMG (using BrainAMP ExG, impedance level below 10 kΩ) and an 8 channel EGG (participants can opt-out of the EGG if they are unable to sleep with it, impedance level below 25 kΩ). See the supplemental files for electrode placement information. Data will be recorded with a 500 Hz sampling frequency and referenced to the vertex.

Dream Reports (orally and written)

Participants will be asked, "What was going through your mind in the minute prior to awakening?" They are instructed beforehand to include any dreams, thoughts, experiences, imagery, sensations, or emotions. If they don't report anything, they will be asked to take a moment to remember. If after 1 minute they cannot remember a dream, they are asked, «Do you feel as if you had a more detailed dream or specific thoughts, imagery, sensations, or emotions that you have now forgotten?» and if they respond, "no" they will be asked, "Before awakening, did you have a feeling or awareness of being asleep?". If they report a dream/thought/experience/imagery/sensation/emotion, this is recorded and written down. Once they stop reporting, they are asked if they remember anything else (repeated up to 3 times if more content is produced). They are asked to estimate the length of the dream. If the dream is longer than one minute, they are asked to focus first on the last minute. Then the dream report will be rated on several scales (lucidity, voluntary control over dream content, vividness, arousal, valence, accuracy, and completeness) from 1 to 5. Furthermore, participants will indicate if they had any visual, auditory, tactile, olfactory, gustatory, and vestibular perceptions (yes/no/unsure). Then they will be asked to describe the previous dream elements (if the dream was longer than 1 minute) or any other dream between the last awakening and now. If they remember previous dreams, they will be asked to rate them on the same scales.

Sampling plan

Participants

Ninety-two healthy male and female volunteers aged 18-35 will be recruited from the general area around Nijmegen, Gelderland, Netherlands. The inclusion criteria to participate in the study are physically and mentally healthy, a dream recall frequency of more than once a week, high English language proficiency, and the ability to sleep in the sleep laboratory. Exclusion criteria are history of or current sleep disorder, current physical or mental illness, intake of medication that influences sleep/wake cycle and/or memory consolidation, frequent coffee consumption (> 4 cups/day), skin disease at intended electrode sites, chronotype incompatible with the study time window, inability to sleep during adaptation night, contraindications for MRI (including pregnancy/breastfeeding), irregular sleep patterns leading up to experimental sessions. Supplementary Table 1 reports the exact criteria for each inclusion/exclusion as well as the measurement used. Data will be excluded from single experimental nights if less than three hours of sleep are obtained. The specific awakening is excluded if less than 85% of auditory cues are presented in the correct sleep stage or less than 5 minutes of auditory cueing can occur. Any participant replacements, dropouts, and exclusions will be reported.

Project OSF:

DOI 10.17605/OSF.IO/YKUQ5

Supplementary Table 1. Exclusion criteria, measure, and contingency.

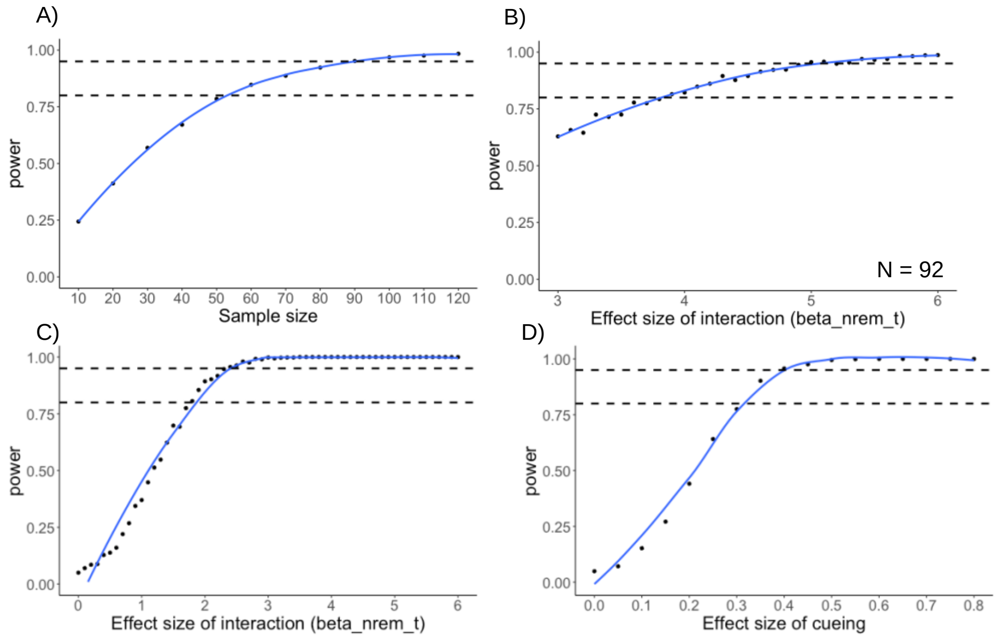
Stage of Assessment	**Testing for**	**Measure**	**Criteria**	**Contingency**
Intake Session	High English language proficiency	Boston Naming Task	< 10 correct	Recruit new participant
Intake Session	Current sleep problems	PSQI	Score > 7	Recruit new participant
Intake Session	Depression	BDI	Score ≥ 20	Recruit new participant
Intake Session	Anxiety	BAI	Score > 15	Recruit new participant
Intake session	Chronotype	MCTQ	Sleep Time after 1 am («I actually get ready to fall asleep at») on the weekdays	Recruit new participant
Intake session	Dream Recall Frequency	MADRE	Dream Recall Frequency < several times a week	Recruit new participant
Intake Session	Mental and Physical Health	General Health Questionnaire	Yes to Sleep Medication Yes to Medication for Mental Health Yes to Medication that is known to influence memory consolidation Yes to Sleep disorder (current or previous) Yes to a current physical or mental health issue	Recruit new participant

			Yes to Skin disease (at electrode location)	
Intake session	Coffee and drug withdrawal	General Health Questionnaire	Yes harder drugs/marijuana daily More than 4 cups of coffee per day	Recruit new participant
Intake session	MRI Incompatibility	MRI questionnaire	Yes to any of the MRI incompatibility questions	Recruit new participant
Adaptation session	Irregular sleep pattern	Actigraphy	Sleep skipped in the six days before	Recruit new participant
Adaptation session	Ability to sleep in the sleep lab	Participant report	Inability to fall asleep with EEG/wanting EEG removed	Recruit new participant
Adaptation session	Inability to sleep in sleep lab or with EEG	EEG	Sleep Efficiency < 70%	Recruit new participant
Any sleep lab session	Influence on sleep and memory	Participant report	Alcohol consumption or coffee consumption after noon (or more than two coffees in the morning)	Reschedule
Any sleep lab session	Inability to fall asleep	Actigraphy	Get up time after 8 am	Reschedule
Experimental nights	Not enough time for sleep-dependent memory consolidation	Sleep Duration	<= 3 hours of sleep	Exclusion experimental night (estimation within the model)

Experimental nights	NREM vs. REM dream	Awakenings	Awakening in wrong sleep stage	Exclusion awakening
Experimental Night: TMR	Correct stimulation	Auditory cues	< 85% incorrect sleep stage (NREM/REM) < 5 minutes	Exclusion awakening
Experimental	Lucidity	Awakening protocol	Lucidity rating = 5	Exclusion awakening
Experimental nights	Missing data	EEG, Memory Task	Data missing due to technical problems	Exclusion experimental night (estimation within the model)
Experimental data	Technical problems	All data	Outlier/missing data due to technical problems	Exclusion and estimation within the model

Experiment day	Questionnaires	Sleep	Dreams	Memory Task	ECG/EGG	Stool sample
Intake Session						
d1-7		T + D	D			
d1-6		T + D	D			
d1-5		T + D	D			
d1-4		T + D	D			
d1-3		T + D	D			
d1-2		T + D	D			
Adaptation Night		EEG	D			
Experimental Session Night		EEG	A + D	Learning		
d1+1		T + D	D	Recall		
d1+2		T + D	D			
d1+3		T + D	D			
d1+4		T + D	D	Recall		
d1+5		T + D	D			
d1+6		T + D	D			
d1+7		T + D	D			
d2-7		T + D	D			
d2-6		T + D	D			
d2-5		T + D	D			
d2-4		T + D	D			
d2-3		T + D	D			
d2-2		T + D	D			
d2-1		T + D	D			
Experimental Session Night		EEG	A + D	Learning		
d2+1		T + D	D	Recall		
d2+2		T + D	D			
d2+3		T + D	D			
d2+4		T + D	D	Recall		
d2+5		T + D	D			
d2+6		T + D	D			
d2+7		T + D	D			

Supplementary Figure 1. Experimental protocol of the study. The study takes part across a whole month. Each participant will visit the institute four times, once for the intake session and three times for the sleep laboratory (1 adaptation night and two experimental sessions). Black indicates data that is collected for each day. T = Tracker, D = Diary, Q = Questionnaire, A = Awakenings, EEG = Electroencephalography (including electrooculography and electromyography), ECG = electrocardiogram, EGG = electrogastrography.



Supplementary Figure 2. Effect size simulations for hypotheses 1 and 2. A) For hypothesis 1, we used effect size estimates from our previous study to simulate 1000 datasets with 10 – 120 participants each. 95% power is reached with 90 participants. **B)** Sensitivity analysis with 92 participants and varying the effect size of the interaction (NREM incorporation*timepoint). With 92 participants, we reach 95% power with an effect size of b ≥ 5 and 80% power with effect size b ≥ 3.9. **C)** Sensitivity analysis with 92 participants and verifying effect size of the interaction (NREM incorporation*timepoint) for the model controlling incorporations for baseline. We reached 95% power with an effect size of b ≥ 2.4 and 80% power with effect size b ≥ 1.8. **D)** For hypothesis 2, we used effect sizes from data on task incorporation into dreams to estimate potential effect sizes for TMR. In the sensitivity analysis with 92 participants and varying the effect size of cueing from 0.0 – 0.8 (0.05 steps), we estimate 95% power with an effect size of b ≥ 0.4 and 80% power with effect size b ≥ 0.3.