A Comparison of Sleep-Based and Retrieval-Mediated Memory Consolidation Using Sigma-Band Activity

Hayley B. Caldwell[a*], Kurt Lushington[b], Alex Chatburn[a]

[a]Cognitive Neuroscience Laboratory – ARCIVE, University of South Australia, St Bernards Road, Magill, South Australia, 5072.

[b]Justice and Society, University of South Australia, St Bernards Road, Magill, South Australia, 5072.

Hayley.Caldwell@mymail.unisa.edu.au[*]

Kurt.Lushington@mymail.unisa.edu.au

Alex.Chatburn@mymail.unisa.edu.au

[*]Corresponding author

Abstract

Sleep and retrieval training both promote memory consolidation, likely via similar neural mechanisms, but they have not yet been directly compared in a counterbalanced design that accounts for potential masking of nap effects by prior retrieval training. The current study aimed to see if we could replicate sleep spindles' enhancement of weakly encoded memories, across both sleep-based and retrieval-mediated memory consolidation, using EEG sigma power (~11–16 Hz) as a proxy for sleep spindles (i.e., sigma-band activity). Thirty subjects (27F, 18–34, *M*=22.17) participated in four separate sessions where they learnt different sets of 104 object-word pairs. Subjects were then tested on their recognition accuracy of the pairs before and after one of four 120 min memory interventions where EEG was recorded: retrieval training (i.e., cued recall practice), restudy (i.e., pair re-exposure), a nap opportunity, or a wakeful rest. Our results did not replicate an enhancement of weakly encoded memories, moderated by either sleep spindles of sigma power. Contrary to expectations, sigma power demonstrated a negative impact on memory outcomes. We also detected a prioritisation to enhance memory outcomes for strongly encoded memories and greater memory outcomes following retrieval training and restudy compared to sleep and wake interventions. These results challenge the assumptions of encoding-strength-dependent enhancements across consolidation. Importantly, our study provides a methodological approach to comparing sleep-based and retrieval-mediated memory consolidation that should be explored across different memory paradigms in future studies.

Keywords: Memory consolidation; Retrieval training; Sleep; Recognition accuracy; Sigma power; Encoding strength.

## 1. Introduction

Memory consolidation involves the incorporation of memories into the neocortical landscape for long-term storage (McClelland et al., 1995). An optimal state for memory consolidation is sleep (Rasch & Born, 2013; Staresina, 2024); but recent evidence suggests that memories can undergo rapid consolidation across wake, via repeated retrieval training (i.e., cued recall; Antony et al., 2017; Ferreira et al., 2019; Ritvo et al., 2019). It is assumed that the mechanisms underlying retrieval-mediated consolidation parallel those for sleep-based consolidation. Findings from both sleep-based and retrieval-mediated consolidation studies report comparable memory changes, including memory accuracy gains (Diekelmann et al., 2009; Hebscher et al., 2019), semanticization (Lifanov et al., 2021; Schapiro et al., 2017), long-term stability (Guran et al., 2020; Lutz et al., 2017), and shifts from hippocampal to neocortical areas (Ferreira et al., 2019; McClelland et al., 1995). Several studies have investigated sleep and retrieval training interventions in combination. These studies typically involve retrieval training followed by sleep, and test memory consolidation across both interventions using the same stimuli (Antony & Paller, 2018; Ashton et al., 2021; Mak & Gaskell, 2023; Schreiner et al., 2018). However, if both interventions act upon the same memories, sleep effects may be masked by prior retrieval training, which makes it difficult to separate their impacts on memories and introduces ceiling effects. To better compare their impacts on memory, it would be useful for sleep and retrieval training to act upon separate stimuli, in a counterbalanced order, and separate the interventions sufficiently in time to minimise crossover effects. Denis and colleagues (2024) have followed this approach, but did not include a restudy control group, making it difficult to isolate if memory changes occur due to retrieval training or memory practice in general. Protocols that do not confound retrieval training and sleep, and employ appropriate control groups, would be useful when comparing brain-based mechanisms of memory consolidation, but this has yet to be investigated.

A well-researched moderator of sleep-based consolidation is encoding strength, such that weakly encoded memories are typically prioritised for consolidation over strongly encoded memories (Baena et al., 2020; Bäuml et al., 2014; Denis et al., 2020; Drosopoulos et al., 2007; Dumay, 2016; Petzka, Charest, et al., 2021; Schapiro et al., 2017, 2018). This encoding-strength-dependent selective enhancement of memories is proposedly due to strongly encoded memories already being somewhat neocortically integrated, such that sleep-based consolidation cannot act upon them further (Himmer et al., 2017). Sleep-based memory

reactivation and the transfer to neocortical stores is thought to be mirrored during retrieval-mediated memory consolidation (Antony et al., 2017). It is possible that retrieval-mediated memory consolidation also selectively enhances weakly encoded memories. Studies have found that memories weakened by interference and directed forgetting are selectively enhanced or saved by retrieval training (Abel & Bäuml, 2016; Halamish & Bjork, 2011; Pastötter et al., 2011). However, these studies have not directly examined encoding strength. A comparison of encoding-strength-dependent selective enhancement across both sleep and retrieval would inform our understanding of mechanisms underlying memory consolidation.

Sleep spindle parameters (i.e., spindle frequency, density, amplitude) have been shown to moderate the enhancement of weakly encoded memories across sleep, with greater sleep spindle activity relating to a greater enhancement of weakly encoded memories over strongly encoded ones (Chatburn et al., 2021; Creery et al., 2015; Denis et al., 2021; Kumral et al., 2023; Petzka, Chatburn, et al., 2021). Sleep spindles are approximately 11–16 Hz oscillations of the EEG signal during non-rapid eye movement (NREM) sleep, that are generated in temporal sequence with slow-wave oscillations and memory reactivations to facilitate the transfer of information into the neocortex (Rasch & Born, 2013; Staresina, 2024). Studies investigating the interaction between sleep spindles and encoding strength have measured memory outcomes using recall accuracy due to its consistent detection of sleep-based consolidation impacts on memories (Diekelmann et al., 2009). However, as retrieval training uses memory recall to consolidate memories, when comparing to sleep an alternate measure is needed, such as recognition accuracy. This will avoid additional study-test matching benefits that retrieval training would have over sleep.

To compare encoding-strength-dependent memory enhancements across sleep and retrieval training, an effective proxy must be used as sleep spindles do not occur during wake. Sigma power has been used as an alternative for spindle parameters to investigate memory changes across sleep-based consolidation (Piosczyk et al., 2013), due to it capturing activity within the sigma band (i.e., ~11–16 Hz). Sigma power has also moderated encoding-strength-dependent forgetting reflecting contextual reinstatement during sleep (Schechtman et al., 2023), which also drives memory changes across retrieval-mediated consolidation (Antony et al., 2017). Some have speculated on the role of sigma-band activity during wake (e.g., Dringenberg, 2019), with Schreiner & Staudigl (2020) even suggesting it could be related to suppression of distractions as spindle-related neurons in the thalamus were associated with

such during wake (Z. Chen et al., 2016). However, no study to date has explored a potential role for sigma power during wakeful memory consolidation.

Overall, if retrieval training and sleep induce the same brain-based mechanisms, it is expected that both consolidation interventions will only enhance weakly encoded memories when sigma power is high. The current study will test subjects across four conditions separated by at least 48 hr (retrieval, restudy, sleep (120 min afternoon nap), and wake) and, using object-word pairs, assess immediate (encoding strength) and delayed (memory outcomes) recognition accuracy. Apart from the wake condition, EEG will used to estimate sigma power during the intervention period.

## 2. Methods

### 2.1. Subjects

Thirty individuals (27 women, one non-binary, two men) aged 18–34 (*M*=22.17, *SD*=4.21) participated in the study. Twenty-five completed all four conditions (retrieval, restudy, sleep, and wake), one subject only completed retrieval, one subject completed restudy, one subject completed wake, and two subjects completed sleep and wake conditions. All collected data was used in analysis. Subjects were right-handed, fluent English speakers, had normal or corrected vision, and self-reported no hearing or sleep problems. Subjects also reported on interview and questionnaire no diagnosed psychiatric or sleep impairments, and were not taking any medications or recreational drugs in the last 6 months impacting EEG (Massand & Bowler, 2013; Nicholls et al., 2013; Weymar et al., 2010). Subjects were recruited via flyers posted around the university campuses, social media, and an online subject recruitment system. Subjects who completed all four conditions received a $200 honorarium. The UniSA Human Research Ethics Committee granted ethics approval for this project (protocol no. 205130).

### 2.2. Materials

#### 2.2.1. Pittsburgh Sleep Quality Index (PSQI)

Sleep habits across the last month were screened using the PSQI (Buysse et al., 1989). The PSQI contains 19 items which are used to generate a total score where scores >5 indicate significant sleep difficulties, which was used as a cutoff for eligibility. The PSQI has demonstrated adequate internal consistency and high test-retest reliability (Backhaus et al., 2002; Buysse et al., 1989; Carpenter & Andrykowski, 1998).

#### 2.2.2. Morningness-Eveningness Questionnaire (MEQ)

The MEQ was used to screen out subjects with extreme circadian types (Horne & Östberg, 1976). The MEQ is a self-report questionnaire involving 19 multiple choice questions on sleep/wake habits. Subjects with scores ≤30 or ≥70 were excluded. The MEQ's items have previously demonstrated a high internal consistency and high test-retest reliability (Li et al., 2011; Pišljar et al., 2019).

2.2.3. Flinders Handedness Survey (FLANDERS)

Handedness was screened for using the FLANDERS (Nicholls et al., 2013). Subjects were instructed to select their hand preference for each of the 10 tasks (e.g., drawing). Selecting "right" corresponds to +1 to their overall score, "left" is –1, and "mixed" is 0. Overall scores between –10 to –5 indicate left-handedness, –4 to +4 indicate mixed-handedness, and +5 to +10 indicate right-handedness. Subjects with scores <5 were excluded. The Flanders has displayed high internal consistency between its items (Nicholls et al., 2013).

2.2.4. Karolinska Sleepiness Scale (KSS)

The KSS was used to control for sleepiness in the analyses (Åkerstedt & Gillberg, 1990). The KSS is a single-question self-report measure where subjects rate their current sleepiness on a scale from 1–10. Answers range from "1 – extremely alert" to "10 – extremely sleepy, cannot keep awake." The KSS has demonstrated higher predictive validity for various measures of alertness than any other sleepiness measure, including reaction time ($r$=.57), number of lapses ($r$=.56), and power spectral density in both alpha ($r$=.40) and theta ($r$=.38) bands (Kaida et al., 2006).

2.2.5. EEG

EEG was collected using a 64-channel passive BrainCap with Ag/AgCl electrodes positioned according to the modified 10–20 system. Electrooculographic activity was measured via two electrodes embedded in the cap, positioned approximately 1cm diagonally away from the outer canthus of each eye. Electromyographic activity was measured via three electrodes (two placed on the skin above each subglottic muscle, and one placed between the subglottic muscles), and electrocardiographic activity was measured by one electrode placed above the heart. Two BrainAmp amplifiers (DC system, sampling rate 500 Hz) amplified the signals, while the BrainVision Recorder software recorded the output. Impedances were kept below 10kΩ throughout all experimental phases.

2.2.6. Stimuli

2.2.6.1. Objects

416 object images (104 per condition) were taken from the Mnemonic Similarity Task (MST) from sets 2–5 and with false alarm rates at or below 33.33% i.e., chance level (Stark et al., 2019). Each object in the learning set was mnemonically similar to another object within the set.

2.2.6.2. Words

Word stimuli were taken from an updated version of the Affective Norms for English Words (ANEW; Warriner et al., 2013). This dataset contains 13915 English words, optimised for use in paired associate tasks, rated on frequency from 0–1000 and 1–9 on valence, arousal, and dominance. This study randomly selected 416 monosyllabic nouns (104 per condition), with valence and arousal ratings between 3.5–6.5. These words were recorded, using the voice of a native English speaker with an Australian accent, for their auditory presentation. These recordings were approximately 500–1000 ms long, each. Words were presented at approximately 70 dB.

2.3. Procedure

Eligible subjects were screened using the PSQI, MEQ, and FLANDERS. Subjects then participated in four conditions, each of 4.5 hr (12:00–16:30; see Fig. 1), separated by at least 48 hr. The night before each condition, subjects were asked to wake up 1 hr earlier than usual to increase the likelihood of an afternoon nap entering NREM sleep, in the sleep condition. Subjects were seated in front of a dedicated computer and were fitted with an appropriately sized EEG cap. Subjects were then given details about the memory task they performed across the rest of the session.

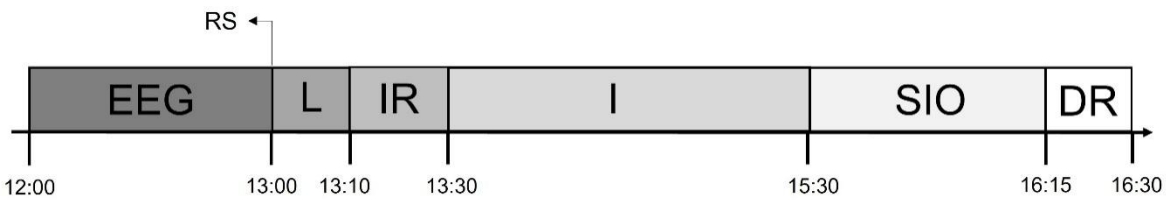


Fig. 1. Procedure timeline. EEG = Setup EEG, RS = Resting State, L = Learning, IR = Immediate Recognition, I = Intervention (retrieval, restudy, sleep, or wake), SI = Sleep Inertia Offset / Retention Period, DR = Delayed Recognition.

The memory task was based on previous protocols which have employed retrieval training and sleep (Ferreira et al., 2019; Liu et al., 2021; Schreiner et al., 2018). Open Sesame v.3.3.10 software was used to create and run each phase of the memory task (Mathôt et al., 2012). The memory task began with a learning phase, then immediate recognition, one of

four intervention conditions (retrieval, restudy, sleep, or wake), and a final delayed recognition. The intervention conditions across sessions were experienced in a pre-determined, counterbalanced order. This counterbalancing was random, except that the subjects always completed one wakeful condition before the sleep condition, to increase their familiarity with the environment and therefore increase their likelihood of napping. The paired associates used in the memory task consisted of randomly paired MST object pictures and auditory words. All subjects were presented with the same object-word pairs, in a randomised order for each condition by subject. A different set of object-word pairs were used in each of the four conditions. For the object in each object-word pair, there was an object in a different object-word pair that was its similar object from the MST. Highly related object-word pairs were manually identified and reshuffled. A visualisation of the memory task can be found in Fig. 2. The following sections detail the memory task protocol.

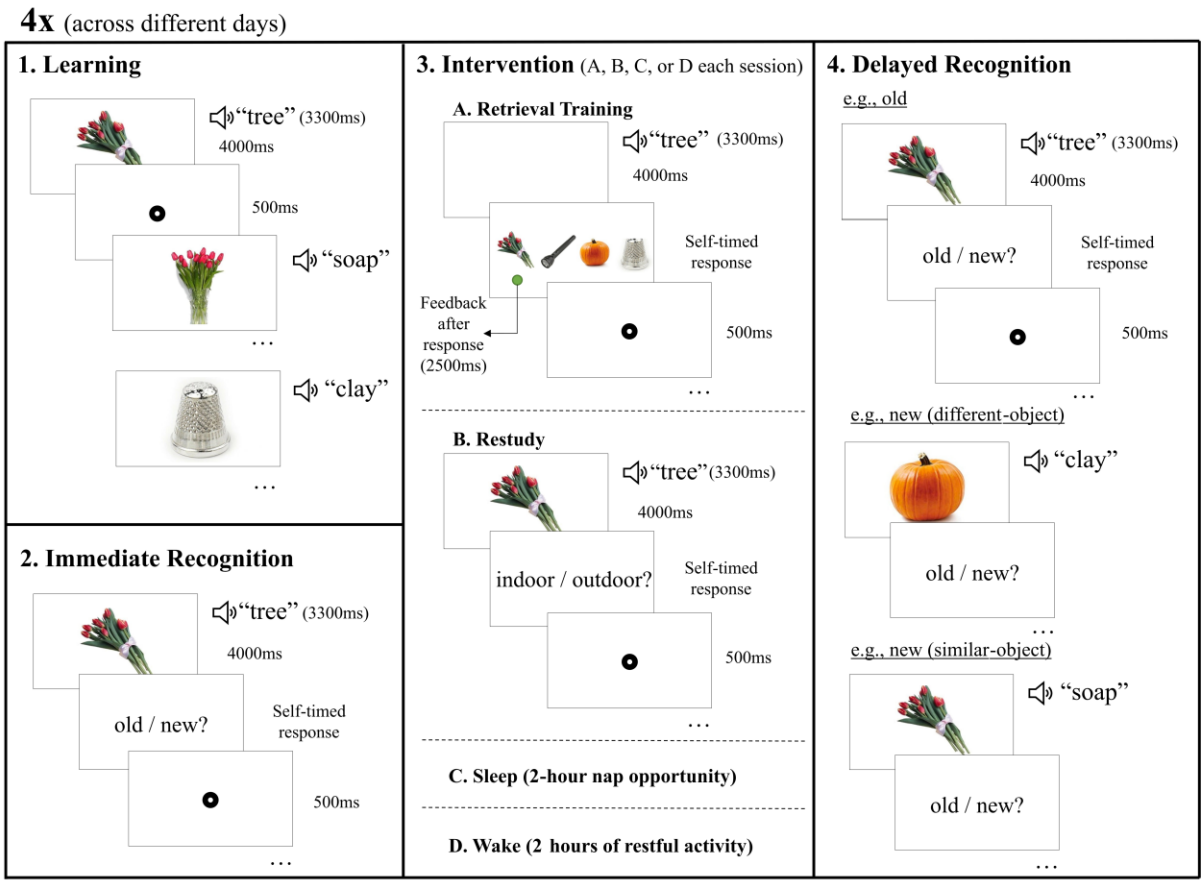


Fig. 2. Memory task. 1. In the learning phase, subjects were exposed once to each of the object-word pairings. 2. Subjects then underwent immediate recognition testing, to distinguish between original (old) and rearranged (new) pairings, based on the pairings learnt during learning. An example of an old pairing is shown here. 3. Subjects underwent one of four intervention conditions. A. Involved retrieval training, where subjects heard each word, recalled the object, then picked it from distractors with feedback after. B. Subjects were

exposed to all the pairs again and had to indicate if the object belonged indoors or outdoors. C. Subjects had a 2 hr nap opportunity. D. Subjects engaged in light activities. 4. Delayed recognition involved the same procedure as immediate recognition, where subjects distinguished between old and new pairings of the objects and words. This includes an example of an old pairing, new pairing with a different object to the object paired with the original word, and a new pairing with a similar object to the object paired with the original word.

2.3.1. Learning

In each session, subjects learnt 104 object-word pairs. This involved viewing the image of an object on a white screen for 4000 ms and a sound of the paired auditory word was played 700 ms after the object onset. The images were displayed in colour, 400 pixels wide, in the centre of the screen. After each object was displayed, a fixation dot (8 pixels wide with a 2-pixel wide centre hole) was presented for 500 ms, before the next object was shown. Each pair was presented once.

2.3.2. Immediate Recognition

Immediately following learning, recognition memory of the pairs was tested by exposing subjects to each object and word again in the same sequence as the learning phase. This was done using an old/new paradigm, where 72 object-word pairs were presented in their correct pairings (old pairs), and 32 in a rearranged pairing (new pairs). To create the rearranged pairings, pairs were shuffled by substituting the object in each pair with another object of a different pair (16 similar objects, 16 different objects). After each pair was presented, a new white screen displayed the prompt “old or new?” Subjects responded via keypress, based on whether they believed that the pairing was presented as it was originally learnt (old / “Q”), or if it was rearranged (new / “P”). Before starting the task, subjects were encouraged to respond quickly, but optimize accuracy over speed. After they responded to the prompt, a fixation dot appeared on the screen for 500 ms before the next object-word pair was shown. No feedback was given to subjects on recognition accuracy performance. To avoid introducing interference effects with the original encoding episodes, the 32 rearranged (new) pairs were not used in subsequent phases.

2.3.3. Intervention

For each of their four visits, subjects completed one of the four interventions, depending on the condition order they were assigned. Each intervention period took 120 min and EEG was recorded apart from one exception. This exception was the wake condition. After the intervention, all conditions included a 45 min retention period, which also served as

a sleep inertia offset period for the nap condition. During this break, subjects remained in the laboratory and engaged in light activities (e.g., reading or scrolling on social media), simulating how they typically spend their free time. The KSS was completed at the end of the intervention. The four interventions are outlined below.

2.3.3.1. Retrieval Training

Subjects completed six rounds of retrieval training practice with the initially learnt object-word pairs. The 72 object-word pairs were randomly divided into two equal length lists, with each list being presented three times. Within each round, the order of presentation for the object-word pairs was randomised per subject. Practice round took 10−12 min. Subjects had 8–10 min breaks between each round so that the practicing was spaced out over the 120 min period.

The retrieval training practice had subjects hear the word of each pair, recall the paired object from their memory, then pick the object that was originally paired with the word from distractors. Subjects were presented a blank white screen and heard the word of one of the object-word pairs that were previously learnt. Subjects were instructed to imagine the paired object on the screen, when they heard the word. After 4000 ms, four object images that were part of previously learnt pairs were displayed 120 pixels wide, equally spaced across the width of the screen. Of the four objects, one of them was the original object that the word was paired with, while the other three were objects from other pairs. The distractor options were pre-determined and consistent across all subjects, but they were different each of the three times that a word was practiced. For every single pair that was practiced, exactly one round included the original object's similar MST object as one of the distractors. Subjects were instructed to choose the object that they believed they previously learnt as paired with the word they just heard. Each distractor object had a black letter displayed above it which indicated the corresponding response key (Q, E, I, or P) required to select the object below it. 500 ms after their response, subjects received feedback on their selection via a green dot (8 pixels wide with a 2-pixel wide centre hole) placed below the correct object for 2000 ms. Subjects were instructed to use this feedback to learn the pairings and improve their memory. After the feedback was presented, a fixation dot was shown in the middle of the screen for 500 ms, before the next word was played on a blank screen. For the breaks between each round, subjects engaged in light activities.

2.3.3.2. Restudy

In the restudy condition, subjects followed the same six-round study-break structure as the retrieval training intervention, practicing half (36) the 72 pairs at a time, three times each. Overall, restudy practice involved being exposed to both constituents of each word pair again. For each object-word pair, subjects were presented the object and word exactly as was done during learning, in a subject-specific randomised order. An object image from each pair was presented in the centre of the screen for 4000 ms, followed by the originally paired word presented auditorily 700 ms later. To maintain subjects' attention, after each pair was presented, a screen prompted subjects to decide if the object of that pair would be commonly found indoors or outdoors. This prompt was displayed in the centre of a white screen, with the words "Indoors / Outdoors?" with the letter "Q" and underneath the word "Indoors," and the letter "P" underneath "Outdoors." These letters underneath the question depicted the keypress required to select the corresponding option. Subjects were instructed to treat this interactivity task as secondary to their practicing of the pairs. After subjects responded, a fixation dot appeared for 500 ms before the next object-word pair was presented.

2.3.3.3. Sleep

Subjects were given a 120 min nap opportunity. Subjects slept in a quiet, darkened bedroom, and to promote sleep white noise was played at approximately 37db during the entire nap opportunity. The researcher remained in the adjacent room, monitoring the EEG signal for signs of sleep and any disturbances to the electrode signals.

2.3.3.4. Wake

In the wake condition, subjects were supervised while they engaged in light activities (e.g., reading a book or watching a movie) for 120 min, to replicate their typical use of free time. During this time, subjects' EEG was not recorded, but they remained in the laboratory with the EEG cap on.

2.3.4. Delayed Recognition

Comparable to the immediate recognition phase, subjects were presented with the same old/new task, but with a new assignment of pairs as either old or new. Subjects were exposed to the 72 object-word pairs they practiced, which were presented using the same sequencing as the immediate recognition phase. For half (36) of the 72 pairs, the word of each pair was presented with its originally paired object from the learning phase. The other half (36) of the object-word pairs were novel (new) pairings. Half of the novel pairings were similar-object lures, whereas the other half were different-object lures. In the exact same structure as the

immediate recognition phase, subjects were prompted, by the screen following the object, to indicate if the pairing was old or new, receiving no feedback on their accuracy. After they answered, a fixation dot appeared for 500 ms before the next pair was presented. Whilst the order of presentation of pairs was randomised for each subject, all shuffled pairings and assignments to old or new categories were pre-determined to remain consistent across all subjects. After the delayed recognition test, subjects had the EEG cap removed, and then left the laboratory.

## 2.4. Data Analysis

### 2.4.1. Recognition Accuracy (*d* prime)

Recognition accuracy in the immediate and delayed recognition tests was represented as *d* prime ($d'$), a sensitivity index of memory performance from Signal Detection Theory that accounts for false alarm rates when calculating accuracy (McNicol, 2005). $D'$ was calculated in R, v4.3.0 (R Core Team, 2022) using the *Psycho* (Makowski, 2018) package, by taking the difference between z transformed probabilities of hits (HI) and false alarms (FA) per subject (i.e., $d'$=z[HI] – z[FA]). Higher $d'$ scores represent a greater recognition memory of old pairs, controlling for potential high false alarm rates. Further, immediate recognition $d'$ scores represent a subjects average encoding strength across the pairs, whereas delayed recognition $d'$ scores represent their post-consolidation memory outcomes. Values were determined to be outliers if they were more than 1.5 times the interquartile range above quartile 3, or below quartile 1, and removed from analyses.

### 2.4.2. EEG Analyses

#### 2.4.2.1. EEG Preprocessing

The EEG signals for the EEG intervention conditions (i.e., retrieval, restudy, and sleep) were processed using MNE-Python v.1.7.0 (Gramfort et al., 2013). Activity from each intervention period was re-referenced to the mastoid electrodes M1 and M2 and down-sampled to 100 Hz (required for spindle detection). A notch filter (50 Hz) removed interference from the Australian mains power supply. Also, a finite impulse response filter using a Hamming window was applied to remove slow signal drifts below 1 Hz and high frequency activity above 40 Hz. Then, an Independent Component Analysis was applied to the retrieval and restudy intervention's data to correct electrooculographic (EOG) artifacts, excluding components which correlated the most strongly with EOG events (via the create_eog_epochs function in MNE-Python). Equal length 30 sec epochs were calculated

across the six study periods for the retrieval and restudy interventions, and during NREM sleep in the sleep intervention. EEG was segmented into 30 sec epochs for consistency across conditions, as this is the length of epochs used during sleep scoring. Sleep scoring of EEG data during the sleep intervention was performed as per established procedures (Berry et al., 2017).

2.4.2.2. Individual Alpha Frequency (IAF)

IAF was estimated using EEG data obtained from a 2 min eyes-closed resting state recording, during which alpha activity is typically most prominent (Grandy et al., 2013). This was represented by the centre of gravity, a weighted average of the power within the alpha band (~7–13 Hz) that circumvents problems with identifying a dominant alpha peak. The alpha centre of gravity was calculated in MNE Python v. 0.23.4, using methods from Corcoran et al. (2018). Subjects' IAF values were used to create subject-specific frequency bands for statistical modelling. This was done using the golden mean-based algorithm described in (Pletzer et al., 2010), whereby the upper and lower limits for the sigma bands was defined based on the subject's IAF.

2.4.2.3. Sigma Power

Time frequency representations were calculated using a Morlet wavelet transformation (Cohen, 2019) within subject-specific sigma bands (~11–16 Hz), for five cycles per 1 Hz frequency step. Absolute power was calculated, across each 30 sec epoch (NREM for sleep, during study for retrieval and restudy conditions), per electrode and sampled time point. These power values were averaged across electrodes, time, and epochs, so that each subject session (one visit to the laboratory per subject) had a single value of average sigma power experienced across the intervention, measured in $\mu V^2/Hz$.

2.4.2.4. Sleep Spindles

The offline detection of sleep spindles in the sleep intervention condition, and the calculation of their parameters (i.e., density, frequency, and amplitude), was done via established methods in MNE-Python v.0.23.4 (Klinzing et al., 2016) using the YASA toolbox (Vallat & Walker, 2021). The data was first narrowed to a subset of electrodes (Fz, F3, F4, Cz, C3, C4, Pz, P3, P4, O1, O2) and NREM epochs. Individual peak spindle frequencies were identified per subject, to customise the detection of spindles events within 2 Hz either side of each subject's peak. To detect spindles, a sliding window of 200 ms was used as a threshold to compute root-mean-square (RMS) values. If the RMS exceeded 1.5 standard

deviations of the filtered signal for 500–3000 ms, the window was considered a sleep spindle (Lacourse et al., 2019). From these spindles, the density (number of spindle events divided the epoch length), frequency (median hertz of the spindles), and amplitude (peak-to-peak microvolt amplitude of detrended spindles) were calculated per epoch, channel, and sleep stage. These values were then averaged to obtain one value of each metric per subject's condition.

### 2.4.3. Statistical analyses

#### 2.4.3.1. Full sample

To test various predictions and interactions between variables in the full sample, three linear mixed-effects models were conducted in R v4.3.0 (R Core Team, 2022) using the *lme4* package (Bates et al., 2015). To determine if significant differences existed between the conditions within each model, we used treatment contrast coding with the base condition set to the sleep condition in each model.

The first model aimed to predict if delayed recognition accuracy differed between all four intervention conditions, controlling for differences in immediate recognition accuracy. This was performed to gage the general improvement of recognition accuracy across all four conditions, as the wake condition could not be included in the subsequent modelling. This model used $d'$ scores at delayed recognition as the outcome, predicted by the fixed effect of intervention condition (i.e., retrieval, restudy, sleep, and wake). This model controlled for $d'$ scores at immediate recognition and KSS scores at delayed recognition to control for the effects of sleepiness on one's behavioural performance. Both $d'$ scores were scaled by subtracting the mean from all values, so the results could be compared with the second model where this scaling was necessary. Lastly, subject ID was used as a random effect.

The second linear mixed-effects model was conducted to test the primary prediction of the study, that greater sigma power would increase the enhancement of weakly encoded memories across the intervention period, across retrieval and sleep only. Only the three EEG intervention conditions (i.e., retrieval, restudy, and sleep) were included in this model and the next, due to no EEG recording occurring in the wake condition. This model used $d'$ scores at delayed recognition as the outcome, predicted by the fixed effects of the EEG intervention conditions, $d'$ scores at immediate recognition (encoding strength), and sigma power. Both $d'$ scores and the sigma power values were scaled by subtracting the mean from all values, to

allow the model to accurately compare values on different scales. Additionally, this model controlled for delayed recognition KSS scores, and subject ID was used as a random effect.

Due to the exploratory nature of aiming to find an analogous neural mechanism involved in retrieval-mediated and sleep-based memory consolidation, another model aimed to test for the EEG intervention condition differences in sigma power. This model used the subject-session-level (one visit to the laboratory) sigma power as the outcome, scaled to centre by subtracting the mean from all values. Sigma power was predicted by a fixed effect of EEG intervention condition, with subject ID entered as a random effect. For both linear mixed-effects models, the modelled effects of the most significant interactions were extracted using the *effects* package (Fox et al., 2022), and used for plotting.

2.4.3.2. Sleep subgroup

A series of exploratory models were conducted on only the nap condition's data ($n$=26). The first set of models aimed to explore if sleep spindle parameters moderated the enhancement of weakly encoded memories. These models were performed to see if our study could replicate the findings of previous research, and therefore support any relationships we observe with sigma power. Three linear models were conducted to predict $d'$ scores at delayed recognition as the outcome, using $d'$ scores at immediate recognition (encoding strength) as a predictor and the KSS scores at delayed recognition as a covariate. Each of the three models then included a different sleep spindle parameter as a predictor: density, frequency, or amplitude. These models were designed to mirror the first full sample model, with the only difference being the ID was not used as a random effect, as the nap group alone did not contain multiple values per subject.

Finally, exploratory models were conducted to give further evidence to whether sigma power was an effective proxy for sleep spindle parameters. To do this, three Pearson-$r$ correlations were conducted to explore if any of our sleep spindle parameters correlated with sigma power, during NREM sleep. Again, each model incorporated one of the three sleep spindle parameters: density, frequency, or amplitude.

3. Results

Table 1 displays the sleep parameters of subjects during the nap intervention. All subjects experienced NREM sleep during their nap opportunity, and only five subjects did not enter slow-wave sleep. Four subjects demonstrated rapid eye movement (REM) sleep during their nap.

Table 1. Sleep characteristics of the nap intervention

| Sleep Measure | Mean (SD) | Range |
|---|---|---|
| TST (min) | 82.44 (26.31) | 25.00–116.00 |
| SOL (min) | 11.90 (11.11) | 1.50–53.00 |
| N1 (%) | 16.70 (9.54) | 5.00–49.50 |
| N2 (%) | 43.90 (14.00) | 19.50–71.00 |
| SWS (%) | 19.00 (18.31) | 0.00–56.00 |
| REM (%) | 2.87 (6.30) | 0.00–20.50 |

*Note.* TST = total sleep time, SOL = sleep onset latency, N1 = stage 1 non-rapid eye movement (NREM) sleep, N2 = stage 2 NREM sleep, SWS = slow wave sleep, REM = rapid eye movement sleep.

3.1. Recognition accuracy improvements were greatest after retrieval and restudy interventions

Table 2 depicts the *d′* recognition accuracy scores at immediate and delayed recognition. These values indicate that recognition accuracy improved somewhat from before to after each intervention.

Table 2. Recognition accuracy *d′* scores across each condition

| Intervention Condition | Mean Immediate Recognition *d′* (SD) | Mean Delayed Recognition *d′* (SD) |
|---|---|---|
| Retrieval | 0.93 (0.78) | 2.49 (0.90) |
| Restudy | 0.95 (0.81) | 2.20 (1.06) |
| Sleep | 1.01 (0.77) | 1.42 (0.91) |
| Wake | 0.75 (0.50) | 0.94 (0.59) |

*Note.* Mean values are reported across subjects, per condition. SD=standard deviation, which is reported in brackets.

Additionally, a linear mixed-effects model tested for delayed recognition accuracy differences between all four intervention conditions. As differences in immediate recognition accuracy were observed between conditions (Table 2), this was controlled for in the model. The model found a significant main effect of condition to predict delayed recognition accuracy, $\chi 2(3)=129.59$, $p<.001$. Based on Fig, delayed recognition accuracy was greatest after retrieval and restudy, compared to sleep and wake interventions. Treatment contrast coding confirmed that the retrieval-sleep ($\beta=1.03$, $p<.001$) and restudy-sleep ($\beta=0.77$, $p<.001$) comparisons were statistically significant.

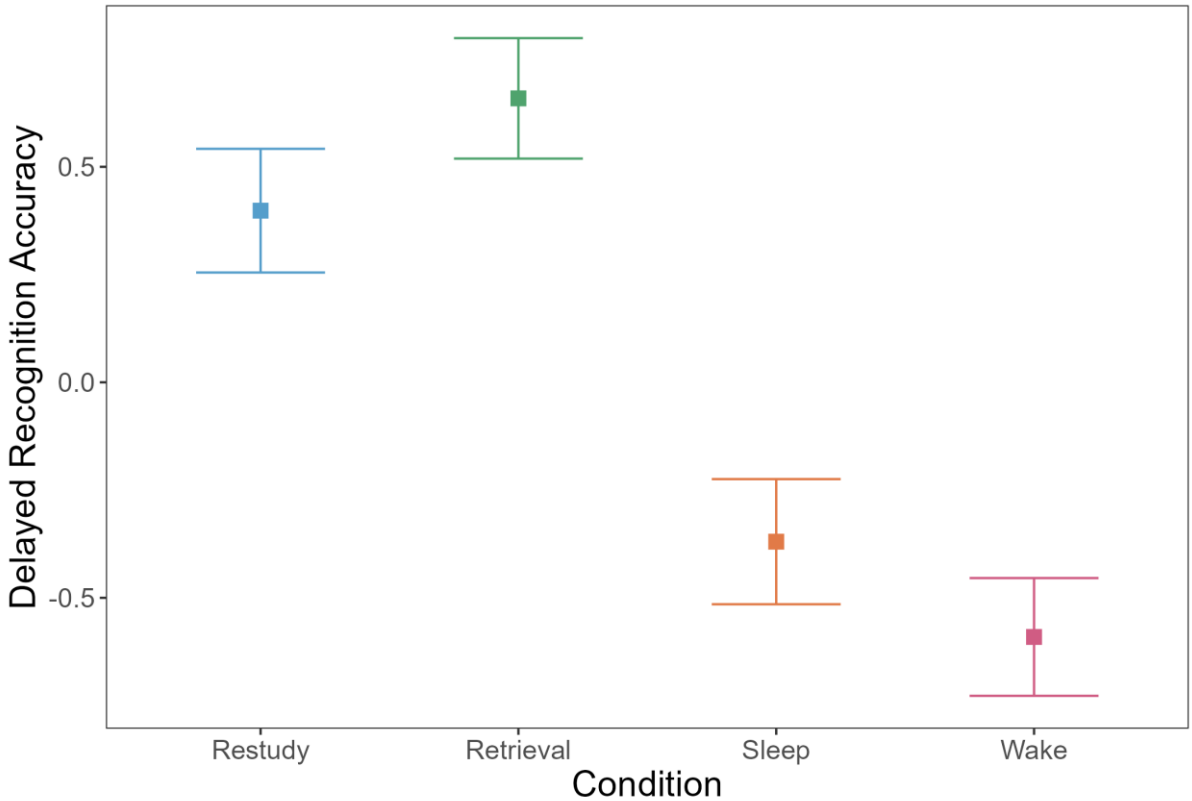


Fig. 3. Condition differences in delayed recognition accuracy. The x-axis represents modelled and centred delayed recognition accuracy, measured via d prime. The error bars represent the 83% confidence intervals.

3.2. Encoding strength, intervention condition, and sigma power individually predict delayed recognition accuracy

A linear mixed-effects model was conducted to detect if sigma power moderates an enhancement of weakly encoded memories over strongly encoded ones. This model found no evidence for a significant interaction between sigma power, immediate recognition $d'$ (encoding strength), and EEG intervention condition to predict delayed recognition $d'$, $\chi 2(1)=0.86$, $p>.05$. The model did reveal three significant main effects of immediate recognition $d'$ ($\chi 2(1)=77.56$, $p<.001$), condition ($\chi 2(2)=49.86$, $p<.001$), and sigma power ($\chi 2(1)=5.27$, $p=.022$) to predict delayed recognition $d'$. These three significant main effects from the model are plotted in Fig. 3. No further main or interaction effects were observed between any predictors in the model. Treatment contrast coding also confirmed the same significant comparisons as the model in section 3.1.

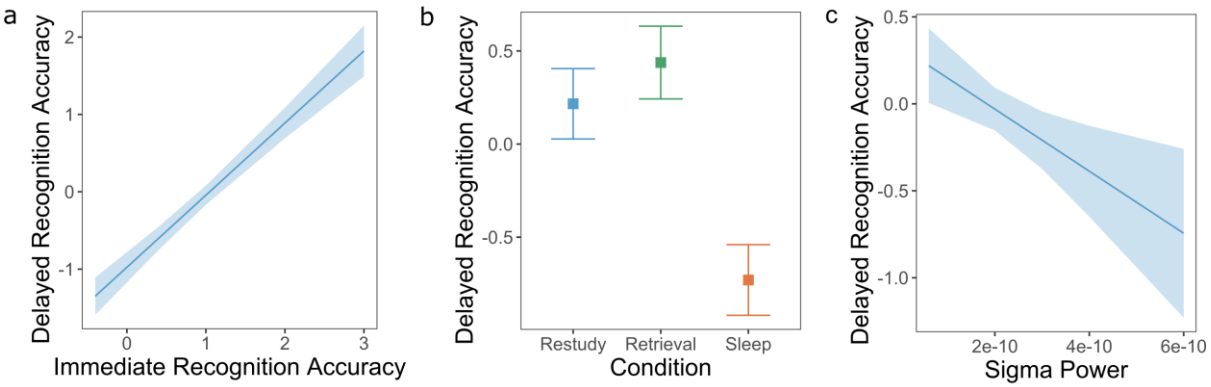


Fig. 3. Main effects of encoding strength, intervention condition, and sigma power to predict delayed recognition accuracy Across all three plots, delayed recognition accuracy is measured as d prime (*d′*), centred around zero, and reflects the modelled effects. a) The main effect of immediate recognition *d′* (encoding strength) to predict delayed recognition *d′*. b) The main effect of intervention condition to predict delayed recognition *d′*. The error bars represent the 83% confidence intervals. c) The main effect of sigma power to predict delayed recognition *d′*. Sigma power units are in $\mu V^2/Hz$.

3.3. Sigma power was greater during retrieval training rounds than during NREM sleep

To initially test for differences in subjects' average sigma power during each intervention condition, a linear mixed effects model was conducted. The model revealed that sigma power was significantly predicted by condition, $\chi 2(2)=6.28$, $p=.043$. These condition differences in sigma power are visualised in Fig. 4. Treatment contrast coding revealed that the retrieval-sleep comparison was statistically significant ($\beta=0.46$, $p=.017$).

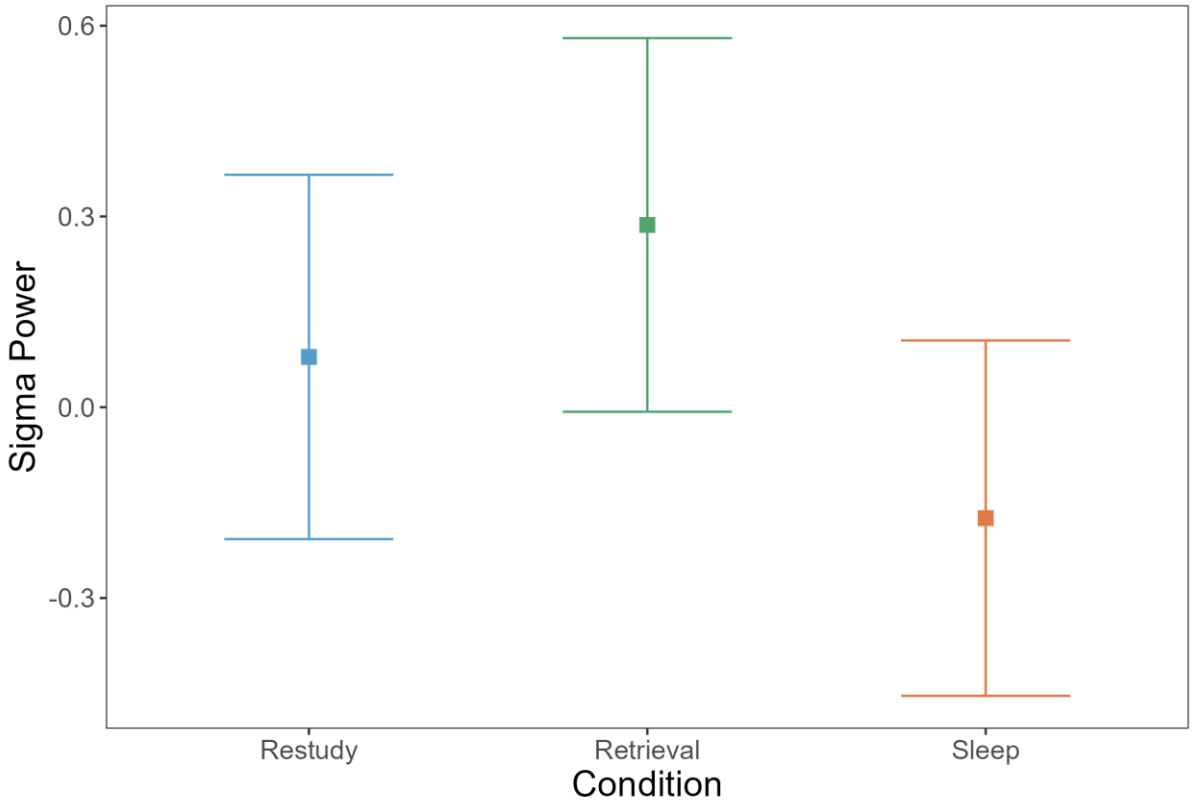


Fig. 4. Differences in sigma power during EEG memory interventions. The y-axis represents the modelled and centred values for sigma power during the three intervention conditions. The error bars represent the 83% confidence intervals.

3.4. Sleep spindle parameters showed no relationship to recognition accuracy, nor sigma power

As we were not able to detect the enhancement of weakly encoded memories, nor a moderating role of sigma power, additional exploration was performed on the sleep group only, to replicate previous findings that sleep spindle parameters moderate weakly encoded memories. This was tested using three linear regression models, replicating the sigma power model described in section 3.2. Each of these models predicted delayed recognition $d'$ from immediate recognition $d'$ (encoding strength), with either spindle density, frequency, or amplitude instead of sigma power as the other predictor. However, no significant interaction effect was found between immediate recognition $d'$ and either spindle density ($F(1)=0.74$, $p>.05$), frequency ($F(1)=0.0009$, $p>.05$), or amplitude ($F(1)=0.005$, $p>.05$), to predict delayed recognition $d'$.

Finally, we tested if sigma power was related to each spindle parameter, and if it therefore served as an effective proxy measurement. To accomplish this, we conducted three Pearson-*r* correlations with sigma power and each of the three spindle parameters (spindle

density, frequency, and amplitude). Each of the three correlations determined there to be a non-significant relationship between sigma power and spindle density ($r(24)=-0.11$, $p>.05$), frequency ($r(24)=-0.13$, $p>.05$), and amplitude ($r(24)=0.2$, $p>.05$).

## 4. Discussion

The broad goal of this study was to test if memory consolidation has a consistent brain-based mechanism across sleep and retrieval training. To test this, we investigated if both sleep and retrieval training interventions enhanced weakly encoded memories over strongly encoded memories, and the role played by sigma power. We also investigated if sigma power and delayed recognition accuracy differed between the intervention conditions. Lastly, we attempted to correlate sleep spindle parameters with sigma power and replicate the enhancement of weakly encoded memories across sleep, moderated by sleep spindle parameters instead. Contrary to previous research, we did not detect an enhancement of weakly encoded memories, moderated by either sigma power or sleep spindles. Intervention condition, encoding strength, and sigma power did not interact to predict memory outcomes. Nonetheless, we did find that intervention condition, encoding strength, and sigma power predicted recognition accuracy, independently. Despite previous assumptions (Piosczyk et al., 2013; Schechtman et al., 2023), our findings also suggest that sigma power may not be an effective proxy for spindle parameters, as we did not detect a relationship between sigma power and sleep spindle parameters.

Consistent with the expectations that practice has the greatest impact on memory outcomes, retrieval and restudy conditions produced significantly higher delayed recognition accuracy than the sleep and wake interventions. This is in line with research that retrieval and restudy similarly boost initial memory accuracy, but the benefits of retrieval over restudy typically only appear after longer delays (Bäuml & Trißl, 2022; Ferreira et al., 2019; Ferreira & Wimber, 2021; Guran et al., 2020). Whilst there was still an enhancement of recognition accuracy scores across the sleep and wake interventions, the magnitude of this was far smaller than either practice intervention. This is not entirely surprising, as greater exposure to memories, and the ability to update them, leads to a better recognition accuracy (Hintzman, 1970). Instead, the slight enhancement of memories across sleep is more in line with research that sleep typically protects memories from short-term decay (Denis et al., 2023; Dumay, 2016; Rasch & Born, 2013).

It is important to consider that condition differences in recognition accuracy do not speak entirely to a difference between sleep-based and retrieval-mediated consolidation mechanisms. This is because recognition accuracy alone does not give absolute insight into consolidation mechanisms. Rather, it is typically the context of behavioural scores that reflect consolidation more reliably (e.g., considering encoding strength and sigma-band activity). More direct comparisons of sleep-based and retrieval-mediated consolidation with other behavioural markers and paradigms (e.g., lure endorsement, recall accuracy, or 1-week follow up testing) will be useful to see if any similarities or differences exist in how they uniquely reflect memory changes. These comparisons will more effectively speak to the fine-grained nature of memory changes across consolidation.

We were unable to detect an enhancement of weakly encoded memories across retrieval or sleep conditions. Instead, we detected a positive relationship between encoding strength and subsequent memory outcomes, regardless of condition. Various studies also report that strongly encoded memories are prioritised for consolidation instead of weaker ones, and that this is also moderated by sleep spindle parameters (Schoch et al., 2017; Tucker & Fishbein, 2008; Wernette & Fenn, 2024; Wislowska et al., 2017). These opposing findings are typically attributed to different studies exploring different halves of an inverted-U-shaped curve, wherein too weakly or too strongly encoded memories are not consolidated, and moderately encoded memories are instead prioritised (Stickgold, 2009). In our study, we expected that learning relatively simple paired associates would place subjects across the right-most half of the inverted-U-shaped curve, wherein weaker encoded memories are prioritised. Alternatively, it is likely that these results are independent from the effects of consolidation, and therefore have no place on either side of the inverted-U-shaped curve. In our study, the enhancement of strongly encoded memories was not moderated by sigma power or any sleep spindle parameters, and the observed relationship did not differ in the restudy and wake conditions where consolidation was not being induced. Instead, it is plausible that our subject-level measure of broad immediate recognition may have been influenced by subjects' trait-like learning ability, which reflects a separate process to their state-like encoding strength (Creery et al., 2015; Denis et al., 2020). Other studies have attempted to avoid this influence by experimentally manipulating encoding strength (Drosopoulos et al., 2007; Schapiro et al., 2017), but research has still found that allowing for a more naturalistic measure of encoding strength, such as ours, still detects these relationships (Chatburn et al., 2021; Djonlagic et al., 2009). Overall, this relationship is likely not due to

consolidation-related changes, but rather trait-like influences of learning ability that do not speak to a commonality unique to sleep and retrieval interventions.

Sigma power during the intervention period was negatively related to memory outcomes, contrary to the beneficial role of sleep spindles on memory outcomes (Antony et al., 2018; Gibson et al., 2022; Helm et al., 2011; Petzka, Chatburn, et al., 2021). This result cannot be explained by a multicollinear relationship between sigma power and condition, as the retrieval condition involved higher sigma power than sleep and greater delayed recognition accuracy, not poorer. Additionally, our custom sigma-band windows were determined based on the individual alpha peaks of each subject, so it is unlikely that we captured activity involved in neighbouring bands (Pletzer et al., 2010). However, this approach is not taken by most other research, which may have contributed to our differential findings. Regardless, for sigma power to relate to memory outcomes without interacting with the intervention condition variable, means that sigma power may not reflect a common mechanism between sleep-based and retrieval-mediated consolidation. Instead, it may reflect an inhibitory mechanism for memory enhancement across multiple interventions and states, and/or a byproduct of poor encoding. If sigma power in this study was reflecting greater sensory gating activity of spindle-related neurons in the thalamus (Z. Chen et al., 2016; Schreiner & Staudigl, 2020) then perhaps a reduced intake of the stimuli during retrieval and restudy practice would have been detrimental to subsequent memory outcomes. Alternatively, if sigma power was related to the reinstatement of initial memory contexts (Schechtman et al., 2023), this continued reinstatement during the interventions may have reflected poor unbinding of relevant memory components from their irrelevant initial contexts, reducing the extent of productive memory pruning. Overall, the precise roles that sigma power plays are unclear at present. To greater understand the role of sigma power during memory interventions, future research should continue to identify in which contexts sigma power predicts memory outcomes, and if it plays a different role during consolidation or across different mental states.

One key aspect that may explain why we did not detect an enhancement of weakly encoded memories moderated by sigma power, is that we estimated sigma power across large epochs during key periods of the interventions (e.g., Piosczyk et al., 2013), rather than around specific reactivation events (e.g., Schechtman et al., 2023). Antony and colleagues (2018) found that only the spindle events that were locked to memory reactivations, predicted memory outcomes. Therefore, measuring sigma power and sleep spindles outside of

reactivation events, whilst unavoidable in our paradigm, may have incorporated a large amount of noise in our measure. This may have abstracted our sigma power measurement beyond the point that it was able to reflect sleep spindle processes. This is supported by the fact that we were unable to detect a relationship between spindle parameters and sigma power. Additionally, we explored if power in other frequency bands (delta, theta, alpha, and beta) calculated across 30 sec epochs predicted memory outcomes. These models did not detect any significant relationships, despite research linking power in these frequency bands to memory outcomes, when power is locked to specific events (Aktürk et al., 2022; Hanslmayr et al., 2009; Klimesch et al., 2006; Vivekananda et al., 2021). Instead, sigma power across sleep and wake may reflect background memory-related processes occurring across the whole intervention that are independent of individual memory events.

Despite previous research, we did not find a role of spindle parameters in enhancing weakly encoded memories. These models were unlikely to be underpowered, as similar sample and stimuli sizes have been used previously to detect effects in similar studies (Baena et al., 2020; Chatburn et al., 2021; Creery et al., 2015; Denis et al., 2021; Johnson et al., 2019; Schreiner et al., 2018; Zeng et al., 2021). These results are also not likely to be influenced by age-related differences in sleep EEG, as our sample did not contain subjects over 35 years of age (Landolt & Borbély, 2001). Additionally, whilst our study used a nap rich in NREM opposed to a full night of sleep, two meta-analyses have found no significant differences in the impact of spindle activity on memories between naps and full-nights of sleep (P. Chen et al., 2024; Kumral et al., 2023). Instead, our inability to detect a role of spindles in this study could be due to consolidation not acting upon weakly encoded memories in such a way that it is detectable through recognition accuracy. Perhaps, the enhancement of weakly encoded memories may be more accurately described as an enhancement of the *recall accuracy* of weakly encoded memories. Consolidation across sleep and retrieval may more specifically strengthen connections between weakly encoded memories and the retrieval cues, thereby increasing their accessibility, which is less necessary for the memory search undergone during recognition memory (Gillund & Shiffrin, 1984). Such a distinction is imperative for future research to investigate further by directly comparing recognition and recall accuracy across consolidation.

Our study pioneers a sound methodological approach for comparing retrieval and sleep interventions. Most previous studies that compare these consolidation interventions, do so with confounding sequential designs that produce ceiling effects, or draw cross-paradigm

comparisons despite the notorious paradigm-specific manifestations of consolidation-related memory changes (Cordi & Rasch, 2021; Diekelmann et al., 2009). Additionally, our use of control conditions for both retrieval and sleep interventions, aides in ruling out general practice effects on memory. To better compare sleep and retrieval training interventions, it would be useful for future studies to take our methodological approach of examining these interventions in parallel and in the same paradigm. Our approach is one of the first to do this, and continuing these direct comparisons will inform a more comprehensive theory of memory consolidation that considers memory changes across multiple states and contexts. More specifically, to compare more fine-grained time-resolved EEG mechanisms between sleep and retrieval training, future comparisons should isolate EEG around key consolidation-related events and track the item-level changes of memories. Additionally, investigating if encoding strength impacts certain aspects of memory outcomes (e.g., recognition, recall, false alarms) more than others, will further identify the fine-grained memory qualities that retrieval training and sleep types prioritise during consolidation.

In summary, while our results did not demonstrate a role of sigma-band activity in enhancing weakly encoded memories across both sleep-based and retrieval-mediated memory consolidation, we did find a potentially inhibitory role for post-learning sigma power on subsequent memory outcomes across sleep and wake. We also demonstrated that the enhancement of weakly encoded memories may not be ubiquitous enough to impact recognition memory. Most crucially, our study offers a successful method for comparing consolidation across different states, using parallel sleep and retrieval training interventions, and adequate control conditions for both, all across different sessions. Continuing to make similar comparisons in future will help understand these two consolidation types, as their impacts manifest in different paradigms.

Acknowledgements

We thank Annaliese Anesbury and Maya Al Safadi for their assistance during data collection. We also thank the subjects for their time.

CRediT authorship contribution statement

Hayley B. Caldwell: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing – original draft, Writing – review and editing.

Kurt Lushington: Conceptualization, Methodology, Project administration, Resources, Supervision, Writing – review and editing.

Alex Charburn: Conceptualization, Methodology, Project administration, Resources, Supervision, Writing – review and editing.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Abel, M., & Bäuml, K.-H. (2016). Retrieval practice can eliminate list method directed forgetting. *Memory & Cognition*, *44*(1), 15–23. https://doi.org/10.3758/s13421-015-0539-x

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, *52*(1–2), 29–37. https://doi.org/10.3109/00207459008994241

Aktürk, T., de Graaf, T. A., Erdal, F., Sack, A. T., & Güntekin, B. (2022). Oscillatory delta and theta frequencies differentially support multiple items encoding to optimize memory performance during the digit span task. *NeuroImage*, *263*, 119650. https://doi.org/10.1016/j.neuroimage.2022.119650

Antony, J. W., Ferreira, C. S., Norman, K. A., & Wimber, M. (2017). Retrieval as a fast route to memory consolidation. *Trends in Cognitive Sciences*, *21*(8), 573–576. https://doi.org/10.1016/j.tics.2017.05.001

Antony, J. W., & Paller, K. A. (2018). Retrieval and sleep both counteract the forgetting of spatial information. *Learning & Memory*, *25*(6), 258–263. https://doi.org/10.1101/lm.046268.117

Antony, J. W., Piloto, L., Wang, M., Pacheco, P., Norman, K. A., & Paller, K. A. (2018). Sleep spindle refractoriness segregates periods of memory reactivation. *Current Biology*, *28*(11), 1736-1743.e4. https://doi.org/10.1016/j.cub.2018.04.020

Ashton, J., Staresina, B., & Cairney, S. (2021). sleep bolsters schematically incongruent memories. *osf.* https://doi.org/10.31234/osf.io/w435f

Backhaus, J., Junghanns, K., Broocks, A., Riemann, D., & Hohagen, F. (2002). Test–retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia.

*Journal of Psychosomatic Research*, *53*(3), 737–740. https://doi.org/10.1016/S0022-3999(02)00330-6

Baena, D., Cantero, J. L., Fuentemilla, L., & Atienza, M. (2020). Weakly encoded memories due to acute sleep restriction can be rescued after one night of recovery sleep. *Scientific Reports*, *10*(1), 1449. https://doi.org/10.1038/s41598-020-58496-4

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bäuml, K.-H. T., Holterman, C., & Abel, M. (2014). Sleep can reduce the testing effect: It enhances recall of restudied items but can leave recall of retrieved items unaffected. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(6), 1568–1581. https://doi.org/10.1037/xlm0000025

Bäuml, K.-H. T., & Trißl, L. (2022). Selective memory retrieval can revive forgotten memories. *Proceedings of the National Academy of Sciences*, *119*(8), e2114377119. https://doi.org/10.1073/pnas.2114377119

Berry, R. B., Brooks, R., Gamaldo, C., Harding, S. M., Lloyd, R. M., Quan, S. F., Troester, M. T., & Vaughn, B. V. (2017). AASM scoring manual updates for 2017 (Version 2.4). *Journal of Clinical Sleep Medicine*, *13*(05), 665–666. https://doi.org/10.5664/jcsm.6576

Bradley, M. M., & Lang, P. J. (1999). Affective norms for English words (ANEW): instruction manual and affective ratings. *Technical Report C-1*, 49.

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(2), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4

Carpenter, J. S., & Andrykowski, M. A. (1998). Psychometric evaluation of the pittsburgh sleep quality index. *Journal of Psychosomatic Research*, *45*(1), 5–13. https://doi.org/10.1016/S0022-3999(97)00298-5

Chatburn, A., Lushington, K., & Kohler, M. J. (2021). Consolidation and generalisation across sleep depend on individual EEG factors and sleep spindle density. *Neurobiology of Learning and Memory*, *179*, 107384. https://doi.org/10.1016/j.nlm.2021.107384

Chen, P., Hao, C., & Ma, N. (2024). Sleep spindles consolidate declarative memory with tags: A meta-analysis of adult data. *Journal of Pacific Rim Psychology*, *18*, 18344909241226761. https://doi.org/10.1177/18344909241226761

Chen, Z., Wimmer, R. D., Wilson, M. A., & Halassa, M. M. (2016). Thalamic circuit mechanisms link sensory processing in sleep and attention. *Frontiers in Neural Circuits*, *9*. https://doi.org/10.3389/fncir.2015.00083

Cohen, M. X. (2019). A better way to define and describe Morlet wavelets for time-frequency analysis. *NeuroImage*, *199*, 81–86. https://doi.org/10.1016/j.neuroimage.2019.05.048

Corcoran, A. W., Alday, P. M., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2018). Toward a reliable, automated method of individual alpha frequency (IAF) quantification. *Psychophysiology*, *55*(7), e13064. https://doi.org/10.1111/psyp.13064

Cordi, M. J., & Rasch, B. (2021). How robust are sleep-mediated memory benefits? *Current Opinion in Neurobiology*, *67*, 1–7. https://doi.org/10.1016/j.conb.2020.06.002

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. https://doi.org/10.5665/sleep.4670

Denis, D., Bottary, R., Cunningham, T. J., Tcheukado, M.-C., & Payne, J. D. (2023). The influence of encoding strategy on associative memory consolidation across wake and sleep. *Learning & Memory*, *30*(9), 185–191. https://doi.org/10.1101/lm.053765.123

Denis, D., DiPietro, C., Spreng, R. N., Schacter, D. L., Stickgold, R., & Payne, J. D. (2024). Sleep and retrieval practice both strengthen and distort story recollection. *Sleep Advances*, *5*(1), zpae083. https://doi.org/10.1093/sleepadvances/zpae083

Denis, D., Mylonas, D., Poskanzer, C., Bursal, V., Payne, J. D., & Stickgold, R. (2021). Sleep spindles preferentially consolidate weakly encoded memories. *Journal of Neuroscience*, *41*(18), 4088–4099. https://doi.org/10.1523/JNEUROSCI.0818-20.2021

Denis, D., Schapiro, A. C., Poskanzer, C., Bursal, V., Charon, L., Morgan, A., & Stickgold, R. (2020). The roles of item exposure and visualization success in the consolidation of memories across wake and sleep. *Learning & Memory*, *27*(11), 451–456. https://doi.org/10.1101/lm.051383.120

Diekelmann, S., Wilhelm, I., & Born, J. (2009). The whats and whens of sleep-dependent memory consolidation. *Sleep Medicine Reviews*, *13*(5), 309–321. https://doi.org/10.1016/j.smrv.2008.08.002

Djonlagic, I., Rosenfeld, A., Shohamy, D., Myers, C., Gluck, M., & Stickgold, R. (2009). Sleep enhances category learning. *Learning & Memory*, *16*(12), 751–755. https://doi.org/10.1101/lm.1634509

Dringenberg, H. C. (2019). Chapter 32 - Sleep and memory consolidation: conceptual and methodological challenges. In H. C. Dringenberg (Ed.), *Handbook of Behavioral Neuroscience* (Vol. 30, pp. 489–501). Elsevier. https://doi.org/10.1016/B978-0-12-813743-7.00032-3

Drosopoulos, S., Schulze, C., Fischer, S., & Born, J. (2007). Sleep's function in the spontaneous recovery and consolidation of memories. *Journal of Experimental Psychology: General*, *136*(2), 169–183. https://doi.org/10.1037/0096-3445.136.2.169

Dumay, N. (2016). Sleep not just protects memories against forgetting, it also makes them more accessible. *Cortex*, *74*, 289–296. https://doi.org/10.1016/j.cortex.2015.06.007

Ferreira, C. S., Charest, I., & Wimber, M. (2019). Retrieval aids the creation of a generalised memory trace and strengthens episode-unique information. *NeuroImage*, *201*, 115996. https://doi.org/10.1016/j.neuroimage.2019.07.009

Ferreira, C. S., & Wimber, M. (2023). The testing effect for visual materials depends on preexisting knowledge.*Journal of Experimental Psychology: Learning, Memory, and Cognition, 49*(10), 1557–1571. https://doi.org/10.1037/xlm0001248

Fox, J., Weisberg, S., Price, B., Friendly, M., Hong, J., Andersen, R., Firth, D., Taylor, S., & R Core Team. (2022). *effects: Effect Displays for Linear, Generalized Linear, and Other Models* (Version 4.2-2) [Computer software]. https://cran.r-project.org/web/packages/effects/index.html

Gibson, T., Cross, Z. R., & Chatburn, A. (2022). Theta activity during encoding interacts with nrem sleep oscillations to predict memory generalization. *Frontiers in Human Neuroscience*, *16*. https://doi.org/10.3389/fnhum.2022.821191

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*(1), 1–67. https://doi.org/10.1037/0033-295X.91.1.1

Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*. https://www.frontiersin.org/articles/10.3389/fnins.2013.00267

Grandy, T. H., Werkle-Bergner, M., Chicherio, C., Lövdén, M., Schmiedek, F., & Lindenberger, U. (2013). Individual alpha peak frequency is related to latent factors of general cognitive abilities. *NeuroImage*, *79*, 10–18. https://doi.org/10.1016/j.neuroimage.2013.04.059

Guran, C.-N. A., Lehmann-Grube, J., & Bunzeck, N. (2020). Retrieval practice improves recollection-based memory over a seven-day period in younger and older adults. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02997

Halamish, V., & Bjork, R. A. (2011). When does testing enhance retention? a distribution-based interpretation of retrieval as a memory modifier. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*(4), 801–812. https://doi.org/10.1037/a0023219

Hanslmayr, S., Spitzer, B., & Bäuml, K.-H. (2009). Brain oscillations dissociate between semantic and nonsemantic encoding of episodic memories. *Cerebral Cortex*, *19*(7), 1631–1640. https://doi.org/10.1093/cercor/bhn197

Hebscher, M., Wing, E., Ryan, J., & Gilboa, A. (2019). Rapid cortical plasticity supports long-term memory formation. *Trends in Cognitive Sciences*, *23*(12), 989–1002. https://doi.org/10.1016/j.tics.2019.09.009

Helm, E. van der, Gujar, N., Nishida, M., & Walker, M. P. (2011). Sleep-dependent facilitation of episodic memory details. *PLOS ONE*, *6*(11), e27421. https://doi.org/10.1371/journal.pone.0027421

Himmer, L., Müller, E., Gais, S., & Schönauer, M. (2017). Sleep-mediated memory consolidation depends on the level of integration at encoding. *Neurobiology of Learning and Memory*, *137*, 101–106. https://doi.org/10.1016/j.nlm.2016.11.019

Hintzman, D. L. (1970). Effects of repetition and exposure duration on memory. *Journal of Experimental Psychology*, *83*(3, Pt.1), 435–444. https://doi.org/10.1037/h0028865

Horne, J. A., & Östberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology*, *4*, 97–110.

Johnson, B. P., Scharf, S. M., Verceles, A. C., & Westlake, K. P. (2019). Use of targeted memory reactivation enhances skill performance during a nap and enhances declarative memory during wake in healthy young adults. *Journal of Sleep Research*, *28*(5), e12832. https://doi.org/10.1111/jsr.12832

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the Karolinska Sleepiness Scale against performance and EEG Variables. *Clinical Neurophysiology*, *117*(7), 1574–1581. https://doi.org/10.1016/j.clinph.2006.03.011

Klimesch, W., Doppelmayr, M., & Hanslmayr, S. (2006). Upper alpha ERD and absolute power: Their meaning for memory performance. In C. Neuper & W. Klimesch (Eds.), *Progress in Brain Research* (Vol. 159, pp. 151–165). Elsevier. https://doi.org/10.1016/S0079-6123(06)59010-7

Klinzing, J. G., Mölle, M., Weber, F., Supp, G., Hipp, J. F., Engel, A. K., & Born, J. (2016). Spindle activity phase-locked to sleep slow oscillations. *NeuroImage*, *134*, 607–616. https://doi.org/10.1016/j.neuroimage.2016.04.031

Kumral, D., Matzerath, A., Leonhart, R., & Schönauer, M. (2023). Spindle-dependent memory consolidation in healthy adults: A meta-analysis. *Neuropsychologia*, *189*, 108661. https://doi.org/10.1016/j.neuropsychologia.2023.108661

Lacourse, K., Delfrate, J., Beaudry, J., Peppard, P., & Warby, S. C. (2019). A sleep spindle detection algorithm that emulates human expert spindle scoring. *Journal of Neuroscience Methods*, *316*, 3–11. https://doi.org/10.1016/j.jneumeth.2018.08.014

Landolt, H.-P., & Borbély, A. A. (2001). Age-dependent changes in sleep EEG topography. *Clinical Neurophysiology*, *112*(2), 369–377. https://doi.org/10.1016/S1388-2457(00)00542-3

Li, S.-X., Li, Q., Wang, X.-F., Liu, L.-J., Liu, Y., Zhang, L.-X., Zhang, B., & Lu, L. (2011). Preliminary test for the Chinese version of the Morningness–Eveningness Questionnaire. *Sleep and Biological Rhythms*, *9*(1), 19–23. https://doi.org/10.1111/j.1479-8425.2010.00480.x

Lifanov, J., Linde-Domingo, J., & Wimber, M. (2021). Feature-specific reaction times reveal a semanticisation of memories over time and with repeated remembering. *Nature Communications*, *12*(1), 3177. https://doi.org/10.1038/s41467-021-23288-5

Lutz, N. D., Diekelmann, S., Hinse-Stern, P., Born, J., & Rauss, K. (2017). sleep supports the slow abstraction of gist from visual perceptual memories. *Scientific Reports*, *7*(1), 42950. https://doi.org/10.1038/srep42950

Mak, M. H., & Gaskell, M. G. (2023). *Effects of sleep and retrieval practice on verbal paired-associate learning across 12 and 24 hours*. OSF. https://doi.org/10.31234/osf.io/phe5j

Massand, E., & Bowler, D. M. (2013). Atypical neurophysiology underlying episodic and semantic memory in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(2), 298–315. https://doi.org/10.1007/s10803-013-1869-9

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457. https://doi.org/10.1037/0033-295x.102.3.419

Nicholls, M. E. R., Thomas, N. A., Loetscher, T., & Grimshaw, G. M. (2013). The Flinders Handedness survey (FLANDERS): A brief measure of skilled hand preference. *Cortex*, *49*(10), 2914–2926. https://doi.org/10.1016/j.cortex.2013.02.002

Pastötter, B., Schicker, S., Niedernhuber, J., & Bäuml, K.-H. T. (2011). Retrieval during learning facilitates subsequent memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(2), Article 2.

Petzka, M., Charest, I., Balanos, G. M., & Staresina, B. P. (2021). Does sleep-dependent consolidation favour weak memories? *Cortex*, *134*, 65–75. https://doi.org/10.1016/j.cortex.2020.10.005

Petzka, M., Chatburn, A., Charest, I., Balanos, G. M., & Staresina, B. P. (2022). Sleep spindles track cortical learning patterns for memory consolidation. *Current Biology*, *32*(11), 2349-2356.e4. https://doi.org/10.1016/j.cub.2022.04.045

Piosczyk, H., Holz, J., Feige, B., Spiegelhalder, K., Weber, F., Landmann, N., Kuhn, M., Frase, L., Riemann, D., Voderholzer, U., & Nissen, C. (2013). The effect of sleep-specific brain activity versus reduced stimulus interference on declarative memory consolidation. *Journal of Sleep Research*, *22*(4), 406–413. https://doi.org/10.1111/jsr.12033

Pišljar, N. T., Štukovnik, V., Kocjan, G. Z., & Dolenc-Groselj, L. (2019). Validity and reliability of the Slovene version of the Morningness-Eveningness Questionnaire. *Chronobiology International*, *36*(10), 1409–1417. https://doi.org/10.1080/07420528.2019.1651326

Pletzer, B., Kerschbaum, H., & Klimesch, W. (2010). When frequencies never synchronize: The golden mean and the resting EEG. *Brain Research*, *1335*, 91–102. https://doi.org/10.1016/j.brainres.2010.03.074

R Core Team. (2022). *The R Foundation*. R-Project.Org. https://www.r-project.org/foundation/

Rasch, B., & Born, J. (2013). About Sleep's Role in Memory. *Physiological Reviews*, *93*(2), 681–766. https://doi.org/10.1152/physrev.00032.2012

Ritvo, V. J. H., Turk-Browne, N. B., & Norman, K. A. (2019). Nonmonotonic plasticity: how memory retrieval drives learning. *Trends in Cognitive Sciences*, *23*(9), 726–742. https://doi.org/10.1016/j.tics.2019.06.007

Schapiro, A. C., McDevitt, E. A., Chen, L., Norman, K. A., Mednick, S. C., & Rogers, T. T. (2017). Sleep benefits memory for semantic category structure while preserving exemplar-specific information. *Scientific Reports*, *7*(1), 14869. https://doi.org/10.1038/s41598-017-12884-5

Schapiro, A. C., McDevitt, E. A., Rogers, T. T., Mednick, S. C., & Norman, K. A. (2018). Human hippocampal replay during rest prioritizes weakly learned information and predicts memory performance. *Nature Communications*, *9*(1), 3920. https://doi.org/10.1038/s41467-018-06213-1

Schechtman, E., Heilberg, J., & Paller, K. A. (2023). Memory consolidation during sleep involves context reinstatement in humans. *Cell Reports*, *42*(4), 112331. https://doi.org/10.1016/j.celrep.2023.112331

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. https://doi.org/10.1016/j.nlm.2017.10.009

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. https://doi.org/10.1016/j.celrep.2018.09.037

Schreiner, T., & Staudigl, T. (2020). Electrophysiological signatures of memory reactivation in humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1799), 20190293. https://doi.org/10.1098/rstb.2019.0293

Staresina, B. P. (2024). Coupled sleep rhythms for memory consolidation. *Trends in Cognitive Sciences*, *28*(4), 339–351. https://doi.org/10.1016/j.tics.2024.02.002

Stark, S. M., Kirwan, C. B., & Stark, C. E. L. (2019). Mnemonic Similarity Task: A tool for assessing hippocampal integrity. *Trends in Cognitive Sciences*, *23*(11), 938–951. https://doi.org/10.1016/j.tics.2019.08.003

Stickgold, R. (2009). How do I remember? Let me count the ways. *Sleep Medicine Reviews*, *13*(5), 305–308. https://doi.org/10.1016/j.smrv.2009.05.004

Tucker, M. A., & Fishbein, W. (2008). Enhancement of declarative memory performance following a daytime nap is contingent on strength of initial task acquisition. *Sleep*, *31*(2), 197–203. https://doi.org/10.1093/sleep/31.2.197

Vallat, R., & Walker, M. P. (2021). An open-source, high-performance tool for automated sleep staging. *eLife*, *10*, e70092. https://doi.org/10.7554/eLife.70092

Vivekananda, U., Bush, D., Bisby, J. A., Baxendale, S., Rodionov, R., Diehl, B., Chowdhury, F. A., McEvoy, A. W., Miserocchi, A., Walker, M. C., & Burgess, N. (2021). Theta power and theta-gamma coupling support long-term spatial memory retrieval. *Hippocampus*, *31*(2), 213–220. https://doi.org/10.1002/hipo.23284

Wernette, E. M. D., & Fenn, K. M. (2024). Factors influencing sleep-dependent consolidation: Sleep strengths memory based on encoding depth but not repetition. *Journal of Sleep Research*. https://doi.org/10.1111/jsr.14091

Weymar, M., Löw, A., Modess, C., Engel, G., Gründling, M., Petersmann, A., Siegmund, W., & Hamm, A. O. (2010). Propranolol selectively blocks the enhanced parietal old/new

effect during long-term recollection of unpleasant pictures: A high density ERP study. *NeuroImage*, *49*(3), 2800–2806. https://doi.org/10.1016/j.neuroimage.2009.10.025

Wislowska, M., Heib, D. P. J., Griessenberger, H., Hoedlmoser, K., & Schabus, M. (2017). *Individual baseline memory performance and its significance for sleep-dependent memory consolidation*. https://doi.org/10.1556/2053.1.2016.001

Zeng, S., Lin, X., Wang, J., & Hu, X. (2021). Sleep's short-term memory preservation and long-term affect depotentiation effect in emotional memory consolidation: Behavioral and EEG evidence. *Sleep*, *44*(11), zsab155. https://doi.org/10.1093/sleep/zsab155