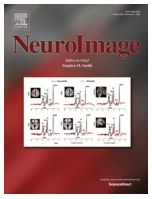
NeuroImage 266 (2023) 119820

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage



# Targeting targeted memory reactivation: Characteristics of cued reactivation in sleep

Mahmoud E.A. Abdellahi[a,*], Anne C.M. Koopman[a], Matthias S. Treder[b], Penelope A. Lewis[a,*]

[a] School of Psychology, Cardiff University Brain Research Imaging Centre (CUBRIC), Cardiff CF24 4HQ, United Kingdom
[b] School of Computer Science and Informatics, Cardiff University, Cardiff CF24 3AA, United Kingdom

ARTICLE INFO

*Keywords:*
Sleep
Targeted memory reactivation
Slow wave sleep
Reactivation
Classification
Machine learning
Slow oscillations
Spindles

ABSTRACT

Targeted memory reactivation (TMR) is a technique in which sensory cues associated with memories during wake are used to trigger memory reactivation during subsequent sleep. The characteristics of such cued reactivation, and the optimal placement of TMR cues, remain to be determined. We built an EEG classification pipeline that discriminated reactivation of right- and left-handed movements and found that cues which fall on the up-going transition of the slow oscillation (SO) are more likely to elicit a classifiable reactivation. We also used a novel machine learning pipeline to predict the likelihood of eliciting a classifiable reactivation after each TMR cue using the presence of spindles and features of SOs. Finally, we found that reactivations occurred either immediately after the cue or one second later. These findings greatly extend our understanding of memory reactivation and pave the way for development of wearable technologies to efficiently enhance memory through cueing in sleep.

## 1. Introduction

Memories are neurally replayed during sleep, and this process is associated with consolidation (Rasch and Born, 2013; Squire et al., 2015; Ólafsdóttir et al., 2018). Targeted memory reactivation (TMR) is a technique in which sensory cues are paired with learned material during wake, then re-presented during subsequent sleep in order to trigger reactivation of the associated material (Cellini and Cappuzo, 2018; Hu et al., 2020). This procedure leads to memory benefits for reactivated material in Non-REM (NREM) sleep (see Hu et al., 2020 for a recent meta-analysis). Importantly, several studies have confirmed the reinstatement of learning related brain activity after TMR cues in NREM sleep (see Lewis and Bendor, 2019 for a review). Studies have looked at the neural structures involved in reactivation (van Dongen et al., 2011; Shanahan et al., 2018), and found both positive (Shanahan et al., 2018; Cairney et al., 2018; Schreiner et al., 2018; Wang et al., 2019), and negative (Murphy et al., 2018) relationships between the extent of reactivation and subsequent memory benefits.

Cortical activity during slow wave sleep (SWS) is characterised by high amplitude slow oscillations (SOs) in which neurones oscillate between hyperpolarization with neuronal silence ("down-state") and depolarisation with sustained firing ("up-state"). Depolarised SO up-states drive memory reactivation in the hippocampus via interactions with thalamic sleep spindles (SS) and hippocampal sharp wave ripples (SWRs) (Diekelmann and Born, 2010). Studies have shown that stimulation during the up-going SO phase is associated with greater memory benefit compared to the down-going phase (Göldi et al., 2019; Shimizu et al., 2018), and that stimulating the up-going phase produces a higher event-related potential (ERP) response compared to the down-going phase (Schabus et al., 2012). This could be due to the fact that neurones are in the process of depolarisation and are thus moving closer to the threshold for firing during the up-going phase. Furthermore, fast spindles, which have been linked to both memory consolidation (Nishida and Walker, 2007) and reactivation (Cairney et al., 2018), typically occur on the up-going phase (Born and Wilhelm, 2012; Siclari et al., 2014).

TMR is thought to prime a memory trace for reactivation (Lewis and Bendor, 2019), and has been shown to trigger SO-spindle complexes (Cairney et al., 2018; Schreiner et al., 2015; Oyarzún et al., 2017). We predict that application of such priming during the up-going phase of the SO just prior to a spindle event may lead to reactivation. On the other hand, application of stimulation during the down-going phase of the oscillation when fast spindles rarely occur and excitability is reduced, is less likely to produce reactivation.

SOs vary in terms of generation locus as well as shape, for instance having different periods, trough depths, and peak to trough slopes. These varied morphologies are thought to relate to the degree of synchronisation across neural populations in the cortex (Siclari et al., 2014; Bernardi et al., 2018). Given these differences, some SOs are likely to facilitate reactivation more efficiently than others. We hypothesise that it may be possible to predict this efficiency based on features of the ongo-

* Corresponding author.
*E-mail addresses:* maheidlahi@gmail.com, abdellahime@cardiff.ac.uk (M.E.A. Abdellahi), LewisP8@cardiff.ac.uk (P.A. Lewis).

https://doi.org/10.1016/j.neuroimage.2022.119820.
Received 2 September 2022; Received in revised form 16 November 2022; Accepted 15 December 2022
Available online 16 December 2022.
1053-8119/© 2022 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

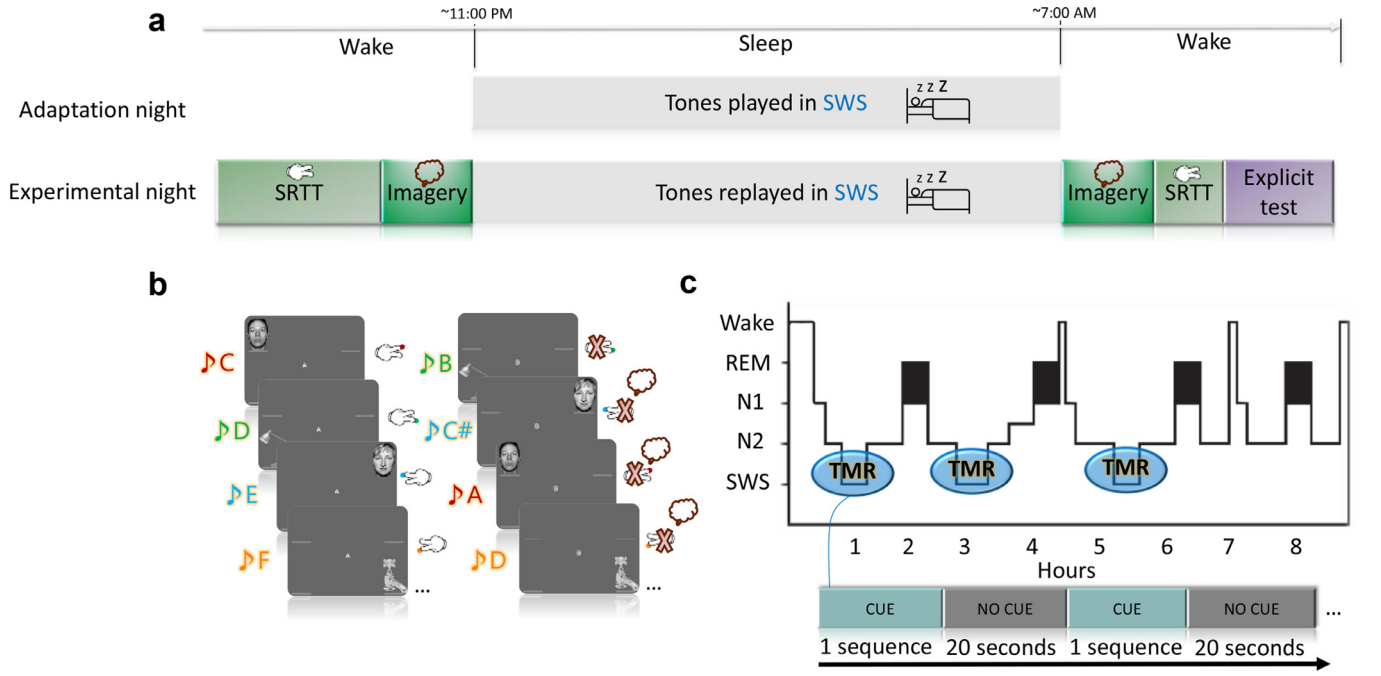



**Fig. 1.** Experimental design. **(a)** Participants spent two nights in the lab, Firstly, an adaptation night during which EEG recordings were acquired and tones were presented to the participants during SWS. Secondly, in the experimental night, participants completed the serial reaction time task (SRTT), followed by motor imagery task. In the imagery task, participants were cued with pictures and sounds but were told to only imagine performing the finger presses without moving. Afterwards, participants slept in the lab and TMR cues were presented during SWS. After waking up, participants completed the motor imagery and then the SRTT, and finally an explicit recall task in which they marked the locations of the images as they had appeared in the sequences as accurately as they could. **(b)** In the SRTT, images are presented in two different sequences each with a different set of tones. Each image is associated with a unique tone and requires a specific button press. In the imagery task, participants hear the tones and see the images as in the SRTT but only imagine pressing the buttons. **(c)** The sounds of only one learned sequence (cued/reactivated sequence) were re-played during SWS sleep, with a 20 s pause between repetitions.

ing oscillatory structure of sleep, with specific reference to SOs and spindles, in the time period directly before stimulation. This would not only optimise stimulation, but also allow more targeted stimulation, minimising the number of sound cues needed to influence consolidation.

In the current work, we set out to characterise memory reactivation after TMR in NREM sleep and to determine whether TMR on the upgoing phase is more likely to elicit reactivation compared to the downgoing phase. Additionally, we ask whether we can predict the optimal time for TMR stimulation using ongoing morphology of SOs and spindles. We use a serial reaction time task (SRTT). In the SRTT, participants respond to audio-visual cues by pressing four buttons, one per finger, with two fingers on each hand (Fig. 1). Each press was cued by a picture-sound pair, and tones associated with the task were replayed during SWS on the post-training night to elicit memory reactivation. Importantly, prior to performing the task, participants were exposed to the tones during an adaptation night. The addition of this adaptation night provided control data during which tones could not have evoked memory reactivation, as they were not yet linked to the task. We then trained a classifier to identify neural responses associated with left- and right-handed presses in wake and applied it after each TMR tone in SWS on both adaptation and experimental nights. We also used the features of the ongoing oscillation to train another classifier to determine whether TMR applied at a given time in the oscillatory sequence would elicit detectable reactivation.

## 2. Results and discussion

### 2.1. TMR improved sequence memory

Prior work has shown that the SRTT is facilitated by TMR in SWS (Cousins et al., 2014; Cousins et al., 2016; Schönauer et al., 2014). Considering only the participants that were included in the classification ($n = 12$), our data show an improvement for the reactivated sequence (Wilcoxon signed rank test, $n = 12$, $p = 0.006$, $z = 2.75$), and no improvement for the non-reactivated sequence (Wilcoxon signed rank test, $n = 12$, $p = 0.084$, $z = 1.73$). However, the difference between the improvements on these two sequences is not significant (Wilcoxon signed rank test, $n = 12$, $p = 0.071$, $z = 1.8$). See (Koopman et al., 2020) for an analysis of the full dataset of 15 participants, which does show a difference between reactivated and non-reactivated sequences.

An explicit test for sequence memory at the end of the data collection adopted the same approach as that in (Cousins et al., 2014) where an item is considered correct if it is in the correct position in the sequence and is part of a segment of more than two such correct items. A Wilcoxon signed-rank test showed no difference in the explicit knowledge between reactivated and non-reactivated sequences (Wilcoxon signed rank test, $n = 12$, $p = 0.932$, $z = -0.085$). This could be due to the fact that participants were much more highly trained on the sequences, as compared to prior investigations of this question, e.g., Cousins et al. (2014) which did not involve an imagery condition.

### 2.2. Multiple reactivations detected after TMR

Prior work (Cairney et al., 2018; Schreiner et al., 2018) has suggested a recurrent pattern of reactivation after a TMR cue, with a reinstatement of the target memory immediately after the cued memory followed by a later reinstatement, see Lewis and Bendor (2019) for a discussion. Building on this work, we examined the time course of classification after TMR for evidence of a similar pattern. Our results revealed a significantly higher classification performance in the experimental night than in the adaptation night, with two different effects described by two clusters after TMR onset (Fig. 2a). An early cluster ($p = 0.02$) occurred immediately after TMR onset and a late cluster ($p = 0.01$) occurred ~1 s later. Results were corrected for multiple comparisons with

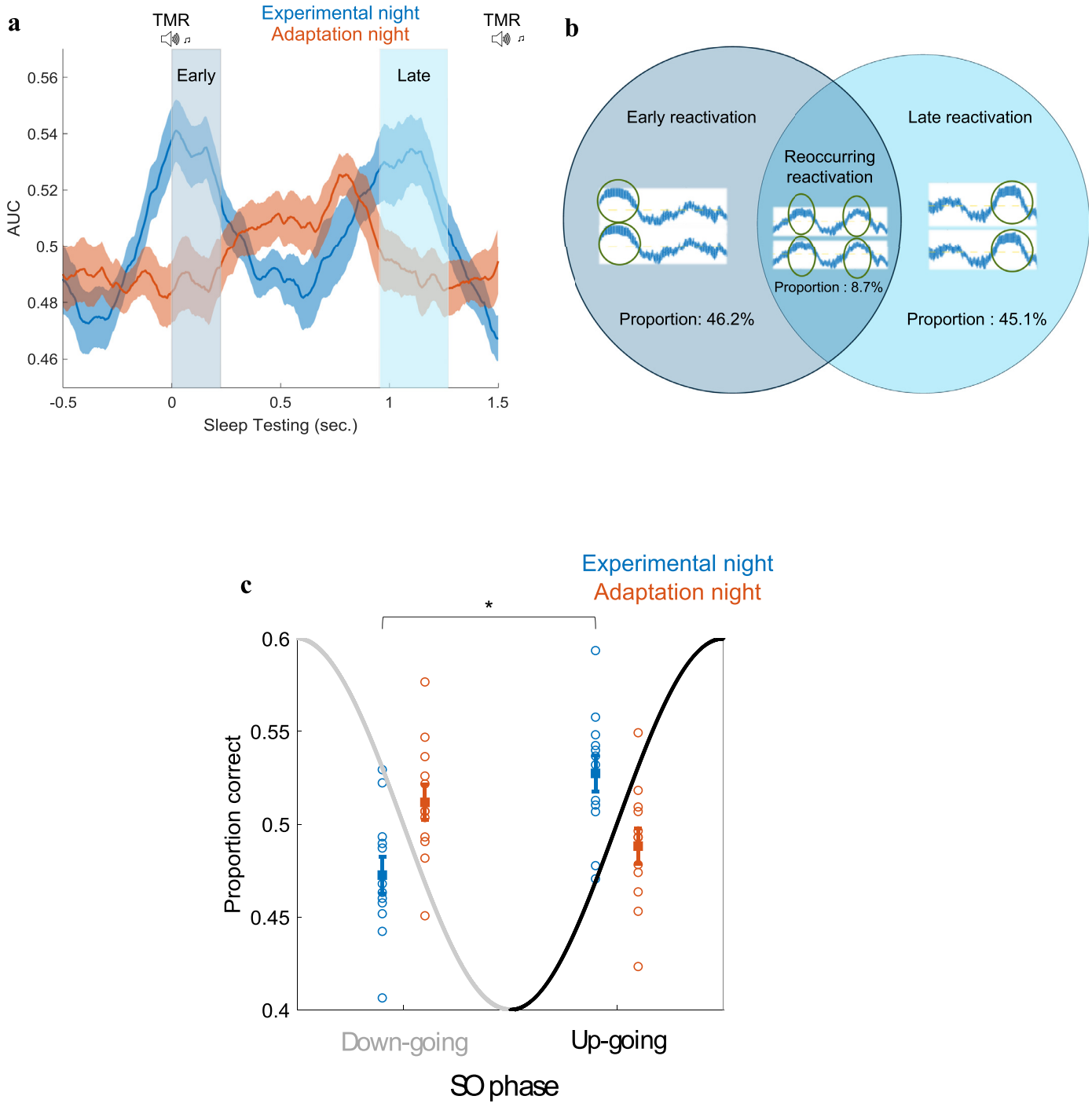



**Fig. 2.** Detected reactivations and their characteristics. **(a)** Classification results for both nights. The blue curve represents the area under the ROC curve (AUC) across time for the experimental night (with the mean represented by a solid curve and standard error shaded around it), red curve represents results of the adaptation night. TMR sounds were presented at the beginning of sleep trials, 'early' and 'late' are used to mark early and late reactivations. Classification results show two significant effects expressed by two clusters, (early cluster, $p = 0.02$, and late cluster, $p = 0.01$). **(b)** Proportions of correct trials with only early reactivation (46.2%), only late reactivation (45.1%), and reoccurring reactivations (8.7%). **(c)** Percentage of correct trials with the TMR cue falling on different SO phase transitions for the two nights. Each circle represents data from one participant and the shown curve is a simplified cartoon representation of the phase of a SO, the two SO phases are marked on the x-axis and the y-axis represents the proportions of correct trials. The preferred phase for early reactivation is when the sound falls on the up-going transition of the SO (Wilcoxon signed rank test, $n = 12$, $P = 0.019$, $Z = 2.4$) compared to down-going.

cluster-based permutation (see methods for details), trial duration in sleep was 1500 ms.

To test whether this was due to recurrent reactivation of the same response, we examined each trial to see whether it included an early reactivation, a late reactivation, or both. We then looked at whether the same trials were classified correctly at both early and late peaks (Fig. 2b). This revealed that the majority of trials contained one reactivation, either early or late, and only 8.7% of trials showed reoccurring reactivation by classifying correctly during both early and late peaks. Comparison of reoccurring reactivation to chance level showed that it was below chance (Wilcoxon signed rank test, $n = 12$, $p = 0.002$, $z = -3.06$) (see methods for details, Fig. 2b).

Overall, these results suggest that the reactivations we are detecting are not recurrent, but instead commonly occur just once after each cue: either early or late within our trial duration. It is possible that this may also have been the case in the prior human studies (Cairney et al., 2018; Schreiner et al., 2018), as they reported the performance across many trials and did not examine individual trials.

### 2.3. Preferred TMR phase for reactivation

There is evidence that TMR may be more effective when applied to the up-going phase of the SO (Göldi et al., 2019; Göldi et al., 2017; Ngo et al., 2013). Moreover, fast rhythms, such as spindle, and gamma activity are more prominent in the SO up-going state than in the SO down-going state (Mölle et al., 2002; Valderrama et al., 2012; Piantoni et al., 2013), also there are changes to the ERP when the auditory stimulation is applied during the up-going phase of the SO (Schabus et al., 2012). Building on the extensive literature relating to reactivation during rodent sharp-wave ripples (Kudrimoti et al., 1999; Nakashiba et al., 2009; O'Neill et al., 2008), data from human epilepsy patients has shown that the SO up-going state shows more sharp-wave ripples and gamma oscillations (Van Quyen et al., 2010). Sharp-wave ripples have been shown to carry reactivation (Zhang et al., 2018), on the other hand, they are suppressed during the SO down-going state (Clemens et al., 2007). Thus, the up-going SO phase appears to be the preferred time for stimulation to improve memory (Göldi et al., 2019).

Given this background, we predicted that TMR would more effectively trigger reactivations if applied to the up-going phase of the SO. We tested this by calculating the percentage of correct trials that got TMR on the up-going phase. Thus, we divided the number of correct trials that received TMR on the up-going phase by the number of correct trials on both up-going and down-going phases. For early reactivation in the experimental night, this showed a significantly higher percentage of correct trials when TMR was applied on the up-going compared to the down-going SO transition, (Wilcoxon signed rank test, $n = 12$, $p = 0.019$, $z = 2.4$, Fig. 2c). As a control, we compared the percentage of correct trials from the adaptation night for up-going and down-going SO phases and chance level (Wilcoxon signed rank test, $n = 12$, $p = 0.24$, $z = -1.18$). We also repeated all analyses for late reactivation but found no difference between up-going and down-going phase transitions for both nights ($p > 0.07$). Raw traces for both correct and incorrect trials are provided as examples in Supplementary Fig. 4.

This analysis shows that TMR cues which fall on the up-going transition of the SO, are more likely to lead to a classifiable early reactivation, than TMR cues that fall on the down-going phase supporting the idea that SOs interact with reactivation in some functional way. This could also be important for optimisation of TMR cueing in order to successfully trigger reactivation.

### 2.4. Predicting reactivation using pre-cue SO features

While the literature suggests that reactivation is modulated by SOs (Rasch and Born, 2013; Inostroza and Born, 2013), the mechanism for this modulation remains to be understood. We were interested to determine whether the features of the ongoing SO prior to stimulation, could predict whether a given TMR cue would produce a classifiable reactivation. In other words, we wanted to know whether some points in the oscillatory pattern are more optimal than others for delivering TMR, and if so, which features of the ongoing oscillatory structure determine this. To examine this, we performed a second classification analysis, this time by training our classifiers on pre-cue SO features. We wanted to see if we could discriminate between correct and incorrect trials (based on the results of our main reactivation classifier, Fig. 2a). To this end, we extracted SO features from Fz electrode during the two seconds of data before the onset of TMR.

The extracted features are described in Supplementary Table 1. These features were fed to decision tree classifiers (Gordon et al., 1984)

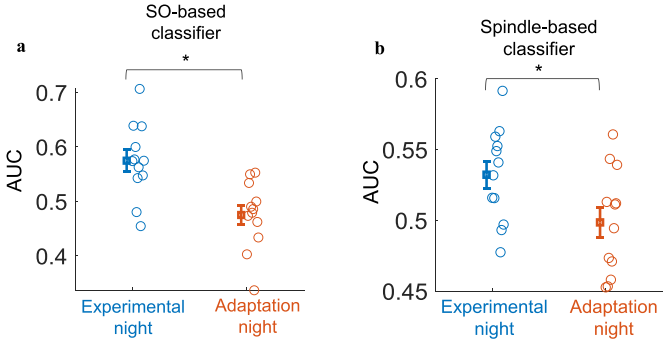



**Fig. 3.** Predicting reactivation using pre-cue features. **(a)** Classification results of the SO-based classifier for the experimental vs. the adaptation night for early reactivation (Wilcoxon signed rank test, $n = 12$, $P = 0.015$, $Z = 2.43$). **(b)** Classification results of the spindle-based classifier for the experimental vs. the adaptation night for late reactivation (Wilcoxon signed rank test, $n = 12$, $P = 0.04$, $Z = 2.04$).

which were trained on two classes: correctly classified, and incorrectly classified from the main classifier, see methods for details. As a control, we compared the results obtained from the SO-based classifier of the experimental night to the SO-based classifier trained and tested using the adaptation night (Fig. 3a). The performance of the experimental night classifier was significantly higher than that of the adaptation night for predicting the early reactivation (Wilcoxon signed rank test, $n = 12$, $p = 0.015$, $z = 2.43$) and higher than chance level (Wilcoxon signed rank test, $n = 12$, $p = 0.006$, $z = 2.75$), but not for predicting the late reactivation ($p > 0.2$). This indicates that it is possible to predict classifiable early reactivation in the experimental night when learned information could actually be reactivated, compared to the control condition in the adaptation night when the task had not been learned yet. This result shows that we can use SO features to predict the optimal time to deliver TMR, in order to maximise the probability of producing a classifiable early reactivation.

In addition to the ongoing pattern of SO oscillations, we were interested in how spindles might impact upon the ability of TMR to elicit classifiable reactivations. We therefore repeated the above analysis, but now using spindle features not SOs. We thus trained a spindle-based classifier to predict whether we could use these higher frequency oscillations to determine whether TMR would produce a correct classification, using features from channels around the motor area (C5, CP3, C6, and CP4). We thus extracted a binary value representing whether there was a spindle in the 1.5 s duration pre-cue (0: no spindle, 1: has spindle), and used this in a decision trees classifier, see methods for details. This showed that we can discriminate correctly classified and incorrectly classified trials, only in the experimental night and not the adaptation night (Wilcoxon signed rank test for experimental vs. adaptation, $n = 12$, $p = 0.04$, $z = 2.04$ and Wilcoxon signed rank test for experimental vs. chance, $n = 12$, $p = 0.015$, $z = 2.43$), Fig. 3b. Subsequently, we analysed the trials of each participant to determine whether it was the presence or absence of spindles that predicted the correct classification by the reactivation classifier. This showed that trials with fewer pre-cue spindles are more likely to have late reactivation (Supplementary Fig. 3). This is in keeping with a study which showed that significant post-cue reactivation was observed in trials with low pre-cue sigma power (Wang et al., 2019). Spindles have been shown to have a periodicity of about 4 s (Antony et al., 2018), thus, it is possible that the occurrence of pre-cue spindles which prevented post-cue spindles and reactivation in that study (Antony et al., 2018) also prevented late reactivation in our study. For early reactivation, there was a difference between classification performance and chance level (Wilcoxon signed rank test, $n = 12$, $p = 0.028$, $z = 2.2$), however, it was not significant compared to the adaptation night (Wilcoxon signed rank test, $n = 12$, $p = 0.14$, $z = 1.5$), generally, the sign of the difference was the same as

that of late reactivation showing that trials with fewer pre-cue spindles are more likely to have early reactivation. Overall, these results suggest that we can use features from spindles to predict when to deliver TMR in order to trigger a classifiable reactivation.

### 2.5. Characteristics of detected reactivations

Because this is a motor task, we wanted to know whether classification of reactivation was derived from the channels over the motor area. We therefore analysed the selected features that were included for classification after the feature selection step (see methods for details). This showed that the selected features always came from the motor area channels (C5, CP3, C6, CP4), with 66.7% of features being chosen from the right motor channels and 33.3% from left. This shows that the activity patterns in wake and sleep arise from the motor area and are related to the motor task.

Because sleep is characterised by relatively low frequencies such as SOs (0.5 – 1.5 Hz), delta waves (1.5 – 4 Hz), and theta (4 – 8 Hz), we hypothesised that these would be the most important for our classification. To investigate this, we applied a low pass filter with cut-off frequency of 10 Hz without smoothing the signals. The resulting classification pattern was similar to the result without this filter in Fig. 2a (early cluster, $p = 0.01$, and late cluster, $p = 0.03$). To determine if classification was driven by a narrower frequency range inside the < 10 Hz band, we employed a filter bank approach and analysed the classification in different frequency ranges. We thus analysed (4 – 8 Hz, 2 – 4 Hz), < 2 Hz, and < 4 Hz. Interestingly, we found a similar classification pattern to that obtained in Fig. 2a when examining just the low frequency ranges: < 2 Hz, and < 4 Hz, Supplementary Fig. 5. This supports the role of SOs in memory reactivation.

As shown in Fig. 2a, we found that reactivation could occur at either of the two different timepoints - either early after the onset of the cue, or approximately one second later. This demonstrates the temporal characteristics of reactivations within trial duration. We also wanted to examine the characteristics of reactivations occurring at these two different times across the time course of stimulation. Our prior work on this task suggested that classification performance decreases as the number of stimulations in a night increases (Belal et al., 2018). We tested whether correct trials preferably occur before or after the middle of the stimulation time. Thus, we indexed the correctly classified trials for early/late reactivation to range from 0 (first trial in stimulation) to 1 (last trial) for every participant. We then compared the indices to 0.5 (middle of stimulation) across participants. This revealed that reactivations could be detected to a similar extent at any time during stimulation and was not more prevalent at the beginning or end of stimulation time. Neither reactivations which occurred right after the TMR tone, nor reactivations which occurred ~1 s after the TMR tone, differed significantly from the middle of the stimulation time (Wilcoxon signed rank test, $n = 12$, $p = 0.39$, $z = 0.86$, and $p = 0.58$, $z = 0.55$ for early and late reactivations, respectively).

We were also interested in determining whether our method could detect temporally compressed reactivation given the rodent literature showing that reactivation in NREM is compressed in comparison to wake (Ji and Wilson, 2007; Lee and Wilson, 2002; Euston et al., 2007). To address this, we compressed the EEG signals of wake data that trained our classification models and then applied the sliding window approach to sleep to check for reactivation. We repeated this for different compression factors (length of wake / length of compressed version of wake). The results are shown in Supplementary Fig. 6. This analysis showed similar patterns of results occur for all compressions of wake trials as when we use the uncompressed data. Interestingly however, when tested against the adaptation night and corrected for multiple comparisons using cluster-based permutations these results are only significant for 5x compression. Notably, when wake trials are compressed and we look for a similar pattern in sleep, any temporal shifts between trials in the timing of reactivation will weaken the effect. This would happen because the length of the sliding window is now shorter compared to no compression. Thus, the significant effect in the compression factor of 5x might be due to the fact that this compression factor was closer to the non-compressed version relative to the other compression factors. This was confirmed by classification of 2.5x compression, which showed a pattern of significance very similar to the non-compressed data, Supplementary Fig. 6i. Given this, it will be difficult to draw firm conclusions about the exact compression of reactivation given the small temporal shifts that may happen in trials.

Finally, we wanted to examine how the performance of early and late reactivations varied across the night of stimulation. Thus, we obtained a performance curve across stimulation time for both early and late peaks by observing the changes of classification performance during the time of that peak throughout trials of stimulation (Supplementary Fig. 1). We used a 50-trial block to calculate classification performance, and slid this forward by one trial, to progress along the stimulation time. We then normalised the stimulation time to have the range (0 to 1), with 0 being the first stimulation in the night and 1 the last stimulation. Interestingly, classification performance between the two peaks differed around approximately 0.6, that is, at 60% of the way through stimulation time, with early reactivation more likely to occur at this time (Supplementary Fig. 1).

### 2.6. The relationship between behaviour and classification performance

Some prior reports have shown a positive relationship between detectable reactivation after TMR tones and the extent of TMR related behavioural benefit (Cairney et al., 2018; Schreiner et al., 2018; Bendor and Wilson, 2012). We searched for this relationship in our data, by testing for correlations between classification and behavioural performance. Because different trials classified correctly at early and late timepoints after the cue, and because such temporally distinct reactivation may potentially also have distinct functional characteristics, we performed correlations twice, using the classification rate at the early peak and then at the late peak. This revealed a negative correlation between pre-sleep reaction time for the reactivated sequence, and classification AUC for the early peak (Spearman $r = -0.60$, uncorrected $p = 0.04$), Fig. 4a. Notably, this did not survive correction for multiple comparisons (Bonferroni correction, $p = 0.16$). In other words, faster pre-sleep performance was associated with a higher classification rate immediately after the TMR cue. This could index a stronger representation that could reactivate more easily, or in a more classifiable form.

Interestingly, the late peak showed quite different associations from the early peak. Here, classification AUC negatively predicted the extent to which responses on the reactivated sequence sped up across the night of sleep (performance just before sleep – performance early post-sleep, Spearman $r = -0.72$, uncorrected $p = 0.01$), Fig. 4b. We refer to this as overnight improvement, however, this change is related to reaction time and not improvement in sequence learning, since our measure of sequence learning involves comparison to the random sequence and this measure does not. The AUC of the late peak also predicted slower reaction times for the non- reactivated sequence after sleep (Spearman $r = 0.68$, uncorrected $p = 0.02$), Fig. 4c. Thus, the stronger the late peak, the slower the non-reactivated sequence was performed immediately after sleep. These results could suggest that when reactivation occurs late (~1 s after the TMR cue), it somehow disrupts both spontaneous and cued consolidation of the task for both the non-reactivated and reactivated sequences. The idea that late reactivation could have this disruptive property is in-line with a study showing a negative correlation between reactivation and improvement (Murphy et al., 2018).

### 2.7. The relationship between lateralized sleep spindles and classification performance

Sleep spindles have been strongly linked with memory reactivation (Rasch and Born, 2013; Klinzing et al., 2019; Antony et al., 2019).

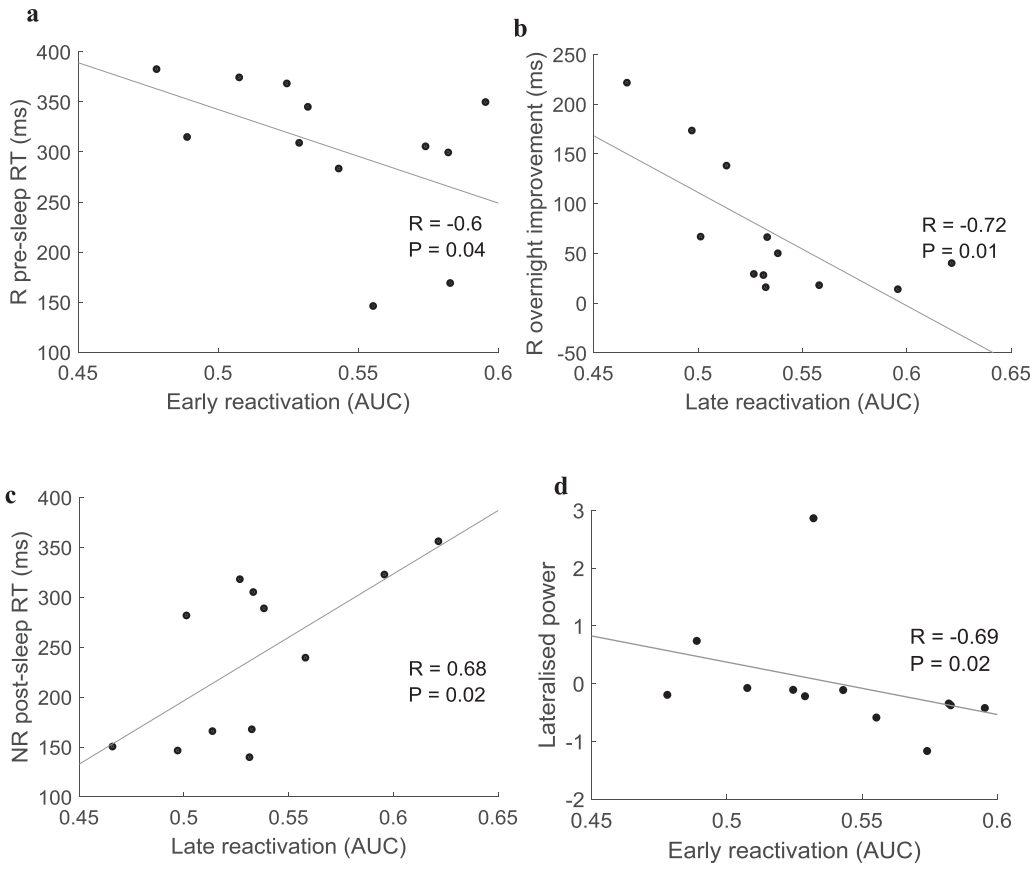



**Fig. 4.** Correlation of classification performance with behavioural results and sigma power. **(a)** Negative correlation between the classification performance (AUC) of the early peak and the average reaction time (RT) of the last four blocks before sleep for the reactivated (R) sequence (spearman correlation = −0.60, uncorrected $p = 0.04$). **(b)** The late peak correlated negatively with the overnight improvement of the reactivated sequence (spearman correlation = −0.72, uncorrected $p = 0.01$). **(c)** The late peak predicted slower reaction times after sleep for the non-reactivated (NR) sequence (spearman correlation = 0.68, uncorrected $p = 0.019$). **(d)** Correlation of lateralized sigma power (z-transformed) with classification performance for the early peak (Spearman correlation = −0.69, $p = 0.016$).

For instance, work in rodents shows that replays correlate with spindles (Peyrache et al., 2012). Furthermore, lateralized spindle density during cue periods has been shown to predict TMR-related benefit (Cousins et al., 2014). We tested for a relationship between sigma power at (11 to 16 Hz) and classification performance. We found that, even though participants used both hands in this task, the lateralized sigma power was negatively associated with the early classification peak, (Spearman $r = -0.69$, $p = 0.016$) as shown in Fig. 4d. Thus, the more the lateralized spindles right before the stimulus compared to after, the more likely we were to classify reactivation immediately after the TMR cue (more details about power calculation in methods). This is interesting in light of a prior analysis of our behavioural data showing TMR-related improvement in the weaker left, but not the stronger right hand over sleep (Koopman et al., 2020).

## 3. Discussion

This study shows that TMR cues are more likely to result in classifiable reactivation when applied during the up-going phase of the SOs. We also show that the pattern of ongoing SOs and spindles before a TMR cue can be used to predict whether each cue will produce a classifiable reactivation. Importantly, the resultant reactivations did not reoccur after the TMR cue, instead occurring either early or late. These findings markedly deepen our understanding of neural reactivations after TMR cues in sleep and may lead to improved methods for efficient boosting of memory via the TMR manipulation.

### 3.1. Timing of reactivation after the cue

The delay between TMR onset and triggered reactivation, is a matter of current interest. Rodent work showing a rapid reverberation of reactivation between cortex and hippocampus at the millisecond scale has led to the idea that replays may 'echo back' again and again after TMR (Rothschild et al., 2017). Other work in rodents (Bendor and Wilson, 2012) suggests that TMR cued replay can continue to repeat for up to 10 s after the offset of the auditory cue, but a second cue can interrupt this replay. In humans, one study showed evidence of reactivation about two seconds after TMR, with a suggestion of earlier reactivation immediately after the cue (Cairney et al., 2018). Another study showed recurrent reactivation after TMR, with the response occurring both immediately after the cue and about two seconds later (Schreiner et al., 2018). Our findings are in keeping with this work, since they suggest that reactivation can occur either immediately after the cue or around one second later. Because our inter-trial interval was only 1500 ms, it is possible that reactivations could occur even later, but the next TMR cue would likely have prevented this in our current design. Notably, although we saw no evidence of recurrent reactivation in the timescale of seconds, this does not rule out the idea of extremely rapid recurrence of reactivation at the millisecond timescale, as suggested in Rothschild et al. (2017). Rapid recurrence would be difficult to see in our current data given that we used EEG and there might be small temporal shifts between trials which yielded a smoothed mean effect. An early reactivation could be related to quick reactivation of the memory in some circumstances. For instance, the negative correlation which we observed between early re-

activation and pre-sleep reaction time (Fig. 4a) suggests that the brains of participants who were quick at the task before sleep also responded quickly to TMR cues with an early reactivation. On the other hand, late reactivation at ~1 s relates to spindles as shown in the analysis of the spindle-based classifiers (Fig. 3b). Furthermore, our analysis of this echo shows that reactivation occurs either late or early (but not both) in the vast majority of trials (~91%).

Since our data show that reactivations can be identified either early or late after the cue, we must ask whether such differences in timing are functionally important. Early reactivation was positively related to pre-sleep behavioural performance, while late reactivation was instead negatively related to the extent of overnight improvement. It is difficult to interpret these findings, but one possible explanation could be that a strongly encoded memory of the task leads to more immediate reactivation after a TMR cue while a weaker memory trace leads to late reactivation which might actually disrupt consolidation of the task if memories become distorted during the delay.

### 3.2. Optimal timing of TMR cues

The exact mechanisms by which TMR triggers reactivation are unknown, but the up-going phase of the SO is more reactive to stimulation than the down-going phase, after all, neurones are preparing to fire as the oscillation approaches its peak and beginning a silent period as it enters the trough. Auditory stimulation after the negative peak of the SO, during the up-going phase, has been shown to produce a higher amplitude ERP than stimulating during the down-going phase (Schabus et al., 2012). TMR can therefore also be expected to have different impact in the up- vs. down- going phase of the SO.

Cortical SOs, thalamo-cortical spindles and sharp wave ripples are thought to be key for memory consolidation, with the up-going SO driving spindle-ripple events with reactivation (Diekelmann and Born, 2010; Sirota and Buzsáki, 2005; Khodagholy et al., 2017). TMR to the up-going phase of the SO has been shown to improve memory (Göldi et al., 2019; Shimizu et al., 2018). This relationship between SO phase and reactivation is clear in the current work, as we show that stimulating the SO up-going phase maximises detectability of reactivation.

SOs are highly heterogeneous, differing both in locus of generation and shape. For instance, SOs differ in period, trough depth, and peak to trough slope (Siclari et al., 2014; Bernardi et al., 2018). Importantly, the SO down-state is thought to be required for the generation of a thalamic down-state which triggers a spindle (Mak-McCully et al., 2017). Given the association between memory reactivation and spindles, along with the fact that spindle initiation requires a sharp SO trough, it is reasonable to suppose that TMR stimulation of some SOs may be more likely to trigger reactivation than TMR stimulation of others. For instance, SOs with a deeper trough or steeper slope, or some combination of these might be more likely to carry reactivation-bearing spindles. Such differences could explain why we were able to predict which stimulations would be successful based on the features of the ongoing SO before the TMR cue, although, notably, the combination of features was necessary and no single SO feature was sufficient for this prediction.

We found that trials with fewer pre-cue spindles are more likely to have late reactivation (Supplementary Fig. 3). This is in good keeping with work from Wang and colleagues showing that lower pre-cue sigma power predicted more post-cue reactivation, and that such reactivation begins around one second after the onset of the cue (Wang et al., 2019). Such predictive analysis could potentially be used to boost the efficacy of TMR by ensuring that stimulation occurs only at the times when it is most likely to be effective. This could minimise any potential disturbance from TMR, which does often lead to arousals when delivered indiscriminately. Such increased precision of cue delivery could be important for translation of the TMR technique from lab to the home environment.

### 3.3. Control analyses

Our control for detecting memory reactivation was two-fold. Firstly, we encapsulated the identity of a memory using its EEG pattern during wake and used this as a guidance for detecting the reactivation of this memory in sleep after TMR cues. We achieved this by training our classification models on EEG from wake during wakeful encoding and testing them in sleep. This procedure ensures that classification strength is caused by the reinstatement of the same encoded memories and related to genuine re-processing of memories during sleep. Our work builds on paradigms showing the discriminability of cued categories in sleep data without the inclusion of wake (Cairney et al., 2018; Schönauer et al., 2017), as well as an approach that includes only the features that caused category discrimination in wake (Wang et al., 2019). Secondly, we use an adaptation night to ensure that our classification results from the experimental night are not caused by sound induced noise in EEG. Notably, we excluded all EEG segments that showed signs of arousals. Furthermore, stimulation late in the night was just as effective at eliciting reactivation as stimulation early in the night as shown in the analysis where we examined classification rates across stimulation time (Supplementary Fig. 1). This suggests that habituation to the tone across multiple stimulations did not reduce the extent to which the tone could elicit reactivation.

### 3.4. Considerations for future studies

The current classification pipeline is suitable for our current dataset with this particular task in mind and the aim of investigating memory reactivation. This pipeline is unfortunately not suitable for all sleep data in general. In the current work, we looked at the prediction of reactivation using SO and spindle features, future studies could also look at the coupling between SOs and spindles given the importance of this coupling in carrying reactivation, e.g., Diekelmann and Born (2010).

### 3.5. Summary

In this study, we show that reactivations can occur either early or late after a sound cue, and appear to have different functional significance depending on this timing. We also show that TMR delivered during the SO up-going transition is more likely to be associated with classifiable reactivation, probably because it heralds the spindle-bearing upstate. Finally, we show that both pre-cue SO morphology and spindle incidence can be used to predict TMR cued reactivation. This method will allow more efficient TMR stimulation, paving the way for the development of wearable devices to effectively stimulate memory consolidation in the home environment.

## 4. Methods

### 4.1. Participants

The current study uses EEG from human participants ($n = 15$), mean age: 23.4 years, 7 females. Participants spent an adaptation night in the lab, then in the experimental night, they completed a SRTT before and after sleep. All participants were right-handed and none of them reported familiarity with the SRTT. All participants had normal or corrected-to-normal vision, normal hearing, and no history of physical, psychological, or neurological disorders. Participants were healthy and had no formally diagnosed sleep disorders. They were asked to answer a shortened version of the Movement Imagery Questionnaire-3 (MIQ-3) to assess their ability to use internal visual and kinaesthetic imagery (Williams et al., 2012) . They were also asked to answer a shortened version of the Edinburgh Handedness Inventory (EHI) to assess their handedness (Veale, 2014). Their alertness was assessed using the Karolinska Sleepiness Scale (KSS) (Åkerstedt and Gillberg, 1990) and the Stanford Sleepiness Scale (SSS) (Hoddes et al., 1973) before they went to bed.

Their responses in a pre-screening questionnaire reported no stressful events and no travel before commencing the study. Participants did not consume alcohol in the 12 h before the study and caffeine in the 24 h prior to the study or perform any extreme physical exercise or nap. This study was approved by the School of Psychology, Cardiff University Research Ethics Committee, and all participants gave written informed consents (Koopman et al., 2020).

### 4.2. Experimental design

Participants were asked to sleep in the lab before doing the SRTT training. During this adaptation night, sounds were played with the same criteria as the actual experiment, but importantly they had not yet been associated with any task. During SWS, TMR cues of the reactivated sequence (12-items) were presented to participants with a 20 s pause between the presentation of each sequence.

Participants performed a SRTT (adapted from (Cousins et al., 2014)). Sounds cued four different finger presses. We delivered the sound cues during SWS. Participants learned two 12-item sequences, A and B, A: 1 2 1 4 2 3 4 1 3 2 4 3 and B: 2 4 3 2 3 1 4 2 3 1 4 1. The location indicated which key on the keyboard needed to be pressed as quickly and accurately as possible: 1 – top left corner = left shift key; 2 – bottom left corner = left Ctrl; 3 – top right corner = up arrow; 4 – bottom right corner = down arrow. Sequences had been matched for learning difficulty; both contained each item three times. The blocks were interleaved so that a block of the same sequence was presented no more than twice in a row, and each block contained three repetitions of a sequence. There were 24 blocks of each sequence (48 blocks in total), and each block was followed by a pause of 15 s wherein feedback on reaction time (RT) and error-rate were presented. The pause could be extended by the participants if they wanted. After the 48 blocks of sequences A and B, participants performed four more blocks that contained random sequences. They contained the same visual stimuli and an 'R' displayed in the centre of the screen. Two of these blocks were paired with the tone group of one sequence (reactivated in sleep), and the other two were paired with the tone group of the other sequence (non-reactivated). Participants were aware that there were two twelve-item sequences, and each sequence was indicated with 'A' or 'B' appearing in the centre of the screen, but they were not asked to learn the sequences explicitly. Counterbalancing across participants determined whether sequence A or B was the first block, and which of the sequences was reactivated during sleep.

Each sequence was paired with a group of pure musical tones, either low tones within the 4th octave (C/D/E/F) or high tones within the 5th octave (A/B/C#/D). These tone groups were counterbalanced across sequences. For each trial, a 200 ms tone was played, and at the same time a visual cue appeared in one of the corners of the screen. Participants were instructed to keep individual fingers of their left and right hand on the left and right response keys, respectively. Visual cues were neutral objects or faces, (Supplementary Fig. 2), used in a previous study (Cousins et al., 2014), which appeared in the same position for each sequence (1 = male face, 2 = lamp, 3 = female face, 4 = water tap). Participants were told that the nature of the cues (objects/faces) was irrelevant. Visual cues stayed on the screen until the correct key was pressed, after which an 880 ms inter-trial interval followed.

After completion of the SRTT, participants were asked to do the same task again, but were instructed to only imagine pressing the buttons. Imagery took place after the SRTT since participants had to be trained on the task before they could meaningfully perform the imagery task. This imagery task consisted of 30 interleaved blocks (15 of each sequence), presented in the same order as during the SRTT. Again, each trial consisted of a 200 ms tone and a visual stimulus, the latter being shown for 270 ms and followed by an 880 ms inter-trial interval. There were no random blocks during the imagery task and no performance feedback was presented during the pause between blocks.

In the morning following the experimental night, participants were asked to perform the task again, motor imagery first, then SRTT. Eventually, they were asked if they remembered the locations of images of the two sequences, to see if one sequence is recalled better than the other one. Motor imagery data set of each participant was used for classification. The adaptation night was useful for eliminating the possibility that a classifier could merely classify sound induced effects on the EEG. Thus, if the classifier can classify the experimental night but not the adaptation night this suggests that it is classifying memory reactivations.

Details regarding the number of stimulations in different conditions and which sequence was reactivated are provided in Supplementary Table 2.

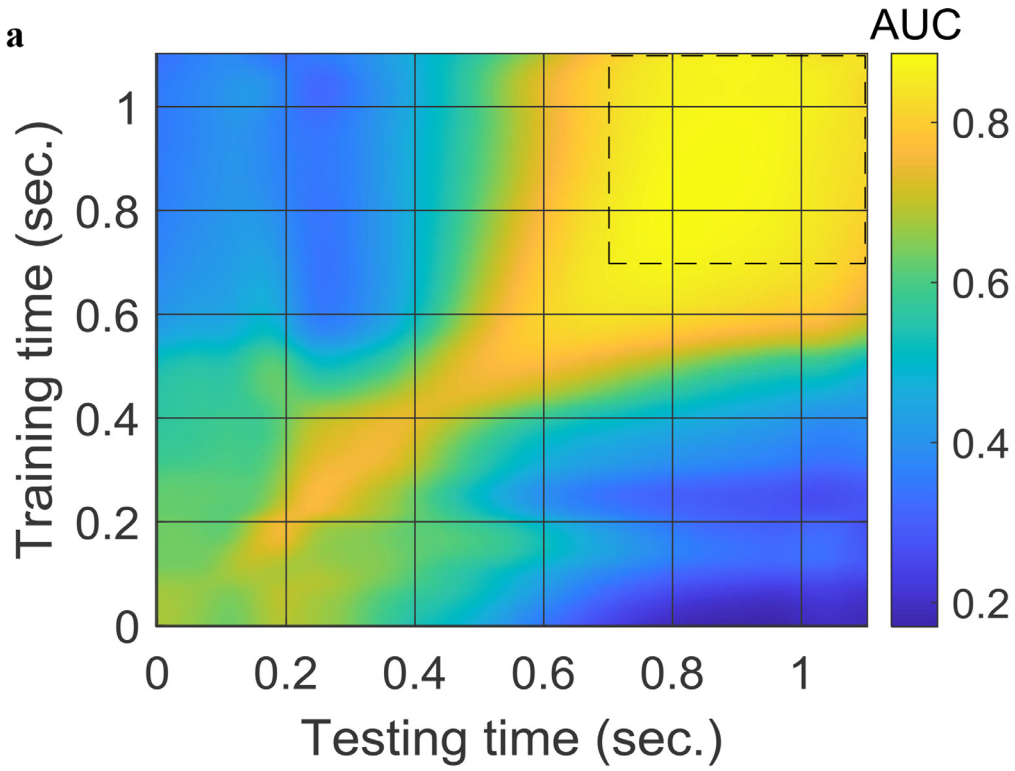
### 4.3. Recording and pre-processing

Data were extracted using 21 electrodes, following the 10–20 EEG system. 13 of the electrodes were placed on standard locations, namely: FZ, CZ, PZ, F3, F4, C5, CP3, C6, CP4, P7, P8, O1, and O2, and they were referenced to the mean of the left and right mastoid electrodes. Other electrodes were the left and right EOG, three EMG electrodes, which were placed on the chin, and the ground electrode placed on the forehead. The impedance was <5 kΩ for each scalp electrode, and <10 kΩ for each face electrode. Recordings were made with an Embla N7000 amplifier and RemLogic 1.1 PSG Software (Natus Medical Incorporated). PSG recordings were scored by two trained sleep scorers and only the parts scored as SWS were kept for further analyses. Data were collected at a 200 Hz sampling rate.

EEG signals were band pass filtered in the frequency range from (0.1 to 30 Hz). Subsequently, trials were cleaned based on statistical measures consisting of variance and mean. Trials were segmented from −0.5 to 3 s relative to the onset of the cue. Trials falling two standard deviations higher than the mean were considered outliers and rejected if they were categorised as outliers in more than 25% of channels. If trials were categorised as outliers in less than 25% of the channels, they were interpolated using triangulation of neighbouring channels. Thus, 11.7% and 11.9% of trials were considered outliers and removed from the experimental night data and the adaptation night, respectively. Analyses were done using FieldTrip (Oostenveld et al., 2011) and Matlab 2018a.

### 4.4. Wake-to-wake classification and determining a time of high discrimination

We started the analysis by performing a wake-to-wake motor imagery classification. This was performed for each participant separately, with trials serving as observations and they were labelled according to the hand they belonged to. EEG signals were pre-processed, and features were extracted by calculating time-domain amplitude averages of 80 ms (40 ms before and 40 ms after every time point). Subsequently, features were fed to a linear discriminant analysis (LDA) classifier (Blankertz et al., 2011). Training and testing were done on wake data in a time x time fashion (King and Dehaene, 2014). The classifier was trained on a specific time point and tested with all time points using a 5-fold cross-validation to build one row in the time x time classification plot, illustrated in Fig. 5a. We assumed that if classification performance is not at a considerably high rate during wake then this would decrease the possibility of classifying sleep reactivation where noise is higher. Consequently, we chose participants that had wake-to-wake classification performance with area under the ROC curve (AUC) >= 0.7, ($n$ = 13). One participant was neglected because of a technical problem during the collection of sleep data. The rest of the data was used for classification ($n$ = 12). We also do realise the rich literature of motor imagery classification with common spatial patterns (CSP) and other methods (Blankertz et al., 2008; Lemm et al., 2005; Pfurtscheller et al., 1997; Pfurtscheller et al., 2006; Ramoser et al., 2000). However, given the differences between wake and sleep data sets and their different nature of noise and oscillations we decided to directly use time domain features with our classifiers.



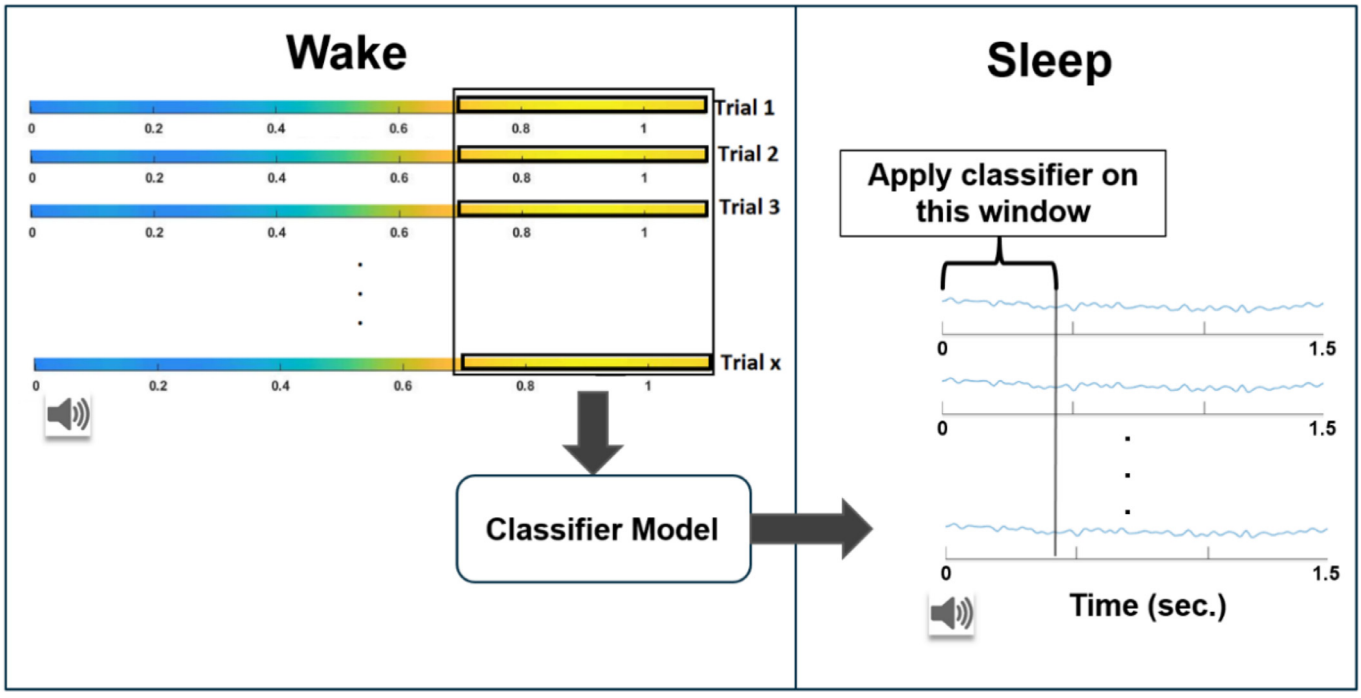




**Fig. 5.** Classification of left- vs. right-hand. **(a)** Grand average AUC values for left- vs. right-hand motor imagery classification using 80 ms smoothing window and LDA classifiers, dashed box represents the time of interest (TOI). **(b)** Illustration of classification procedure (training with wake and testing with sleep) which was applied for both the experimental night and adaptation night. A sliding window approach was used, wherein a classifier was tested on a window from sleep and the classification result replaced the centre of that window then the window was slid by one time point to construct a performance curve across trial time in sleep.

Initial investigations revealed a higher classification performance for left- vs right-hand (where both fingers were aggregated into one class) than for faces vs objects. Therefore, we conducted the analysis on right- vs left-hand imagery. The trial length was defined as the duration between cue onsets (1.1 s in wake). Sound cues had a duration of 200 ms and were played from time 0 of trials. During sleep, trial length was 1500 ms.

Classification of motor imagery during wake showed a time period with maximum classification performance (marked with dashed box in Fig. 5a). This time region should be useful for discriminating left hand and right hand. We defined this time period as the time of interest (TOI). A TOI is a time window that has a high classification rate, indicating its ability to discriminate between classes. It acts as a temporal marker of expected discrimination. To locate this window, we used a threshold of 0.85 on the average classification AUC from all participants.

### 4.5. Wake-sleep classification

Once we had built a classifier on wake data, we tested it on sleep data. We applied it on sleep using a sliding window approach, as shown in Fig. 5b. We employed a sliding window approach, classification was applied on the first testing window in sleep, for example: [0 to 0.38] s, which matched the length of the TOI. Then, classification performance was placed at the centre time of this window, i.e., at 0.190 s. Subsequently, the sliding sleep window was shifted by one time point, and the process was repeated. Thus, the results of classification consisted of AUC values across time.

The wake-to-sleep classifiers used the concatenated averages inside the TOI as features. These concatenated time points were reduced to the most informative contiguous time points using mutual information on wake data for each participant. We reduced the features to the most informative time points since the reactivation might be temporally short compared to wake activation. Consequently, we slide a shorter window that contains the most informative features which enables the classifier to detect the reactivation if it was temporally short or long. The most informative time points were chosen such that the time points were contiguous and contained the highest 10% of the mutual information values.

We devised a method for removing noisy trials with no TMR effect. We assume noisy trials belong to a new 'no effect' class which does not contain discriminative features for right- vs left-hand. The feature values of those trials should fall near the decision boundary in the feature space, in a region where the classifier is uncertain (corresponding to a maximum posterior probability close to 0.5). Thus, we define trials as 'no effect' if they fall in that area. We rejected noisy trials falling in the area of uncertainty and used three hundred clean trials from every participant. Importantly, to avoid any bias, this cleaning process was unsupervised, meaning that the information of the ground truth class labels of sleep data was not used. Moreover, we verified that this cleaning process would not improve classification performance if the data we were trying to clean was random and contained no discriminative information, which was the case with the adaptation night. It would also not improve classification performance if sleep data was not scattered in the feature space in a similar way to wake samples because the decision boundary position and orientation which are determined using wake will then be meaningless for sleep samples. Thus, this cleaning process only works if sleep data contains discriminative information. Importantly, the exact same cleaning procedure was performed for both the experimental and adaptation nights for completeness.

Classification was assessed using the area under the ROC curve (AUC). Because each point in the AUC represents a different certainty value, and measures the true positive and false positive rates at that certainty value, this provided an indication of the performance of classification at each level of certainty. The average certainty of classification was represented by the posterior probabilities, this gave 0.86 with 0.1 standard deviation.

### 4.6. Preferred TMR phase analysis

Phase information was extracted using Hilbert transform on the band pass filtered signal (0.5 to 2 Hz) using FZ electrode. We divided phase values into two ranges: [0 to $\pi$] and ($\pi$ to $2\pi$], indicating the two transitions: down-going and up-going, respectively. For each participant, we determined the number of correctly classified trials in which TMR fell on either phase range in each night, then divided this by the number of trials on the up-going and down-going transitions. Trials were deemed 'correct' when the prediction of the classifier was the same as the respective trial category. The same process was repeated for the adaptation night.

### 4.7. Lateralized sleep sigma power analysis

The lateralized sigma power [11 16] Hz was calculated using the short time Fourier transform during the duration: [0 to 0.5] *sec*. relative to cue onset which was around the early reactivation. Lateralized power was calculated as the difference between left and right motor channels (C6, CP4, C5, CP3) and was baseline corrected ([−0.2 0] s. relative to cue onset). Consequently, percentage change from baseline was calculated.

### 4.8. Reoccurrence of reactivation

We statistically tested if one reactivation (early or late) was more likely to happen or whether reactivation was reoccurring after a sound cue. Thus, we took the accuracy for recurring reactivation (i.e., the ratio of correct trials during the time of both early and late reactivation simultaneously) and compared it to the probability of both reactivations happening simultaneously after a sound cue (the accuracy for early reactivation multiplied by the accuracy for late reactivation) as a chance level. We performed this analysis for every participant and compared the accuracy of reoccurring reactivation to chance level.

### 4.9. SO-based classification

The SO-based classification consisted of 200 decision trees ensemble. Leave one out classification was used, wherein data from all participants except one was used to train the classifier and the left-out participant was used for testing the classifier. This gave a classification result for the left-out participant and the process was then repeated until the classification performance was calculated for all participants. Every decision tree was trained on a random subset of trials from the training set and tested on the testing set and the final result was the aggregated votes from all decision trees. The extraction of SO and spindle events followed the implementation in (Navarrete et al., 2020). Briefly, for spindles, signals were filtered using a two-pass bandpass FIR filter between 11 and 17 Hz, subsequently, the root mean square (RMS) power was calculated using a window of 200 ms. A threshold on power of 86.639 (equivalent to 1.5 standard deviation from the mean of a normal distribution) was applied and spindles were defined as segments that exceeded that threshold and had duration $> 300$ ms and $< 3$ s. For SOs, two-pass FIR bandpass filtering was applied to the signals between 0.3 and 3 Hz. SOs were then extracted as segments with negative deflections that had consecutive zero crossings between 0.25 and 1 s. Non-SOs that had $< 75$ microV were excluded before classification.

### 4.10. Statistical testing

To assess the statistical significance of the classification results, we compared the classification performance of the experimental night against the adaptation night. Sounds played during the adaptation night were the same sounds used in the experimental night but because the adaptation night was before participants had been trained on the experimental task, these sounds were not yet associated with any memories.

This control was used to make sure that classification was not derived due to sound induced features/noise in EEG.

Statistical analysis was performed using the classification results from the two nights with cluster based permutation using FieldTrip (Oostenveld et al., 2011). Monte Carlo was used with a sample-specific test statistic threshold of 0.05, permutation test threshold for clusters of 0.05, and 10,000 permutations. The correction window used in the test was the whole length of sleep trial, i.e., [0 to 1.5] s.

## Data and code availability statement

All relevant data generated or analysed are available at the Open Science Framework (OSF) upon request, including the EEG data, behavioural files, and all MATLAB scripts for classification and other analyses.

## Ethics statement

This study was approved by the School of Psychology, Cardiff University Research Ethics Committee, and all participants gave written informed consents. Participants private identifications are all anonymized.

## Declaration of Competing Interest

The authors declare no competing interests.

## Credit authorship contribution statement

**Mahmoud E.A. Abdellahi:** Formal analysis, Writing – original draft. **Anne C.M. Koopman:** Visualization, Methodology, Data curation, Writing – original draft. **Matthias S. Treder:** Writing – original draft. **Penelope A. Lewis:** Visualization, Methodology, Writing – original draft.

## Data availability

I have shared a link to data and code.

## Acknowledgments

This work was supported by the ERC grant SolutionSleep to P.A.L and ERC funded the Ph.D. of M.E.A.A. We would like to thank Miguel Navarrete for his help in the extraction of slow oscillation features for the slow oscillation-based classification and also Lorena Santamaria, Martyna Rakowska and members of our group NaPs for the useful advice.

## Funding

This work was funded by the ERC grant SolutionSleep, 681607, to P.A.L.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2022.119820.

## References

Åkerstedt, T., Gillberg, M., 1990. Subjective and objective sleepiness in the active individual. Int. J. Neurosci. doi:10.3109/00207459008994241.
Antony, J.W., et al., 2018. Sleep spindle refractoriness segregates periods of memory reactivation. Curr. Biol. doi: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 Neurosci. doi:10.1016/j.tins.2018.09.012.
Belal, S., et al., 2018. Identification of memory reactivation during sleep by EEG classification. Neuroimage 176, 203–214.
Bendor, D., Wilson, M.A., 2012. Biasing the content of hippocampal replay during sleep. Nat.Neurosci. 15, 1439–1444.
Bernardi, G., Siclari, F., Handjaras, G., Riedner, B.A., Tononi, G., 2018. Local and widespread slow waves in stable NREM sleep: evidence for distinct regulation mechanisms. Front. Hum. Neurosci. doi:10.3389/fnhum.2018.00248.
Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., Müller, K.R., 2008. Optimizing spatial filters for robust EEG single-trial analysis. IEEE Signal Process. Mag. doi:10.1109/MSP.2008.4408441.
Blankertz, B., Lemm, S., Treder, M., Haufe, S., Müller, K.R., 2011. Single-trial analysis and classification of ERP components - a tutorial. Neuroimage doi:10.1016/j.neuroimage.2010.06.048.
Born, J., Wilhelm, I., 2012. System consolidation of memory during sleep. Psychol. Res. doi:10.1007/s00426-011-0335-6.
Cairney, S.A., Guttesen, A.Á.V., El Marj, N., Staresina, B.P., 2018. Memory consolidation is linked to spindle-mediated information processing during sleep. Curr. Biol. 28, 948–954 e4.
Cellini, N., Cappuzo, A., 2018. Shaping memory consolidation via targeted memory. Ann. N. Y. Acad. Sci. 1426, 52–71.
Clemens, Z., et al., 2007. Temporal coupling of parahippocampal ripples, sleep spindles and slow oscillations in humans. Brain doi:10.1093/brain/awm146.
Cousins, J.N., El-Deredy, W., Parkes, L.M., Hennies, N., Lewis, P.A., 2014. Cued memory reactivation during slow-wave sleep promotes explicit knowledge of a motor sequence. J. Neurosci. 34, 15870–15876.
Cousins, J.N., El-Deredy, W., Parkes, L.M., Hennies, N., Lewis, P.A, 2016. Cued reactivation of motor learning during sleep leads to overnight changes in functional brain activity and connectivity. PLoS Biol. 14, e1002451.
Diekelmann, S., Born, J., 2010. The memory function of sleep. Nat. Rev. Neurosci. doi:10.1038/nrn2762.
Euston, D.R., Tatsuno, M., McNaughton, B.L., 2007. Fast-forward playback of recent memory sequences in prefrontal cortex during sleep. Science doi:10.1126/science.1148979, (80-.).
Göldi, M., Poppel, E.Van, Rasch, B., Schreiner, T, 2017. Cueing memory during sleep is optimal during slow-oscillatory up-states. bioRxiv 185264. doi:10.1101/185264.
Göldi, M., van Poppel, E.A.M., Rasch, B., Schreiner, T., 2019. Increased neuronal signatures of targeted memory reactivation during slow-wave up states. Sci. Rep. doi:10.1038/s41598-019-39178-2.
Gordon, A.D., Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and regression trees. Biometrics doi:10.2307/2530946.
Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., Dement, W.C., 1973. Quantification of sleepiness: a new approach. Psychophysiology doi:10.1111/j.1469-8986.1973.tb00801.x.
Hu, X., Cheng, L.Y., Chiu, M.H., Paller, K.A., 2020. A meta-analysis of targeted memory reactivation. Psychol. Bull. 146 (3), 218–244.
Inostroza, M., Born, J., 2013. Sleep for preserving and transforming episodic memory. Annu. Rev. Neurosci. 36, 79–102.
Ji, D., Wilson, M.A., 2007. Coordinated memory replay in the visual cortex and hippocampus during sleep. Nat. Neurosci. 10, 100–107.
Khodagholy, D., Gelinas, J.N., Buzsáki, G., 2017. Learning-enhanced coupling between ripple oscillations in association cortices and hippocampus. Science doi:10.1126/science.aan6203, (80-.).
King, J.R., Dehaene, S., 2014. Characterizing the dynamics of mental representations: the temporal generalization method. Trends Cogn. Sci. 18, 203–210.
Klinzing, J.G., Niethard, N., Born, J., 2019. Mechanisms of systems memory consolidation during sleep. Nat. Neurosci. doi:10.1038/s41593-019-0467-3.
Koopman, A.C.M., et al., 2020. Targeted memory reactivation of a serial reaction time task in SWS, but not REM, preferentially benefits the non-dominant hand. bioRxiv doi:10.1101/2020.11.17.381913, 2020.11.17.381913.
Kudrimoti, H.S., Barnes, C.A., McNaughton, B.L., 1999. Reactivation of hippocampal cell assemblies: effects of behavioral state, experience, and EEG dynamics. J. Neurosci. 19, 4090–4101.
Lee, A.K., Wilson, M.A., 2002. Memory of sequential experience in the hippocampus during slow wave sleep. Neuron 36, 1183–1194.
Lemm, S., Blankertz, B., Curio, G., Müller, K.R., 2005. Spatio-spectral filters for improving the classification of single trial EEG. IEEE Trans. Biomed. Eng. doi:10.1109/TBME.2005.851521.
Lewis, P.A., Bendor, D., 2019. How targeted memory reactivation promotes the selective strengthening of memories in sleep. Curr. Biol. 29, R906–R912.
Mölle, M., Marshall, L., Gais, S., Born, J., 2002. Grouping of spindle activity during slow oscillations in human non-rapid eye movement sleep. J. Neurosci. doi:10.1523/jneurosci.22-24-10941.2002.
Mak-McCully, R.A., et al., 2017. Coordination of cortical and thalamic activity during non-REM sleep in humans. Nat. Commun. 8, 15499.
Murphy, M., Stickgold, R., Parr, M.E., Callahan, C., Wamsley, E.J., 2018. Recurrence of task-related electroencephalographic activity during post-training quiet rest and sleep. Sci. Rep. 8, 1–10.
Nakashiba, T., Buhl, D.L., McHugh, T.J., Tonegawa, S., 2009. Hippocampal CA3 output is crucial for ripple-associated reactivation and consolidation of memory. Neuron doi:10.1016/j.neuron.2009.05.013.
Navarrete, M., et al., 2020. Examining the optimal timing for closed-loop auditory stimulation of slow-wave sleep in young and older adults. Sleep doi:10.1093/sleep/zsz315.
Ngo, H.V.V., Martinetz, T., Born, J., Mölle, M., 2013. Auditory closed-loop stimulation of the sleep slow oscillation enhances memory. Neuron doi:10.1016/j.neuron.2013.03.006.
Nishida, M., Walker, M.P., 2007. Daytime naps, motor memory consolidation and regionally specific sleep spindles. PLoS One doi:10.1371/journal.pone.0000341.
O'Neill, J., Senior, T.J., Allen, K., Huxter, J.R., Csicsvari, J., 2008. Reactivation of experience-dependent cell assembly patterns in the hippocampus. Nat. Neurosci. doi:10.1038/nn2037.

Ólafsdóttir, H.F., Bush, D., Barry, C., 2018. The role of hippocampal replay in memory and planning. Curr. Biol. doi:10.1016/j.cub.2017.10.073.
Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. doi: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. J. Neurosci. doi:10.1523/JNEUROSCI.3537-16.2017.
Peyrache, A., Dehghani, N., Eskandar, E.N., Madsen, J.R., Anderson, W.S, 2012. Spatiotemporal dynamics of neocortical excitation and inhibition during human sleep. Proc. Natl. Acad. Sci. 109, 1731–1736.
Pfurtscheller, G., Neuper, C., Flotzinger, D., Pregenzer, M., 1997. EEG-based discrimination between imagination of right and left hand movement. Electroencephalogr. Clin. Neurophysiol. doi:10.1016/S0013-4694(97)00080-1.
Pfurtscheller, G., Brunner, C., Schlögl, A., Lopes da Silva, F.H., 2006. Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. Neuroimage doi:10.1016/j.neuroimage.2005.12.003.
Piantoni, G., et al., 2013. Modulation of gamma and spindle-range power by slow oscillations in scalp sleep EEG of children. Int. J. Psychophysiol. doi:10.1016/j.ijpsycho.2013.01.017.
Ramoser, H., Müller-Gerking, J., Pfurtscheller, G., 2000. Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Trans. Rehabil. Eng. doi:10.1109/86.895946.
Rasch, B., Born, J., 2013. About sleep's role in memory. Physiol. Rev. 93, 681–766.
Rothschild, G., Eban, E., Frank, L.M., 2017. A cortical-hippocampal-cortical loop of information processing during memory consolidation. Nat. Neurosci. doi:10.1038/nn.4457.
Schönauer, M., Geisler, T., Gais, S., 2014. Strengthening procedural memories by reactivation in sleep. J. Cogn. Neurosci. 26, 143–153.
Schönauer, M., et al., 2017. Decoding material-specific memory reprocessing during sleep in humans. Nat. Commun. doi:10.1038/ncomms15404.
Schabus, M., et al., 2012. The fate of incoming stimuli during NREM sleep is determined by spindles and the phase of the slow oscillation. Front. Neurol. doi:10.3389/fneur.2012.00040.
Schreiner, T., Lehmann, M., Rasch, B., 2015. Auditory feedback blocks memory benefits of cueing during sleep. Nat. Commun. doi:10.1038/ncomms9729.
Schreiner, T., Doeller, C.F., Jensen, O., Rasch, B., Staudigl, T., 2018. Theta phase-coordinated memory reactivation reoccurs in a slow-oscillatory rhythm during nrem sleep. Cell Rep. 25, 296–301.
Shanahan, L.K., Gjorgieva, E., Paller, K.A., Kahnt, T., Gottfried, J.A., 2018. Odor-evoked category reactivation in human ventromedial prefrontal cortex during sleep promotes memory consolidation. Elife 7, 1–21.
Shimizu, R.E., et al., 2018. Closed-loop targeted memory reactivation during sleep improves spatial navigation. Front. Hum. Neurosci. doi:10.3389/fnhum.2018.00028.
Siclari, F., et al., 2014. Two distinct synchronization processes in the transition to sleep: a high-density electroencephalographic study. Sleep doi:10.5665/sleep.4070.
Sirota, A., Buzsáki, G., 2005. Interaction between neocortical and hippocampal networks via slow oscillations. Thalamus Relat. Syst. doi:10.1017/S1472928807000258.
Squire, L.R., Genzel, L., Wixted, J.T., Morris, R.G., 2015. Memory consolidation. Cold Spring Harb. Perspect. Biol. doi:10.1101/cshperspect.a021766.
Valderrama, M., et al., 2012. Human gamma oscillations during slow wave sleep. PLoS One doi:10.1371/journal.pone.0033477.
van Dongen, E.V., Takashima, A., Barth, M., Fernandez, G., 2011. Functional connectivity during light sleep is correlated with memory performance for face-location associations. Neuroimage 57, 262–270.
Van Quyen, M.L., et al., 2010. Large-scale microelectrode recordings of high-frequency gamma oscillations in human cortex during sleep. J. Neurosci. doi:10.1523/JNEUROSCI.5049-09.2010.
Veale, J.F., 2014. Edinburgh handedness inventory - short form: a revised version based on confirmatory factor analysis. Laterality doi:10.1080/1357650X.2013.783045.
Wang, B., et al., 2019. Targeted memory reactivation during sleep elicits neural signals related to learning content. J. Neurosci. 39, 6728–6736.
Williams, S.E., et al., 2012. Further validation and development of the movement imagery questionnaire. J. Sport Exerc. Psychol. doi:10.1123/jsep.34.5.621.
Zhang, H., Fell, J., Axmacher, N., 2018. Electrophysiological mechanisms of human memory consolidation. Nat. Commun. 9, 4103.