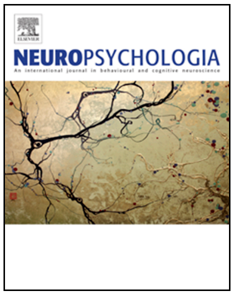

No Evidence for a Targeted Memory Reactivation Effect on Word-Meaning Priming

Lewis V. Ball, Eva Kimel, Vanessa G. Keller, Eloise Ward, Scott A. Cairney, Matthew H.C. Mak, Lu Li, Jennifer M. Rodd, M.Gareth Gaskell

PII: S0028-3932(25)00199-X

DOI: https://doi.org/10.1016/j.neuropsychologia.2025.109264

Reference: NSY 109264

To appear in: *Neuropsychologia*

Received Date: 12 February 2025

Revised Date: 15 July 2025

Accepted Date: 27 August 2025

Please cite this article as: Ball, L.V., Kimel, E., Keller, V.G., Ward, E., Cairney, S.A., Mak, M.H.C., Li, L., Rodd, J.M., Gaskell, M.G., No Evidence for a Targeted Memory Reactivation Effect on Word-Meaning Priming, *Neuropsychologia*, https://doi.org/10.1016/j.neuropsychologia.2025.109264.

No Evidence for a Targeted Memory Reactivation Effect on Word-Meaning Priming

Lewis V. Ball[1], Eva Kimel[1], Vanessa G. Keller[1], Eloise Ward[1], Scott A. Cairney[1], Matthew H.C. Mak[2], Lu Li[3], Jennifer M. Rodd[4], M. Gareth Gaskell[1]

1: Department of Psychology, University of York, UK
2: Department of Psychology, University of Warwick, UK
3: Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong, China
4: Division of Psychology and Language Sciences, University College London, UK

Corresponding author at Department of Psychology, University of York, Heslington, York YO10 5DD, UK; Email address: lewis.ball@york.ac.uk (Lewis V. Ball)

Abstract

The subordinate meaning of a homonym becomes temporarily more accessible after it is encountered, an effect termed word-meaning priming. Over the longer-term, word-meaning priming is better maintained across periods of sleep compared with wakefulness. This has been explained as sleep actively consolidating episodic memories related to recent linguistic events (Gaskell et al., 2019). Here, we tested this hypothesis by investigating whether word-meaning priming can be boosted following sleep using targeted memory reactivation (TMR), a technique of biassing specific memories for sleep-based consolidation by presenting information-associated sensory cues during sleep. In an exposure phase, 40 (of 80) homonyms were primed toward their subordinate meaning via a sentence, which was also associated with an auditory cue (the homonym) for TMR. Participants then took a ~2 hour nap, where half of the cues from exposure (memory cues) were replayed with the aim of strengthening the subordinate sentence meaning, along with 20 cues that had not been encountered previously (control cues). After sleep, there was an overall word-meaning priming effect, however there was no additional benefit of TMR on priming, nor did TMR benefit the recall of contextual information. Interestingly, there was an increased sleep spindle/beta band power response to memory cues relative to control cues, indicating cue-evoked memory reprocessing during sleep. These findings are consistent with a bounded role of sleep in actively consolidating linguistic-related memories.

Keywords: Episodic memory, Language processing, Sleep, Memory reactivation, Semantic ambiguity

Data availability: The research data and analysis code is available on the Open Science Framework (https://osf.io/frkms/?view_only=d754fa5eb708479c95adffdc243e87ef).

1. Introduction

Language is strongly modulated by the context in which it is used, and individuals must make use of the available contextual information to decode incoming language and communicate appropriately. Consider, for example, a scenario where two individuals are walking through a forest. One member of the couple suddenly announces their admiration for the different styles of bark on display. In this instance, they are very likely to be referring to the outer coverings of the trees in the forest as opposed to the noise made by a dog.

The word *bark* is labelled as a homonym in the psycholinguistic domain, defined as a word with more than one distinct meaning. Typically, the meaning that is used most frequently in language (the dominant meaning) is more accessible and processed more efficiently than lower-frequency (subordinate) meanings (Rayner & Duffy, 1986; Rodd, 2020; Vitello & Rodd, 2015). For instance, when completing a word-association task that involves generating an associated word of a homonym, participants tend to provide responses that relate to the dominant meaning (*dog, woof*) more frequently than the subordinate meaning (*tree, branch*) (Gilbert & Rodd, 2022; Twilley et al., 1994). Interestingly, these preferences are not set in stone and are sensitive to recent language experiences. For example, when homonyms are encountered in a sentence context that supports the subordinate interpretation (e.g., "*the branches and the bark had been damaged by the storm*"), participants in a subsequent word-association test tend to generate a greater proportion of responses that are consistent with the primed subordinate meaning, compared to a control condition without prior sentence exposure (Betts et al., 2018; Blott et al., 2022; Gaskell et al., 2019; Gilbert et al., 2018; 2021; Rodd et al., 2013; 2016; see Parker et al., 2023 for a similar processing advantage in eye-tracking data). This effect of exposure on subsequent homonym interpretation is termed *word-meaning priming*.

Initial accounts of word-meaning priming assumed that accessing the subordinate interpretation triggered long-term changes to pre-existing lexical-semantic representations. In the case of homonymy, phonological/orthographic representations are assumed to map onto multiple different semantic representations according to connectionist accounts. When a homonym is recognised, activity within the semantic network initially represents an unhelpful 'blend state' of the multiple meanings (Rodd, 2020). To rectify this, recurrent connections between semantic units help the network to settle onto a single meaning which, based on accumulated experiences with the word, likely corresponds to the dominant meaning (Armstrong & Plaut, 2016; Rodd, 2020; Rodd et al., 2004). However, a recent encounter with the subordinate meaning can strengthen the connections between the relevant semantic units, increasing the likelihood of the network settling onto that same meaning later (Gilbert et al., 2018; Rodd et al., 2016). This has been described as the *immediate alteration account* of word-meaning priming (Gaskell et al., 2019).

The immediate alteration account posits that language exposure leads to prompt and lasting changes to the lexical representations of word meanings. Effectively, the long-term balance between the likelihood of different meanings of an ambiguous word is altered as a consequence of the recent experience. However, this interpretation is challenged by the fact that word-meaning priming tends to decay relatively quickly, particularly within the first hour of exposure to the subordinate meaning (Rodd et al., 2016). This is problematic because the immediate alteration account does not obviously contain a mechanism to accommodate a gradual return to the balance of meaning likelihoods that was evident prior to sentence exposure. Furthermore, the rate of decay following exposure is mediated by periods of sleep. This was explored in Gaskell et al. (2019), who measured word-meaning priming 2 and 12 hours after exposure. Some participants slept during the intervening period whilst others spent an equivalent amount of time awake. In both delay intervals, priming was only evident in participants who slept. In a second experiment, Gaskell and colleagues measured priming 24 hours after exposure. One group of participants slept overnight, soon after exposure in the evening, and then spent the remainder of the 24 hours awake (sleep-wake group), whilst the other group of participants had their exposure in the morning, meaning that they had a day awake prior to a night's sleep (wake-sleep group). Importantly, word-meaning priming was only observed in the sleep-wake group. This finding is important, because it implies that sleep promotes the continued support of linguistic memories, particularly if it occurs quite soon after encoding (see Mak et al., 2023; 2024 for similar sleep-maintained priming effects in other aspects of language).

To account for the supporting role of sleep on word-meaning priming, Gaskell et al. (2019) proposed the *episodic context account* (see also Curtis et al., 2022; Mak et al., 2025), which argues that episodic memory is routinely involved in supporting language comprehension (Duff & Brown-Schmidt, 2012; 2017). Specifically, it argues that an experience of language processing, such as reading a sentence, gives rise to a new *context-specific representation* in episodic memory that binds together and extracts the core lexical-semantic information portrayed in the discourse. Alongside long-term semantic knowledge, this representation can support how similar linguistic material is processed in the future, and may bias lexical processing toward the information expressed in the representation. Importantly, the episodic nature of context-specific representations makes them susceptible to forgetting during wake as well as sleep-related memory consolidation, in which sleep actively consolidates and integrates new episodic memories into long-term storage, improving their retention (Born & Wilhelm, 2012; Rasch & Born, 2013). On a systems level, this may occur via a shift in memory representation from an initial episodic memory trace stored in the hippocampus to representation in cortical networks which supports long-term semantic memory (Davis & Gaskell, 2009; Frankland & Bontempi, 2005; McClelland et al., 1995; Winocur & Moscovitch, 2011). The redistribution of a memory is mediated by hippocampal memory reactivation, particularly during non-REM (NREM) sleep, which reinstates and replays cortical neural activity associated with a memory encoded prior to sleep (Denis & Cairney, 2023; Schreiner et al., 2021; Sterpenich et al., 2021), supporting memory transfer to the cortex. As with other episodic memory traces, then, the long-term utility of a context-specific representation in supporting language processing is facilitated by sleep (Gaskell et al., 2019). That is, periods of sleep can stabilise and consolidate the memory for long-term storage, increasing the likelihood that the representation will provide a continued source of support to lexical processing. However, time spent awake can prove detrimental to the memory through memory decay (Hardt et al., 2013) and/or interference from new sensory input.

To date, empirical support for the episodic context account partly stems from studies comparing language exposure priming effects in participants who slept soon after exposure (i.e., a sleep group) compared with those who spent an equivalent amount of time awake (Gaskell et al., 2019; Mak et al., 2023; 2024), and by manipulating sleep onset (Gaskell et al., 2019). However, the extent to which these manipulations can implicate a *causal* role of sleep in word meaning priming is limited. For instance, sleep vs. wake designs must often configure memory tests at different times of day for respective participant groups. As memory performance is influenced by the time of day that a task is performed (Barner et al., 2019), between group differences in language exposure priming effects (Gaskell et al., 2019; Mak et al., 2023; 2024) could be confounded by such circadian factors. To overcome this issue and provide stronger causal evidence for an active role of sleep in consolidating linguistic-related memories, a *targeted memory reactivation (TMR)* design provides an appealing solution of biasing specific memories for sleep-associated consolidation. Generally speaking, TMR involves the covert presentation of a sensory cue, such as an odour or sound, during sleep that was associated with a specific piece of information before sleep. A common finding is superior post-sleep memory performance for information that was cued during sleep via the presentation of the associated sensory cue, compared with uncued information (Carbone & Diekelmann, 2024; Hu et al., 2020). The memory benefits of TMR have been observed across a range of memory phenomena including location memory (Cairney et al., 2014; 2016; Creery et al., 2015; Oyarzún et al., 2017; Rasch et al., 2007; Rudoy et al., 2009; Shimizu et al., 2018), emotional memory (Cairney et al., 2014), associative memory (Cairney et al., 2017; 2018; Joensen et al., 2022) and language learning (Batterink & Paller, 2015; Göldi & Rasch, 2019; Neumann et al., 2020). In the language domain, recall of a German word is superior after being presented with its (learned) Dutch word translation during sleep, as compared to words that were not presented (Schreiner & Rasch, 2015). The reintroduction of a memory cue during sleep is thought to trigger the hippocampus into reactivating the associated memory, therefore facilitating its integration into the cortical networks (Lewis & Bendor, 2019; Schreiner & Rasch, 2017). In support of this, several studies have reported increased theta (~4-8 Hz) and sleep spindle (~11-16 Hz) activity following the onset of a memory cue (Antony et al., 2018; Cairney et al., 2018; Farthouat et al., 2017; Göldi et al., 2019; Groch et al., 2017; Guttesen et al., 2024; Joensen et al., 2022; Laventure et al., 2018; Oyarzún et al., 2017; Silfuentes-Ortega & Peigneux, 2024; Schreiner & Rasch, 2017), oscillatory activity associated with sleep-based memory reprocessing. According to Schreiner and Rasch (2017), theta activity represents the reinstatement of memory-related neural activity while (later) spindle activity is necessary for redistributing the memory to the cortex (though according to Antony et al. (2019), it is spindles that encompass memory reinstatement). Hence, the fact that memory cues elicit enhanced oscillatory power in these frequency bands provides support that the reintroduction of information-associated sensory cues during sleep can bias reactivation of the associated material and improve its long-term retention. Furthermore, since TMR can be manipulated within-subjects, circadian confounds associated with between-subjects sleep vs. wake designs are also overcome. Hence, a worthy extension to the episodic context account literature is to incorporate a TMR manipulation into an experiment measuring the effect of language exposure and sleep on subsequent lexical processing, such as word-meaning priming. If word-meaning priming is enhanced by TMR, then this would provide causal evidence for an active role of sleep in consolidating context-specific memories related to recent linguistic experiences (Gaskell et al., 2019). We explored this proposition in the current paper.

1.1 Present experiment

Our experiment consisted of a three-phased design which is illustrated in Figure 1. Our stimuli consisted of 80 *prime sentences*, each containing a unique homonym, with the global meaning of the sentence referring to the homonym's subordinate meaning. Each prime sentence was also paired with a *cue word*, which consisted of an auditory recording of the homonym.

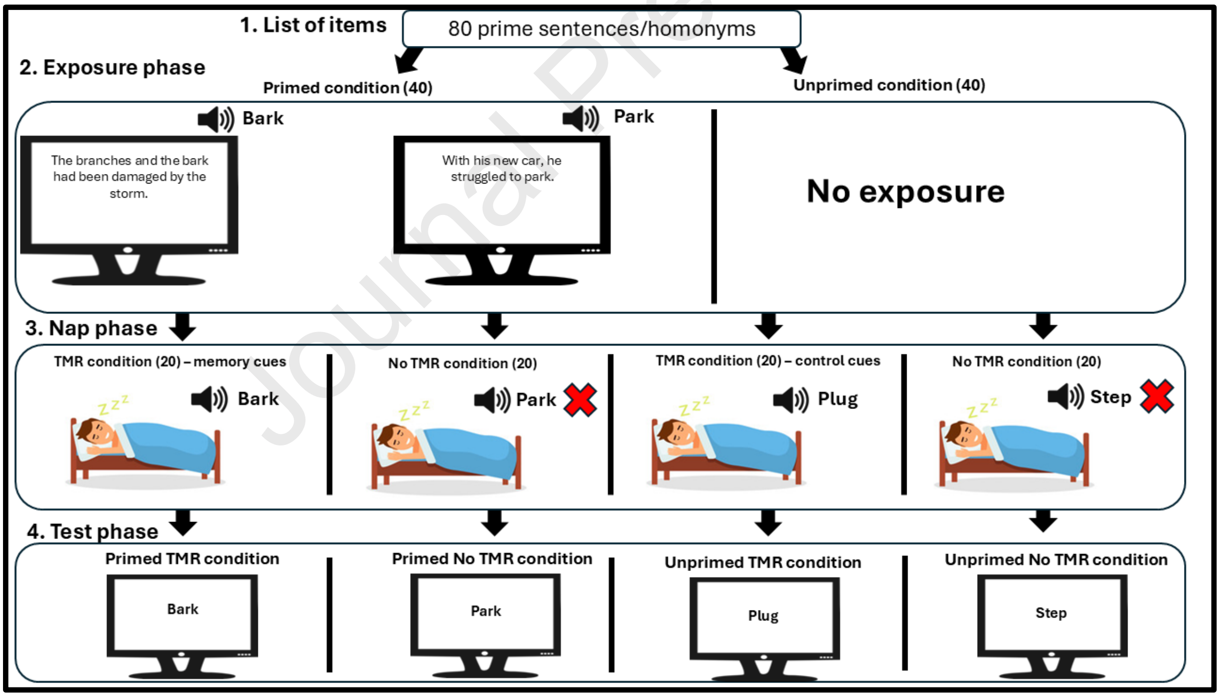

Figure 1: A schematic overview of the phases and experimental conditions in the experiment. The monitors are intended to illustrate the presentation of a homonym or prime sentence in a certain phase and do not illustrate task procedures (see Figure 2 for this information). The assignment of homonyms/prime sentences to each condition was counterbalanced across participants.

In the exposure phase, participants read 40 of the 80 prime sentences with the intention of priming the homomyn's subordinate meaning. The remaining 40 prime sentences were not encountered and therefore made up the unprimed condition. During sentence reading, the associated cue word was also presented, and participants were encouraged to associate this with the sentence. As a subset of the cue words would be replayed to participants during sleep with the aim of reactivating the associated memory (prime sentence), these instructions were intended to result in a relatively strong association between the cue word and prime sentence for the purpose of TMR (Schreiner & Rasch, 2017). To accommodate this, however, it was necessary to alter the tasks of the exposure phase (see *Procedure*) that are typical of prior word-meaning priming studies, where participants read a sentence and then judge whether the meaning of the sentence is semantically related or not to a probe word (Betts et al., 2018; Curtis et al., 2022; Gaskell et al., 2019; Gilbert et al., 2018; 2021; Rodd et al., 2013). Before proceeding with our TMR experiment, we first examined in an online behavioural experiment with 40 participants (no TMR was performed) whether word-meaning priming is elicited over ~20 minutes with these altered encoding procedures. This behavioural experiment can be found on OSF (https://osf.io/frkms/?view_only=d754fa5eb708479c95adffdc243e87ef). Indeed, a significant word-meaning priming effect was revealed (numerically speaking, there was a 0.16 increase in the proportion of subordinate meaning responses in word association following subordinate meaning exposure). We also utilise and present this data in an exploratory analysis (section 3.2.2.3.2) comparing the magnitude of our behavioural effects between the online experiment and TMR experiment.

The second stage of the experiment consisted of the nap phase, where 20 of the cue words from the exposure phase were replayed to the participant during NREM sleep. We refer to these as *memory cues* since they were associated with a memory (prime sentence) before sleep, and served to trigger reprocessing of the associated memory for consolidation (i.e., TMR). The remaining 20 cues from exposure were not replayed (No TMR). Additionally, we also presented 20 auditory cues from the unprimed condition (cues not encountered at exposure). As these cues were not encountered nor associated with a memory before sleep, we refer to these as *control cues*. We examined the EEG-evoked response to memory cues for signs of memory (re)processing by comparing such responses with EEG activity elicited by the lexically-matched words (i.e., the control cues).

The final stage of the experiment consisted of the test phase, which included word-association and fill-in-the-blank tasks. In word association, participants generated lexical associates for all 80 homonyms. We compared the proportion of responses referring to the subordinate meaning between the primed and unprimed conditions as an index of word-meaning priming. More importantly with respect to our research questions, we also analysed the effect of TMR (TMR vs. No TMR) within the primed condition, to test whether the context-specific representations that may support word-meaning priming were consolidated via TMR. As far as we are aware, the effect of TMR has not been examined in a lexical processing task such as word association before. We therefore included the fill-in-the-blank task as an explicit memory test of cued recall, given prior evidence that TMR positively impacts cued recall (Carbone & Diekelmann, 2024; Hu et al., 2020). In this task, participants encountered all 80 prime sentences with the homonym absent. For sentences the participant recognised, they were asked to recall the missing word, or use guesswork based on sentence semantics for sentences they did not recognise. The experiment was pre-registered ahead of data collection (https://osf.io/84m2p/?view_only=e4bf3cad7df64e0891a7e92dacee73ff) and we made the following predictions:

1. In the word-association task, **a)** there will be a greater proportion of responses that are consistent with the subordinate meaning in the primed than the unprimed condition (main effect of priming). **b)** Within the primed condition, we expect that responses will be more consistent with the subordinate meaning for primed contexts that were cued during sleep compared to uncued sentences (a simple effect of TMR within the primed condition).
2. In the fill-in-the-blank task, **a)** the correct homonym will be generated more frequently in the primed than the unprimed condition (main effect of priming). **b)** WIthin the primed conditions, we expect the correct homonym will be generated more frequently for sentences that were cued during sleep compared to uncued sentences (a simple effect of TMR within the primed condition).
3. We expect to observe a significant increase in spindle (i.e., sigma band) power and theta power (Cairney et al. 2018; Guttesen et al., 2024) in response to cues associated with a prime sentence (i.e., memory cues) relative to control cues.

2.1 Methods

2.1.1 Participants

Seventy participants took part in the experiment and received monetary compensation (~£10/hour) or course credits for taking part. Participants were recruited from a university study participation scheme. Five participants were excluded from data analysis for failing to reach NREM sleep, whilst seven participants were excluded for meeting our pre-registered criteria of failing to achieve one full round of TMR. This yielded a final sample of 58 participants who contributed data to the analysis (49 females; *M* *age* = 19.31 years, *SD* age = 1.40 years). As pre-registered, we specified to cease recruitment after a certain date (30th April, 2024)[1] or after achieving a target sample size of 80 participants, which was based on Brysbaert and Steven's (2018) recommendation of attaining ~1600 observations per condition in a within-subjects design. For this experiment, we reached our cut-off date first, meaning we fell short of our target sample size. All participants reported English to be their native language and reported no known history of language, sleep, or attentional-related disorders. To facilitate sleep onset, participants were asked to refrain from consuming alcohol or caffeine 24 hours before the experiment, to wake up two hours earlier than normal on the day of the study, and to consume a relatively substantial meal before they arrived at the laboratory. Participants verified compliance with these instructions at the start of the experiment by ticking a checkbox. We pre-registered that we would exclude participants who answered more than two of the five attention checks in the recall task (exposure phase) incorrectly. However, all participants passed the attention checks. This research was approved by the Research Ethics Committee of the Department of Psychology, XXX.

2.1.2 Materials

Our critical stimuli, taken from Gaskell et al. (2019), consisted of 80 prime sentences. Each prime sentence contained a target homonym (e.g., *bark*) that primed the subordinate meaning of the word (e.g., "*The branches and bark had been damaged by the storm.*"). Prior to the experiment, we performed a norming study on a group of independent participants (*n* = 20) who read all of the prime sentences with the homonym missing and were asked to

---

[1] The cut-off date coincided with the end of our funding that supported this research.

type the missing word. We then modified sentences in which the homonym was correctly produced by > 3 participants. This was done to reduce the predictability of the homonym, which might otherwise act as an additional source of performance (recall) in the fill-in-the-blank task. Each prime sentence was also paired with an audio recording of the target homonym, which we refer to as the *cue word* (*M length of audio* = 607 ms; *SD* = 108 ms). The cue words were produced by a female, native English speaker in a sound-attenuated booth. The items were split into two sets of 40 which were matched on homonym length, sentence length, and length of the cue word recording for the purpose of counterbalancing across priming conditions. The full list of stimuli is available on the Open Science Framework (https://osf.io/frkms/?view_only=d754fa5eb708479c95adffdc243e87ef).

There were also 11 filler prime sentences, taken from Curtis et al. (2022). These sentences did not include a homonymic word (e.g., "*I soaked in the bathtub and listened to music for hours*") and were included as fillers in the exposure phase to reduce the salience of the ambiguity associated with the homonyms. These sentences were also paired with a cue word that consisted of a word contained in the sentence (e.g., "bathtub"). Finally, a further ten prime sentences containing an ambiguous homonym (and associated cue word) were selected to serve as attention checks in the exposure phase (see *Procedure*, section 2.1.4).

2.1.3 Design

There were two within-participant independent variables: 1) Priming (whether the subordinate meaning of a homonym was primed in the exposure phase or not) and 2) TMR (whether a cue word was presented during sleep or not). A fully-crossed factorial design was used, giving rise to four experimental conditions: Primed TMR; Unprimed TMR; Primed No TMR; Unprimed No TMR (see Figure 1). The assignment of stimuli to each condition was counterbalanced across participants.

2.1.4 Procedure

The task procedures are illustrated in Figure 2. All experimental tasks were programmed in MATLAB using *Psychtoolbox* (Kleiner et al., 2007).

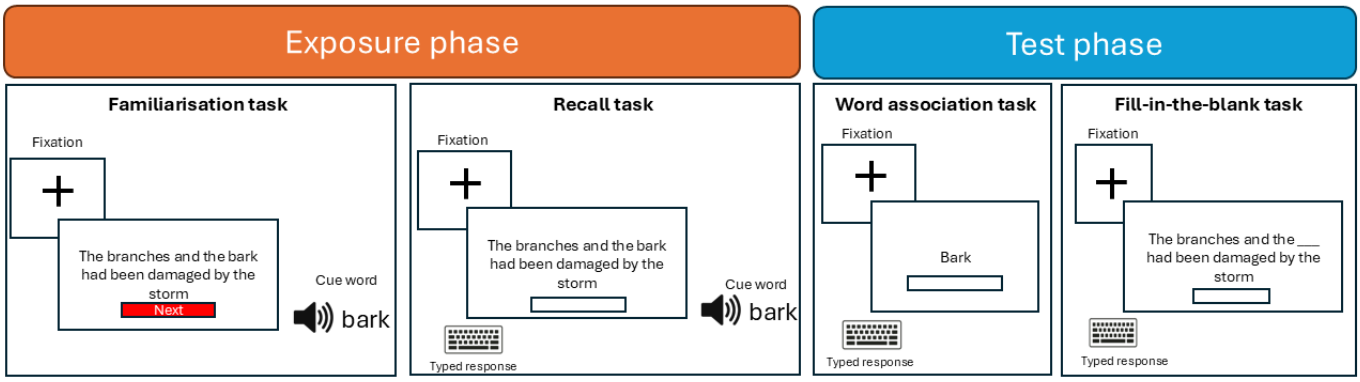

Figure 2: A schematic illustration of the experimental tasks completed in the exposure and test phases. In the familiarisation task, participants generated mental images of 40 sentences and associated each one with the cue word. In the recall task, participants recalled the cue word of each sentence (and heard the word for a second time). In the word association task, participants generated associates of all 80 homonyms. In the fill-in-the-blank task, participants recalled/guessed missing words from sentences. Task instructions and stimulus timings are provided in text.

EEG set up

Participants arrived at the lab at ~12 pm and were initially set up with the EEG electrodes. Sleep EEG was monitored using a Natus Embla NDx recording system with RemLogic 4.0 recording software. Scalp EEG electrodes were positioned according to the 10-20 system at frontal (F3 and F4), central (C3 and C4), parietal (P3 and P4), occipital (O1 and O2), and mastoid (M1 and M2) locations, referenced online to Cz with a ground electrode positioned on the forehead, and a sampling rate of 256 Hz. Bilateral electrooculography electrodes were positioned, as were electromyography electrodes positioned on the mentalis and submentalis (bilaterally). Electrode impedances were kept <5 Ω.

Exposure phase

After EEG setup was completed, participants completed the Stanford Sleepiness Scale (SSS) to provide a subjective measure of sleepiness before sleep (SSS1) before completing the exposure phase, which consisted of two tasks: the familiarisation task and recall task (see Figure 2). On a single trial in the *familiarisation task*, a fixation cross was presented for 500 ms, which was replaced by one of the prime sentences located in the centre of the screen. Participants were asked to read and develop a mental image of the sentence. After they had read and developed a mental image, participants pressed a ‘Next’ button located below the sentence, which, 1000 ms after being pressed, triggered the audio presentation of the associated cue word. Upon hearing the cue word, participants were asked to integrate this word into their mental image, such that they can ‘hear’ the word in the image. The purpose of these instructions was to facilitate the encoding of an association between learned information (i.e., the prime sentence) and the corresponding cue word. Before beginning the task, participants were shown a visual example to facilitate their understanding of the instructions. The prime sentence remained on screen for 3000 ms after the onset of the cue word before disappearing in preparation for the next trial. There were five attention checks throughout the familiarisation task in which two questions followed the cue word. First, participants were asked to type the cue word that they had just heard, and were also asked a comprehension question related to the previous sentence. These questions were designed to ensure that participants had heard and processed the cue word, and had read and processed the prime sentence.

Following the familiarisation task, participants completed a *recall task*. This task was designed to strengthen the association between learned information (i.e., the prime sentence) and the cue word developed in the familiarisation task. Participants were informed that on any given trial, they would be presented with one of the sentences from the familiarisation task and were instructed to type the word that they had heard with each sentence. Participants were encouraged to retrieve their mental images to help with recall. A single trial began with a 500 ms fixation cross, which was replaced by one of the prime sentences located in the centre of the screen. After 1500 ms, a text box appeared below the prime sentence allowing participants to provide a typed response. Participants were asked to press the ‘Return’ key to submit their response, and were not allowed to proceed without giving a response. Feedback was presented immediately after the response was submitted, either in the form of a green tick for a ‘correct’ response or a red cross for an ‘incorrect’ response. The correct, corresponding cue word was simultaneously played regardless of their response, and the sentence remained on-screen 3000 ms following the onset of the audio. On average, participants recalled the cue word on 92% of trials (SD = 26.60%). On attention check trials, participants were asked to type a specific word from the sentence. Within both the familiarisation task and recall task there were 58 trials: 40 critical prime trials, 11 filler non-homonym trials, and 5 filler homonym trials which served as attention checks. Each task was split into two blocks of 28 trials, and participants could take a short break in between blocks. For each participant, the order of trials within tasks was randomised.

Nap phase and TMR

Immediately after the exposure phase at ~1 pm, participants got into bed and were left to sleep for ~2 hours. Polysomnographically-recorded sleep was monitored online by the experimenter in a nearby control room. The experimenter commenced TMR when the participants exhibited ~2 minutes of continuous N3/slow-wave sleep (SWS) which was immediately paused if the participant displayed signs of arousal, waking up or entering a different sleep stage. Cue words were presented through a speaker positioned ~1.5 m above the bed and were separated by a 5 second interval (jittered at 0.2 seconds). A single round of TMR consisted of 40 cue words, 20 of which were presented in the preceding exposure phase and associated with a prime sentence (Primed TMR condition - *memory cues*) whilst the other 20 were not previously encountered (Unprimed TMR condition - *control cues*). A new randomised order of cue words was established for each TMR round and participant. On average, 4.5 TMR rounds were completed (*M total cue words* = 181; *SD* = 124; *M memory cues* = 91; *SD* = 63; *M control cues* = 90; *SD* = 62).

Test phase

At ~3 pm, participants were woken up and the electrodes were removed. After ~20 minutes had passed from waking up in order to alleviate the effect of sleep inertia, participants completed the SSS for a second time (SSS2) before completing the test phase, which consisted of two tasks: word-association and fill-in-the-blank. In word association, participants provided an associate for each of 80 homonyms, 40 of which were primed to the subordinate meaning in the exposure phase. Participants were informed that they would see a single word on-screen and should type the first word that came to mind after reading the word; “tennis-wimbledon” was used as an example. The task began with a single practice trial, followed by the 80 critical trials. Any given trial began with a fixation cross for 500 ms, followed by a target homonym in the centre of the screen. Below the homonym was a text box allowing participants to provide a typed response. Participants were asked to press the ‘Return’ key to submit their response, and were not allowed to proceed without giving a response. There were also five attention checks, in which participants, following their response, were asked to type the on-screen word (homonym) from the previous trial. Participants were allowed to take a short break halfway through the task.

In the fill-in-the-blank task, participants encountered all 80 of the critical homonym prime sentences, 40 of which had been encountered in the exposure phase. Critically, the homonym from the prime sentence was absent (e.g., “The branches and the ___ had been damaged by the storm”). Participants were informed that some of the sentences had been encountered in the exposure phase, and that they should try to recall the correct word for sentences they believe they have encountered. If they didn’t recognise the sentence, they were asked to provide what they thought the missing word might be, based on the meaning of the sentence. A trial began with a fixation cross for 500 ms, followed by one of the prime sentences (with the homonym absent). A response cue appeared below the sentence, allowing the participant to provide a typed response. Participants were asked to press the ‘Return’ key to submit their response, and were not allowed to proceed without giving a response. After submitting their response, participants indicated whether their response was sourced by recalling the word from memory, or whether their response was a guess based on the context of the sentence. We recorded the mechanism of response to allow for an exploratory analysis on the fill-in-the-blank data (see section 3.2.2.3.1 for details). For each participant, the order of trials within tasks was randomised.

EEG Preprocessing

EEG preprocessing and statistical analysis (below) was carried out in *FieldTrip* (Oostenveld et al., 2011). Offline, the EEG data were re-referenced to the average of the mastoid electrodes, high-pass filtered at 0.5 Hz, notch-filtered at 49-51 Hz and segmented into trials (-1 to 3.5 seconds around the cue word). Artefact rejection utilised a combined approach of automated and manual methods. Automated artefact rejection was carried out using FieldTrip’s *ft_artifact_zvalue* function in which the EEG signal is filtered, z-transformed, and thresholded to detect artefacts. We filtered the data at 15-32 Hz which prior research has shown to comprise sleep-associated muscle artefacts (Brunner et al., 1996; *M trials automatically excluded* = 3.53; *SD* = 3.36). We then visually inspected the remaining data and performed manual artefact rejection (*M* trials manually excluded = 8.12; *SD* = 8.31). Finally, we excluded EEG trials which fell outside of N2 or N3/SWS (*M trials outside N2 orN3/SWS* = 5.40; *SD* = 7.21) using the automated sleep scoring tool *SomnoBot* (https://somnobot.fh-aachen.de/; Guillot & Thorey, 2021). We restricted the EEG analysis to N2 and N3/SWS epochs as the effect of TMR is strongest in these sleep stages (Hu et al., 2020). In all, 10.72% of trials were rejected (*SD* = 8.39%).

3.2 Results

3.2.1 Sleep data and alertness

Table 1 presents mean SSS scores and time spent in different sleep stages.

Table 1: Mean SSS scores and time spent in minutes in different sleep stages (SD) as determined by *SomnoBot*[2]. Higher SSS scores indicate greater perceived sleepiness.

SSS scores		Sleep stages				
SSS1	SSS2	N1	N2	N3/SWS	REM	Total sleep
2.86 (0.93)	2.48 (1.01)	14.85 min (8.23)	52.98 min (17.16)	6.11 min (9.23)	12.47 min (12.82)	86.39 min (22.64)

Concerning SSS scores, a paired-samples t-test revealed a statistically significant difference between SSS1 and SSS2 (*t* (57) = 2.41, *p* = .019, *d* = 0.32). The lower scores in SSS2 suggests participants were subjectively more alert while completing the test phase than the exposure phase, following a period of sleep.

3.2.2 Behavioural results

As pre-registered, data from the word-association and fill-in-the-blank tasks were separately analysed using generalised linear mixed-effects modelling (GLMM) in R using the *lme4* package (Bates et al., 2015). The models included fixed-effects of priming and TMR (both sum-coded), and their interaction. To construct our random-effects structure, we used the *Buildmer* package (Voeten, 2022). This allowed us to identify the most parsimonious random

---

[2] Time spent in N3/SWS is lower than in some TMR studies using a nap design (e.g., Cairney et al., 2014). When comparing a subset of 10 manually-scored data sets against *SomnoBot* scores, we found that *SomnoBot* appeared to overestimate N2 at the expense of N3/SWS (e.g., 42% of manually-scored SWS epochs were labelled as N2 by SomnoBot).However, because our EEG analysis focused on TMR trials occurring in either N2 or N3/SWS, we did not view this as a concern. Whatsmore, of the 10 manually-scored data sets, there was substantial agreement between the automatic and manual scores (Cohen's kappa = 0.76), which is very similar to the level of agreement achieved between human raters (Lee et al., 2022).

effects structure that was capable of converging, given our data (by setting the direction argument to “order”)[3]. For both models, we report on the main effect of priming, to test the predictions that the primed condition will respectively generate more responses that are consistent with the (primed) subordinate meaning in word association and a superior ability to generate the absent homonym in fill-in-the-blank, than the unprimed condition. To test the predictions that TMR will respectively generate more subordinate meaning responses (word-association task) and superior cued recall (fill-in-the-blank task), *within the primed condition* (a simple effect of TMR), we used the emmeans package (Lenth et al., 2021) to compare the Primed TMR and Primed No TMR conditions, separately for each task. We also report on the interactions between Priming and TMR, since we acknowledged in our pre-registration that our hypothesised effect(s) could also be qualified by a significant interaction.

3.2.2.2 Word association

All participants passed the attention check trials in the word-association task. Obvious spelling errors were corrected before the word-association responses were coded in two stages. First, we automatically coded responses using Gilbert and Rodd’s (2022) automated coding scheme, which assigns a meaning code to a given response by comparing the interpretation of the same response within a coded data set of nine prior studies. In total, 81.4% of all responses were automatically coded. This method was unable to assign a meaning code if a) a given response had not been produced in the coded data set or b) a given homonym was not contained in the coded data set. Uncoded responses were manually coded by an experimenter blind to experimental conditions (LB), and any uncertainties were discussed with a second experimenter (GG). If, after discussion, uncertainty remained as to the interpretation of a response, it received a score of ‘Unclear’ and was discarded from the analysis. In all, responses were classified into the following categories:

1. Dominant meaning (2765, 60%)
2. Subordinate (primed) meaning (1484, 32%)
3. Other meaning (203, 4%)

---

[3] This model building procedure was used for all GLMMs analysing behavioural data in this paper.

4. Unclear - ambiguous with respect to two or more possible meanings or not clear what the response refers to (188, 4%).
5. Spelling error (0, 0%)

Our analysis focused on the first two categories, with the dependent variable reflecting whether a given response was consistent with the primed, subordinate meaning (1) or the dominant meaning (0). The left-hand side of Table 2 summarises the model output, and Figure 3 presents condition means.

Table 2: Model summaries for the word-association and fill-in-the-blank tasks. Asterisks denote statistically significant terms at alpha <.05.

	Word association				***Fill-in-the-blank***			
	B	***SE***	***z***	***p***	***B***	***SE***	***z***	***p***
Intercept	-0.82	0.14	-5.70	<.001*	-0.55	0.19	-2.88	.004*
Priming	0.61	0.07	9.06	<.001*	2.42	0.11	21.05	<.001 *
TMR	0.01	0.04	0.33	.740	-0.06	0.05	-1.03	.305
Priming x TMR	-0.01	0.04	-0.26	.793	0.05	0.07	0.83	.407

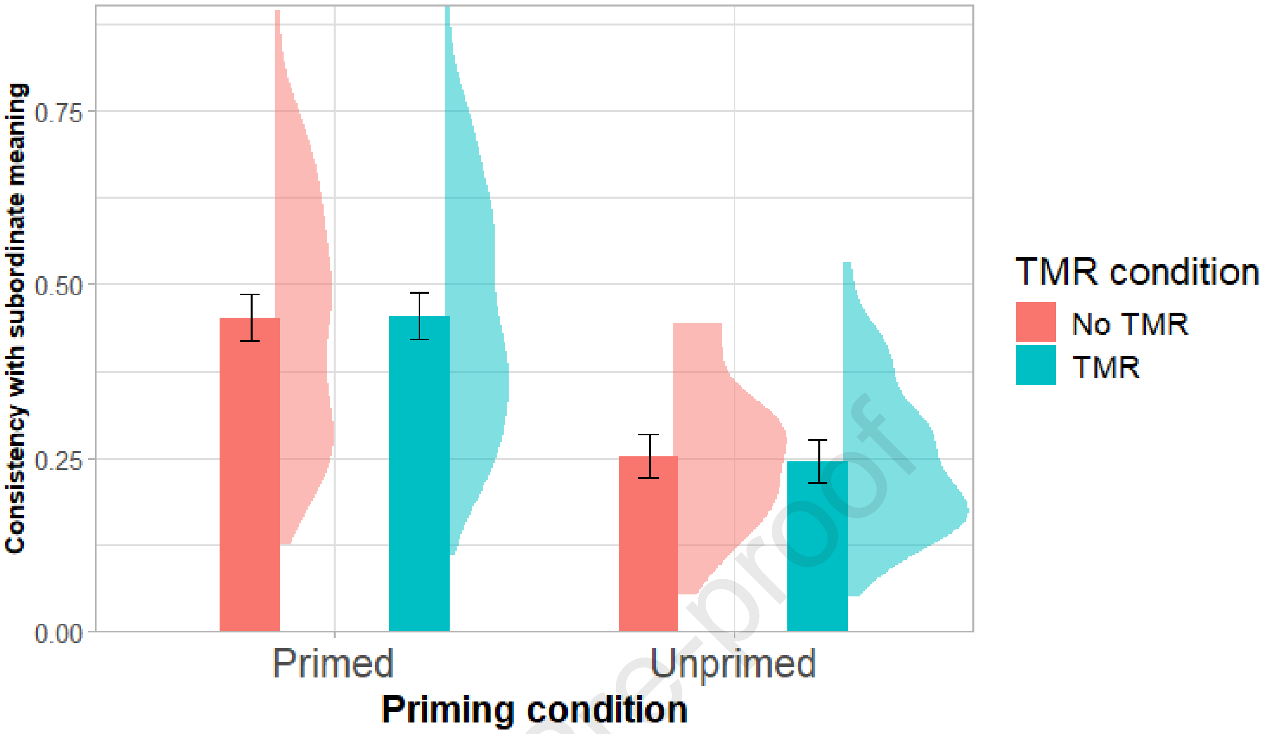

Figure 3: Proportion of responses in the word-association task that were consistent with the subordinate (primed) meaning. Error bars represent 95% within-subjects confidence intervals. The density function represents the distribution of participant mean scores, calculated using a Gaussian kernel.

The main effect of priming was statistically significant (*p* <.001) revealing an overall word-meaning priming effect. That is, the primed condition elicited a greater proportion of responses that were consistent with the subordinate meaning (*M proportion of subordinate meaning responses* = 0.45, *SD* = 0.50) than the unprimed condition (*M* = 0.25, *SD* = 0.43) in the word-association task. Within the primed conditions, however, there was no significant simple effect of TMR (*p* = .959), with equivalent performance between the Primed TMR (*M* = 0.45, *SD* = 0.50) and Primed No TMR (*M* = 0.45, *SD* = 0.50) conditions. Further, the interaction between priming and TMR was non-significant (*p* = .793). Thus, we found no significant TMR effect on word-meaning priming.

3.2.2.2 Fill-in-the-blank

The dependent variable reflected whether a given response was the correct homonym (1) or not (0). Obvious spelling errors were corrected. The right-hand side of Table 2 summarises the model, and Figure 4 presents condition means.

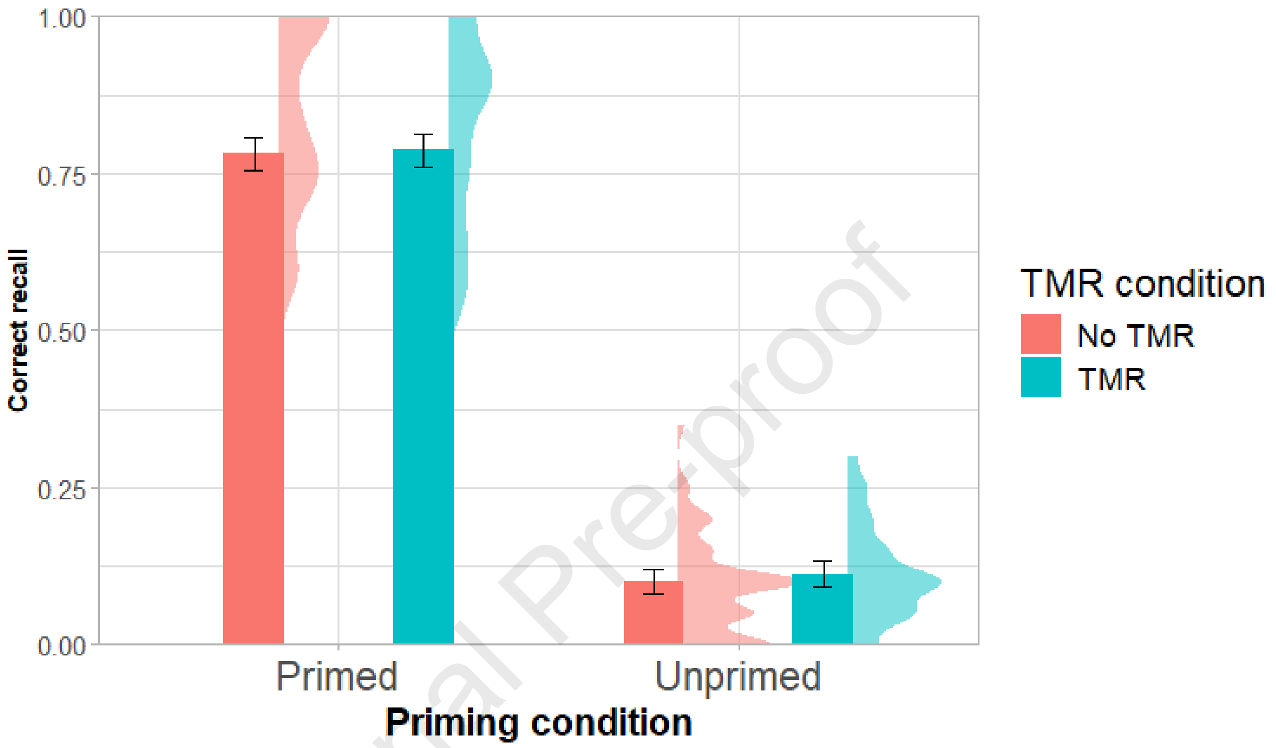

Figure 4: Proportion of correct recall in the fill-in-the-blank task. Error bars represent 95% within-subjects confidence intervals. The density function represents the distribution of participant mean scores, calculated using a Gaussian kernel.

The main effect of priming was statistically significant (*p* <.001), with a superior ability to generate the correct, absent homonym in the primed (*M proportion of correct recall* = 0.78, *SD* = 0.41) than the unprimed condition (*M* = 0.11, *SD* = 0.31). Within the primed conditions, there was no significant simple effect of TMR (*p* = .980), with comparable performance between the Primed TMR (*M* = 0.79, *SD* = 0.41) and Primed No TMR (*M* = 0.78, *SD* = 0.41) conditions. Further, the interaction between priming and TMR was non-significant (*p* = .407).

3.2.2.3 Exploratory analyses

3.2.2.3.1 Fill-in-the-blank exploratory analysis: Mechanism of response

We performed a pre-registered exploratory analysis in which we restricted the data to responses that were indicated to have been sourced from recalling the missing word from memory. Our rationale for this analysis was that TMR could have a stronger effect in instances where memory recall was the primary source of behaviour. As with the main analysis however, there was again a main effect of priming ($B = 2.16$, $SE = 0.30$, $z = 7.30$, $p < .001$) and no significant simple effect of TMR within the primed condition ($p = .785$; *M Primed TMR* = 0.91; *SD* = 0.29; *M Primed No TMR* = 0.91; *SD* = 0.28). The interaction between priming and TMR was also non-significant ($B = 0.16$, $SE = 0.25$, $z = 0.64$, $p = .525$). Thus, across the two analyses reported here, we found no evidence of a TMR effect in the fill-in-the-blank task.

3.2.2.3.2 Analysis across experiments

As a reminder, before proceeding with our TMR experiment, we first conducted an online behavioural experiment with 40 participants (https://osf.io/frkms/?view_only=d754fa5eb708479c95adffdc243e87ef) to determine whether word-meaning priming is elicited with encoding instructions that are necessary for a TMR manipulation. Word-meaning priming was indeed elicited, and fill-in-the-blank performance was also superior in the primed condition.

Whilst the encoding and test phase procedures were identical across experiments (aside the online vs in-lab difference), the nature of the intervening period differed. The online experiment had a ~20 minute period during which participants watched a video of "Shaun the Sheep" (chosen for its minimal linguistic content), whilst in the TMR experiment the two phases were separated by a ~2 hour period of sleep (which elicited roughly a 2.5 hour delay from the final exposure of a homonym in the test phase and its presentation in the test phase). This led us to perform a (non pre-registered) exploratory analysis to determine if the magnitude of our behavioural effects differed across experiments, potentially due to differences in time and/or cognitive states related to the intervening periods. Table 3 presents condition means for the two behavioural tasks across experiments. For the TMR experiment, we have averaged across TMR conditions, and did not include this variable in our statistical models since this variable was absent in the online experiment.

Table 3: Condition means for word association (proportion of subordinate meaning responses) and fill-in-the-blank (proportion of correct recall) across experiments. Values in parentheses represent standard deviations.

	Word association		Fill-in-the-blank	
	Online experiment	TMR experiment	Online experiment	TMR experiment
Primed	0.40 (0.49)	0.45 (0.50)	0.69 (0.46)	0.78 (0.41)
Unprimed	0.24 (0.43)	0.25 (0.43)	0.10 (0.30)	0.11 (0.31)

For these analyses, we built GLMMs that included the within-subjects factor “priming” and the between-subjects factor “experiment”. Concerning the word-association data, there was a main effect of priming (*B* = 0.56, *SE* = 0.04, *z* = 13.87, *p* < .001), revealing an overall word-meaning priming effect (*M Primed proportion of subordinate meaning responses* = 0.43; *SD* = 0.50; *M Unprimed* = 0.24; *SD* = 0.43). This was further qualified by a significant interaction between priming and experiment (*B* = -0.07, *SE* = 0.03, *z* = -2.33, *p* = .020), showing that the word-meaning priming effect was stronger in the TMR experiment (0.20 increase in the proportion of subordinate meaning responses in the primed (vs. unprimed) condition; *Odds Ratio* = 3.57, *95% CI* [2.97 - 4.28]) than the online experiment (0.16 increase; *OR* = 2.68, *95% CI* [2.16 - 3.32]).

Concerning the fill-in-the-blank task, there was a main effect of priming (*B* = 2.18, *SE* = 0.08, *z* = 25.96, *p* < .001), indicating that the ability to generate the correct, absent homonym was superior in the primed (*M proportion of correct recall* = 0.75; *SD* = 0.44) than the unprimed (*M* = 0.10; *SD* = 0.31) condition. This was again further qualified by a significant interaction between priming and experiment (*B* = -0.24, *SE* = 0.08, *z* = -3.05, *p* = .002), showing that recall was superior in the TMR experiment (0.67 increase in the correct homonym being provided in the primed, relative to the unprimed, condition; *OR* = 125.97, *95% CI* [81.40 - 194.94]) than the online experiment (0.59 increase; *OR* = 48.72, *95% CI* [30.73 - 77.27].

These findings imply that the magnitude of our behavioural effects were significantly larger in Experiment 2 compared to Experiment 1 despite the longer lag between exposure and test. We return to this result in the Discussion.

3.2.3 EEG results

The EEG data from one participant was not included in the analysis due a technical issue which prevented the onset of the cue word from being tagged in the EEG recording. Hence, the EEG analyses, including the analyses exploring the relationship between cue-evoked EEG response and behaviour (see below), were carried out in 57 participants.

We performed a time-frequency analysis over the EEG data to compare cue-evoked EEG responses between memory and control cues. First, we established time-frequency representations (TFRs) for each participant and condition for frequencies ranging from 4-30 Hz, using a 5-cycle Hanning taper in 0.5 Hz and 5 ms steps with an adaptive window-length. TFRs were converted to a % power change relative to a -300 ms to -100 ms pre-cue baseline window, with the -100 ms offset intended to prevent spectral leakage (Guttesen et al., 2024). The TFRs for memory cues and control cues are presented in Figure 5A and 5B, respectively.

The TFRs for memory and control cues were then compared, at frequencies 4-30 Hz and a time window of 0 ms – 2500 ms, using a non-parametric cluster-based permutation analysis to control for multiple comparisons (Maris & Oostenveld, 2007) with 1000 permutations. We adopted and pre-registered the frequency range 4-30 Hz based on prior TMR work that observed TMR-induced oscillatory activity in this range (Cairney et al., 2018; Guttesen et al., 2024). Clusters were defined as channel x frequency x time samples (channel neighbours were defined according to the 10-20 system). As a first step, experimental conditions (memory vs. control cue) were compared at each sample via a dependent-sample t-test, yielding a t-value. Samples at $\alpha < .05$ were retained and clustered (defined above), creating a combined cluster-level (test) statistic. Next, on each permutation, this process was repeated with condition labels randomly shuffled, creating a distribution of clusters under the null hypothesis. We then calculated a Monte-Carlo significance probability by calculating the proportion of permutation statistics that were larger than the observed test statistic(s), with our critical alpha-level set at .05 to indicate statistical significance.

The permutation analysis revealed one significant cluster, identifying a significant increase in power across the spindle and beta bands (~11-22 Hz) in response to memory > control cues, from approximately 1.2 - 2.2 seconds post-cue onset (see Figure 5C) over the right hemisphere (F4, C4, P4, O2 - see Figure 5D).

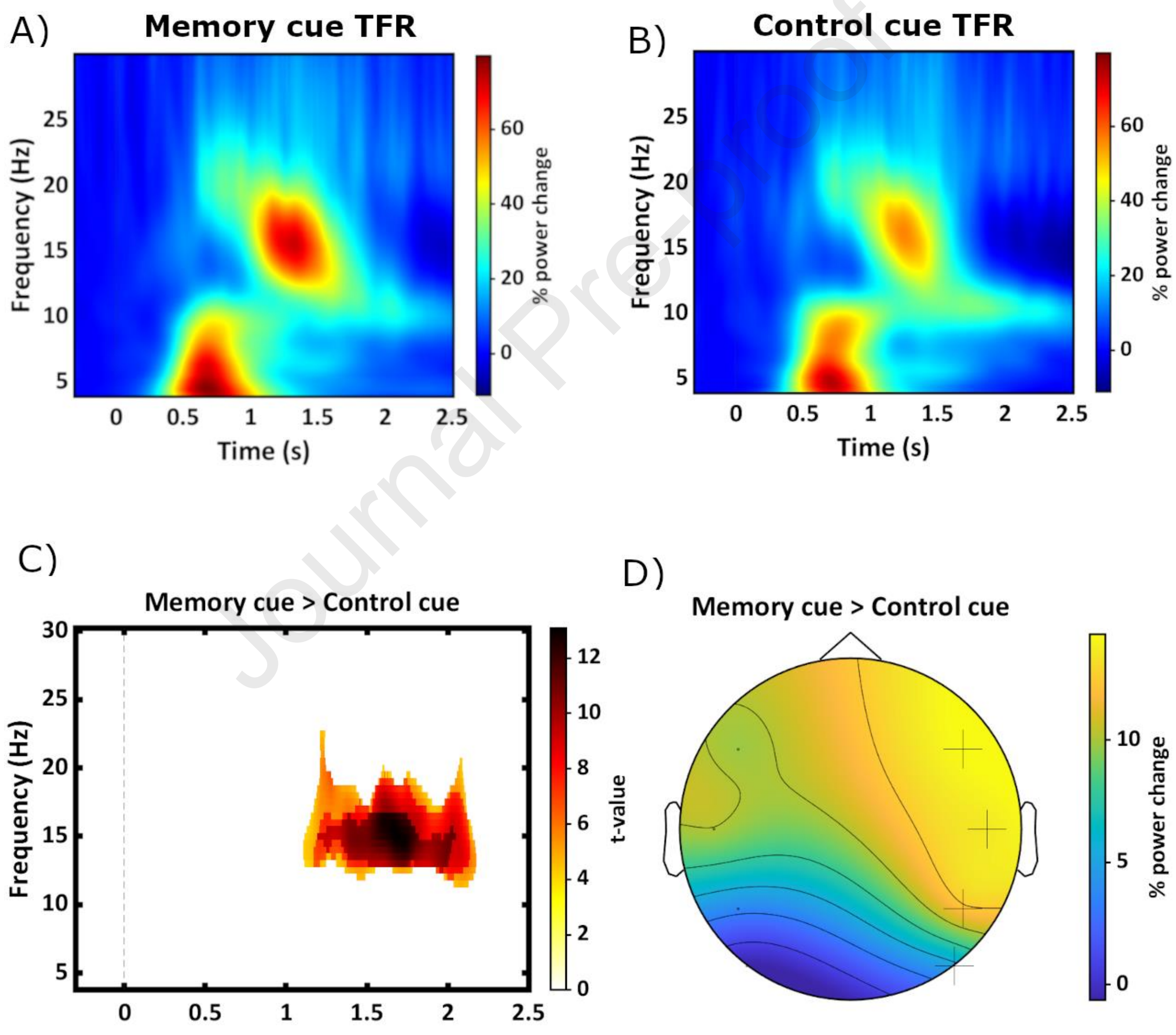

Figure 5: EEG activity: A) Memory cue and B) control cue grand-average time frequency representations (averaged to the right hemisphere electrodes); C) T-value map of the significant

cluster (memory cue > control cue) in time-frequency space; D) The topography of memory cue > control cue, averaged to the time window and frequency range of the significant cluster. Electrodes that contributed to the significant cluster are marked with a cross.

3.2.4 Relationship between behaviour and EEG activity

In our pre-registration, we stated that we would perform a participant-wise correlation exploring the relationship between behavioural performance (i.e., word-meaning priming and cued recall) and cue-evoked EEG activity, should we observe significant differences in EEG activity between memory and control cues as determined by the permutation analysis. Thus, we proceeded with this correlational analysis given the results of the permutation analysis. This analysis was motivated by findings in the TMR literature that the magnitude of cue-evoked EEG activity often correlates with the benefit of TMR on behaviour (Antony et al., 2018; Groch et al., 2017).

We quantified behavioural performance in two ways, separately for each task and for each participant. First, we calculated a “TMR simple effect” score which subtracted the mean in the Primed No TMR condition from the mean in the Primed TMR condition. This metric aimed to mirror our hypothesised effect of observing a significant simple effect of TMR in the primed condition. In an attempt to quantify the benefit of TMR more directly by taking into consideration performance in the unprimed conditions, we also calculated a “TMR benefit” score as follows: (Primed TMR - Unprimed TMR) - (Primed No TMR - Unprimed No TMR). Here, larger values indicate that the effect of priming was greater following TMR than No TMR.

Concerning EEG activity, for each participant, we calculated a single value of average EEG power, reflecting memory cue > control cue difference averaged to the temporal, spatial, and spectral windows of significant differences between memory and control cues as determined by the permutation analysis.

Four correlations were performed, correlating TMR simple effect and TMR benefit scores against EEG power, separately for each task. Concerning the word-association task, there was no significant association between TMR simple effect and average EEG power (*r*(55) = -.12, *p* = .365; Figure 6A) or between TMR benefit scores and average EEG power *r*(55) = .05, *p* = .720; Figure 6B). Similarly, for the fill-in-the-blank task, there was no significant association between TMR simple effect scores and average EEG power (*r*(55) = -

.05, *p* = .730; Figure 6C) or between TMR benefit scores and average EEG power (*r*(55) = -.12, *p* = .388; Figure 6D).

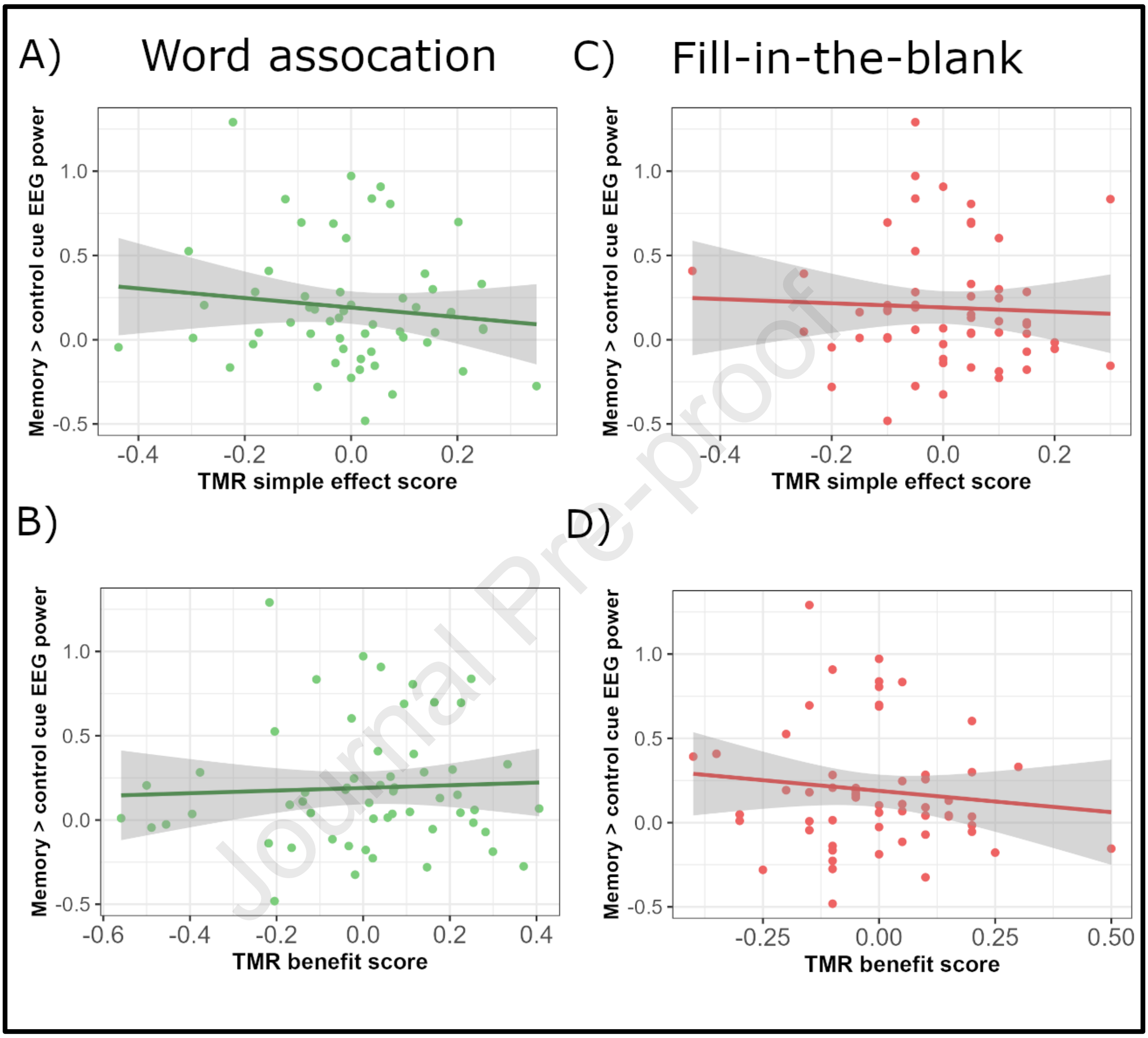

Figure 6: Correlation between average EEG power and TMR simple effect and TMR benefit scores in the word-association (A, B) and fill-in-the-blank (C, D) task.

4. Discussion

According to the episodic context account (Gaskell et al., 2019), word-meaning priming, and language processing more generally, is supported by the formation of a context-specific representation in episodic memory which biases language processing toward previously exposed contextual information. The continued support of these representations over the long-term is supported by periods of sleep. Consistent with an active role of sleep on memory (Born & Wilhelm, 2012; Rasch & Born, 2013), sleep-based memory reactivation can facilitate the consolidation and integration of a newly-formed episodic representation into long-term storage in cortical networks. Supporting evidence for this hypothesis derives from studies comparing behaviour between sleep and wakefulness conditions. In Gaskell et al. (2019), word-meaning priming over 12 hours was only evident after a period of overnight sleep relative to a day of wakefulness (see also Mak et al., 2023; 2024). Over 24 hours, priming was maintained (relative to an immediate test) in participants who slept overnight soon after language exposure but declined in those who experienced daytime wakefulness in between exposure and overnight sleep (Gaskell et al., 2019). These results are not easily explained in terms of sleep playing a passive role in memory by providing a temporary barrier from subsequent linguistic input (Ellenbogen et al., 2006), since both groups would have experienced roughly equal amounts of sensory interference following language exposure. However, these results still fall short of offering evidence for a *causal* role of sleep in word meaning priming. The goal of this research was to probe the causal role of sleep more directly by measuring word-meaning priming in a TMR paradigm. If sleep is actively involved in consolidating context-specific representations related to recent linguistic encounters (Gaskell et al., 2019), then reintroducing associated sensory cues of these memories during sleep may theoretically bias them for consolidation, maintaining the long-term utility of the memory in supporting the future processing of similar material.

In our experiment, participants were first exposed to the subordinate meanings of several homonyms through sentences and associated each sentence with a cue word. Participants then took a ~2 hour long nap, during which a subset of the cue words from exposure were replayed with the aim of triggering reprocessing of the associated memory (sentence) during sleep. In the test phase, participants completed word-association and fill-in-the-blank tasks to index word-meaning priming and cued recall, respectively. Consistent with hypothesis 1a, we found an overall word-meaning priming effect, reflecting a greater proportion of word association responses being consistent with the subordinate meaning in the primed compared to the unprimed condition. Importantly, however, TMR did not enhance the proportion of subordinate meaning responses in the primed condition, inconsistent with hypothesis 1b. Similarly, although the correct homonym was provided more frequently in the primed than the unprimed condition in the fill-in-the-blank task (consistent with hypothesis 2a), TMR did not enhance cued recall of the homonym in the primed condition, inconsistent with hypothesis 2b. Our EEG results suggest that some aspect of TMR-induced memory reprocessing took place in the nap phase. Specifically, we identified a significant cluster of TMR-induced spindle/beta band activity in response to memory relative to control cues, with spindle activity implicated in sleep-based memory reprocessing (Antony et al., 2019; Schreiner & Rasch, 2017). However, on the participant level, the relative increase in sleep spindle activity (to memory cues) did not correlate with TMR-enhanced behaviour.

In the following text, we address the implications of our findings in more detail. We begin by discussing the EEG results and what they can inform us about memory reprocessing in our experiment. We then discuss the behavioural findings, with a particular focus on the null TMR effects.

4.1 Memory cues induced sleep-based memory reprocessing

The increased sleep spindle/beta band power response to memory cues, relative to control cues, provides neurocognitive evidence for cue-induced memory reprocessing during sleep. Sleep spindles are heavily implicated in sleep-based memory consolidation (Born & Wilhelm, 2012; Cairney et al., 2018) and are argued to play a particularly important role in reinstating and reprocessing memory-related neural activity for consolidation (Antony et al., 2019). Consistent with this, several TMR studies, including our results, have observed an increased spindle response to memory cues (Antony et al., 2018; Cairney et al., 2018; Farthouat et al., 2017; Göldi et al., 2019; Groch et al., 2017; Guttesen et al., 2024; Laventure et al., 2018; Oyarzún et al., 2017). Interestingly, this appears to be the first TMR study to report a significant cue-induced spindle response specifically with semantically-ambiguous (homonym) cue words, supporting the generality of verbal cues in inducing spindle activity/sleep-based memory reprocessing (Guttesen et al., 2024). Although we had no a-priori predictions regarding the latency of the TMR-induced spindle response, it is worth comparing the timing of our spindle response to prior work. For example, Guttesen et al. (2024), who also used a memory vs. control cue contrast with verbal (word) stimuli, reported an earlier spindle response beginning around 0.8 seconds after cue word onset, whilst in Cairney et al. (2018), the spindle response began later at 1.7 seconds. Our effect, beginning ~1.2 seconds after cue word onset, therefore lies somewhat in the middle. The earlier response in Guttesen and colleagues could be explained by the use of a single control cue, which was repeated within a single TMR round to match the number of memory cue presentations. Potentially, this could have attenuated neural activity associated with the control cue, revealing earlier differences in spindle activity with memory cues. The later response in Cairney and colleagues is less clear. One speculative possibility is that cue words were roughly 100 ms longer in Cairney et al. compared to our cue words (*M length of audio* = 704 ms vs. 607 ms), which could have delayed both lower (i.e., sensory) and higher (i.e., memory) order processing of the cues.

We also predicted an increased theta power response to memory cues (Groch et al., 2017; Guttesen et al., 2024; Joensen et al., 2022; Laventure et al., 2018; Schreiner & Rasch, 2015; Sifuentes-Ortega & Peigneux, 2024). Schreiner and Rasch (2017) propose that cue-induced theta activity is indicative of successful memory reinstatement, while a subsequent spindle response is necessary for its redistribution to the cortex. Based on Figure 5, theta activity was quite comparable between memory and control cues in our experiment, with the permutation analysis revealing no significant differences within this spectral range. One explanation for the lack of differences in theta activity may concern our use of lexically matched control cues, which consisted of 20 unprimed homonyms. In the language domain, theta activity is related to lexical-semantic processing, with increased theta responses to words relative to pseudowords (Krause et al., 2006; Marinkovic et al., 2012; Schmidg et al., 2024) and open-class relative to closed-class words (Bastiaansen et al., 2005). Thus, semantically-rich words are associated with enhanced theta activity. During sleep, although mental cognition and awareness is generally reduced relative to wakefulness, the brain remains responsive to lexical information, such as detecting semantic incongruencies (Ibáñez et al., 2006). The similar theta response to memory and control cues may, therefore, reflect similar lexical-semantic processes, since both sets of cues consisted of homonyms, which could have obscured a possible memory-related theta response (Schreiner & Rasch, 2017). This could partly explain the similar dissociation in oscillatory activity in Cairney et al. (2018) —enhanced spindle activity following memory cues but no enhancement in theta activity relative to control cues–given that memory and control cues were adjectives in that study. Part of Schreiner and Rasch's (2017) rationale behind the involvement of theta in sleep-based memory processing stems from TMR studies reporting enhanced theta activity (during sleep) for cued information that was gained over sleep (not remembered before sleep but successfully remembered after sleep) compared to cued information that was lost over sleep (Groch et al., 2017; Schreiner & Rasch, 2015; see also Göldi et al., 2019 for a similar comparison). Hence, although contrasting verbal memory and control cues appears sufficient at revealing memory-related spindle activity (Cairney et al., 2018; Guttesen et al., 2024), probing and revealing memory-related theta activity may require a more fine-grained contrast that considers item-level memory before/after sleep to circumvent language-related activity. Intriguingly, this may be less relevant when using non-verbal cues since they carry less lexical-semantic information, and indeed memory vs. control cue contrasts have revealed differences in theta activity with environmental sounds (Sifuentes-Ortega & Peigneux, 2024; Oyarzún et al., 2017) and odours (Laventure et al., 2018). An interesting avenue for future TMR research is to identify optimal conditions for delineating between memory and language-related, cue-evoked theta activity, particularly with verbal cues.

4.2 No evidence of a TMR effect on behaviour

Despite neurocognitive evidence for memory reprocessing during sleep in response to memory cues, there was no evidence for the predicted TMR effect on behaviour. Reassuringly, the lack of a TMR effect cannot be attributed to a failure of eliciting a behavioural effect in the first place, since there was an overall word-meaning priming effect (and recall performance was superior in the primed condition). The results from our cross-experiment exploratory analysis also offer a speculative suggestion that the mechanism behind word-meaning priming was enhanced by sleep, since word-meaning priming and recall were both significantly larger in the TMR experiment over ~2.5 hours compared to our initial online behavioural experiment that measured behaviour over ~20 minutes of wakefulness. We should acknowledge methodological differences across experiments that could have contributed to differences in results, particularly the fact the TMR experiment was conducted in the lab. These participants may have been more engaged than participants in the online experiment and less susceptible to distractions by virtue of the more controlled setting. That said, online experiments can produce comparable data to lab-based experiments (Germine et al., 2012; McConnell et al., 2025), although researchers should consider the use of appropriate attention (Rodd, 2024) which we employed. However, these findings could also implicate a theoretical explanation. Generally speaking, word-meaning priming decays over time, particularly within the first few hours of language exposure (Rodd et al., 2012; 2016). However, if we treat the online experiment as a rough baseline of word-meaning priming estimated soon after exposure, then observing *numerically stronger* priming ~2.5 hours later after a period of sleep (TMR experiment) is consistent with trends seen in the literature. This was demonstrated in Gaskell et al. (2019) where priming increased numerically (relative to an immediate test) over 2 and 12 hours of sleep. In Mak et al. (2023), priming involving non-homonyms and word-class ambiguous words also increased over 12 hours of sleep (at least when the behavioural measure required internally-generated responses such as word association). Conversely, word-meaning priming over 2 and 12 hours of wakefulness tends to decay (Gaskell et al., 2019; Rodd et al., 2012; 2016; see also Mak et al., 2023). Hence, the increased strength in priming seen over sleep reverses the expected reduction in priming observed over wake. Similar to past work, we speculate that these exploratory results may be partly driven by a sleep-enhanced memory effect, perhaps acting upon context-specific representations in episodic memory (Gaskell et al., 2019). This finding is difficult to explain by an immediate alteration account of word-meaning access, since alterations to lexical representations following language exposure are assumed to be prompt and therefore insusceptible to further changes over sleep (Gilbert et al., 2018; Rodd et al., 2016).

Nonetheless, episodic memories developed in the exposure phase did not experience a consolidating effect of TMR. To unpack this null result, it is first worthy to consider whether we had sufficient statistical power to detect a TMR effect. Following Brysbaert and Stevens' (2018) sample size recommendation for within-subjects designs, we established that 80 participants would provide sufficient power for detecting a simple effect of TMR within the primed condition (our hypothesised effect in both behavioural tasks). Our final sample of 58 participants, despite being relatively large for a TMR study (Hu et al., 2020), therefore fell short of this recommended estimate[4], meaning the null TMR effect could reflect a Type-II error (see Wick & Rasch, 2023, for a similar discussion of statistical

---

[4] As a reminder, we pre-registered that we would stop recruitment if we attained 80 usable data sets or once we reached a certain cut-off date, and it was the latter which was reached first.

power in TMR). The issue of statistical power is made particularly relevant in TMR experiments given that TMR effects tend to be relatively small (Hu et al., 2020).

We explored these null results further by calculating Bayes Factors to inform us of the amount of evidence in favour of the null hypothesis, reflecting no TMR effect on behaviour. Specifically, we calculated Bayes Factors (BF_{10}) according to the main effect of TMR after restricting the data to the primed condition[5]. We obtained BF_{10} of 0.19 and 0.20 for the main effect of TMR in word-association and fill-in-the-blank, respectively, indicating moderate evidence in favour of the null hypothesis (Lee & Wagenmakers, 2013).

What can these results inform us about the role of sleep and memory on language processing? If we assume a (true) null TMR effect on word-meaning priming, then the role of sleep in consolidating linguistic-related memories may be more passive than the episodic context account predicts (Gaskell et al., 2019). That is, rather than actively consolidating new episodic memories through sleep-based memory reactivation, sleep could instead enhance memory by providing a period of sufficiently minimal sensory input that could otherwise interfere with new memories (Ellenbogen et al., 2006). For instance, according to Yonelinas et al.'s (2019) contextual-binding account, episodic memories remain dependent on hippocampal processing rather than undergo cortical consolidation. After initial encoding, memories are subsequently prone to interference from other memories that share similar contextual, temporal, or cognitive details. Importantly, sleep reduces such interference by limiting subsequent memory encoding. Can a passive account of sleep explain the present results, and prior findings of sleep-maintained lexical priming effects? To date, the strongest evidence to be taken as evidence against a passive account is the 24-hour experiment in Gaskell et al. (2019). As a reminder, word-meaning priming was enhanced in participants who slept soon after exposure to homonyms in the context of their subordinate meanings before being tested 24 hours later (sleep-wake group), but decayed in those who experienced a day of wakefulness after exposure, before sleeping later in the day and were tested the following morning (wake-sleep group). As both participant groups

---

[5] To the best of our knowledge, it is not possible to calculate a Bayes Factor for a simple effect of one factor (TMR vs No TMR) within a specific level of another factor (primed condition) from a GLMM, which was the key comparison in our main analyses to test our hypotheses (using the *emmeans* package). To calculate our Bayes Factors, we instead filtered the data to the primed condition, and subsequently built a Bayesian ANOVA (with default priors) exploring the main effect of TMR, following Wills et al. (2020). Hence, we deviated slightly from the inferential contrasts used in our main analyses.

experienced roughly equal amounts of wakefulness throughout the 24 hour period, a passive account of sleep would predict similar patterns of priming due to equivalent levels of sensory interference in between language exposure and the final test. However, Yonelinas et al.'s (2019) contextual-binding account takes a more nuanced assumption, and assumes that sensory interference will be strongest when it occurs closer in time to exposure, particularly from information that shares a similar temporal context with newly-encoded memories. Hence, experiencing sleep soon after exposure will provide stronger protection against temporal interference than if sleep occurred later in the day, which could provide an alternative explanation for the 24 hour results in Gaskell et al. (2019).

Nonetheless, prior research has identified an important link between active sleep consolidation and language in some cases. For example, the extent to which newly-learned words display behavioural hallmarks of cortical consolidation is positively correlated with overnight sleep spindle activity (Tamminen et al., 2010; 2013), and new words are generally processed with greater involvement from cortical networks one day after learning (Takashima et al., 2014; 2017). In addition, the amount of SWS after learning - the stage of sleep most strongly associated with sleep-related memory consolidation (Born & Wilhelm, 2012; Rasch & Born, 2013) - is positively correlated with the retention of novel phonotactic constraints (Gaskell et al., 2014) and grammatical rules related to novel articles (Batterink et al., 2014).

The current results, in contrast, found no evidence for an active role of sleep in consolidating memories for familiar words encountered in certain semantic contexts (e.g., such as the sentence "*The branches and bark had been damaged by the storm.*"). This raises the possibility that the perceived novelty of linguistic information might serve as a prerequisite for active sleep-related memory consolidation in the language domain, with memory consolidation being preferentially recruited for relatively *new* linguistic material. The longer-term retention of memories related to familiar linguistic information may therefore depend more heavily on passive sleep roles, such as reduced interference (Yonelinas et al., 2019). An episodic context account that encompassess elements of both active and passive roles of sleep, offering "multicausal" routes of sleep-induced language/memory benefits (Berres & Erdfelder, 2021), may therefore be most parsimonious with the current literature. This would fit with recent evidence rejecting a "strong" version of the episodic context account based on measures of discourse memory. For example, in

Mak et al. (2024), sleep-enhanced memory of passages in short stories, relative to wake, was more nuanced than predicted under the episodic context account and dependent on certain factors, such as the type of retrieval process. Future research is nonetheless needed to further tease apart the respective role of sleep in consolidating memories related to new and familiar linguistic information.

Stepping aside from the issue of sleep's role in memory, the absent TMR effect on word-meaning priming might also be a consequence of TMR having a weak influence on the mechanisms underpinning word-association responses, compared with other types of memory processing. Although TMR effects have been reported in the language domain, we did not find a TMR-enhanced effect on generated lexical associates. Speculatively, it could be that TMR effects, at least in the language domain, are more sensitive to tasks that measure experiment-specific learning. In Schreiner and Rasch (2015), for example, participants learned Dutch-German word pair translations and were tested on these specific pairings in the post-sleep test phase, with word recall enhanced by TMR (see also Göldi & Rasch, 2019; Neumann et al., 2020). In Batterink & Paller (2017), the effect of TMR was measured in an artificial grammar task, with the same words and phrase structures used across learning and test phases. Hence, across these studies, performance in the test phase was specifically guided by learned information accumulated from the pre-sleep phase. For instance, assuming no prior understanding, correctly recalling the German translation when cued with a Dutch word will almost exclusively make use of memory traces experiencing these words together. This contingency, however, is weaker in the word-association task used in the current studies since there are at least two sources of information that guide responses: previous (learning) experiences and long-term semantic knowledge. Potentially, this interplay between competing information sources could dampen a TMR effect, as there is no guarantee that the source of a generated associate for a (primed) homonym originates from an episodic memory trace that would benefit from TMR. Future research could look to measure the effect of TMR on word-meaning priming while isolating learning/exposure effects on homonym interpretation as much as possible, such as using a semantic judgement task (Curtis et al., 2022; Gilbert et al., 2018).

It is also worth noting the magnitude of word-meaning priming in our experiment. We reported a 0.2 increase in the proportion of subordinate meaning responses in the primed compared to the unprimed condition which, numerically speaking, appears to be the largest word-meaning priming effect to be reported in the literature to date. For example, in Experiment 1 of Gaskell et al. (2019) which measured priming with similar materials over a similar interval of sleep, a 0.04 increase was reported. The relative magnitude of our effect could be explained by alterations to our encoding procedures. Rather than compare the meaning of a sentence against a probe word that is typical in word-meaning priming experiments (Betts et al., 2018; Curtis et al., 2022; Gaskell et al., 2019; Gilbert et al., 2018; 2021; Rodd et al., 2013), we instructed participants to develop a mental image of each prime sentence and to associate this with the homonymic cue word. Sentence encoding was therefore less incidental in our experiment, which could have encouraged more input from memory systems, particularly the hippocampus that is traditionally implicated in relational binding (Cohen et al., 1997). In theory, this would lead to stronger memory traces for the prime sentences which could, therefore, provide a stronger source of support to subsequent homonym interpretation. Importantly, however, the large effects reported here could reflect an “upper-limit” of word-meaning priming which could have made it difficult to detect a TMR effect. Presumably, prior linguistic experience plays only a bounded role on lexical processing given independent and competing influences from other information sources such as long-term semantic knowledge. If we had drawn out the maximum (or close to the maximum) effect of exposure, then any additional benefit of TMR on word-meaning priming would be negligible. This is particularly important given that the effect of TMR on behaviour tends to be small-to-medium sized (Hu et al., 2020). Although the formation of an association between contextual information and the cue word was necessary to perform TMR, future research could seek to achieve the same prerequisite while also attempting to decrease the overall effect of exposure.

A second puzzling finding was the null TMR effect on cued recall in the fill-in-the-blank task. This was unexpected given that TMR is known to enhance memory recall (Carbone & Diekelmann, 2024; Hu et al., 2020), which the fill-in-the-blank task served to measure. Similar to the word-meaning priming results, we found a main effect of priming such that the ability to generate the absent homonym was superior in the primed condition, and this difference was significantly larger in Experiment 2. This suggests that memories of the sentences were sufficiently encoded, and these memories were more influential on fill-in-the-blank performance following sleep. One explanation for the null TMR effect concerns retrieval practice. In the recall task completed in the exposure phase, participants recalled the word that they had heard with the sentence in the preceding familiarisation task, which was the same word that was to be recalled in the fill-in-the-blank task. A robust finding in the memory literature is that the long-term retention of learned information is facilitated if some aspect of the information was actively retrieved during learning (Roediger & Butler, 2011). Recent accounts of this retrieval practice effect posit that retrieval prompts consolidation of the associated information on a systems level (Antony et al., 2017). As a result, retrieved information may be less susceptible to general sleep-related consolidation (Bäuml et al., 2014; Mak & Gaskell, 2024) and TMR (Joensen et al., 2022). Although the recall task served to strengthen the association between the prime sentences and memory cues, it could have inadvertently consolidated memories of the sentences enough to reduce a consolidating effect of TMR. Whilst this would neutralise a TMR effect on cued recall, it could feasibly have contributed to the null TMR effect on word-meaning priming.

5. Conclusion

According to the episodic context account, effects such as word-meaning priming are underpinned by episodic memories that represent contextual information from recent linguistic events. This framework further proposes that sleep actively consolidates these memories, enhancing their longer-term support. We examined this prediction by attempting to bias specific linguistic-related memories for consolidation using TMR in a word-meaning priming paradigm. Although we observed an overall word-meaning priming effect and very good recall of contextual information, there was no evidence for an additional benefit of TMR on behaviour. However, there was an enhanced sleep spindle response to memory relative to control cues, implying that some aspect of sleep-based memory reprocessing had taken place. These results suggest that linguistic-related, episodic memories contribute to the processing of similar materials in the future. However, the lack of a TMR effect provides evidence that seems to be consistent with sleep playing a bounded role in actively consolidating such memories, with sleep-related memory consolidation perhaps preferentially recruited for relatively novel linguistic information. Instead, sleep-enhanced memory for familiar information may be better explained by a more passive role of sleep, such as reducing interference from related information.

Acknowledgements: We thank XXX for help recording the stimuli, XXX for TMR assistance, XXX for advice on EEG preprocessing, and members of the XXX at XXX for their general advice and suggestions on the research. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

Funding sources: This research was funded by an Economic and Social Research Council grant (ES/T008571/1) awarded to XXX and XXX.

Declaration of competing interest: We have nothing to declare.

References

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., 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. https://doi.org/10.1016/j.cub.2018.04.020.

Antony, J. W., Schönauer, M., Staresina, B. P., & Cairney, S. A. (2019). Sleep spindles and memory reprocessing. *Trends in Neurosciences, 42*(1), 1-3.

Armstrong, B. C., & Plaut, D. C. (2016). Disparate semantic ambiguity effects from semantic processing dynamics rather than qualitative task differences. *Language, Cognition and Neuroscience, 31*(7), 940-966. https://doi.org/10.1080/23273798.2016.1171366.

Barner, C., Schmid, S. R., & Diekelmann, S. (2019). Time-of-day effects on prospective memory. *Behavioural Brain Research, 376*, 112179. https://doi.org/10.1016/j.bbr.2019.112179.

Bastiaansen, M. C., Van Der Linden, M., Ter Keurs, M., Dijkstra, T., & Hagoort, P. (2005). Theta responses are involved in lexical—Semantic retrieval during language processing. *Journal of Cognitive Neuroscience, 17*(3), 530-541. https://doi.org/10.1162/0898929053279469.

Bates D, Mächler M, Bolker B, Walker S (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1–48. 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.

Berres, S., & Erdfelder, E. (2021). The sleep benefit in episodic memory: An integrative review and a meta-analysis. *Psychological Bulletin, 147*(12), 1309-1353.

Batterink, L. J., & Paller, K. A. (2017). Sleep-based memory processing facilitates grammatical generalization: Evidence from targeted memory reactivation. *Brain and Language, 167*, 83-93. https://doi.org/10.1016/j.bandl.2015.09.003.

Batterink, L. J., Oudiette, D., Reber, P. J., & Paller, K. A. (2014). Sleep facilitates learning a new linguistic rule. *Neuropsychologia, 65*, 169-179. https://doi.org/10.1016/j.neuropsychologia.2014.10.024.

Betts, H. N., Gilbert, R. A., Cai, Z. G., Okedara, Z. B., & Rodd, J. M. (2018). Retuning of lexical-semantic representations: Repetition and spacing effects in word-meaning priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(7), 1130–1150. https://doi.org/10.1037/xlm0000507.

Blott, L. M., Hartopp, O., Nation, K., & Rodd, J. M. (2022). Learning about the meanings of ambiguous words: evidence from a word-meaning priming paradigm with short narratives. *PeerJ, 10*, e14070. https://doi.org/10.7717/peerj.14070.

Born, J., & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological research, 76*, 192-203. https://doi.org/10.1007/s00426-011-0335-6.

Brunner, D., Vasko, R., Detka, C., Monahan, J., Reynolds, C III., & Kupfer, D. (1996). Muscle artifacts in the sleep EEG: Automated detection and effect on all-night EEG power spectra. *Journal of Sleep Research, 5*(3), 155-164. https://doi.org/10.1046/j.1365-2869.1996.00009.x.

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1*(1). https://doi.org/10.5334/joc.10.

Cairney, S. A., Durrant, S. J., Hulleman, J., & Lewis, P. A. (2014). Targeted memory reactivation during slow wave sleep facilitates emotional memory consolidation. *Sleep, 37*(4), 701-707. https://doi.org/10.5665/sleep.3572.

Cairney, S. A., Guttesen, A. V., El Marj, N., & Staresina, B. P. (2018). Memory consolidation is linked to spindle-mediated information processing during sleep. *Current Biology, 28*(6), 948-954. https://doi.org/10.1016/j.cub.2018.01.087.

Cairney, S. A., Lindsay, S., Sobczak, J. M., Paller, K. A., & Gaskell, M. G. (2016). The benefits of targeted memory reactivation for consolidation in sleep are contingent on memory accuracy and direct cue-memory associations. *Sleep, 39*(5), 1139-1150. https://doi.org/10.5665/sleep.5772.

Cairney, S. A., Sobczak, J. M., Lindsay, S., & Gaskell, M. G. (2017). Mechanisms of memory retrieval in slow-wave sleep. *Sleep, 40*(9), zsx114. https://doi.org/10.1093/sleep/zsx114.

Carbone, J., & Diekelmann, S. (2024). An update on recent advances in targeted memory reactivation during sleep. *npj Science of Learning, 9*(1), 31. https://doi.org/10.1038/s41539-024-00244-8.

Cohen, N. J., Ryan, J., Hunt, C., Romine, L., Wszalek, T., & Nash, C. (1999). Hippocampal system and declarative (relational) memory: summarizing the data from functional neuroimaging studies. *Hippocampus, 9*(1), 83-98.

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.

Curtis, A. J., Mak, M. H., Chen, S., Rodd, J. M., & Gaskell, M. G. (2022). Word-meaning priming extends beyond homonyms. *Cognition, 226,* 105175. https://doi.org/10.1016/j.cognition.2022.105175.

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1536), 3773-3800. https://doi.org/10.1098/rstb.2009.0111.

Denis, D., & Cairney, S. A. (2023). Neural reactivation during human sleep. *Emerging Topics in Life Sciences, 7*(5), 487-498. https://doi.org/10.1042/ETLS20230109.

Duff, M. C., & Brown-Schmidt, S. (2012). The hippocampus and the flexible use and processing of language. *Frontiers in Human Neuroscience, 6.* https://doi.org/10.3389/fnhum.2012.00069.

Duff, M. C., & Brown-Schmidt, S. (2017). Hippocampal Contributions to Language Use and Processing. In D. E. Hannula, & M. C. Duff (Eds.), *The hippocampus from cells to systems: Structure, connectivity, and functional contributions to memory and flexible cognition* (pp. 503–536). Springer International Publishing. https://doi.org/10.1007/978-3-319-50406-3_16.

Ellenbogen, J. M., Payne, J. D., & Stickgold, R. (2006). The role of sleep in declarative memory consolidation: passive, permissive, active or none?. *Current Opinion in Neurobiology, 16*(6), 716-722. https://doi.org/10.1016/j.conb.2006.10.006.

Farthouat, J., Gilson, M., & Peigneux, P. (2017). New evidence for the necessity of a silent plastic period during sleep for a memory benefit of targeted memory reactivation. *Sleep Spindles & Cortical Up States, 1*(1), 14-26. https://doi.org/10.1556/2053.1.2016.002.

Frankland, P. W., & Bontempi, B. (2005). The organization of recent and remote memories. *Nature Reviews Neuroscience, 6*(2), 119-130. https://doi.org/10.1038/nrn1607.

Gaskell, M. G., Cairney, S. A., & Rodd, J. M. (2019). Contextual priming of word meanings is stabilized over sleep. *Cognition, 182*, 109-126. https://doi.org/10.1016/j.cognition.2018.09.007.

Gaskell, M. G., Warker, J., Lindsay, S., Frost, R., Guest, J., Snowdon, R., & Stackhouse, A. (2014). Sleep underpins the plasticity of language production. *Psychological Science, 25*(7), 1457-1465. https://doi.org/10.1177/09567976145359.

Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review, 19*, 847-857. https://doi.org/10.3758/s13423-012-0296-9.

Gilbert, R. A., Davis, M. H., Gaskell, M. G., & Rodd, J. M. (2018). Listeners and readers generalize their experience with word meanings across modalities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(10), 1533-1561. https://doi.org/10.1037/xlm0000532.

Gilbert, R. A., Davis, M. H., Gaskell, M. G., & Rodd, J. M. (2021). The relationship between sentence comprehension and lexical-semantic retuning. *Journal of Memory and Language, 116*, 104188. https://doi.org/10.1016/j.jml.2020.104188.

Gilbert, R. A., & Rodd, J. M. (2022). Dominance norms and data for spoken ambiguous words in British English. *Journal of Cognition, 5*(1). https://doi.org/10.5334/joc.194.

Göldi, M., & Rasch, B. (2019). Effects of targeted memory reactivation during sleep at home depend on sleep disturbances and habituation. *NPJ Science of Learning, 4*(1), 5. https://doi.org/10.1038/s41539-019-0044-2.

Göldi, M., van Poppel, E. A. M., Rasch, B., & Schreiner, T. (2019). Increased neuronal signatures of targeted memory reactivation during slow-wave up states. *Scientific Reports, 9*(1), 2715. https://doi.org/10.1038/s41598-019-39178-2.

Groch, S., Schreiner, T., Rasch, B., Huber, R., & Wilhelm, I. (2017). Prior knowledge is essential for the beneficial effect of targeted memory reactivation during sleep. *Scientific Reports*, 7(1), 39763. https://doi.org/10.1038/srep39763.

Guillot, A., & Thorey, V. (2021). RobustSleepNet: Transfer learning for automated sleep staging at scale. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29, 1441-1451. https://doi.org/10.1109/TNSRE.2021.3098968.

Guttesen, A. V., Denis, D., Gaskell, M. G., & Cairney, S. A. (2024). Delineating memory reactivation in sleep with verbal and non-verbal retrieval cues. *Cerebral Cortex, 34*(5), bhae183. https://doi.org/10.1093/cercor/bhae183.

Hu, X., Cheng, L. Y., Chiu, M. H., & Paller, K. A. (2020). Promoting memory consolidation during sleep: A meta-analysis of targeted memory reactivation. *Psychological Bulletin, 146*(3), 218-244. https://doi.org/10.1037/bul0000223.

Ibáñez, A., López, V., & Cornejo, C. (2006). ERPs and contextual semantic discrimination: degrees of congruence in wakefulness and sleep. *Brain and Language, 98*(3), 264-275. https://doi.org/10.1016/j.bandl.2006.05.005.

Joensen, B. H., Harrington, M. O., Berens, S. C., Cairney, S. A., Gaskell, M. G., & Horner, A. J. (2022). Targeted memory reactivation during sleep can induce forgetting of overlapping memories. *Learning & Memory, 29*(11), 401-411. https://doi.org/10.1101/lm.053594.122.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3?. *Perception, 36*(14).

Krause, C. M., Grönholm, P., Leinonen, A., Laine, M., Säkkinen, A. L., & Söderholm, C. (2006). Modality matters: the effects of stimulus modality on the 4-to 30-Hz brain electric oscillations during a lexical decision task. *Brain Research, 1110*(1), 182-192. https://doi.org/10.1016/j.brainres.2006.06.100.

Laventure, S., Pinsard, B., Lungu, O., Carrier, J., Fogel, S., Benali, H., Lina, J-M., Boutin, A., & Doyon, J. (2018). Beyond spindles: interactions between sleep spindles and boundary frequencies during cued reactivation of motor memory representations. *Sleep, 41*(9), zsy142. https://doi.org/10.1093/sleep/zsy142.

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course* (pp. xiii, 264). Cambridge University Press. https://doi.org/10.1017/CBO9781139087759.

Lee, Y. J., Lee, J. Y., Cho, J. H., & Choi, J. H. (2022). Interrater reliability of sleep stage scoring: a meta-analysis. *Journal of Clinical Sleep Medicine, 18*(1), 193-202. https://doi.org/10.5664/jcsm.9538.

Lenth, R. V. (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. *R package version, 1.7*.1-1. https://CRAN.R-project.org/package=emmeans.

Lewis, P. A., & Bendor, D. (2019). How targeted memory reactivation promotes the selective strengthening of memories in sleep. *Current Biology, 29*(18), PR906-R912. https://doi.org/10.1016/j.cub.2019.08.019.

Mak, M. H. C, & Gaskell, M. (2023, March 1). Effects of sleep and retrieval practice on verbal paired-associate learning across 12 and 24 hours. https://doi.org/10.31234/osf.io/phe5j.

Mak, M.H.C., Ball, L, O'Hagan, A., Walsh, C. R., & Gaskell, M. G. (2025). Involvement of Episodic Memory in Language Comprehension: Naturalistic Comprehension Pushes Unrelated Words Closer in Semantic Space for at Least 12 Hours. http://dx.doi.org/10.2139/ssrn.4975599.

Mak, M. H. C., Curtis, A. J., Rodd, J. M., & Gaskell, M. G. (2023). Episodic memory and sleep are involved in the maintenance of context-specific lexical information. *Journal of Experimental Psychology: General, 52*(11), 3087–3115. https://doi.org/10.1037/xge0001435.

Mak, M. H. C., Curtis, A. J., Rodd, J. M., & Gaskell, M. G. (2024). Recall and recognition of discourse memory across sleep and wake. *Journal of Memory and Language, 138*, 104536. https://doi.org/10.1016/j.jml.2024.104536.

Marinkovic, K., Rosen, B. Q., Cox, B., & Kovacevic, S. (2012). Event-related theta power during lexical-semantic retrieval and decision conflict is modulated by alcohol intoxication:

anatomically constrained MEG. *Frontiers in Psychology, 3,* 121. https://doi.org/10.3389/fpsyg.2012.00121.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177-190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

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. https://doi.org/10.1037/0033-295X.102.3.419.

McConnell, K., Hintz, F., & Meyer, A. S. (2025). Individual differences in online research: Comparing lab-based and online administration of a psycholinguistic battery of linguistic and domain-general skills. *Behavior Research Methods, 57*(1), 1-19. https://doi.org/10.3758/s13428-024-02533-x.

Neumann, F., Oberhauser, V., & Kornmeier, J. (2020). How odor cues help to optimize learning during sleep in a real life-setting. *Scientific Reports, 10*(1), 1227. https://doi.org/10.1038/s41598-020-57613-7.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*(1), 156869. https://doi.org/10.1155/2011/156869.

Oyarzún, J. P., Morís, J., Luque, D., de Diego-Balaguer, R., & Fuentemilla, L. (2017). Targeted memory reactivation during sleep adaptively promotes the strengthening or weakening of overlapping memories. *Journal of Neuroscience, 37*(32), 7748-7758. https://doi.org/10.1523/JNEUROSCI.3537-16.2017.

Parker, A. J., Taylor, J., & Rodd, J. M. (2023). Readers use recent experiences with word meanings to support the processing of lexical ambiguity: Evidence from eye movements.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*(3), 191-201.

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.

Rasch, B., Büchel, C., Gais, S., & Born, J. (2007). Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science, 315*(5817), 1426-1429. https://doi.org/10.1126/science.1138581.

Rodd, J. M. (2024). Moving experimental psychology online: How to obtain high quality data when we can't see our participants. *Journal of Memory and Language, 134*, 104472. https://doi.org/10.1016/j.jml.2023.104472.

Rodd, J. M. (2020). Settling into semantic space: An ambiguity-focused account of word-meaning access. *Perspectives on Psychological Science, 15*(2), 411-427. https://doi.org/10.1177/1745691619885860.

Rodd, J. M., Cai, Z. G., Betts, H. N., Hanby, B., Hutchinson, C., & Adler, A. (2016). The impact of recent and long-term experience on access to word meanings: Evidence from large-scale internet-based experiments. *Journal of Memory and Language, 87*, 16-37. https://doi.org/10.1016/j.jml.2015.10.006.

Rodd, J. M., Cutrin, B. L., Kirsch, H., Millar, A., & Davis, M. H. (2013). Long-term priming of the meanings of ambiguous words. *Journal of Memory and Language, 68*(2), 180-198. https://doi.org/10.1016/j.jml.2012.08.002.

Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2004). Modelling the effects of semantic ambiguity in word recognition. *Cognitive science, 28*(1), 89-104. https://doi.org/10.1207/s15516709cog2801_4.

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*(1), 20-27. https://doi.org/10.1016/j.tics.2010.09.003.

Rudoy, J. D., Voss, J. L., Westerberg, C. E., & Paller, K. A. (2009). Strengthening individual memories by reactivating them during sleep. *Science, 326*(5956), 1079-1079. https://doi.org/10.1126/science.1179013.

Shimizu, R. E., Connolly, P. M., Cellini, N., Armstrong, D. M., Hernandez, L. T., Estrada, R., Aguilar, M., Weisend, M. P., Mednick, S. C., & Simons, S. B. (2018). Closed-loop targeted memory reactivation during sleep improves spatial navigation. *Frontiers in Human Neuroscience, 12*, 28. https://doi.org/10.3389/fnhum.2018.00028.

Schmidig, F. J., Ruch, S., & Henke, K. (2024). Episodic long-term memory formation during slow-wave sleep. *Elife, 12*, RP89601. https://doi.org/10.7554/eLife.89601.3.

Schreiner, T., Petzka, M., Staudigl, T., & Staresina, B. P. (2021). Endogenous memory reactivation during sleep in humans is clocked by slow oscillation-spindle complexes. *Nature Communications, 12*(1), 3112. https://doi.org/10.1038/s41467-021-23520-2.

Schreiner, T., & Rasch, B. (2015). Boosting vocabulary learning by verbal cueing during sleep. *Cerebral Cortex, 25*(11), 4169-4179. https://doi.org/10.1093/cercor/bhu139.

Sifuentes-Ortega, R., & Peigneux, P. (2024). Does targeted memory reactivation during slow-wave sleep and rapid eye movement sleep have differential effects on mnemonic discrimination and generalization?. *Sleep*, *47*(10), zsae114. https://doi.org/10.1093/sleep/zsae114.

Sterpenich, V., van Schie, M. K., Catsiyannis, M., Ramyead, A., Perrig, S., Yang, H. D., ... & Schwartz, S. (2021). Reward biases spontaneous neural reactivation during sleep. *Nature Communications, 12*(1), 4162. https://doi.org/10.1038/s41467-021-24357-5.

Takashima, A., Bakker, I., Van Hell, J. G., Janzen, G., & McQueen, J. M. (2014). Richness of information about novel words influences how episodic and semantic memory networks interact during lexicalization. *NeuroImage, 84,* 265-278. https://doi.org/10.1016/j.neuroimage.2013.08.023.

Takashima, A., Bakker, I., Van Hell, J. G., Janzen, G., & McQueen, J. M. (2017). Interaction between episodic and semantic memory networks in the acquisition and consolidation of novel spoken words. *Brain and Language, 167*, 44-60. https://doi.org/10.1016/j.bandl.2016.05.009.

Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, M. G. (2010). Sleep spindle activity is associated with the integration of new memories and existing knowledge. *Journal of Neuroscience, 30*(43), 14356-14360. https://doi.org/10.1523/JNEUROSCI.3028-10.2010.

Tamminen, J., Ralph, M. A. L., & Lewis, P. A. (2013). The role of sleep spindles and slow-wave activity in integrating new information in semantic memory. *Journal of Neuroscience, 33*(39), 15376-15381. https://doi.org/10.1523/JNEUROSCI.5093-12.2013.

Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory & Cognition, 22*(1), 111–126. https://doi.org/10.3758/BF03202766.

Vitello, S., & Rodd, J. M. (2015). Resolving semantic ambiguities in sentences: Cognitive processes and brain mechanisms. *Language and Linguistics Compass, 9*(10), 391-405. https://doi.org/10.1111/lnc3.12160.

Voeten, C. C. (2022). buildmer: Stepwise Elimination and Term Reordering for MixedEffects Regression. *R package version, 2*, 6. https://CRAN.R-project.org/package=buildmer.

Wick, A., & Rasch, B. (2023). Targeted memory reactivation during slow-wave sleep vs. sleep stage N2: no significant differences in a vocabulary task. *Learning & Memory, 30*(9), 192-200. https://doi.org/10.1101/lm.053683.122.

Wills, A., Walsh, C., Sharpe, P., & Mitchell, C. (2020). More on Bayesian ANOVA. Retrieved from https://www.andywills.info/rminr/anova4.html.

Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. *Journal of the International Neuropsychological Society, 17*(5), 766-780. https://doi.org/10.1017/S1355617711000683.

Yonelinas, A. P., Ranganath, C., Ekstrom, A. D., & Wiltgen, B. J. (2019). A contextual binding theory of episodic memory: systems consolidation reconsidered. *Nature Reviews Neuroscience, 20*(6), 364-375. https://doi.org/10.1038/s41583-019-0150-4.

Highlights (85 characters each, max of 5)

- We primed homonyms towards their subordinate meanings before participants took a nap.
- Such meanings were more accessible after the nap (word-meaning priming).
- Targeted memory reactivation, initiated during the nap, did not boost priming.
- Memory-related auditory cues were associated with enhanced sleep spindle activity.
- Results imply a bounded role of sleep in actively supporting language processing.