Does slow oscillation-spindle coupling contribute to sleep-dependent memory consolidation? A Bayesian meta-analysis

Thea Ng[1,2], Eunsol Noh[3], and Rebecca M. C. Spencer[3,4,5†]

Abstract

The active system consolidation theory suggests that information transfer between the hippocampus and cortex during sleep underlies memory consolidation. Neural oscillations during sleep, including the temporal coupling between slow oscillations (SO) and sleep spindles (SP), may play a mechanistic role in memory consolidation. However, differences in analytical approaches and the presence of physiological and behavioral moderators have led to inconsistent conclusions. This meta-analysis, comprising 23 studies and 297 effect sizes, focused on four standard phase-amplitude coupling measures including coupling phase, strength, percentage, and SP amplitude, and their relationship with memory retention. We developed a standardized approach to incorporate non-normal circular-linear correlations. We found strong evidence supporting that precise and strong SO-fast SP coupling in the frontal lobe predicts memory consolidation. The strength of this association is modulated by memory type, aging, and dynamic spatio-temporal features, including SP frequency and cortical topography. In conclusion, SO-fast SP coupling should be considered as a general physiological mechanism for memory consolidation.

1 Main

Over the past three decades, accumulating evidence supports the role of sleep neural oscillations and their cross-frequency coupling in the spatiotemporal coordination across brain regions, supporting sleep-dependent memory consolidation[1–3]. During nREM sleep, oscillatory activities including cortical slow oscillations (SO; 0.16-4 Hz, Figure 1A), thalamocortical sleep spindles (SP; 8-16 Hz), and hippocampal sharp-wave ripples (SWR; 80-300 Hz), along with their long-distance coordination, are widely believed to be closely associated with the process of transferring temporary encoded memory traces to cortical networks for consolidation[4–6].

Given that non-invasive recordings (e.g., EEG) cannot accurately detect SWRs in deep brain structures, electrophysiological studies in humans primarily focus on the role of SOs and SPs in memory consolidation. Consistent evidence indicates that, following intensive learning, SP density and its co-occurrence with SOs significantly increase compared to baseline nights and control groups[7–11]. Correspondingly, measures of SPs and their coupling with SOs predict over-sleep retention performance of newly acquired memories[12–16].

The role of sleep oscillations in memory consolidation is supported by consistent neurobiological foundations. SOs are generated and dominant in the prefrontal cortex during slow-wave sleep (SWS) and propagate anteriorly to posteriorly as traveling waves[17–20]. The corticothalamic input of SOs during their depolarization up-state phase organizes the synchronous occurrence of SPs in the thalamic reticular nucleus. Subsequently, SPs propagate back widely to cortical areas through synchronized thalamocortical projections, resulting in EEG-measured SPs[21–26]. The peak discharge of SPs in the cortex is associated with increased dendritic Ca^{2+} synchronization[20,27,28]. This precise association enhances synaptic plasticity, leading to long-term changes in synaptic connections between cortical neurons, a mechanism for memory consolidation[29–32].

Despite extensive research on the functions of these oscillations during sleep and their causal sequences, it is unclear how the spatiotemporal coordination mechanisms facilitate specific temporal sequences of the peak discharge and the associated memory consolidation. Neuronal activity exhibits various forms of cross-frequency coupling, mainly including phase-phase coupling (PPC) and phase-amplitude coupling (PAC). The coupling is considered crucial in regulating the integration of information across multiple spatial and temporal scales[33,34]. Studies related to memory often focus on the unidirectional PAC regarding the strong amplitude modulation of faster oscillations (FO) driven by the phase modulation of relatively slower oscillations in a hierarchical order. The phase of SOs influences rhythmic spikes of FOs near the up-state peak or down-state trough of SOs[35–37]. In memory networks, PAC is considered a crucial mechanism supporting neural communication and plasticity[36].

The active system consolidation theory supports the subdivi-

*[1]Neuroscience & Behavior Program, Mount Holyoke College. [2]Department of Mathematics & Statistics, Mount Holyoke College. [3]Neuroscience & Behavior Program, University of Massachusetts, Amherst. [4]Department of Psychological & Brain Sciences, University of Massachusetts, Amherst. [5]Institute of Applied Life Sciences, University of Massachusetts, Amherst. † Corresponding author: rspencer@umass.edu

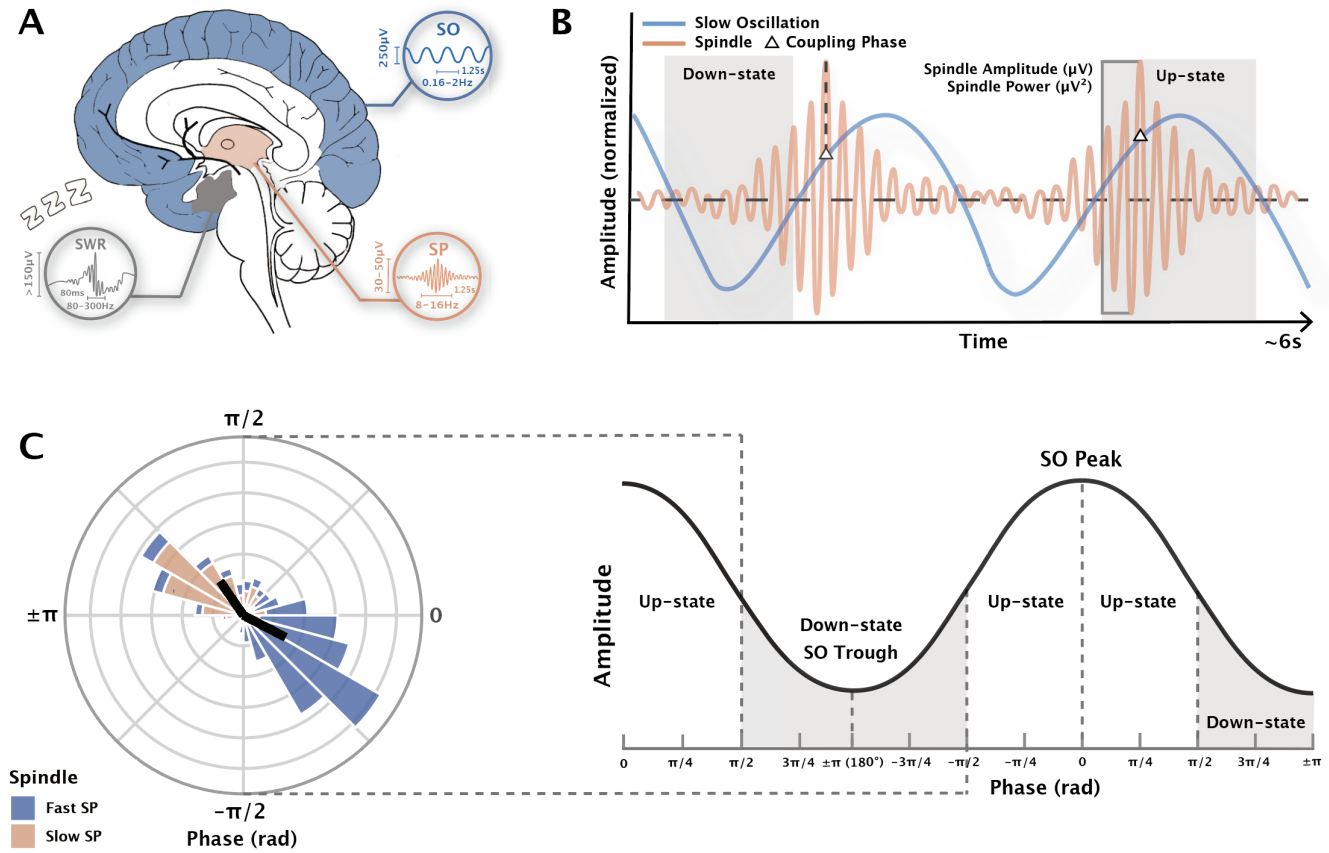

Figure 1: Measurement of phase-amplitude coupling (PAC) in slow oscillation and spindle events

Notes. (A) The origin of neural oscillations during sleep. *SO* slow oscillation; *SP* sleep spindle; *SWR* sharp wave ripple. In each subgraph, the vertical line indicates the typical amplitude of that sleep wave, while the horizontal line indicates the typical frequency and duration. Note that SPs also propagate along the cortex, and the figure only displays the origin. (B) Electrophysiology representation diagram of the SO-SP coupling. SO and SP amplitudes are normalized. The phase of SOs when SPs are at their maximum amplitude are recorded as the coupling phase. The occurrence of SPs and SO-SP coupling are not necessarily continuous as shown in the diagram. (C) Coupling preferred phase and strength diagram. (Left) The phase and strength used in the circular plot are simulated data from existing dataset for visualization purposes only. At the group level, the mean circular direction shows the preferred SO phase, while the mean vector length shows the strength of the precise coupling. (Right) The phase of SO peaks is noted as 0°, while the phase of SO troughs is noted as ±180°.

sion of memory consolidation processes into three coordinated steps, which occur in the precise hierarchical temporal structure of SO-FO coupling involving SOs, SPs, and SWRs[38–40]: 1) Hippocampal SWRs driven by cortical SOs facilitate the repeated replay of newly encoded memories, 2) The maximum amplitude of SWRs precisely lock with the trough phase of SPs during the depolarized up-state of SOs (SP-SWR coupling), 3) The maximum amplitude of SPs couple with the up-state of SOs (SO-SP coupling). These phase-locked coupling dynamics are considered integral to consistent communication between the hippocampus and neocortex, constituting a key mechanism for the reactivation and redistribution of temporary memory traces across cortical areas[41,42].

Pharmacological and stimulation research has provided evidence for causal relationships between coupling and memory consolidation. For example, the increasing level of inhibitory neurotransmission driven by GABAergic drug zolpidem improves the phase precision and strength of SO-SP coupling, and, in turn, contribute to enhanced memory retention performance[55–58]. Studies using calcium imaging and SP stimulation explained the significance of the precise coupling phase for synaptic plasticity: SP spike discharges extracted through electrical stimulation during SO up-states, efficiently modify excitatory neocortical synapses[27]. Recently, Niethard and colleagues[20] observed that only SP spikes occurring around up-state peaks of SOs were accompanied by amplified calcium activity patterns to optimize synaptic plasticity. This spatial-temporal mechanism represents a critical perspective in the study of SO-SP coupling in the memory consolidation.

When quantifying cross-frequency coupling, the spike-timing-dependent transfer theory[41,59] signified the importance of coupling phases and inconsistency in FO amplitudes. Cou-

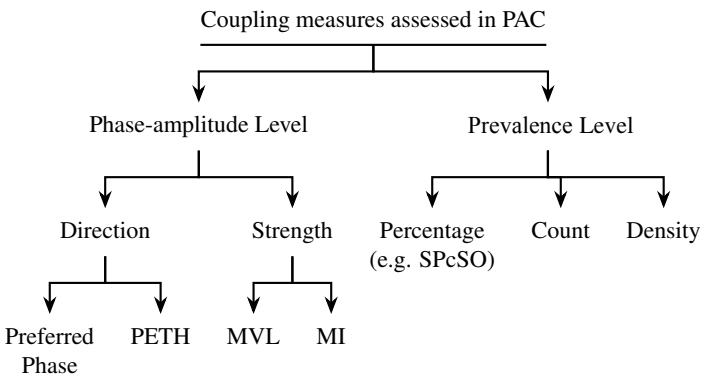

Figure 2: Hierarchical diagram of coupling measures

Notes. *PETH* peri-event time histogram; *MV*L mean vector length; *MI* modulation index. *SPcSO* Percentage of SPs coupled with SOs in all SP events. In contrast to the term "frequency" used throughout the text in reference to the neural oscillation, the coupling prevalence in the diagram indicates the occurrence of SO-SP coupling events.

pling phase has been widely shown to reflect dynamic functional configurations, indicating remote communication and precise timing changes in synaptic activity among neural populations sharing cognitive functions[60,61]. Comparing to the depolarization state (positive half-wave) of cortical SOs which drives FOs including SPs, its inhibitory hyperpolarization state exhibits relative silence of neurons[24,62]. The preferred phase represents the phase of SOs at the maximum amplitude of FOs (Figure 1B, C). Another common method is the peri-event time histogram (PETH), representing the proportion of FOs centered (i.e., maximum trough) at different SO phase bins[51,63].

Phase alone does not reveal the inconsistent strength of FO amplitude between coupled and non-coupled SO phases within each time window. Therefore, the analysis of phase-amplitude distribution constitutes the so-called coupling strength, also defined as coupling inconsistency. In memory research, two commonly used quantification methods for coupling strength include mean vector length[35] (MVL) and modulation index[64] (MI). Both determine the strength of PAC by measuring the uneven distribution of FO amplitude in different SO phase directions. Past simulation studies have proved the effectiveness of these two methods in extracting coupling strength under different conditions and the existence of noise[65–67].

Compared to the coupling phase and strength, which reflect the contrast on the phase-amplitude level, recent research also focuses on the overall occurrence of coupling events across the night. This involves assessing the prevalence of co-occurrence, including the coupling percentage[45,47,48,51] (i.e., co-occurrence rate, % coupling events / oscillation events), coupling counts[57], and the coupling density[68]. It is reasonable to categorize all measures related to co-occurrence under a category defined as "coupling prevalence", in contrast to the phase and strength (see Figure 2).

In SO-SP coupling studies, although most of them supported the importance of coupling phase in predicting memory retention, this association varies in magnitude across memory types, sleep stages, and age[43–51,53,54]. In contrast, the significance of coupling strength and percentage on the memory consolidation is inconsistent across studies[52–54,82].

Furthermore, previous meta-analyses of SP function consistently indicate that SP amplitude (and power) is the most effective predictor of memory consolidation and cognitive abilities[14,69]. The measure of neural oscillation amplitude is reported based on changes compared to mean amplitude, and the absolute amplitude difference between peak and trough reflects the intensity of synchronized neural activity near that specific scalp region[70]. Although overall SP amplitude is not a direct measure of SO-SP coupling, its modulation by SO phase is the core mechanism. Prior studies report mixed findings on the relation between SP amplitude and both coupling phase and strength[124,175]. Including SP amplitude reflects group differences in SP activity overnight, and allows comparison of the roles of coupling metrics and SP alone, as examined in many of the included studies[51,57,116]. Also, only 4 studies reported SP amplitude separately for coupled events, which limits targeted analyses. Therefore, we also include the mean peak-to-trough amplitude and power of all SP events in the meta-analysis (Figure 1B).

Electrophysiology evidence from studies included in our meta-analysis[45,47,68] and others[16,63,172] reported that the association between memory consolidation and SO-SP coupling is influenced by a variety of behavioral and physiological factors under different conditions. Among the moderators that have received significant attention in recent research are memory types, development and aging, pharmacological manipulations, disorders, and sleep stages. The early dual-process hypothesis proposed a dissociation between hippocampus-dependent and non-dependent memories reinforced during SWS and REM periods, respectively[59]: The standard active consolidation theory supported the role of hippocampal-cortical nesting during nREM sleep for the consolidation of hippocampus-dependent memories, including verbal, visual, and spatial declarative memories that are episodic-related[59,71,72]. In contrast, some research has linked the processing of non-hippocampus-dependent memory, such as emotional and procedural memories, with REM sleep[73–76].

This outdated theory may have led early SO-SP coupling studies to focus primarily on the experimental design of declarative memory inferences. Recent updates to this foundational theory propose that the hippocampus plays a crucial role in the consolidation of non-hippocampus-dependent memories during nREM sleep through the reactivation of spatiotemporal contextual features[77–81]. This has sparked discussions and research on whether hippocampal involvement in SO-SP-SWR coupling constitutes a general physiological principle in all types of memory, or at least in memories with spatiotemporal contexts. Particularly in the association between SO-SP coupling and non-hippocampal-dependent consolidation, recent studies have obtained conflicting results[11,13,14,44,53,68,82–84].

In addition, early behavioral experiments have demonstrated

the sensitivity of sleep-dependent memory consolidation to age-related changes in the brain, from development to aging[85,86]. The most significant changes of sleep neural oscillations during development include the maturation and dominance of frontal-originated global SOs and frontal detected SPs[87–89]. This change is associated with the improvement of memory consolidation from childhood to adolescence. In older adults, the time spent in SWS sharply decreases, followed by a reduction in the number and amplitude of SOs[90,91]. Their prefrontal SP events decrease by over 40% compared to young adults and are correlated with their weakened memory performance[92,93].

Taking a step further, Helfrich et al[50]. and Hahn et al[47]. have respectively identified the modulating roles of aging and development in the frontal SO-SP coupling, including changes in the precision of coupling phase, shifts in coupling topography, as well as improvements and impairments in memory retention performance. Without exception, all age-related studies on sleep-dependent memory consolidation emphasize the involvement of the prefrontal cortex and posterior hippocampus, which contribute to episodic memory processing and consolidation. Differences in their gray matter volumes, and changes in structural integrity during development and aging, are related to the representation of neural oscillations[92,94,95].

Besides these moderators, overlooked in most studies is the region-specificity of different memory types and the frequency of SPs[96]. The topography and frequency distribution coupling are neglected or subject to misleading interpretations when correlating other measures. SP frequency has traditionally been defined between 12-16 Hz[97]. However, recent research has found that SPs of different frequencies dominate at different phases of the SO cycle and in distinct cortical areas[98,99]. Therefore, studies started to analyze the SO-SP coupling by splitting SPs into fast and slow subtypes[11,44,100]. Since some studies found that fast SPs predominate in the centroparietal region, while slow SPs are more common in the frontal region, a significant amount of studies selectively extracted specific types of SPs from limited electrodes[100,117,173]. Some studies even averaged all electrodes in their spectral and/or time-series analysis to estimate metrics of oscillations and their couplings[13,98,121] (see more details in Section 3.2 and 3.4). This narrowed measure falls into the pitfall of considering all SPs as the regional oscillation, and overlooking the out-of-phase distribution of SOs in different cortical areas[101].

Acknowledging the existence of regional SO and SP events, recent evidence supports that global SO and SPs, as traveling waves, constitute the majority of their oscillation cycles[17,102,103]. Traveling waves are considered to play a crucial role in the transmission of information, including memory, between specific brain areas[104]. Once projected onto the cortex, global SPs exhibit rotational characteristics, looping from the temporal lobe to the parietal and frontal lobe sequentially (TPF) and then travel back to the temporal lobe[103,105]. This results in phase gradients in the burst of SP spikes across different cortical areas, which represents the direction of asymmetric cortical propagation and might explain the phase shifts in coupling[106,107]. It is worth noting that in other types of oscillations, the traveling waves have been found to also follow descending frequency gradients[108].

Dynamic spatiotemporal features of global SPs are believed to create necessary conditions for the synaptic plasticity[102]. Recently, relevant clues have emerged in research on SO-SP coupling. Hahn et al[47]. and Helfrich et al[50]. found only the phase and strength of coupling detected in frontal electrodes had sufficient predictive power for the success of memory retention. Denis et al.[45] reported a phase gradient from anterior to posterior area in SO-fast SP coupling by channels. Although all this evidence indicates the necessity for investigation of the spatiotemporal specificity of SO-SP coupling, current between-study differences and the ambiguity in defining SP types (see Section 3.4) introduce uncertainty to the measure of coupling topography. To our knowledge, there is no previous study that has systematically analyzed the dynamic spatiotemporal distribution of SO-SP coupling and their association with the frequency of oscillations and memory consolidation.

In summary, current research on the SO-SP coupling and memory consolidation faces multiple methodological challenges. Disadvantages of polysomnographic (PSG) studies include long recording times, limitations on subject recruitment, and variations in signal processing approaches. Researchers have the flexibility to choose analytical methods and selectively report results that favor their hypotheses[110,111], thereby increasing the risk of false positives and lack of reproducibility. Considering methodological limitations, varied analytical approaches, disparate quality of reports, and conflicting evidence, there is a clear urgency of a meta-analysis to consolidate the true effect sizes reported in studies, as well as propose standardized and targeted research and analysis methods.

To incorporate a broader range of studies without increasing between-study heterogeneity and publication bias, we requested unreported or selectively omitted effect sizes, as well as individual-level coupling data that were pre-processed but not included in correlation analysis. This approach helped us include more studies adopting a similar experimental design but using incomparable analysis methods, which is common in memory research. In addition, these data provide us with sufficient statistical power to compare spatio-temporal shifts in coupling phase and its association with memory consolidation.

The purpose of the current meta-analysis and review is to aggregate studies of SO-SP coupling to understand inconsistent results across studies. To be more specific, we used Bayesian hierarchical models to examine whether the cross-frequency coupling between SOs and SPs is associated with memory consolidation, and determine which specific measure(s) of SO-SP coupling most accurately predict memory retention perfor-

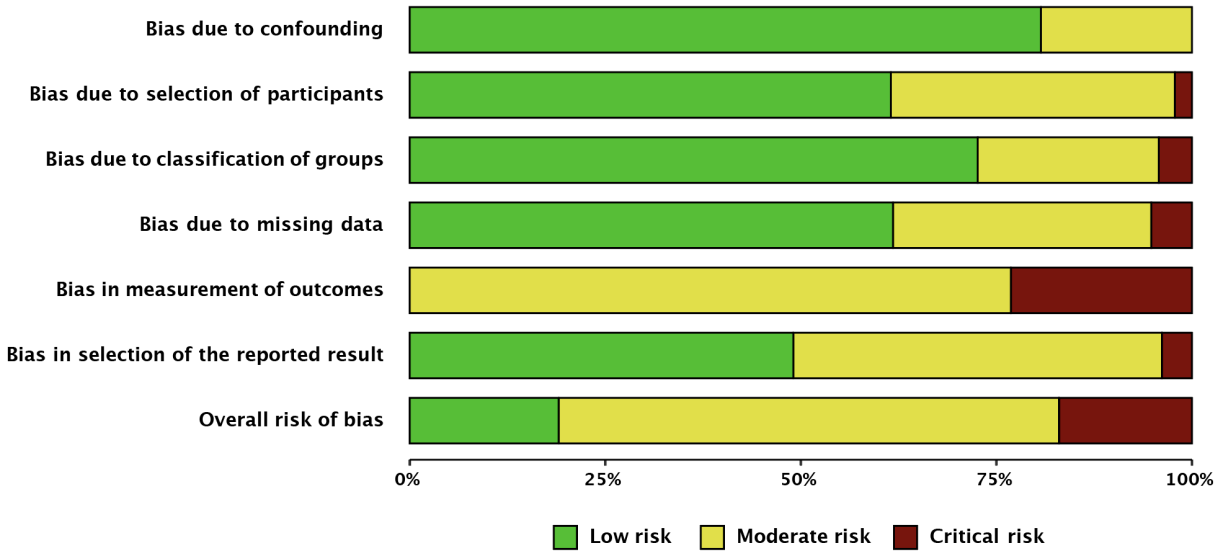

Figure 3: Risk of bias assessment summary plot adapted from ROBINS-I[109]

Notes. The most significant heterogeneity is revealed in the measurement of outcome, while the overall assessment indicated a moderate risk of bias across studies after requesting unreported results and data transformation. Based on the complexity of the type of measures involved in phase amplitude coupling analysis, we believe that this degree of risk of bias is acceptable. Specific evaluations for each study were reported in the Supplemental Material 3.

mance. In addition, we conducted moderator analyses to understand if physiological and behavioral factors modulate the strength of the relationship between SO-SP coupling and memory consolidation. We propose five questions for moderators: 1) Is SO-SP coupling a general physiological mechanism for memory consolidation or targeted at specific type(s) of memory?; 2) How do development and aging affect the association between coupling and memory consolidation?; 3) Is the coupling detected under specific (functional) region(s) more indicative of the memory retention?; 4) In which frequency range(s) of SPs is coupling more likely to be linked to memory consolidation?; 5) Does the stage and bout of sleep affect the measure of coupling-memory association? Besides the main study, we also analyzed the preferred distribution of coupling phases under different cortical areas and SP frequencies. In the rest of the discussion, we focus on summarizing conflicts and misinterpretations in current research, proposing strategies for research standardization, as well as discussing the implications of our results for future research focus and clinical applications.

2 Results

2.1 Study characteristics

In the final dataset of 23 studies included in the data analysis, the participants had an average age of 24.8 years, ranging from 7.3 to 78.0. The average female representation among participants was 49.5%. The average sample size for each study was 31.7 participants, ranging from 10 to 151. Each study contributed 4 effect sizes for each coupling measure on average. The included studies consist of 17 overnight and 6 nap studies, in which a total of 19 tasks for declarative memory and 6 tasks for procedural memory were measured. In addition, we assessed the risk of bias for each study before conducting data analysis (see Figure 3).

2.2 Coupling phase

We first assessed the association between SO-SP coupling preferred phase and memory retention following sleep. 23 studies ($k = 90$) were included in the Bayesian hierarchical model. We transformed the circular linear correlation to standardized coefficients (see Section 5.5.2). Forest and regression plots for overall and moderation models are reported in Figure 4. In addition, the results of hypothesis tests for the overall and moderator models are reported in Table 1 using the Bayes factor and posterior probability. Consistent with the funnel plot (Figure S11A), neither Egger regression ($p = 0.59$) nor rank correlation test ($p = 0.52$) found the existence of publication bias.

2.2.1 Overall model

The analysis of a random effect model on the overall phase-memory association revealed that, without considering moderating factors, very strong and consistent evidence supports a small-sized association between the preferred coupling phase and memory consolidation across studies, $r_{zpooled} = 0.07$ [0.01,

Table 1: Result of directional hypothesis tests for each pair of factor levels (conditions) in overall and each moderation model of the coupling phase-memory association

Moderator	Overall	Memory Task			Age	Spindle	PSG Channel			Stage	Period
Condition	H_1	Verbal	Emotional	Spatial	Younger	Fast	Frontal	Frontal	Central	N2	Night
Control	H_0	Motor	Motor	Motor	Older	Slow	Central	Posterior	Posterior	SWS	Nap
BF_{10}	58.35	1.30	1.09	0.07	160.94	11.39	13.58	6.13	0.86	1.74	2.23
Probability	0.98	0.57	0.52	0.06	0.99	0.92	0.93	0.86	0.46	0.63	0.69

Notes. *Condition* conditions hypothesized to be associated with stronger phase-memory association than other factor levels; *Control* Variables hypothesized to be associated with weaker phase-memory association; BF_{10} Bayes factor in favor of H_1 over H_0; *N2* nREM2 stage; *SWS* slow-wave sleep.

0.13], BF_{10} = 58.35, probability = 0.98. This implies that the likelihood of H_1 is over 58 times greater than the H_0, covering approximately 98% of posterior samples.

Multilevel model analysis indicates a heterogeneity similar to typical correlational meta-analyses[112] ($M_g = 0.13$), including a between-study heterogeneity of g = 0.07 [0.00, 0.16], as well as a within-study heterogeneity of g = 0.04 [0.00, 0.11]. Focal analysis showed no difference from the original overall model, with a effect size $r_{zpooled}$ = 0.07 [-0.06, 0.19]. Sensitivity analysis regarding prior robustness also revealed similar results (see Supplemental Material 9).

2.2.2 Moderator and sensitivity models

Sufficient evidence shows that moderators including memory tasks, age, spindle types, and PSG channels provide additional predictive power with strong favor for each hypothesis (all contain $BF_{10} \leq 0.1$ or ≥ 10). Figure 4C indicates that most mixed conditions have a considerably higher uncertainty compared to other normal conditions.

Memory Task Contrary to the assumption, there is no evidence to support a difference of phase-memory association between and motor memory (k = 36) and verbal memory retention (k = 35, $r_z = \Delta 0.01$, BF_{10} = 1.30, probability = 0.57), as well as between motor and emotional memory retention (k = 12, $r_z = \Delta 0.01$, BF_{10} = 1.09, probability = 0.52). However, strong evidence supports that spatial tasks have a considerably lower phase-memory association compared to motor tasks (k = 7, $r_z = \Delta 0.19$, BF_{10} = 0.07, probability = 0.06). The task model has a weight of 0.29, which indicates that the moderator effect of memory task is relatively weak and the task model performed worse than the overall model.

Participant Age As shown in Figure 4B, extremely strong evidence supports that with the increase of age, the slope of correlation exhibits a strongly decreasing trend ($r_{z\beta} = \Delta$-0.006 [-0.010, -0.001], BF_{10} = 160.94, probability = 0.99). Each ten-year increase in age is associated with a decrease in effect size of 0.06, which is consistent with our hypothesis that precise SO-SP coupling becomes less predictive of memory retention with the increase of age. The moderator effect of age becomes less pronounced during development after excluding the older adult data that represent aging effects, $r_{z\beta} = \Delta$ -0.005 [-0.013, 0.004], BF_{10} = 5.51, showing a moderate effect. Accounting for the factor of age group significantly increased the predictive power compared to the intercept-only model, given a stacked weight of 1.00.

Spindle Frequency Moderation model with SP frequency range highlights a stronger association between memory retention and SO-fast SP coupling phase (k = 43) rather than slow SPs (k = 28, $r_z = \Delta - 0.07$, BF_{10} = 11.39, probability = 0.92), consistent with the hypothesis based on studies comparing the role of different SP types on the memory retention[43,98]. The SP model performed as well as the overall model, weight = 0.48.

PSG Channel The highest pooled correlation was observed in the frontal electrode clusters (k = 29), compared to the central (k = 36, $r_z = \Delta - 0.08$, BF_{10} = 13.58, probability = 0.93) and posterior channels (k = 22, $r_z = \Delta - 0.07$, BF_{10} = 6.13, probability = 0.86), which represents a moderate-to-strong moderation effect. The frontal area has the largest phase-memory association, r_z = 0.12 [0.03, 0.21]. Effect sizes in posterior areas does not differ from central areas ($r_z = \Delta - 0.01$, BF_{10} = 0.86, probability = 0.46).

Sleep Stage and Bout The sleep stage appeared to be a weak predictor with a weight of 0.00, in which the nREM2 stage (k = 18) could not predict a higher phase-memory association than SWS (k = 30, $r_z = \Delta - 0.03$, BF_{10} = 1.74, probability = 0.63). The sleep bout also revealed no difference in phase-memory association between overnight sleep (k = 73) and nap condition (k = 17, $r_z = \Delta - 0.03$, BF_{10} = 2.23, probability = 0.69), with no predictive role (weight = 0.00) compared to the overall model.

2.2.3 Spatio-temporal analysis of coupling phase

By aggregating individual-level data across studies, we found that phase of SO-fast SP coupling has a significant quadratic association with memory retention in frontal regions (see Figure 5A), r = 0.26, r_z = 0.20, $p < 0.01$. However, this association is not significant in either central (r = 0.11, r_z = 0.05, p = 0.22) or posterior regions (r = 0.08, r_z = 0.01, p = 0.47).

Another important finding is a considerable shift of the pre-

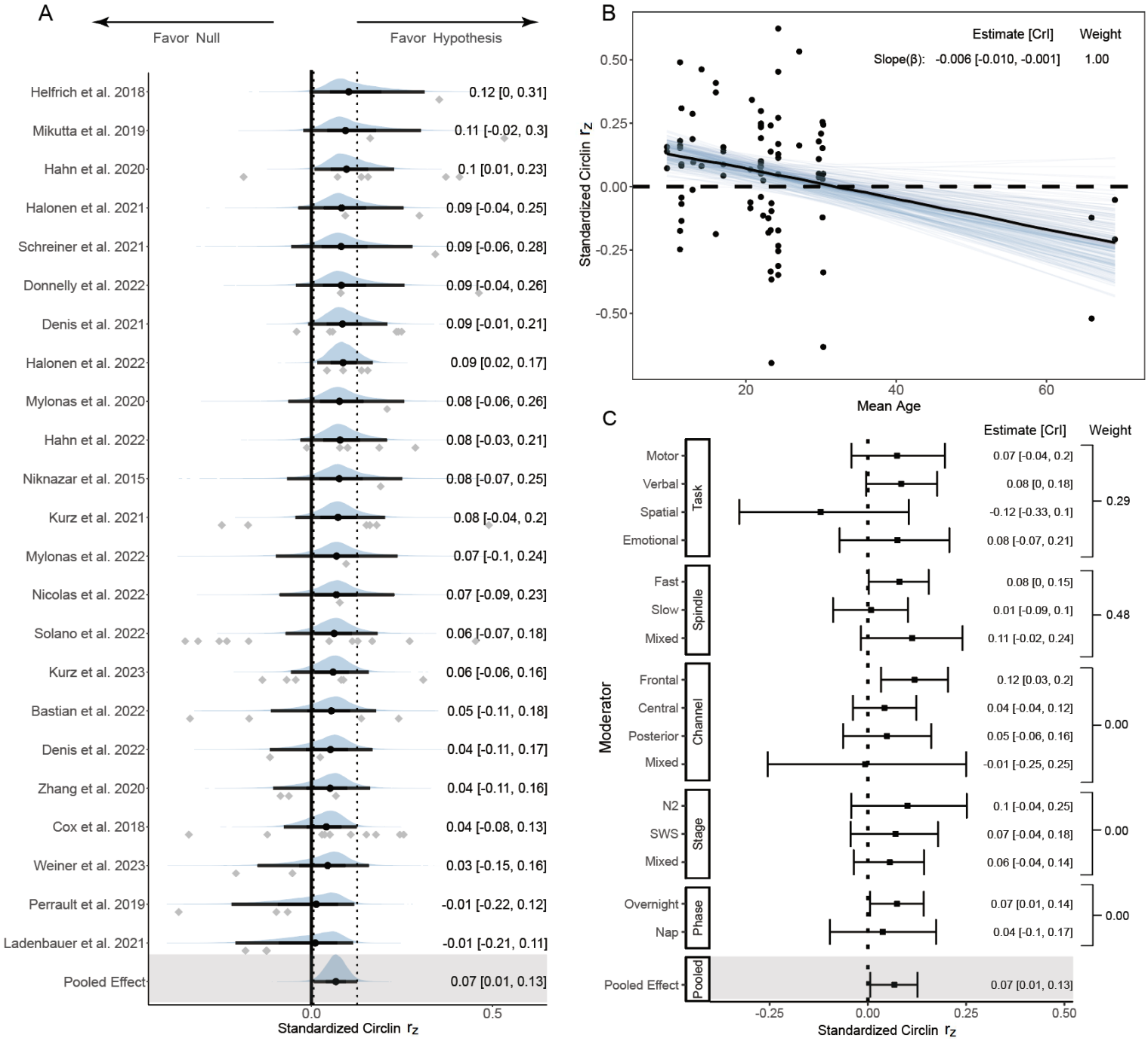

Figure 4: Forest and regression plots for the association between SO-SP coupling phase and memory retention.

Notes. (A) Overall model forest plot at study-level. Dashed lines indicate the 95% credible interval (*CrI*) of the pooled effect size. The black point and error bar for each study shows the adjusted estimation of effect size and 95% *CrI* combining data and prior information. The gray dots under each distribution show raw effect sizes of each study. Effect size-level plots can be found in Figure S5.1. (B) Meta regression plot with age as moderator. Blue lines represent 200 overplotted spaghetti fit lines to visualize predictions. (C) Moderator-level forest plot. Each box represents a type of moderator. Mixed Effect sizes with mixed conditions from different factor levels listed above. *Weight* Stacked weight of each moderation model in the paired model performance comparison between the moderator and overall (intercept-only) model. The stacked weight of the overall model in each pair of comparisons can be calculated as 1 - weight of the moderation model.

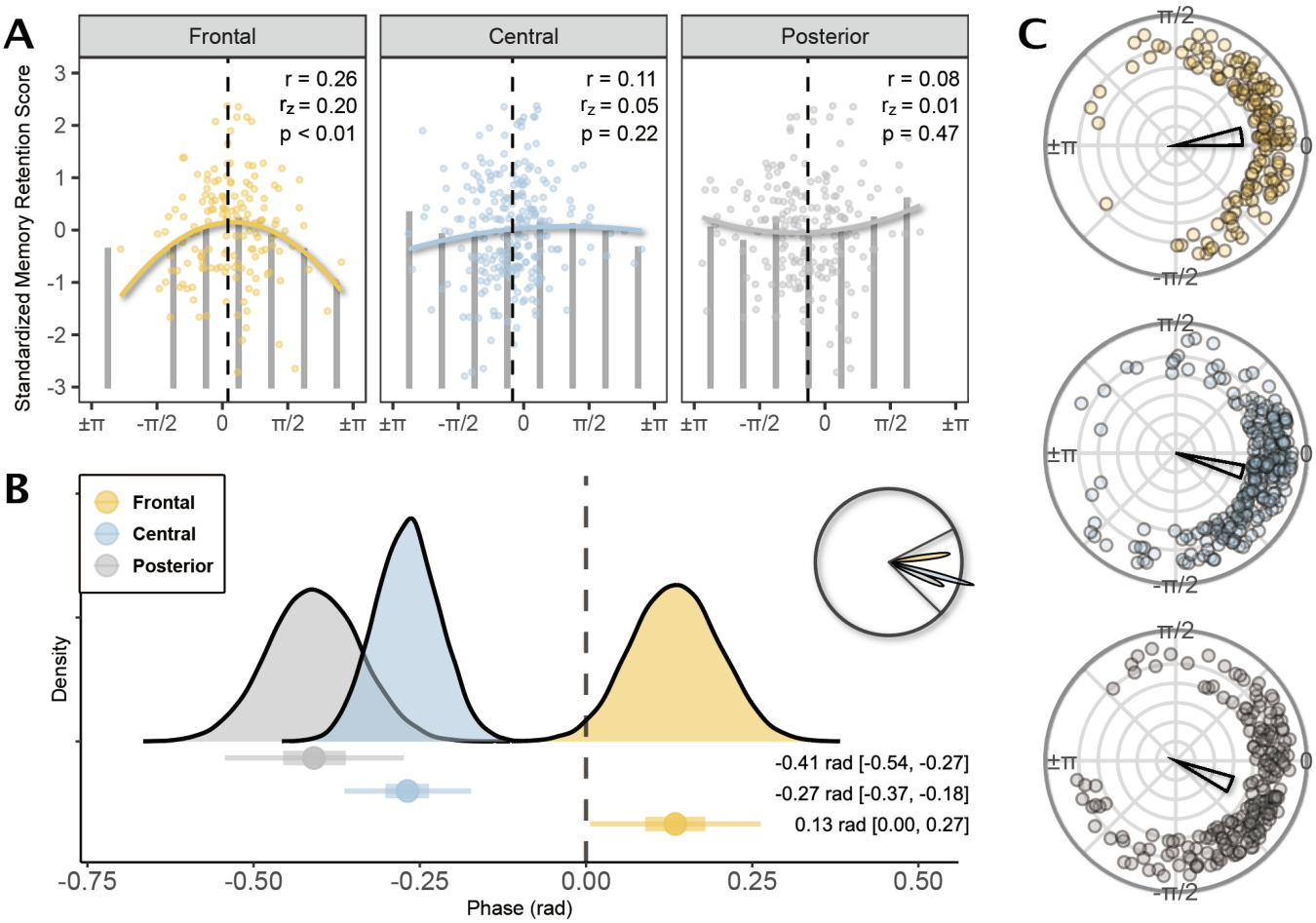

Figure 5: Preferred slow oscillation-fast spindle coupling phase and its association with the memory retention

Notes. (A) Quadratic regression of the phase-memory association under different regions of PSG channels aggregated from studies included in the meta-analysis. *0* peak of SO upstate; $\pm\pi$ trough of SO downstate; *r* circular-linear correlation coefficient; r_z standardized circular-linear correlation coefficient. Bars represent the mean memory retention scores per $\pi/4$ radian (45°). Dashed vertical line represents the mean preferred phase across studies. Colored quadratic fit line represents the direction of the relationship. The direction of their relationship gradually flips as the PSG channel moves from the front to the posterior area. Only under the frontal and central channels, fit lines display a quadratic relationship, with a peak of memory score near the up-state peak of SOs. In posterior channels, in contrast, the relationship is convex although it is not significant. Non-significant quadratic regression between SO-slow SP coupling and memory are reported in Supplemental Material 13. (B) Posterior distributions of mean preferred phases from the Bayesian circular mixed-effect model. The circular posterior distribution has been reported in the top-right corner, and the area between two black lines has been projected in a linear scale in the main graph. The vertical line reflects the up-state peak of slow oscillation. Dots and error bars reflect the mean and 95% credible intervals of phases detected from each channel cluster. Phase values have been reported as radians. (C) Circular plot of the preferred coupling phase. Arranged from top to bottom in the order of frontal, central, and posterior. The direction of each colored dot represents the preferred coupling phase of each subject recorded from PSG channels in each cluster. The direction of the mean resultant vector indicates the mean preferred coupling phase across subjects, the width indicates the 95% credible interval of the mean coupling phase, while the length from 0 (center) to 1 (circumference) indicates the consistency of coupling phase across subjects.

ferred coupling phase from frontal to posterior regions, similar to Kurz et al[51]. After taking into account the repeated measurement, the frontal area has a preferred phase around 0.13 rad [0.00, 0.27], which is the closest to the peak of SO (0 rad). The coupling consistently happened earlier when moving towards further dorsally, reflected by the phase in central (-0.27 rad [-0.37, -0.18]) and posterior area (-0.41 rad [-0.54, -0.27]).

Extremely strong evidence supports a considerable difference of the preferred coupling phase between frontal and central areas, Δ - 0.40 rad, $BF_{10} = +\infty$[1], probability = 1.00; between frontal and posterior areas, Δ - 0.54 rad, $BF_{10} = +\infty$, probability = 1.00; and between central and posterior areas, Δ - 0.14 rad, $BF_{10} = 34.4$, probability = 0.97. Differences can be observed from the shift of posterior distribution in Figure 5B and location of data clusters in Figure 5C. Moreover, the consistency of coupling phase across subjects also decreases from frontal ($z = 0.69$, $p < 0.05$, Rayleigh test) to posterior ar-

[1] An "infinite" BF_{10} value indicates that all posterior samples are overwhelmingly compatible with H_1, given the data and priors. It reflects a reporting convention where BF_{10} becomes unbounded as the likelihood under H_0 approaches zero.

Table 2: Result of directional hypothesis tests for each pair of factor levels (conditions) in overall and each moderation model of the SP amplitude-memory association

Moderator	Overall	Memory Task			Age	Spindle	PSG Channel			Stage	Period
Condition	H_1	Verbal	Emotional	Spatial	Younger	Fast	Frontal	Frontal	Central	N2	Night
Control	H_0	Motor	Motor	Motor	Older	Slow	Central	Posterior	Posterior	SWS	Nap
BF_{10}	8.28	22.19	1.50	9.30	3.70	12.21	1.88	24.67	13.65	5.37	2.54
Probability	0.89	0.96	0.60	0.90	0.79	0.92	0.65	0.96	0.93	0.84	0.72

Notes. *Condition* conditions hypothesized to be associated with stronger amplitude-memory association than other factor levels; *Control* Variables hypothesized to be associated with weaker amplitude-memory association; BF_{10} Bayes factor in favor of H_1 over H_0; *N*2 nREM2 stage; *SWS* slow-wave sleep.

eas ($z = 0.59, p < 0.05$, Rayleigh test). The differences of coupling phase can explain part of the discrepancy of the timing of coupling occurrence among previous studies[43,47,63,113]

In summary, we observed that SO-fast SP coupling in the frontal region occurred at the latest phase observed across all regions, characterized by the highest precision and strength of phase-locking. We also found the strongest phase-memory association in frontal regions following a typical quadratic relationship.

2.3 Spindle amplitude

We next assessed the association between the amplitude/power of SP and memory retention during the learning night. 18 original studies (k = 78) were included in the Bayesian hierarchical model. The mean amplitude of fast SPs (in μv) is 33.99 [26.44, 41.55], while the mean amplitude of slow SPs is 39.41 [33.61, 45.22]. Forest and regression plots for overall and moderation models are reported in Figure 6. In addition, the results of hypothesis tests for the overall and moderator models are reported in Table 2. The publication bias of the amplitude-memory association studies is unclear but potentially provides evidence of asymmetry. Egger's regression test ($p = 0.02$), rank correlation test ($p = 0.11$), and funnel plot (Figure S11C) did not demonstrate significant publication bias.

2.3.1 Overall model

We observed that without accounting for moderation effects, there was moderate evidence supporting a positive association between SP peak-to-trough amplitude and memory consolidation, r_{pooled} = 0.07 [-0.04, 0.18], BF_{10} = 8.28, probability = 0.89. The likelihood of our hypothesis being true is about 8 times greater than the null, covering approximately 89% of posterior samples. However, the pooled posterior distribution spans a wide range due to a large standard error, 95% credible intervals of most studies and the pooled effect size included 0. Three-level analysis has found a relatively high between-study heterogeneity of g = 0.20 [0.12, 0.32], and a normal level of within-study heterogeneity of g = 0.04 [0.00, 0.12]. The focal model indicated a slightly larger effect size compared to the overall model, with a pooled effect size r_{pooled} = 0.12 [-0.03, 0.29]. Prior sensitivity analysis did not reveal different patterns of the association (see Supplemental Material 9).

2.3.2 Moderator and sensitivity models

Although evidence shows that memory type, spindle type, and PSG channel location modulates the magnitude of correlation between SP amplitude and memory ($BF_{10} \leq 0.1$ or ≥ 10), it is worth noting that all moderation models did not provide enough additional information to increase the performance of model prediction relative to the overall model (all weight $\leq$ 0.31).

Memory Task The result of memory task moderation model shows that overnight retention of declarative tasks, including verbal tasks (k = 34, $r = \Delta$ 0.21, BF_{10} = 22.19, probability = 0.96); emotional tasks (k = 12, $r = \Delta$ 0.04, BF_{10} = 1.50, probability = 0.60); and spatial tasks (k = 14, $r = \Delta$ 0.19, BF_{10} = 9.30, probability = 0.90), have a higher association with the SP amplitude compared to the motor memory retention ($k = 36$). It is worth noting that two types of hippocampus-dependent memory, including verbal and spatial memory, both have larger estimates compared to the non-hippocampus-dependent memory. However, only the comparison between verbal and motor tasks provided strong favor for our hypothesis. Strong evidence supports a positive association between the SP amplitude and verbal declarative memory retention, r = 0.15 [0.01, 0.29].

Spindle Frequency The model results using the SP type as the moderator indicate a stronger positive relationship between fast SP amplitude (k = 36) and memory retention compared to slow SP amplitude (k = 20, $r = \Delta$ -0.09, BF_{10} = 12.21, probability = 0.92). A higher fast SP amplitude tends to be associated with a higher memory retention ability, r = 0.1 [-0.04, 0.23].

PSG Channel Strong evidence supports differences of pooled correlation between the frontal (k = 29) and posterior region (k = 16, $r = \Delta$ -0.13, BF_{10} = 24.67, probability = 0.96), as well as between the central (k = 30) and posterior region ($r = \Delta$ -0.11, BF_{10} = 13.65, probability = 0.93), which is consistent with our assumptions. Frontal regions have the largest positive amplitude-memory association, r = 0.11 [-0.02, 0.25]. In

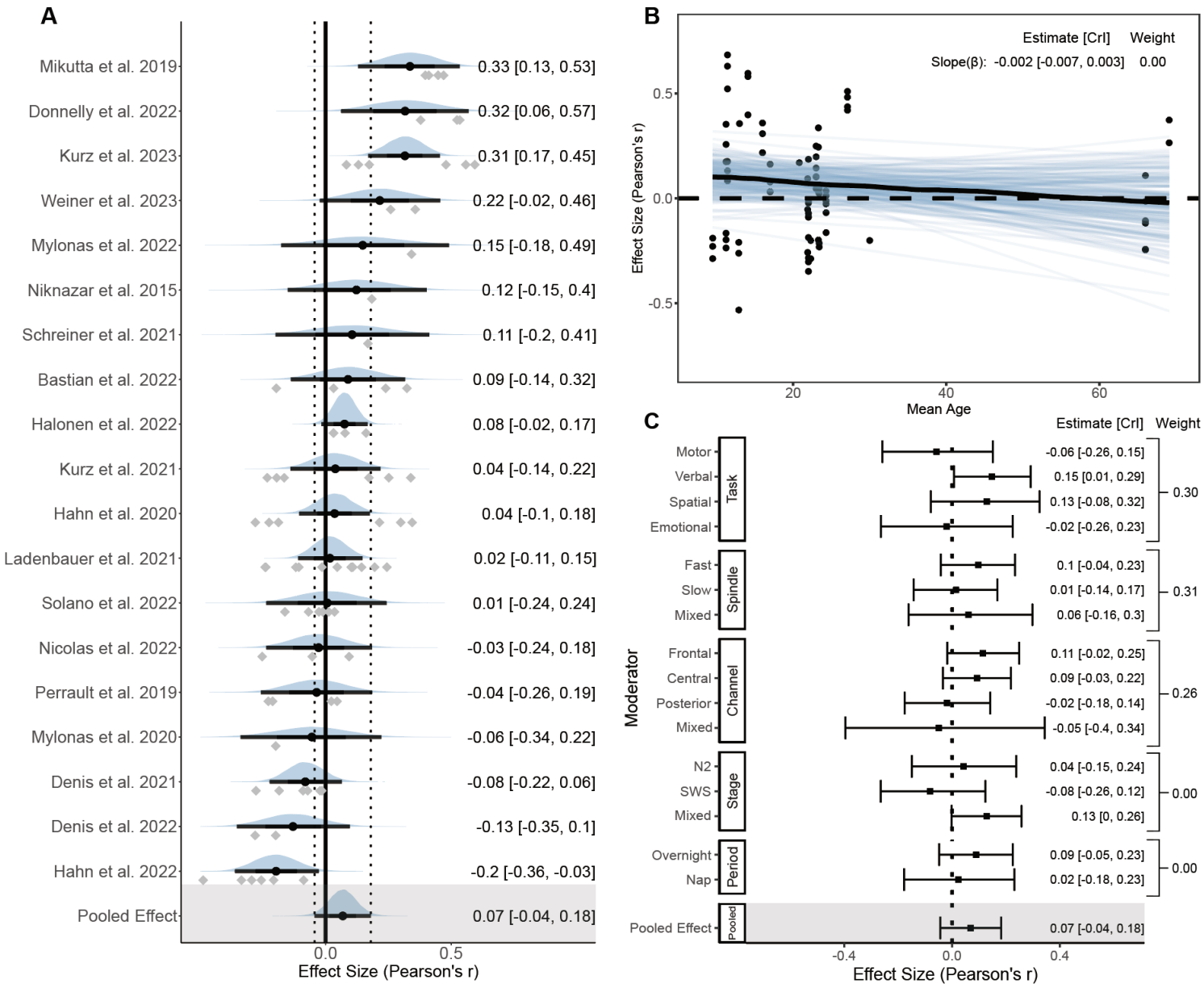

Figure 6: Forest and regression plots for the association between SP amplitude and memory retention.

Notes. (A) Overall model forest plot at study-level. The solid vertical line represents the mean Pearson correlation coefficient under the null hypothesis. Dashed lines indicate the 95% credible interval (*CrI*) of the pooled effect size. The black point and error bar for each study shows the adjusted estimation of effect size and 95% *CrI* combining data and prior information. The gray dots under each distribution show raw effect sizes of each study. Effect size-level plots can be found in Figure S5.2. (B) Meta regression plot with age as moderator. Blue lines represent 200 overplotted spaghetti fit lines to visualize predictions. (C) Moderator-level forest plot. Each box represents a type of moderator. Mixed Effect sizes with mixed conditions from different factor levels listed above. *Weight* Stacked weight of each moderation model in the paired model performance comparison between the moderator and overall (intercept-only) model. The stacked weight of the overall model in each pair of comparisons can be calculated as 1 - weight of the moderation model.

contrast, we did not observe a notable difference in amplitude-memory association between frontal and central regions ($r = \Delta$ -0.02, BF_{10} = 1.88, probability = 0.65).

Other Moderators There is a moderate evidence supporting a difference in the amplitude-memory association between the N2 and SWS stage ($r = \Delta - 0.12$, BF_{10} = 5.37, probability = 0.84), but no reliable difference between overnight and nap conditions ($r = \Delta$ -0.07, BF_{10} = 2.54, probability = 0.72). Also, the effect of age on the prediction of amplitude-memory asso-

Table 3: Result of directional hypothesis tests for each pair of factor levels (conditions) in overall and each moderation model of the coupling strength-memory association

Moderator	Overall	Memory Task			Age	Spindle	PSG Channel			Stage	Period
Condition	H_1	Verbal	Emotional	Spatial	Younger	Fast	Frontal	Frontal	Central	N2	Night
Control	H_0	Motor	Motor	Motor	Older	Slow	Central	Posterior	Posterior	SWS	Nap
BF_{10}	111.04	2.77	10.62	1.10	9.97	1.91	2.81	1.57	0.70	0.33	2.66
Probability	0.99	0.73	0.91	0.52	0.91	0.66	0.74	0.61	0.41	0.25	0.73

Notes. *Condition* conditions hypothesized to be associated with stronger strength-memory association than other factor levels; *Control* Variables hypothesized to be associated with weaker strength-memory association; BF_{10} Bayes factor in favor of H_1 over H_0; *N*2 nREM2 stage; *SWS* slow-wave sleep.

ciation is highly limited ($r_\beta = \Delta$ -0.002 (-0.007, 0.003), BF_{10} = 3.70, probability = 0.79).

2.4 Coupling strength

Next, we measured the association between coupling strength and memory retention. 21 original studies (k = 86) were included in the Bayesian hierarchical model. The mean SO-fast SP coupling strength measured by mean vector length is 0.33 [0.27, 0.39], while the mean SO-slow SP coupling strength is 0.23 [0.19, 0.27]. Forest and regression plots for overall and moderation models are reported in Figure 7. In addition, the results of hypothesis tests for the overall and moderation models are reported in Table 3. Neither Egger regression ($p = 0.53$) nor rank correlation test ($p = 0.67$) showed evidence of potential publication biases.

2.4.1 Overall Model

Consistent with results of the overall phase-memory association, extremely strong evidence supports a positive coupling strength-memory association in the intercept-only model (r_{pooled} = 0.08 [0.02, 0.15], BF_{10} = 111.04, probability = 0.99), which strongly favors our hypothesis. The likelihood of data under our hypothesis is over 110 times greater than the negative direction, covering above 99% of posterior samples.

Similar to typical correlational meta-analyses, the overall model consists of a between-study heterogeneity of g = 0.08 [0.01, 0.17], and a within-study heterogeneity of g = 0.05 [0.00, 0.14]. The result of focal analysis showed no difference from the overall model but has a wider credible interval due to a smaller number of effect sizes, r_{pooled} = 0.09 [-0.02, 0.20]. Sensitivity analysis regarding prior robustness revealed consistent results with the overall model (see Supplemental Material 9).

2.4.2 Moderator and sensitivity model

Surprisingly, there is no enough evidence supporting a difference between each pair of factor levels for almost all moderators (all $0.1 < BF_{10} < 10$, except emotional versus motor tasks), and there was no single moderator model that had better a performance than the intercept-only model (all weight < 0.5). This result represents a relatively consistent strength-memory association, regardless of the impact of moderators.

Memory Task With the exception of the emotional task condition (k = 12) which predicts a larger strength-memory association than the motor task condition (k = 32, $r = \Delta$ -0.20, BF_{10} = 10.62, probability = 0.91), the type of memory task only had a weak impact on the association (all other $0.33 < BF_{10} < 3$). In addition, given the emotional task condition has a relatively small number of effect sizes, the evidence of group differences also need to be interpreted with caution.

Age Model The strongest predictive power among moderation models has been found with participant's age, and performs similar in the trade-off between complexity and power with the intercept-only model (weight = 0.44). As the age increases, the slope of the correlation exhibits a decreasing trend that is nearly in strong favor of our hypothesis ($r_\beta = \Delta$ -0.003 [-0.007, 0.001], BF_{10} = 9.97, probability = 0.91). The strength-memory association becomes weaker with the increase of age. The effect of age remains moderate after removing older adults from the analysis, $r_\beta = \Delta$ -0.005 [-0.015, 0.008], BF_{10} = 4.05.

Other Moderators There was no strong evidence of a difference of strength-memory association between the fast (k = 41) and slow SPs (k = 26, $r = \Delta$ -0.02, BF_{10} = 1.91, probabil-ity = 0.66). Also, there is no evidence for a group difference among any pair of conditions of PSG channels, sleep stages, or sleep bout type (all $0.33 < BF_{10} < 3$). However, consistent with the results for the coupling phase and SP amplitude, the largest positive strength-memory association moderated by PSG channels was still found in frontal channels (r = 0.10 [0.00, 0.19]).

2.5 Coupling percentage

The last part of our analysis focused on the association between the percentage of SPs coupled with SOs, and memory consolidation. 11 studies (k = 43) were included in the Bayesian hierarchical model. The mean SO-fast SP coupling percentage (%) is 21.17 [15.95, 26.39], while the SO-slow SP coupling percentage (%) is 24.07 [17.59, 30.56]. Forest and regression plots for

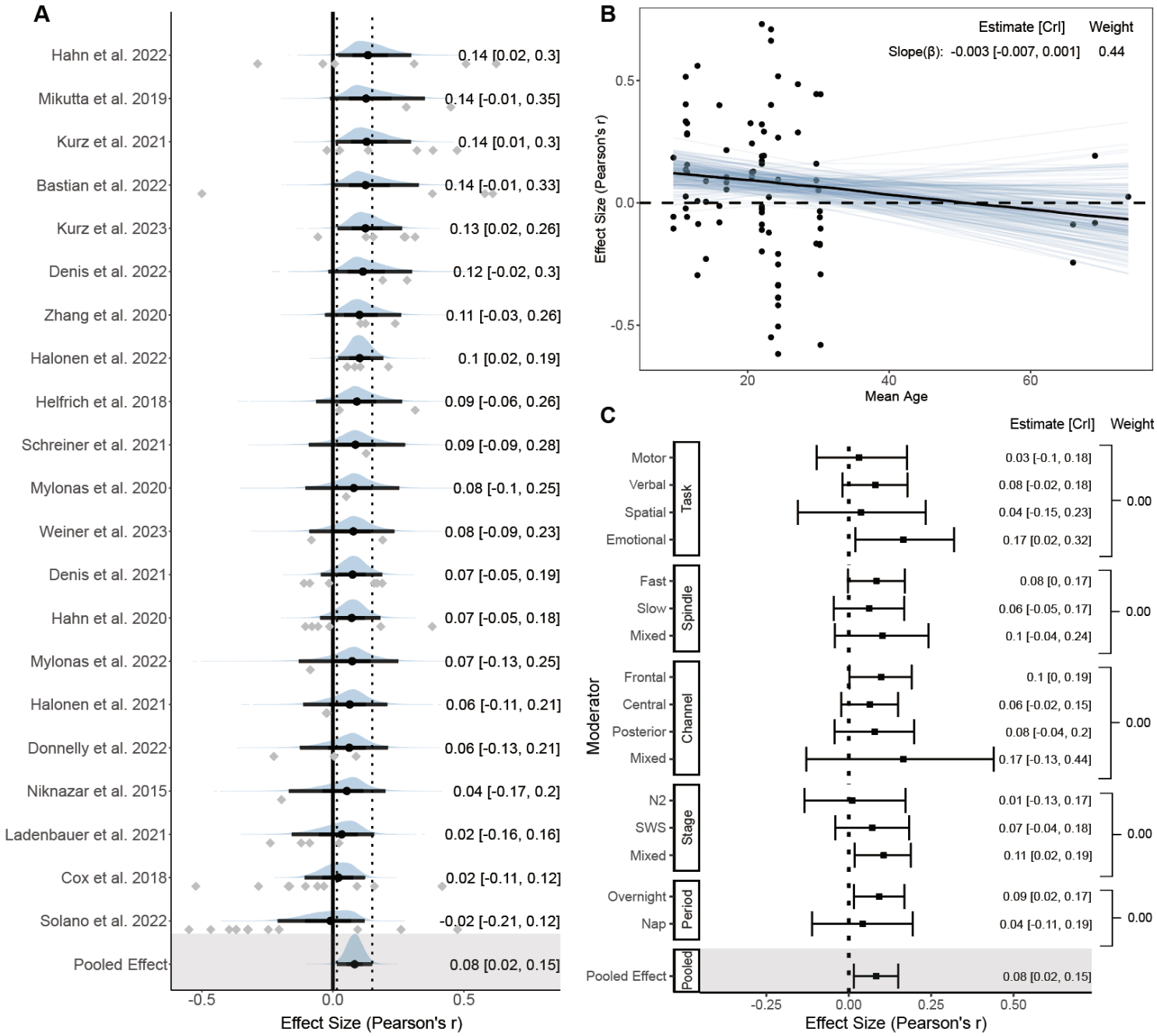

Figure 7: Forest and regression plots for the association between coupling strength and memory retention.

Notes. (A) Overall model forest plot at study-level. The solid vertical line represents the mean Pearson correlation coefficient under the null hypothesis. Dashed lines indicate the 95% credible interval (*CrI*) of the pooled effect size. The black point and error bar for each study shows the adjusted estimation of effect size and 95% *CrI* combining data and prior information. The gray dots under each distribution show raw effect sizes of each study. Effect size-level plots can be found in Figure S5.3. (B) Meta regression plot with age as moderator. Blue lines represent 200 overplotted spaghetti fit lines to visualize predictions. (C) Moderator-level forest plot. Each box represents a type of moderator. Mixed Effect sizes with mixed conditions from different factor levels listed above. *Weight* Stacked weight of each moderation model in the paired model performance comparison between the moderator and overall (intercept-only) model. The stacked weight of the overall model in each pair of comparisons can be calculated as 1 - weight of the moderation model.

Table 4: Result of directional hypothesis tests for each pair of factor levels (conditions) in overall and each moderation model of the coupling percentage-memory association

Moderator	Overall	Memory Task			Age	Spindle	PSG Channel			Stage	Period
Condition	H_1	Verbal	Emotional	Spatial	Younger	Fast	Frontal	Frontal	Central	N2	Night
Control	H_0	Motor	Motor	Motor	Older	Slow	Central	Posterior	Posterior	SWS	Nap
BF_{10}	0.38	0.05	0.17	0.22	2.16	0.28	5.15	3.38	1.24	2.02	0.44
Probability	0.28	0.05	0.14	0.18	0.68	0.22	0.84	0.77	0.55	0.67	0.31

Notes. *Condition* conditions hypothesized to be associated with stronger percentage-memory association than other factor levels; *Control* Variables hypothesized to be associated with weaker percentage-memory association; BF_{10} Bayes factor in favor of H_1 over H_0; $N2$ nREM2 stage; SWS slow-wave sleep.

overall and moderator models are reported in Figure 8. In addition, the results of hypothesis tests for the overall and moderator models are reported in Table 4. Consistent with the funnel plot (Figure S11E), both Egger regression ($p = 0.78$) and rank correlation test ($p = 0.87$) showed no publication bias.

2.5.1 Overall Model

Compared to all other coupling measures, we observed a weak association between coupling percentage and memory consolidation (r_{pooled} = -0.03 [-0.15, 0.07], BF_{10} = 0.38, probability = 0.28). The overall model consists of a moderate between-study heterogeneity of g = 0.11 [0.01, 0.25], and a within-study heterogeneity of g = 0.05 [0.00, 0.14]. 95% credible intervals of all studies and the pooled effect size include 0. Due to the limited number of effect sizes, we did not perform a focal analysis. Prior sensitivity analysis revealed consistent results with the overall model (see Supplemental Material 9).

2.5.2 Moderator and Sensitivity Models

Few moderators had an influential impact on the percentage-memory association (All $0.1 < BF_{10} < 10$, except verbal versus motor tasks, all weight < 0.40), which is not surprising given the weak association in the overall model. For memory types, we observed a moderate trend of declarative memory retention towards a negative association with coupling percentage, in contrast to the positive trend of percentage-motor memory association. Therefore, the association measured by motor memory tasks (k = 13) is considerably higher than verbal tasks, k = 13, $r = \Delta$ -0.25, BF_{10} = 0.05, probability = 0.05. All other moderators did not trend toward a specific direction and all 95% credible intervals include 0.

3 Discussion

As the first meta-analysis focusing on the coupling between slow oscillation and spindle events, our results, combining 297 effect sizes, provide reliable evidence for the involvement of thalamocortical SO-SP coupling in memory consolidation. In particular, the precision and strength of coupling, represented by the preferred phase and coupling strength, showed significant effect sizes in their correlation with memory retention performance. Moderators including age, memory type, cortical area, and SP frequency modulated the magnitude of these associations. The conclusive results are summarized in Table 5.

3.1 Precision and strength of SO-SP coupling as strong predictors of memory consolidation

In previous meta-analyses focusing on spindle events, the amplitude and power of SPs have been considered the most predictive measures for memory consolidation[14] and cognitive abilities[69]. However, our findings indicate that, the SP amplitude-memory association is subject to high variabilities. Only the association between fast SP amplitude and hippocampal dependent memory consolidation is supported by strong evidence. In contrast, the precision and strength of coupling between the fast SP peak amplitude and the up-state peak of SOs, is more crucial indicators for predicting memory retention performance. Our result confirmed the importance of cross-frequency coupling in the hierarchical temporal nesting within the hippocampus-thalamus-cortex information transmission loop, proposed in the active system consolidation theory[3,6,38,114]. Together with other studies included in the review[47,50,57,63,115], our results suggest a crucial role of coupling but did not support the role of spindle events alone in memory consolidation.

Moreover, we found that the phase and strength of coupling between fast SPs and SOs, retains strong predictive ability for memory retention performance in most sub-groups of the moderator analysis. This result confirms the robustness of the connection between coupled SO-fast SP phase and memory consolidation as a general physiological mechanism. Similar results were not observed with slow SPs. Since we did not find publication bias in both analyses regarding coupling phase and strength, it reduces the likelihood of overestimating true effect sizes. Given we have identified predictive powers for memory retention in both coupling phase and strength, and Weiner and colleagues[54] reported a synergistic interaction between phase and strength, future studies are necessary to further investigate

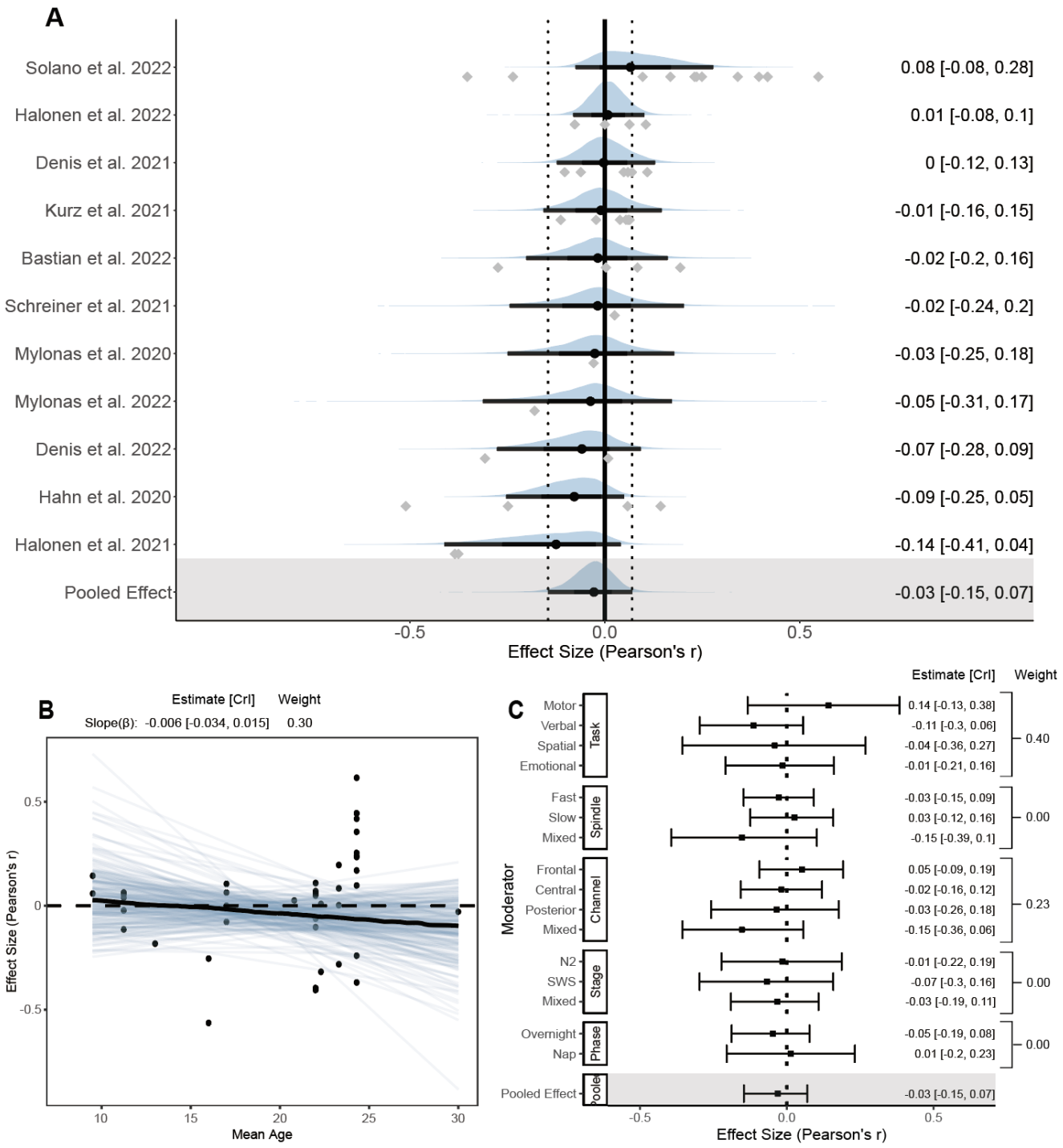

Figure 8: Forest and regression plots for the association between coupling percentage and memory retention.

Notes. (A) Overall model forest plot at study-level. The solid vertical line represents the mean Pearson correlation coefficient under the null hypothesis. Dashed lines indicate the 95% credible interval (*CrI*) of the pooled effect size. The black point and error bar for each study shows the adjusted estimation of effect size and 95% *CrI* combining data and prior information. The gray dots under each distribution show raw effect sizes of each study. Effect size-level plots can be found in Figure S5.4. (B) Meta regression plot with age as moderator. Blue lines represent 200 overplotted spaghetti fit lines to visualize predictions. (C) Moderator-level forest plot. Each box represents a type of moderator. Mixed Effect sizes with mixed conditions from different factor levels listed above. *Weight* Stacked weight of each moderation model in the paired model performance comparison between the moderator and overall (intercept-only) model. The stacked weight of the overall model in each pair of comparisons can be calculated as 1 - weight of the moderation model.

Table 5: Descriptive table of the meta-analysis result of measures of SO-SP coupling characteristics

Measure	n	k	Pooled Effect Size	BF_{10}	Influential Moderators
Coupling Phase	23	90	0.07 [0.01, 0.13]	58.35	Memory Type, Age, Channel, Spindle
Spindle Amplitude	18	78	0.07 [-0.04, 0.18]	8.28	Memory Type, Age, Channel, Spindle
Coupling Strength	22	86	0.08 [0.02, 0.15]	111.04	Age
Coupling Percentage	11	43	$-0.03[-0.15, 0.07]$	0.38	None

Notes. n number of studies included; k number of effect sizes included; BF_{10} Bayes factor in favor of H_1 over H_0 (see Table 6 for the interpretation).

the relationship between these phase-amplitude level measures.

3.2 The modulation role of cortical area and oscillation frequency

Most of the studies included in the meta-analysis supported the significance of the association between the coupling phase and memory consolidation. However, conflicting conclusions have been reported regarding the direction of the phase. In terms of the preferred coupling phase we find two predominant views, with a subset of studies reporting that coupling events are concentrated before the up-state peak (-π/2-0) of SOs[44,51,57,100,116,117]. Other authors found that the preferred phase is when fast SPs precisely coupled with the up-state peak (0)[47,48,50,53,68,136]. In either case, however, most of the studies agreed that coupling occurring closer to the up-state peak of SOs could predict better memory retention performance.

In addition to differences in signal processing approaches and age leading to discrepancies in phases reported, we believe that another reason is different PSG electrodes and SP frequencies filtered to detect coupling events (see Table 6). Research found that fast SPs occur more frequently in centro-parietal regions, while slow SPs are predominately in frontal regions[98,118–120]. This evidence led some research to apply this conclusion to coupling studies, and detecting fast SPs only from centro-parietal electrodes or slow SPs only from frontal electrodes[43,57,100,117]. However, we suspect that this approach restricts researchers from considering the origin and spread of oscillations. The prefrontal cortex has been proposed to be responsible for the generation of posterior-propagating global SOs[6,17,23,123], while global SPs, the majority of SP events, propagate in an rotating direction following gradients[102,103,105]. The significance of this dynamic interaction of oscillations for spatiotemporal coordination remains poorly understood.

SO-SP coupling can be widely detected across the fronto-parietal cortex[122], and the interaction between centroparietal dominated fast SPs and frontal dominated SOs, along with their propagations, are proposed to be crucial for long-range transmission of memory-related information from thalamus to neocortex[5,23,42,47,124]. Our spatial-temporal analysis provides extremely strong evidence that the post-peak SO-fast SP coupling phase recorded from frontal regions occurs significantly closer to the up-state peak of SOs, and considerably later than the pre-peak phase in centro-parietal regions, which supports phase shifts across electrodes found in previous studies[44,45,121], emphasizing the importance of coupling precision but not intensity. Moderator models of all four coupling measures consistently indicate that coupling detected from frontal regions has the largest association with the memory consolidation. Moreover, a significant quadratic phase-memory relationship is only observed in frontal regions. In addition, the frontal region has the strongest and most active SOs as its origin site[17-20], which may contribute to the role of frontal coupling. To the best of our knowledge, our study is the first to provide consistent evidence supporting that successful memory consolidation is modulated by spatiotemporal specificity, emphasizing the precise coupling of frontal fast SPs targeting the up-state peak of SOs.

3.3 Further moderation through memory types, aging, and sleep conditions

The predictive ability of coupling for memory consolidation has been studied across various types of memory, including emotional, spatial, verbal, and motor tasks. Surprisingly, our results indicate that the coupling phase has a widespread predictive role for both declarative and procedural memory. In contrast, evidence only supports a correlation between coupling strength and hippocampus-dependent memory. In addition, Hahn et al.[47,82] provided a region-specific view, claiming significant associations of SO-SP coupling with declarative and procedural memory exists in the frontal lobe and motor cortex, respectively. Therefore, it is worthwhile for future research to study the both shared mechanisms and region-specificity of coupling across different memory types, including those traditionally considered non-hippocampus-dependent. However, we must exercise caution that the number of effect sizes extracted for emotional and spatial tasks is limited, making it susceptible to the influence of any single study.

One disadvantage in our analysis of age as a moderator is the limited number of studies on children and older adults. The linear meta-regression as a function of age did not capture the moderating pattern of development but only reflected aging through the decreasing trend of the coupling-memory association. For older adults, the meta-regression showed that all four measures approached null hypothesis levels of non-

significance, indicating a diminishing effect in predicting memory consolidation through coupling phase or strength with aging. These results support previous studies on the memory consolidation impairment in older adults[50,63]. While our results did not directly reflect the influence of development on the coupling-memory association, evidence from past research suggested that the strength of SO-SP coupling in frontal lobes increases with developmental age, indicating the development of long-term memory networks[47]. We suggest future studies to further compare SO-SP coupling across different age groups, contrasting the regional distribution of coupling changes during development and aging.

However, evidence did not support the moderation role of stage conditions, including sleep stages and bouts, in the coupling-memory association, which is somewhat surprising. SOs are predominant in SWS[125], while sleep SPs are the primary oscillation feature in N2[119]. Our included studies showed that the co-occurrence rate of SO-SP coupling is higher during SWS compared to N2. One possible explanation, as suggested by our meta-analysis results, is that the precision and strength of coupling is more predictive of memory consolidation than the coupling percentage. Thus considering the contribution of different stages to memory consolidation through co-occurrence rates may not be meaningful. Regarding sleep bouts (naps versus overnight sleep), no difference were found in any coupling measures between overnight sleep and naps, consistent with the results of a meta-analysis focusing on spindle-memory association[14,126]. It implies the potential benefits of napping in memory consolidation and the clearance of hippocampal traces for storing new knowledge during the daytime.

3.4 Challenges of current statistical approaches in measuring EEG-behavior associations

One of the crucial factors limiting further interpretation of our results is the variation in statistical methods across studies, including differences in the definition of behavioral and physiological measures, the event detection and analysis methods employed, and limitations in estimating nonlinear relationships. Firstly, we agree with McConnell and colleagues[127] that there has been confusion in previous studies, where the term "slow spindle" is inconsistently used to refer to either frontal SPs (~10-14 Hz) or low-frequency slow SPs (~7-11 Hz). Compounding this issue, individual differences in spindle frequency are often overlooked, leading to challenges in reliably distinguishing between slow and fast spindles. Our result reveals that frontal fast SPs (~10-14 Hz) occurring near the up-state peak of frontal-predominant global SOs with lower frequency than the centro-parietal fast SPs (~14-17 Hz), which becomes predominant with development[87,88,127,129]. Some studies have also reported difficulty in separating these two types of spindles[47]. In contrast, slow SP occurs before the trough of the SO down-state. Our findings support a detailed categorization of SP types proposed in previous studies[127–130], involving separate measures between pre-peak early-fast SPs, post-peak (or peak) late-fast SPs (see Figure 5B), and pre-trough slow SPs (see Figure S13.1B). Alternatively, an effective approach could involve extracting signals from electrodes across different cortical areas for each type of SP for comparison. In addition, the current definition of moderators is also quite vague and conflicted. Some studies conducted analyses across mixed conditions, including using a global frequency range (e.g., 10-16 Hz) for SP detection, as well as reporting only one shared effect size for all conditions.

Moreover, we observed significant between-study differences in the measures for memory retention performance. Specifically, multiple retention measures already exist when solely considering verbal tasks for measuring declarative memory, including 1) the value difference in the number of correctly recalled words (Post-sleep – Pre-sleep number of correct words)[54,57,100]; 2) the value difference in the percentage of correctly recalled words (Post-sleep – Pre-sleep % of correct words)[47,58,116]; and 3) the ratio difference in the percentage of correctly recalled words (Post-sleep / Pre-sleep % of correct words)[45,53,117]. Similar situations exist in measures for other memory retention tasks. When we tested the sleep-memory association using these three formulas separately, we found that in some extreme cases, different measures could even alter the direction of the association. Thus, we strongly suggest the consistency in memory retention measures in future studies[131].

In addition, misinterpretation of phase-memory associations in some studies poses a threat to the validity of results. Most studies use circular-linear correlation to measure this association, but we observed a prevalence of exaggeration when explaining the effect size. Circular-linear correlations lack directionality[132], making it challenging to precisely estimate the improvement or decay of memory consolidation at specific SO phases. It is likely to be influenced by fluctuations in memory scores in any segment of SO phases, so associating it with hypotheses targeting specific phases might lead to incorrect conclusions. An effective solution is to visualize the regression and superimpose quadratic fit lines to assess whether it follows a typical quadratic relationship around the global maximum or minimum. In addition, through simulation studies (see Supplemental Material 12), we observed that circular-linear correlation coefficients exhibit weak robustness and severe deviation from normal distribution for small to medium sample sizes ($n < 100$), commonly encountered limitations in PSG studies.

To suppress the bias introduced by non-linear relationships, we believe there are two available solutions: 1) Hahn et al.,[47] calculated the absolute distance of the preferred phase of each participant from the upstate peak (0), while Kruz[52] and Weiner[54] applied this method to transform the phase-memory association into a linear relationship for subsequent testing. Our results provided a solid support for using SO up-state peaks as the center of transformation. 2) We developed a method to standardize the circular-linear correlation coefficient by transforming the sampling distribution of correlation to be normally distributed and centered at 0 under null, which takes into ac-

count of the sample size to eliminate the overestimation of effect sizes, and becomes comparable with Pearson's correlation to enhance the comparability across studies (see Section 5.5.2).

In reporting results, despite the tendency to introduce multiple comparison issues due to the presence of multiple time points and electrodes during PSG recording[133], less than half of the included studies conducted corrections for multiple comparisons. We suggest that researchers adopt a combination of cluster-based permutation, hierarchical models, surrogate testing[176], and ROI to address the complexity of data structures across temporal, spectral, and spatial domains, and report both corrected and uncorrected results. Furthermore, we found that the majority of studies tend to selectively report results for significant electrodes, memory types, or coupling measures. These practices increase the likelihood of false positives and the overestimation of effect size[134]. We advocate for the standardization in reporting **1.** all three coupling metrics, including coupling phase, strength, and prevalence, **2.** their interactions with each other, and **3.** their associations with memory performance. Each metric captures a distinct property of the coupling process and may interact with one another[54], so it is necessary to provide a more comprehensive understanding of the coupling mechanism. We suggest that researchers should at least report exact values of effect sizes (e.g. correlation coefficients, standardized mean differences, or odds ratios), sample sizes, test statistics, p-values, and standard errors or confidence intervals for both significant and insignificant results to allow effective comparisons between studies. While only 2 out of the 23 studies included in our analysis disclosed all processed data and analysis code in their publication, we appreciate the responses of almost all authors who provided valid data or clarifications upon email requests that helped us mitigate heterogeneity and publication bias in subsequent analyses.

3.5 Limitations and future research directions

The between-study discrepancy in measuring cross-frequency coupling presents a potential challenge to the comprehensive inclusion of studies. Due to methodological disparities and the limited prevalence of PETH studies, our current analysis only included studies using the preferred phase to measure the precision of coupling. Additionally, due to the expectation of high heterogeneity, we excluded gray literature without peer-review and memory measures with low comparability to other studies, such as the targeted memory reactivation[135]. While the inclusion of a restricted number of published studies may introduce a potential publication bias, we effectively addressed this concern through data requests, focal and sensitivity analysis, and meta-regression, finally constraining the overall risk of bias of most studies to a small to moderate level.

One of the advantages of our study is attributable to a substantial number of effect sizes and sample sizes. However, only a few studies reached a sample size of 30-50. In addition, our analysis focused on studies with healthy populations. Future research should consider adopting more homogeneous analysis methods and openly sharing processed sleep and memory data, thereby aggregating sufficiently large and representative sample sizes in larger-scale studies. Moreover, our findings of no differences in the coupling-memory association between naps and overnight sleep suggest that naps paradigms may be efficient and thereby allow for a larger number of subjects.

Besides our suggestion for the classification of memory types and SP frequency, our results offer insights in studying causal sequences through more precise detection and stimulation of neural oscillations, and offer insights for enhancing consolidation under aging or pathological conditions. We believe tracing dynamic distributions of SO-SP coupling and considering memory type and age-related regional changes would be intriguing, and might be crucial for the understanding of the selective mediation of the spike-timing-dependent synaptic plasticity. High-density PSG has been instrumental in enhancing spatial sampling density for localizing SOs, subcortical areas, and even analyzing intra- and inter-regional PAC[17,137–143]. Moreover, integrating techniques with higher spatial resolutions, such as MRI and iEEG, will enhance the interpretative capacity of PSG results.

Finally, efforts must be made in future sleep research towards open science. To the best of our knowledge, as the first meta-analysis conducted on cross-frequency couplings, we have made our analysis code and data publicly available. Our posterior parameters predicted for each coupling measure can serve as priors in future meta-analyses, allowing for the integration of new data to update the model. We recommend that future sleep research should complete pre-registration on platforms such as OSF, and clearly delineate pre-registered analyses from exploratory analyses in their reports to enhance methodological consistency and reproducibility. We also encourage original studies to share individual-level data and code. These efforts will contribute to enhancing reproducibility and comprehensive inclusion of results in future meta-analyses.

4 Conclusion

This meta-analysis revealed the crucial role of the precise and strong SO-fast SP cross-frequency coupling in promoting memory consolidation with data demonstrating sufficient statistical power for the first time. By aggregating effect sizes of associations between four commonly used coupling measures and memory retention performance, we observed a very strong evidence that the precision and strength of coupling can both predicts enhanced memory retention scores across almost all conditions as a general physiological mechanism, while fast SP amplitudes detected in the frontal lobe also showed associations with memory consolidation. Although the effect sizes of main models are relatively small, our moderator analyses demonstrated the dynamic nature of coupling-memory relationships. Stronger associations were observed in subgroups including young adults, frontal regions, fast spindles, and declarative memory, all of which have historically demonstrated relationships with better memory performance. Therefore, our results provide important information regarding complex memory consolidation mechanisms.

This evidence provides insights for future research on how the contribution of coupling to memory consolidation is modulated by behavioral and physiological factors, as well as meaningful information for the simultaneous implementation of MRI and electrical stimulation in the SO-SP coupling analysis. We believe that SO-SP coupling can offer more precise predictions for memory consolidation within the context of spatiotemporal specificity. This meta-analysis emphasizes the need for standardization and replication studies and points to those measures to focus on for future research in the field.

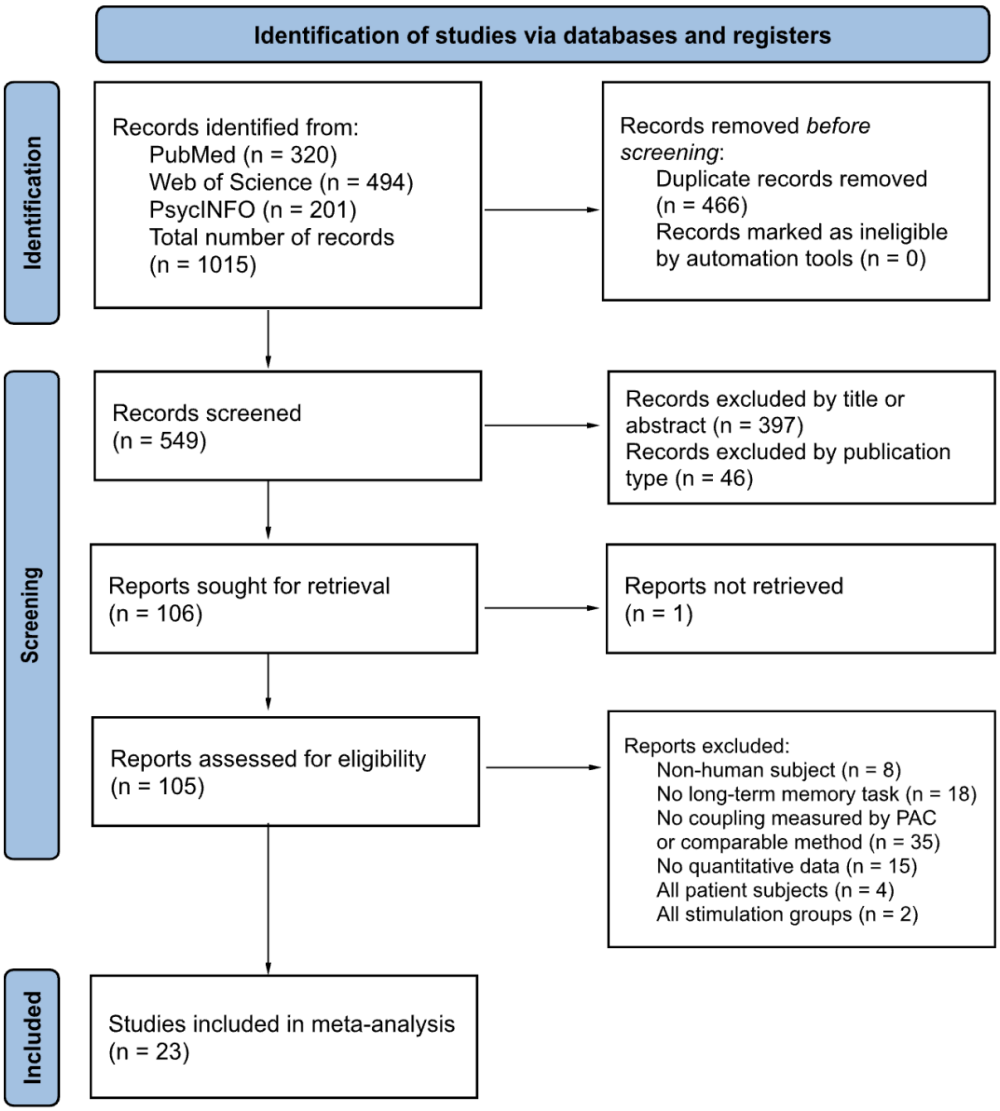

Figure 9: PRISMA flow diagram of literature search, screening, and inclusion for systematic review and meta-analysis

5 Method

5.1 Literature search

The retrieval and screening of relevant studies were conducted in accordance with the 2020 PRISMA statement[144]. A comprehensive literature search was performed in three databases based on the retrieval qualities evaluation[145,146]: PubMed, Web of Science, and PsycINFO, covering the period up to July 1, 2023. Only studies which have undergone peer review and have been published were included in the meta-analysis. Boolean operators were utilized to combine the following search terms: (sleep OR nap) AND (slow oscillat* OR slow wave OR sleep oscillat* OR slow-wave OR SO) AND (spindle OR sigma OR SP) AND (coupl* OR pair* OR lock* OR coordinat* OR interact* OR synchro*) AND (motor learning OR memory OR cogniti*). After filtering the papers based on inclusion criteria, a citation search was conducted and all papers identified from article citations underwent manual screening.

5.2 Inclusion criteria

We applied the following criteria[147] to identify articles eligible for the meta-analysis as of July 1, 2023. Studies were included if the study 1) had memory encoding task before and recall measures after a sleep interval; 2) was published in English; 3) measured at least one of the standard SO-SP coupling measures (coupling phase, strength, percentage, and/or SP amplitude) by phase-amplitude coupling (PAC) or comparable meth-

ods; 4) assessed the correlation between memory retention and SO-SP coupling during N2 and/or SWS stage(s); 5) was published as an original research paper in a peer-reviewed journal. Studies were excluded if they 1) only measured SO-SP coupling influenced by the impact of medication or interventions (e.g. brain stimulation method) with no control condition; 2) only included non-human subjects, patient groups, or clinical groups (no control healthy group); 3) was a group under the wake condition. If a study included both groups that meet the inclusion and exclusion criteria and those that do not, only data from the group(s) that meet the criteria (e.g., control condition, control group) were included in the following analysis.

5.3 Study selection

For the study selection, data retrieved from the three databases listed above were imported into EndNote 21 for automatic and manual duplication removal. The screen process was reported in Figure 9. After preliminary screening for titles, abstracts and article types, 105 original research papers were screened by independent full-text review. Reviewers discussed all papers with discrepancies regarding their inclusion and ultimately achieved consistency. Finally, we included 23 eligible studies into the meta-analysis, comprising a total of 297 effect sizes from four types of coupling measurements with 730 samples.

5.4 Data extraction

In the studies included in the meta-analysis, relevant effect sizes and specified study characteristics were extracted by two reviewers independently in accordance with PRISMA guidelines[144] to minimize the bias introduced by subjective judgments. The consistency of data extraction between reviewers is 99.6%, and all discrepancies were resolved through discussion. Four specific measures of SO-SP coupling (Figure 1) were examined for their association with memory retention performance: 1) preferred coupling phase; 2) SP peak-to-trough amplitude; 3) coupling strength; and 4) coupling percentage. After extracting effect sizes, phase-radian alignment was corrected (Figure 1C). Extracted study characteristics included: 1) sample size, age, gender distribution; 2) pre-specified moderators (see Section 5.6 for more details); 3) publication details. A summary of main information for each study is outlined in Table 6. The detailed data for each effect size can be found in Table S2.1-S2.5 and the publicly available repository: https://osf.io/9mh5d/. The overall quality of studies is assessed in accordance with the Robins-I[juni2016risk, 109] and NIH criteria[148], with adjustments (detailed in Table S3.1) made to accommodate the specific attributes of the study.

Due to relatively large heterogeneity and discrepancy in methodology and results reported between studies, as well as the non-parametric nature of the direction in circular-linear correlations, we improved comparability and credibility across studies by requesting both unreported effect sizes and processed individual-level memory and physiological data to reduce the publication bias. If a study measured relevant sleep and memory features but either: 1) did not report any extractable or convertible effect sizes; 2) reported imprecise p-values due to insignificance; or 3) reported only one effect size for cross-group data; 4) reported effect sizes that were not comparable to other studies, coupled with the absence of data in supplementary material, we requested missing data or processed individual data from corresponding authors via email. 21 out of 23 authors responded positively and provided the requested data. Studies without a response after two months or declined our requests were excluded or partially excluded from the meta-analysis for the part where insufficient information was provided for calculating effect sizes. For studies reporting correlations only through scatterplots, linear correlations were estimated using software from the ShinyDigitise package[149,150] in R, while circular-linear correlations were first assessed by extracting estimated individual data from the plots by the online software WebPlotDigitizer[151], followed by the circular linear correlation analysis in R. Additionally, key missing study characteristic data was requested via email at the same time to broaden the scope of moderator analysis.

5.5 Effect size calculations

To standardize effect sizes for comparability across studies, we chose to standardize the bounded circular-linear correlation coefficient[132,152] to unbounded r_z to examine the association between coupling phase (in radians) and memory consolidation. We used the Pearson product-moment correlation coefficient r[153] (hereinafter referred to as Pearson's r) and transformed fisher's z to report the linear relationships between SP amplitude (in μV), coupling strength (mean vector length or modulation index), coupling percentage (%), with memory consolidation separately. All transformable effect size measures including t-statistic, p-value, β statistic, and η^2, were extracted and converted to Pearson's r. In cases where effect measures were reported as non-parametric correlation including Spearman's rho (ρ, i.e.; r_s) and no author response could reanalyze for parametric correlation, we used it as an imperfect estimation of Pearson's r[154] and excluded from sensitivity analyses.

5.5.1 Pearson product-moment correlation coefficient

The constrained range of values between -1 and 1 measured by Pearson's r restricts the selection of an ideal unbounded prior distribution for Bayesian hierarchical models[155] (discussed in Section 5.7). The deviation of its sampling distribution from the assumptions of normal, especially when the effect size is large, might provide inaccurate estimation for the sampling variance[156]. To address this limitation, we used the metafor package[157] in R to transform Pearson's r into normalized and unbounded Fisher's z by formula[156]:

$$z = \frac{1}{2}\ln(\frac{1+r}{1-r}) = tanh^{-1}r$$

When reporting and interpreting the results, we reversed the transformation from Fisher's z back to Pearson's r to present each estimated effect size and credible interval (CrI) by formula:

$$r = (\frac{e^{2z}-1}{e^{2z}+1})$$

5.5.2 Standardized circular-linear correlation coefficient

Meanwhile, the circular-linear correlation coefficient r (hereinafter referred to as circlin r) is a type of pearson correlation coefficient (PCC) distributed from 0 to 1 without direction, achieved by transforming the phase into $\sin\theta$ and $\cos\theta$ to create linear parameters[132,152], to assess the relationship between a random unit vector and another linear random variable:

$$Circlin\ r = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2 \cdot r_{12} \cdot r_{13} \cdot r_{23}}{1 - r_{23}^2}}$$

$$r_{12} = \text{corr}(x, \cos\theta),\ r_{13} = \text{corr}(x, \sin\theta),\ r_{23} = \text{corr}(\cos\theta, \sin\theta)$$

Since there is no existing method to standardize the circlin r, we developed a approximation approach to transform the bounded and non-linear distribution to unbounded standardized normal distribution $\mathcal{N}(0,1)$. The verification and performance can be found in Supplemental Material 12.

The population distribution of circular linear correlation ρ with n samples can be approximate to a chi-square distribution with 2 degree of freedom[158], χ_2^2. We derived that the circular-linear correlation coefficient has the following mean and variance in an approximate form and verified in R:

$$E(\rho) = \sqrt{\frac{\pi}{2n}},\ Var(\rho) = \frac{4-\pi}{2n},\ 0 \le r \le 1,\ n \ge 2$$

We can observe that different than pearson's r, which always has a population mean $H_0 : \rho = 0$ under the null hypothesis, the circlin r has a population mean between 0 and 1, which tends to approach 1 and display left-skewness as the sample size n decreases, while approach 0 and exhibit right-skewness as the n increases (also see Figure S12.1). The sampling distribution of correlated circlin r also shows the same property of skewness when approaching lower and upper bounds (Figure S12.2-12.3). Therefore, it is clearly not appropriate to use raw coefficients in the meta-analysis.

To transform its sampling distribution to be unbounded, normally distributed and centered at 0 for our meta-analysis, we first scaled the circlin r as a form of $n\,r^2$ to approximate the χ_2^2 distribution[158]. Here the sample size has been weighted thus eliminate the bias introduced by large r under small sample sizes. We can find the quantile of circlin r in the upper tail.

$$P(nr^2 \le n\rho^2) = 1 - e^{-nr^2/2}, \quad 0 \le r < 1$$

Finally, by transforming non-normal chi-square deviates to standardized normal distribution and scale the distribution by the sample size, circlin r can be transformed into a standardized normal scale r_z:

$$\Phi(r_z\sqrt{n}) = P(nr^2 \le n\rho^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{r_z\sqrt{n}} e^{-y^2/2}\,dy$$

The sampling distribution of transformed circlin r_z follows a approximate normal distribution when $n > 15$. When $n < 15$ it would be slightly right skewed. Its population distribution can be approximate to Pearson's r and it is bounded between -1 (null) and 1 (alternative). The interpretation the strength of r_z is comparable with Pearson's r (see Table 6), and now r_z can also be transformed to Fisher's z and include in the meta-analysis. It is worth noting that small effect sizes are common in neuroscience and meta-analyses due to the complexity of underlying mechanisms and the presence of numerous confounding variables and hierarchical structures, so small correlations may carry substantial and meaningful information to interpret. Monte Carlo simulation results and the code used for transformations were also reported in Supplemental Material 12.

Table 6: Interpretation of the standardized circular-linear correlation coefficient

r_z	**Strength**
< 0	No Effect
0	Null
0.1	Small
0.3	Moderate
0.5	Strong

5.6 Methodological characteristics and moderators

In previous studies, researchers have applied different experimental designs to investigate coupling and summarized divergent conclusions regarding each coupling parameter. To develop a more systematic understanding about the heterogeneity of approaches, six moderators were selected in the moderator analysis. Table 1 provides a complete list of moderators in each study, and categories that define moderators are listed below, the range represents the number of effect size included in the analysis across four types of coupling measures:

1. **Sleep Stage**: Sleep stages were grouped by N2 (nREM2; k = 11-18) and SWS (slow-wave sleep; k = 9-30), while uncategorized sleep stages were encoded as ”mixed” (k = 23-49).

2. **Sleep Phase**: Nighttime sleep lasting typically more than 6 hours was encoded as “overnight” (k = 32-73), while short naps lasting around 1-2 hours during the early morning or afternoon were classified in the ”nap” group (k = 11-27).

3. **Age**: Due to the limited number of studies on children, adolescents and older adults, age was encoded as a continuous variable based on the mean age of each study for the meta-regression.

4. **PSG Channel**: Effect size has been categorized into three clusters based on the cortical area under PSG electrodes, including frontal (F3, Fz, and F4; k = 15-36), central (C3, Cz, and C4; k = 13-29), and posterior (P3, Pz, P4, O1, and O2; k = 8-22), and mixed (k = 3-7). Other electrodes within the same area but exhibiting excessive deviation from the midline or containing insufficient information have been excluded (E.g. C1, PO8). Sleep parameters were computed by averaging SP amplitude, coupling phase, strength, and percentage across electrodes within each cluster.

5. **Spindle Type**: To include a wider range of comparable studies, we specified frequency boundaries for fast SPs (12-16 Hz; k = 22-43) and slow SPs (8-12 Hz; k = 16-28). Effect sizes that transversed frequency boundaries or reported together are classified as “mixed” (k = 16-28).

6. **Memory Type**: We classified comparable memory tasks into four domains presented in Table 2. Verbal memory can also be interpreted as neutral non-spatial memory, in contrast to the other two types of declarative memory.

Table 7: Summary of memory tasks

Memory Task Domain	Memory Task Modality
Verbal Tasks (n = 12)	Word-Pair (WP)
	Word List
	Novel Metaphor
	Word-Image Pair (WIP)
Emotional Task (n = 3)	Picture-recognition (IMG)
Spatial Tasks (n = 4)	Spatial Memory
	2D Object Location (2DL)
	Visuo-spatial (VS)
Motor Tasks (n = 6)	Motor Sequence (MST)
	Gross-motor Juggle
	Mirror-tracing Task (MTT)
	Visuomotor Adaptation (VMA)

5.7 Statistical analysis

We chose to fit overall and subgroup (moderator) models using Bayesian hierarchical random-effects and mixed-effects models, respectively, due to the common occurrence of multiple effect sizes for the same set of participants reported within the same study, violating the assumption of independence of effect sizes, as well as the limited and unequal group size and potential high heterogeneity. In comparison to frequentist models, Bayesian models incorporate prior probabilities for each parameter by considering likelihood information (i.e., the effect sizes we extracted) to establish a model for predicting the posterior probability distribution of parameters, which provide transparent inferences with lower risks of false positives. In the posterior distribution, confidence intervals (CIs) reported by frequentist methods are replaced by credible intervals (CrI), which can be interpreted as “there is a 95% probability that the parameter lies within the interval”, thereby providing a more precise prediction of true probabilities. Additionally, Bayesian methods could more effectively build models that account for multiple sources of heterogeneity and allow the analysis of the impact of moderator variables representing different measures and participant groups on the pooled effect size[159]. As Kruschke & Liddell[174] described, the shrinkage property of Bayesian models helps prevent false alarms by pulling extreme values toward the group mean, which is especially valuable when accounting for potentially outlying estimates. This explains the observed differences between model distributions and the raw effect sizes in our forest plots. All statistical analyses for Bayesian models were conducted in R using the brms package[160] along with supporting packages bayesplot[161], metaviz[162], and customized codes for visualization. Contrary to the misconception that Bayesian models are overly complex or opaque, they are increasingly valued for their accuracy and transparent inferences[174]. We also recognize that some researchers may prefer frequentist approaches. To support transparency and comparability, we also provided the traditional meta-analytic results in Supplementary Material 10, which demonstrate consistency with our Bayesian findings.

For each overall and subgroup model (Table 4), we applied the Markov chain Monte Carlo (MCMC) method, a general family of algorithms used to approximate complex probability distributions, with the no-U-turn Hamiltonian Monte Carlo (HMC) sampler, the default algorithm in Stan, to set up four chains, with each chain undergoing 12,000 iterations (including 2,000 warm-ups), the minimum requirement for calculating accurate Bayes Factors (BF)[163], which reflects how much more likely the data are under one hypothesis. The target average acceptance probability was set to 0.99 with the maximum tree depth to 15 to ensure robustness in the posterior distribution.

Convergence of the MCMC was checked following the suggestion of WAMBS-Checklist[164] through 1. graphical posterior predictive checks that assess how well the model predicted data replicated the observed data; and 2. trace plots and 3. Gelman and Rubin diagnostic that assess the convergence of Markov chains. Ideally, the posterior distribution should overlap with the distribution of the test data generated, the trace plot distribution should resemble a uniformly undulating wave with high overlap between chains (see Figure S4.2), and the Potential Scale Reduction Factor (Rˆ) should be less than 1.1. Autocorrelation plots were used to ensure low temporal dependency between successive samples. Example of diagnostic plots were reported in the Supplemental Material S4.1-4.3. Any non-convergence at the aforementioned stages led to reconfiguration of chains and iterations for analysis. Estimation of each intercept, moderator and heterogeneity in the posterior distribution were extracted, and the estimation of mean, distribution and 95% credible interval of each effect size and the pooled effect size were reported in forest plots.

5.7.1 Overall model

We fitted random-effects models that predict the relationships between coupling phase, SP amplitude, coupling strength, and coupling percentage with memory consolidation, considering only heterogeneity and sampling errors. The three-level Bayesian model superimposed random effects for sampling error (first level), between-study heterogeneity (second level), and within-study heterogeneity (third level):

$$z \mid \sigma_z \sim 1 + (1|\text{Effect}) + (1|\text{Study})$$

By introducing priors and likelihood information to model intercept and heterogeneity parameters, our random-effects model can be represented using the following formula:

$$\hat{\theta} \sim N(\theta_{\text{ij}}, \sigma_{\text{ij}}^2)$$
$$\theta_{\text{ij}} \sim N(\theta_i, \sigma_i^2)$$
$$\theta_i \sim N(\mu_\theta, \tau^2)$$
$$\hat{\theta} \sim N(\mu_\theta, \sigma_{\text{ij}}^2 + \sigma_i^2 + \tau^2)$$
$$(\mu_\theta, \tau^2, \sigma_i^2) \sim p(\cdot), \quad \tau^2, \sigma_i^2 > 0$$
$$\mu_\theta \sim N(z_{\text{mean}}, z_{\text{sd}})$$
$$\tau, \sigma_i \sim HC(z_{\text{lp}}, z_{\text{sp}})$$

where z represents the fisher's z-transformed correlation coefficient, τ^2 represents between-study heterogeneity, σ_i^2 represents within-study heterogeneity, and ${\sigma_{ij}}^2$ represents sampling error. Since the distribution of Fisher's z and heterogeneity[165] all follow certain probability functions, and non-informative priors could not provide reasonable estimates, we selected weak informative priors in the model to tradeoff between allowing collected real data to influence posterior distributions and excluding extreme outliers. For the intercept of all models, we chose a standardized normal distribution $\mathcal{N}(0,1)$ as the prior distribution, assuming no effect to control false positives.

The prior for heterogeneity used the Half-Cauchy distribution, known for its heavy-tailed property, improving it to be more tolerant of high heterogeneity[166]. Also, half-cauchy truncated at 0, which is consistent with the fact that heterogeneity cannot be less than 0. For between-study heterogeneity τ^2, we obtained a mean between-study heterogeneity of $\mu = 0.13$ from a dataset of heterogeneity reported in 498 correlational meta-analyses[112] using Fisher's z or Pearson's r as the measure of effect size, which represent a replication of the methodology by McKinney et al.[167] regarding prior selection. Additionally, due to differences in measurement methods and memory tasks across studies, we held a prior belief in some degree of between-study heterogeneity. For reasons discussed above, we set the prior using half-Cauchy distribution $HC(0, 0.5)$ by extending the scale parameter from 0.13 to 0.5 to account for more extreme heterogeneity. Lastly, the prior for within-study heterogeneity is also set as to be $HC(0, 0.5)$ to balance the uncertainty arising from different measurement approaches or sample groups. This combination of priors reduced the risk of overestimation, accounted for substantial uncertainty, and increased transparency by explicitly encoding all assumptions.

5.7.2 Moderator models

For moderator analysis, we introduced continuous and categorical variables into the overall model as fixed effects, forming a following mixed-effects model:

$$z \mid \sigma_z \sim 1 + \text{Moderator} + (1|\text{Effect}) + (1|\text{Study})$$
$$\hat{\theta} \sim N(\mu_\theta + \sum_{k=1}^{n} \beta_k \cdot X_{ijk}, \sigma_{\text{ij}}^2 + \sigma_i^2 + \tau^2)$$

in which X denotes random variables for moderators, k is the random variable of moderators, and n represents the number of moderators specified in the model. In the subgroup model, we chose to use the weak informative prior $\mathcal{N}(0,1)$ for each moderator as fixed effects to account for the uncertainty in the magnitude of the aggregate effects caused by moderators. In the meta-regression model accounting for the effect of age, we chose $\mathcal{N}(0, 0.1)$ to model the change over years. Except for the exploratory analysis, all statistical models used in data analysis are summarized in Table 6.

To assess directional hypotheses proposed for each moderator model, we conducted non-linear hypothesis testing within the Bayesian framework, obtaining Bayesian Factors (BF_{10}, defined as Evidence Ratio (ER) in the package *brms*) for different levels within each model, which performed tests of evidence for the ratio of marginal likelihoods between two levels. In addition, we conducted hypothesis testing on the overall model compared to null. Table 8 presents the strength of evidence represented by different Bayesian Factors, which is analogous to the one-tailed t-test in the interpretation. In directional hypotheses, $BF_{10} = 3$ means the hypothesis is 3 times more likely than alternative. Additionally, we used the posterior probability, also defined as credibility score[168], to report the percentage of posterior samples consistent with the direction of the hypothesis. A posterior probability of 1 indicates that all posterior sample draws align with the hypothesis.

Table 8: Interpretation of Bayes Factor (BF_{10}) for the strength of evidence

BF_{10}	**Direction**
< 0.1	Strong, Favor alternative
$0.1 - 0.33$	Moderate, Favor alternative
$0.33 - 1$	Weak, Favor alternative
$1 - 3$	Weak, Favor hypothesis
$3 - 10$	Moderate, Favor hypothesis
> 10	Strong, Favor hypothesis

The predictive power of models were compared in pairs by computing model weights via leave-one-out (LOO) stacking of posterior predictive distributions[169], including the comparison between the overall model and a single moderator model, as well as a single moderator model and a moderator model with an additional moderator and interaction term. Models with additional predictor(s) that had higher stacked weights indicate that it enhanced the overall predictive power of the model. This approach could help identify moderators that influence the correlation between SO-SP coupling and memory retention. Model weights have also been computed via Pseudo-BMA method as the controlled analysis.

5.7.3 Sensitivity analysis

Finally, in sensitivity testing, we conducted separate assessments for 1) publication bias; 2) focal models excluding studies that introduced significant heterogeneity, and 3) the impact of priors. Regarding publication bias, although efforts were made to minimize potential controllable biases through the data request (especially for effect size computed but not reported by studies), effect size transformation, and multi-level models, differences in data analysis approaches and experimental conditions could not be fully addressed at the meta-analysis level. Therefore, we additionally used a frequentist approach by using Restricted Maximum Likelihood (REML) using the metafor package[157] to fit three-level models to extract data for generating funnel plots and quantifying publication bias. Frequentist results of each model were also reported in Supplemental Material 10 for readers unfamiliar with Bayesian statistics. Beside the Egger's regression test and rank correlation test, we chose counter enhanced funnel plot[170] and superimposed the Egger's regression line. We also performed the time-lag bias test to quantify the impact of heterogeneity and effects of the year of publication.

In the focal analysis, we excluded studies focused only on non-declarative memory tasks, those with a high risk of bias (see Supplemental Material 3) or significant methodological differences, those only tested post-sleep memory retention, or with a Pareto k diagnostic value larger than 0.7. We also removed studies that adopted modulation index as their measures in the focal model of coupling strength. In addition, the method of prior sensitivity test, as well as the posterior predictive check, were reported in the Supplemental Material 4.

5.7.4 Phase spatiotemporal analysis

The non-negative and non-linear distribution of circular linear correlation makes it impossible to determine the direction of correlation. This is one of the reasons why we have requested individual-level processed data from most of authors of included studies, in addition to requesting the effect size. By standardizing memory scores of each individual in each study using z-scores, we overlaid the data from most of comparable studies to fit and visualize the nonlinear relationship between the coupling phase and memory consolidation using the best fit second-degree quadratic line for combination of each SP type and channel location. It could effectively improve the interpretation of nonlinear relationships and circular direction of the SO-SP coupling phase across studies.

For the comparison of coupling phase across PSG channel clusters, we used the Bayesian circular mixed-effects model[171] with 12,000 iterations to account for repeated measurement and differences of sample sizes between channels. PSG channel location has been set as a fixed effect, while subject has been taken into account as a random effect. 95% credible interval and the posterior distribution of each circular direction was reported in the circular plot to compare the timing of occurrence of SO-SP coupling across different cortical areas. Similar to the main analysis, the Bayes factor and posterior probability has been used to evaluate evidence for the circular mean difference. In addition, the inconsistency of coupling phase has been assessed by the Rayleigh test.

Data availability

All effect size level and study level data used for analysis are publicly released in the Open Science Framework repository: https://osf.io/9mh5d/, and shared in the supplemental dataset.

Code availability

All code for data analysis, custom functions, and Bayesian models are publicly released in the Open Science Framework repository: https://osf.io/9mh5d/

Acknowledgements

We would like to express our sincere gratitude to all authors of included studies who have generously shared demographic, physiological, behavioral data, or effect sizes, or clarified experimental methods with us through email communication. Work of R.M.C.S was funded in part by NIH R01 AG040133.

Declaration of competing interest

The authors have declared that no competing interests exist.

Table 9: Main characteristics for each study included in the meta analysis

Author (Year)	N	M_{age}	Stage	Condition	PSG Channels	Spindles	Task	Memory Type	Measures Included
Bastian et al. (2022)	15	23.3	N2, SWS	nap	frontal, central	fast, slow	Spatial	declarative	phase, amp, str, pct
Cox et al. (2018)	24	30.2	N2, SWS	overnight	frontal, central, posterior	fast, slow	MST	motor	phase, str
Denis et al. (2022)	31	22.3	N2, SWS	overnight	frontal, central	fast	IMG	declarative	phase, str, pct
Denis et al. (2021)	34	22	N2, SWS	nap	frontal, central, posterior	fast, slow	WP	declarative	phase, amp, str, pct
Donnelly et al. (2022)	16	14.1	N2, SWS	overnight	frontal, central, posterior	fast, slow	2DL	declarative	phase, amp, str
Hahn et al. (2020)	33	CH: 9.5 AD: 16	SWS	overnight	frontal, central, posterior	fast, slow	WP	declarative	phase, amp, str, pct
Hahn et al. (2022)	42	AD: 12.9 YA: 22.0	SWS	overnight	frontal, central, posterior	fast, slow	Juggle	motor	phase, amp, str
Halonen et al. (2021)	27	22	N2, SWS	overnight	frontal, central	fast	Metaphor	declarative	phase, str, pct
Halonen et al. (2022)	151	17	N2, SWS	overnight	frontal, central	fast, slow	IMG	declarative	phase, amp, str, pct
Helfrich et al. (2018)	52	YA: 20.4 OA: 73.8	SWS	overnight	frontal	fast	WP	declarative	phase, str
Kurz et al. (2021)	19	11.24	N2, SWS	overnight	frontal, central, posterior	fast, slow	IMG	declarative	phase, amp, str, pct
Kurz et al. (2023)	30	11.43	N2, SWS	overnight	frontal, central, posterior	fast, slow	Word List	declarative	phase, amp, str
Ladenbauer et al. (2021)	43	YA: 23 OA: 66	N2, SWS	nap	frontal, central	fast	WP, VS	declarative	phase, amp, str
Mikutta et al. (2019)	20	27.1	N2, SWS	overnight	central	fast, slow	Word List, MTT	declarative, motor	phase, amp, str
Mylonas et al. (2020)	28	30	N2	overnight	central	fast	MST	motor	phase, amp, str, pct
Mylonas et al. (2022)	14	13	N2	overnight	central	fast, slow	Spatial	declarative	phase, amp, str, pct
Nicolas et al. (2022)	24	21.9	N2, SWS	nap	frontal, central, posterior	fast	MST	motor	phase, amp
Niknazar et al. (2015)	28	22	N2	nap	central	fast	WP	declarative	phase, amp, str
Perrault et al. (2019)	16	23.4	N2, SWS	overnight	frontal, posterior	fast, slow	WP	declarative	phase, amp
Schreiner et al. (2021)	20	20.8	N2, SWS	nap	central	fast	WIP	declarative	phase, amp, str, pct
Solano et al. (2022)	10	24.3	N2, SWS	overnight	frontal, central, posterior	fast, slow	VMA	motor	phase, amp, str, pct
Weiner et al. (2023)	25	69.1	N2, SWS	overnight	frontal, central	fast, slow	WP	declarative	phase, amp, str
Zhang et al. (2020)	28	20.6	N2, SWS	overnight	frontal, central, posterior	fast, slow	WP	declarative	phase, str

Notes. *N* sample size, only includes groups included in the meta analysis; *n2* Stage 2 nREM sleep; *sws* slow wave sleep; *amp* spindle amplitude; *str* coupling strength; *pct* coupling percentage; *CH* Children; *AD* adolescents; *YA* young adults; *OA* older adults; PSG posterior channels include both parietal and occipital electrodes. Details about memory task types are listed in Table 8; additional detailed information and data are listed in Tables S2.1- S2.5 in the supplemental material.

References

1. Buzsaki, G. & Draguhn, A. Neuronal oscillations in cortical networks. *Science* **304,** 1926–1929 (2004).
2. Hyafil, A., Giraud, A.-L., Fontolan, L. & Gutkin, B. Neural cross-frequency coupling: connecting architectures, mechanisms, and functions. *Trends in neurosciences* **38,** 725–740 (2015).
3. Klinzing, J. G., Niethard, N. & Born, J. Mechanisms of systems memory consolidation during sleep. *Nature neuroscience* **22,** 1598–1610 (2019).
4. Clemens, Z. *et al.* Temporal coupling of parahippocampal ripples, sleep spindles and slow oscillations in humans. *Brain* **130,** 2868–2878 (2007).
5. Maingret, N., Girardeau, G., Todorova, R., Goutierre, M. & Zugaro, M. Hippocampo-cortical coupling mediates memory consolidation during sleep. *Nature neuroscience* **19,** 959–964 (2016).
6. Staresina, B. P. *et al.* Hierarchical nesting of slow oscillations, spindles and ripples in the human hippocampus during sleep. *Nature neuroscience* **18,** 1679–1686 (2015).
7. Gais, S., Mölle, M., Helms, K. & Born, J. Learning-dependent increases in sleep spindle density. *Journal of Neuroscience* **22,** 6830–6834 (2002).
8. Mölle, M., Marshall, L., Gais, S. & Born, J. Learning increases human electroencephalographic coherence during subsequent slow sleep oscillations. *Proceedings of the National Academy of Sciences* **101,** 13963–13968 (2004).
9. Mölle, M., Eschenko, O., Gais, S., Sara, S. J. & Born, J. The influence of learning on sleep slow oscillations and associated spindles and ripples in humans and rats. *European Journal of Neuroscience* **29,** 1071–1081 (2009).
10. Schmidt, C. *et al.* Encoding difficulty promotes postlearning changes in sleep spindle activity during napping. *Journal of Neuroscience* **26,** 8976–8982 (2006).
11. Solano, A., Riquelme, L. A., Perez-Chada, D. & Della-Maggiore, V. Visuomotor adaptation modulates the clustering of sleep spindles into trains. *Frontiers in Neuroscience* **16,** 803387 (2022).
12. Holz, J. *et al.* EEG sigma and slow-wave activity during NREM sleep correlate with overnight declarative and procedural memory consolidation. *Journal of sleep research* **21,** 612–619 (2012).
13. Nicolas, J. *et al.* Sigma oscillations protect or reinstate motor memory depending on their temporal coordination with slow waves. *eLife* **11,** e73930 (2022).
14. Kumral, D., Matzerath, A., Leonhart, R. & Schönauer, M. Spindle-dependent memory consolidation in healthy adults: A meta-analysis. *Neuropsychologia,* 108661 (2023).
15. Kurdziel, L., Duclos, K. & Spencer, R. M. Sleep spindles in midday naps enhance learning in preschool children. *Proceedings of the National Academy of Sciences* **110,** 17267–17272 (2013).
16. Rodheim, K., Kainec, K., Noh, E., Jones, B. & Spencer, R. M. Emotional memory consolidation during sleep is associated with slow oscillation–spindle coupling strength in young and older adults. *Learning & Memory* **30,** 237–244 (2023).
17. Massimini, M., Huber, R., Ferrarelli, F., Hill, S. & Tononi, G. The sleep slow oscillation as a traveling wave. *Journal of Neuroscience* **24,** 6862–6870 (2004).
18. Kurth, S. *et al.* Traveling slow oscillations during sleep: a marker of brain connectivity in childhood. *Sleep* **40,** zsx121 (2017).
19. Malerba, P., Whitehurst, L. N., Simons, S. B. & Mednick, S. C. Spatio-temporal structure of sleep slow oscillations on the electrode manifold and its relation to spindles. *Sleep* **42,** zsy197 (2019).
20. Niethard, N., Ngo, H.-V. V., Ehrlich, I. & Born, J. Cortical circuit activity underlying sleep slow oscillations and spindles. *Proceedings of the National Academy of Sciences* **115,** E9220–E9229 (2018).
21. Contreras, D., Destexhe, A., Sejnowski, T. J. & Steriade, M. Spatiotemporal patterns of spindle oscillations in cortex and thalamus. *Journal of Neuroscience* **17,** 1179–1196 (1997).
22. Kim, U., Bal, T. & McCormick, D. A. Spindle waves are propagating synchronized oscillations in the ferret LGNd in vitro. *Journal of Neurophysiology* **74,** 1301–1323 (1995).
23. Marshall, L., Helgadóttir, H., Mölle, M. & Born, J. Boosting slow oscillations during sleep potentiates memory. *Nature* **444,** 610–613 (2006).
24. Neske, G. T. The slow oscillation in cortical and thalamic networks: mechanisms and functions. *Frontiers in neural circuits* **9,** 88 (2016).
25. Oyanedel, C. N., Durán, E., Niethard, N., Inostroza, M. & Born, J. Temporal associations between sleep slow oscillations, spindles and ripples. *European Journal of Neuroscience* **52,** 4762–4778 (2020).
26. Steriade, M. *et al.* The corticothalamic system in sleep. *Front Biosci* **8,** d878–d899 (2003).

27. Rosanova, M. & Ulrich, D. Pattern-specific associative long-term potentiation induced by a sleep spindle-related spike train. *Journal of Neuroscience* **25,** 9398–9405 (2005).

28. Seibt, J. *et al.* Cortical dendritic activity correlates with spindle-rich oscillations during sleep in rodents. *Nature communications* **8,** 684 (2017).

29. Lindemann, C., Ahlbeck, J., Bitzenhofer, S. H., Hanganu-Opatz, I. L., *et al.* Spindle activity orchestrates plasticity during development and sleep. *Neural Plasticity* **2016** (2016).

30. Martin, S. J., Grimwood, P. D. & Morris, R. G. Synaptic plasticity and memory: an evaluation of the hypothesis. *Annual review of neuroscience* **23,** 649–711 (2000).

31. Miyamoto, D., Hirai, D. & Murayama, M. The roles of cortical slow waves in synaptic plasticity and memory consolidation. *Frontiers in neural circuits* **11,** 92 (2017).

32. Steriade, M. & Timofeev, I. Neuronal plasticity in thalamocortical networks during sleep and waking oscillations. *Neuron* **37,** 563–576 (2003).

33. Canolty, R. T. & Knight, R. T. The functional role of cross-frequency coupling. *Trends in cognitive sciences* **14,** 506–515 (2010).

34. Palva, J. M., Palva, S. & Kaila, K. Phase synchrony among neuronal oscillations in the human cortex. *Journal of Neuroscience* **25,** 3962–3972 (2005).

35. Canolty, R. T. *et al.* High gamma power is phase-locked to theta oscillations in human neocortex. *Science* **313,** 1626–1628 (2006).

36. Fell, J. & Axmacher, N. The role of phase synchronization in memory processes. *Nature reviews neuroscience* **12,** 105–118 (2011).

37. Jensen, O. & Colgin, L. L. Cross-frequency coupling between neuronal oscillations. *Trends in cognitive sciences* **11,** 267–269 (2007).

38. Born, J. & Wilhelm, I. System consolidation of memory during sleep. *Psychological research* **76,** 192–203 (2012).

39. Takehara-Nishiuchi, K. Neurobiology of systems memory consolidation. *European Journal of Neuroscience* **54,** 6850–6863 (2021).

40. Winocur, G. & Moscovitch, M. Memory transformation and systems consolidation. *Journal of the International Neuropsychological Society* **17,** 766–780 (2011).

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

42. Helfrich, R. F. *et al.* Bidirectional prefrontal-hippocampal dynamics organize information transfer during sleep in humans. *Nature communications* **10,** 3572 (2019).

43. Bastian, L. *et al.* Spindle–slow oscillation coupling correlates with memory performance and connectivity changes in a hippocampal network after sleep. *Human Brain Mapping* **43,** 3923–3943 (2022).

44. Cox, R., Mylonas, D. S., Manoach, D. S. & Stickgold, R. Large-scale structure and individual fingerprints of locally coupled sleep oscillations. *Sleep* **41,** zsy175 (2018).

45. Denis, D. *et al.* Sleep spindles preferentially consolidate weakly encoded memories. *Journal of Neuroscience* **41,** 4088–4099 (2021).

46. Donnelly, N. A. *et al.* Sleep EEG in young people with 22q11. 2 deletion syndrome: A cross-sectional study of slow-waves, spindles and correlations with memory and neurodevelopmental symptoms. *eLife* **11,** e75482 (2022).

47. Hahn, M. A., Heib, D., Schabus, M., Hoedlmoser, K. & Helfrich, R. F. Slow oscillation-spindle coupling predicts enhanced memory formation from childhood to adolescence. *eLife* **9,** e53730 (2020).

48. Halonen, R., Kuula, L., Antila, M. & Pesonen, A.-K. The overnight retention of novel metaphors associates with slow oscillation–spindle coupling but not with respiratory phase at encoding. *Frontiers in Behavioral Neuroscience* **15,** 712774 (2021).

49. Halonen, R., Kuula, L., Lahti, J., Räikkönen, K. & Pesonen, A.-K. The association between overnight recognition accuracy and slow oscillation-spindle coupling is moderated by BDNF Val66Met. *Behavioural Brain Research* **428,** 113889 (2022).

50. Helfrich, R. F., Mander, B. A., Jagust, W. J., Knight, R. T. & Walker, M. P. Old brains come uncoupled in sleep: slow wave-spindle synchrony, brain atrophy, and forgetting. *Neuron* **97,** 221–230 (2018).

51. Kurz, E.-M. *et al.* How do children with autism spectrum disorder form gist memory during sleep? A study of slow oscillation–spindle coupling. *Sleep* **44,** zsaa290 (2021).

52. Kurz, E.-M., Zinke, K. & Born, J. Sleep electroencephalogram (EEG) oscillations and associated memory processing during childhood and early adolescence. *Developmental Psychology* **59,** 297 (2023).

53. Mikutta, C. *et al.* Phase-amplitude coupling of sleep slow oscillatory and spindle activity correlates with overnight memory consolidation. *Journal of sleep research* **28,** e12835 (2019).

54. Weiner, O. M. *et al.* Slow oscillation-spindle cross-frequency coupling predicts overnight declarative memory consolidation in older adults. *European Journal of Neuroscience* **59,** 662–685 (2024).

55. Carbone, J. *et al.* The effect of zolpidem on targeted memory reactivation during sleep. *Learning & Memory* **28,** 307–318 (2021).

56. Kersanté, F., Purple, R. J. & Jones, M. W. The GABAA receptor modulator zolpidem augments hippocampal-prefrontal coupling during non-REM sleep. *Neuropsychopharmacology* **48,** 594–604 (2023).

57. Niknazar, M., Krishnan, G. P., Bazhenov, M. & Mednick, S. C. Coupling of thalamocortical sleep oscillations are important for memory consolidation in humans. *PloS one* **10,** e0144720 (2015).

58. Zhang, J., Yetton, B., Whitehurst, L. N., Naji, M. & Mednick, S. C. The effect of zolpidem on memory consolidation over a night of sleep. *Sleep* **43,** zsaa084 (2020).

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

60. Jiang, H., Bahramisharif, A., van Gerven, M. A. & Jensen, O. Measuring directionality between neuronal oscillations of different frequencies. *Neuroimage* **118,** 359–367 (2015).

61. Sauseng, P. & Klimesch, W. What does phase information of oscillatory brain activity tell us about cognitive processes? *Neuroscience & Biobehavioral Reviews* **32,** 1001–1013 (2008).

62. Mölle, M., Marshall, L., Gais, S. & Born, J. Grouping of spindle activity during slow oscillations in human non-rapid eye movement sleep. *Journal of Neuroscience* **22,** 10941–10947 (2002).

63. Muehlroth, B. E. *et al.* Precise slow oscillation–spindle coupling promotes memory consolidation in younger and older adults. *Scientific reports* **9,** 1940 (2019).

64. Tort, A. B. *et al.* Dynamic cross-frequency couplings of local field potential oscillations in rat striatum and hippocampus during performance of a T-maze task. *Proceedings of the National Academy of Sciences* **105,** 20517–20522 (2008).

65. Hülsemann, M. J., Naumann, E. & Rasch, B. Quantification of phase-amplitude coupling in neuronal oscillations: comparison of phase-locking value, mean vector length, modulation index, and generalized-linear-modeling-cross-frequency-coupling. *Frontiers in neuroscience* **13,** 421376 (2019).

66. Samiee, S. & Baillet, S. Time-resolved phase-amplitude coupling in neural oscillations. *NeuroImage* **159,** 270–279 (2017).

67. Tort, A. B., Komorowski, R., Eichenbaum, H. & Kopell, N. Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *Journal of neurophysiology* **104,** 1195–1210 (2010).

68. Mylonas, D. *et al.* The effects of eszopiclone on sleep spindles and memory consolidation in schizophrenia: a randomized clinical trial. *Neuropsychopharmacology* **45,** 2189–2197 (2020).

69. Ujma, P. P. Sleep spindles and general cognitive ability–A meta-analysis. *Sleep Spindles & Cortical Up States* **2,** 1–17 (2021).

70. Teplan, M. *et al.* Fundamentals of EEG measurement. *Measurement science review* **2,** 1–11 (2002).

71. Gais, S. & Born, J. Declarative memory consolidation: mechanisms acting during human sleep. *Learning & memory* **11,** 679–685 (2004).

72. Marshall, L. & Born, J. The contribution of sleep to hippocampus-dependent memory consolidation. *Trends in cognitive sciences* **11,** 442–450 (2007).

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

74. Plihal, W. & Born, J. Effects of early and late nocturnal sleep on declarative and procedural memory. *Journal of cognitive neuroscience* **9,** 534–547 (1997).

75. Sara, S. J. Sleep to remember. *Journal of Neuroscience* **37,** 457–463 (2017).

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

77. Ackermann, S. & Rasch, B. Differential effects of non-REM and REM sleep on memory consolidation? *Current neurology and neuroscience reports* **14,** 1–10 (2014).

78. Boutin, A. & Doyon, J. A sleep spindle framework for motor memory consolidation. *Philosophical Transactions of the Royal Society B* **375,** 20190232 (2020).

79. King, B. R., Hoedlmoser, K., Hirschauer, F., Dolfen, N. & Albouy, G. Sleeping on the motor engram: the multifaceted nature of sleep-related motor memory consolidation. *Neuroscience & Biobehavioral Reviews* **80,** 1–22 (2017).

80. Sawangjit, A. *et al.* The hippocampus is crucial for forming non-hippocampal long-term memory during sleep. *Nature* **564,** 109–113 (2018).

81. Spencer, R. M., Sunm, M. & Ivry, R. B. Sleep-dependent consolidation of contextual learning. *Current Biology* **16,** 1001–1005 (2006).

82. Hahn, M. A. *et al.* Slow oscillation–spindle coupling strength predicts real-life gross-motor learning in adolescents and adults. *eLife* **11,** e66761 (2022).

83. Nishida, M. & Walker, M. P. Daytime naps, motor memory consolidation and regionally specific sleep spindles. *PloS one* **2,** e341 (2007).

84. Wei, Y., Krishnan, G. P., Komarov, M. & Bazhenov, M. Differential roles of sleep spindles and sleep slow oscillations in memory consolidation. *PLoS computational biology* **14,** e1006322 (2018).

85. Kopasz, M. *et al.* Sleep and memory in healthy children and adolescents–a critical review. *Sleep medicine reviews* **14,** 167–177 (2010).

86. Spencer, R. M., Gouw, A. M. & Ivry, R. B. Age-related decline of sleep-dependent consolidation. *Learning & Memory* **14,** 480–484 (2007).

87. Hahn, M. *et al.* Developmental changes of sleep spindles and their impact on sleep-dependent memory consolidation and general cognitive abilities: A longitudinal approach. *Developmental science* **22,** e12706 (2019).

88. Shinomiya, S., Nagata, K., Takahashi, K. & Masumura, T. Development of sleep spindles in young children and adolescents. *Clinical Electroencephalography* **30,** 39–43 (1999).

89. Timofeev, I. *et al.* Spatio-temporal properties of sleep slow waves and implications for development. *Current Opinion in Physiology* **15,** 172–182 (2020).

90. Harand, C. *et al.* How aging affects sleep-dependent memory consolidation? *Frontiers in neurology* **3,** 18473 (2012).

91. Petit, D., Gagnon, J.-F., Fantini, M. L., Ferini-Strambi, L. & Montplaisir, J. Sleep and quantitative EEG in neurodegenerative disorders. *Journal of psychosomatic research* **56,** 487–496 (2004).

92. Mander, B. A. *et al.* Prefrontal atrophy, disrupted NREM slow waves and impaired hippocampal-dependent memory in aging. *Nature neuroscience* **16,** 357–364 (2013).

93. Martin, N. *et al.* Topography of age-related changes in sleep spindles. *Neurobiology of aging* **34,** 468–476 (2013).

94. Ishii, R. *et al.* Healthy and pathological brain aging: from the perspective of oscillations, functional connectivity, and signal complexity. *Neuropsychobiology* **75,** 151–161 (2018).

95. Saletin, J. M., van der Helm, E. & Walker, M. P. Structural brain correlates of human sleep oscillations. *Neuroimage* **83,** 658–668 (2013).

96. Geva-Sagiv, M. & Nir, Y. Local sleep oscillations: implications for memory consolidation. *Frontiers in neuroscience* **13,** 475266 (2019).

97. Rechtschaffen, A. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects, 1–55 (1968).

98. Mölle, M., Bergmann, T. O., Marshall, L. & Born, J. Fast and slow spindles during the sleep slow oscillation: disparate coalescence and engagement in memory processing. *Sleep* **34,** 1411–1421 (2011).

99. Zeitlhofer, J. *et al.* Topographic distribution of sleep spindles in young healthy subjects. *Journal of sleep research* **6,** 149–155 (1997).

100. Perrault, A. A. *et al.* Whole-night continuous rocking entrains spontaneous neural oscillations with benefits for sleep and memory. *Current Biology* **29,** 402–411 (2019).

101. Nir, Y. *et al.* Regional slow waves and spindles in human sleep. *Neuron* **70,** 153–169 (2011).

102. Dickey, C. W. *et al.* Travelling spindles create necessary conditions for spike-timing-dependent plasticity in humans. *Nature communications* **12,** 1027 (2021).

103. Muller, L. *et al.* Rotating waves during human sleep spindles organize global patterns of activity that repeat precisely through the night. *eLife* **5,** e17267 (2016).

104. Muller, L., Chavane, F., Reynolds, J. & Sejnowski, T. J. Cortical travelling waves: mechanisms and computational principles. *Nature Reviews Neuroscience* **19,** 255–268 (2018).

105. O'Reilly, C. & Nielsen, T. Assessing EEG sleep spindle propagation. Part 2: experimental characterization. *Journal of Neuroscience Methods* **221,** 215–227 (2014).

106. Ermentrout, G. B. & Kleinfeld, D. Traveling electrical waves in cortex: insights from phase dynamics and speculation on a computational role. *Neuron* **29,** 33–44 (2001).

107. Hindriks, R., van Putten, M. J. & Deco, G. Intra-cortical propagation of EEG alpha oscillations. *Neuroimage* **103,** 444–453 (2014).

108. Zhang, H., Watrous, A. J., Patel, A. & Jacobs, J. Theta and alpha oscillations are traveling waves in the human neocortex. *Neuron* **98,** 1269–1281 (2018).

109. Sterne, J. A. *et al.* ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *bmj* **355** (2016).

110. Cordi, M. J. & Rasch, B. How robust are sleep-mediated memory benefits? *Current Opinion in Neurobiology* **67,** 1–7 (2021).

111. Reverberi, S., Kohn, N. & Fernández, G. No evidence for an effect of explicit relevance instruction on consolidation of associative memories. *Neuropsychologia* **143,** 107491 (2020).

112. Van Erp, S., Verhagen, J., Grasman, R. P. & Wagenmakers, E.-J. Estimates of between-study heterogeneity for 705 meta-analyses reported in Psychological Bulletin from 1990–2013. *Journal of Open Psychology Data* **5** (2017).

113. Joechner, A.-K., Hahn, M. A., Gruber, G., Hoedlmoser, K. & Werkle-Bergner, M. Sleep spindle maturity promotes slow oscillation-spindle coupling across child and adolescent development. *eLife* **12,** e83565 (2023).

114. Klinzing, J. G. *et al.* Spindle activity phase-locked to sleep slow oscillations. *Neuroimage* **134,** 607–616 (2016).

115. Mölle, M. & Born, J. Slow oscillations orchestrating fast oscillations and memory consolidation. *Progress in brain research* **193,** 93–110 (2011).

116. Ladenbauer, J., Ladenbauer, J., Külzow, N. & Flöel, A. Memory-relevant nap sleep physiology in healthy and pathological aging. *Sleep* **44,** zsab002 (2021).

117. Schreiner, T., Petzka, M., Staudigl, T. & Staresina, B. P. Endogenous memory reactivation during sleep in humans is clocked by slow oscillation-spindle complexes. *Nature communications* **12,** 3112 (2021).

118. Anderer, P. *et al.* Low-resolution brain electromagnetic tomography revealed simultaneously active frontal and parietal sleep spindle sources in the human cortex. *Neuroscience* **103,** 581–592 (2001).

119. De Gennaro, L. & Ferrara, M. Sleep spindles: an overview. *Sleep medicine reviews* **7,** 423–440 (2003).

120. Urakami, Y. Relationships between sleep spindles and activities of cerebral cortex as determined by simultaneous EEG and MEG recording. *Journal of Clinical Neurophysiology* **25,** 13–24 (2008).

121. Denis, D. *et al.* Slow oscillation-spindle coupling is negatively associated with emotional memory formation following stress. *European Journal of Neuroscience* **55,** 2632–2650 (2022).

122. Fernandez, L. M. & Lüthi, A. Sleep spindles: mechanisms and functions. *Physiological reviews* **100,** 805–868 (2020).

123. Achermann, P. & Borbély, A. Low-frequency (¡ 1 Hz) oscillations in the human sleep electroencephalogram. *Neuroscience* **81,** 213–222 (1997).

124. Baena, D., Fang, Z., Ray, L. B., Owen, A. M. & Fogel, S. M. Brain activations time locked to slow wave-coupled sleep spindles correlates with intellectual abilities. *Cerebral Cortex* **33,** 5409–5419 (2023).

125. Dang-Vu, T. T. *et al.* Spontaneous neural activity during human slow wave sleep. *Proceedings of the National Academy of Sciences* **105,** 15160–15165 (2008).

126. Schmid, D., Erlacher, D., Klostermann, A., Kredel, R. & Hossner, E.-J. Sleep-dependent motor memory consolidation in healthy adults: A meta-analysis. *Neuroscience & Biobehavioral Reviews* **118,** 270–281 (2020).

127. McConnell, B. V. *et al.* The aging slow wave: a shifting amalgam of distinct slow wave and spindle coupling subtypes define slow wave sleep across the human lifespan. *Sleep* **44,** zsab125 (2021).

128. Bernardi, G., Siclari, F., Handjaras, G., Riedner, B. A. & Tononi, G. Local and widespread slow waves in stable NREM sleep: evidence for distinct regulation mechanisms. *Frontiers in human neuroscience* **12,** 248 (2018).

129. McConnell, B. V. *et al.* The rise and fall of slow wave tides: vacillations in coupled slow wave/spindle pairing shift the composition of slow wave activity in accordance with depth of sleep. *Frontiers in neuroscience* **16,** 915934 (2022).

130. Siclari, F. *et al.* Two distinct synchronization processes in the transition to sleep: a high-density electroencephalographic study. *Sleep* **37,** 1621–1637 (2014).

131. Németh, D. *et al.* Optimizing the methodology of human sleep and memory research. *Nature Reviews Psychology* **3,** 123–137 (2024).

132. Mardia, K. Linear-circular correlation coefficients and rhythmometry. *Biometrika,* 403–405 (1976).

133. Yang, J., Zhu, H. & Tian, X. Group-level multivariate analysis in EasyEEG toolbox: examining the temporal dynamics using topographic responses. *Frontiers in neuroscience* **12,** 369892 (2018).

134. Ioannidis, J. P. *et al.* Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet* **383,** 166–175 (2014).

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

136. Mylonas, D. *et al.* Dyscoordination of non-rapid eye movement sleep oscillations in autism spectrum disorder. *Sleep* **45,** zsac010 (2022).

137. Krishnaswamy, P. *et al.* Sparsity enables estimation of both subcortical and cortical activity from MEG and EEG. *Proceedings of the National Academy of Sciences* **114,** E10465–E10474 (2017).

138. Gong, R. *et al.* Spatiotemporal features of β-γ phase-amplitude coupling in Parkinson's disease derived from scalp EEG. *Brain* **144,** 487–503 (2021).

139. Riedner, B. A. *et al.* Sleep homeostasis and cortical synchronization: III. A high-density EEG study of sleep slow waves in humans. *Sleep* **30,** 1643–1657 (2007).

140. Roehri, N., Bréchet, L., Seeber, M., Pascual-Leone, A. & Michel, C. M. Phase-amplitude coupling and phase synchronization between medial temporal, frontal and posterior brain regions support episodic autobiographical memory recall. *Brain Topography* **35,** 191–206 (2022).

141. Routier, L. *et al.* Plasticity of neonatal neuronal networks in very premature infants: Source localization of temporal theta activity, the first endogenous neural biomarker, in temporoparietal areas. *Human Brain Mapping* **38,** 2345–2358 (2017).

142. Seeber, M. *et al.* Subcortical electrophysiological activity is detectable with high-density EEG source imaging. *Nature communications* **10,** 753 (2019).

143. Siebenhühner, F., Wang, S. H., Palva, J. M. & Palva, S. Cross-frequency synchronization connects networks of fast and slow oscillations during visual working memory maintenance. *eLife* **5,** e13451 (2016).

144. Page, M. J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International journal of surgery* **88,** 105906 (2021).

145. Gusenbauer, M. & Haddaway, N. R. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research synthesis methods* **11,** 181–217 (2020).

146. Martín-Martín, A., Thelwall, M., Orduna-Malea, E. & Delgado López-Cózar, E. Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations. *Scientometrics* **126,** 871–906 (2021).

147. Meline, T. Selecting studies for systemic review: Inclusion and exclusion criteria. *Contemporary issues in communication science and disorders* **33,** 21–27 (2006).

148. National Heart, L., Institute, B., *et al. Study quality assessment tools* 2019.

149. Pick, J. L., Nakagawa, S. & Noble, D. W. Reproducible, flexible and high-throughput data extraction from primary literature: The metaDigitise r package. *Methods in Ecology and Evolution* **10,** 426–431 (2019).

150. Ivimey-Cook, E. R., Noble, D. W., Nakagawa, S., Lajeunesse, M. J. & Pick, J. L. Advice for improving the reproducibility of data extraction in meta-analysis. *Research Synthesis Methods* **14,** 911–915 (2023).

151. Rohatgi, A. WebPlotDigitizer user manual version 3.4. *URL http://arohatgi.info/WebPlotDigitizer/app,* 1–18 (2014).

152. Fisher, N. I. *Statistical analysis of circular data* (cambridge university press, 1995).

153. Lipsey, M. W. & Wilson, D. B. *Practical meta-analysis.* (SAGE publications, Inc, 2001).

154. Myers, L. & Sirois, M. J. Spearman correlation coefficients, differences between. *Encyclopedia of statistical sciences* **12** (2004).

155. Thompson, C. G. & Semma, B. An alternative approach to frequentist meta-analysis: A demonstration of Bayesian meta-analysis in adolescent development research. *Journal of Adolescence* **82,** 86–102 (2020).

156. Fisher, R. A. *et al.* 014: On the" Probable Error" of a Coefficient of Correlation Deduced from a Small Sample. (1921).

157. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *Journal of statistical software* **36,** 1–48 (2010).

158. Fisher, N. I., Lewis, T. & Embleton, B. J. *Statistical analysis of spherical data* (Cambridge university press, 1993).

159. Sutton, A. J. & Abrams, K. R. Bayesian methods in meta-analysis and evidence synthesis. *Statistical methods in medical research* **10,** 277–303 (2001).

160. Bürkner, P.-C. brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software* **80,** 1–28 (2017).

161. Gabry, J., Simpson, D., Vehtari, A., Betancourt, M. & Gelman, A. Visualization in Bayesian workflow. *J. R. Stat. Soc. A* **182,** 389–402 (2019).

162. Wagner, J. *et al.* Metaviz: interactive statistical and visual analysis of metagenomic data. *Nucleic acids research* **46,** 2777–2787 (2018).

163. Kass, R. E. & Raftery, A. E. Bayes factors. *Journal of the american statistical association* **90,** 773–795 (1995).

164. Depaoli, S. & Van de Schoot, R. Improving transparency and replication in Bayesian statistics: The WAMBS-Checklist. *Psychological methods* **22,** 240 (2017).

165. Röver, C. *et al.* On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. *Research Synthesis Methods* **12,** 448–474 (2021).

166. Williams, D. R., Rast, P. & Bürkner, P.-C. Bayesian meta-analysis with weakly informative prior distributions (2018).

167. McKinney, K., Stefan, A. M. & Gronau, Q. F. Developing prior distributions for bayesian meta-analyses (2021).

168. Cox, C. *et al.* A systematic review and Bayesian meta-analysis of the acoustic features of infant-directed speech. *Nature Human Behaviour* **7,** 114–133 (2023).

169. Vehtari, A., Gelman, A. & Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and computing* **27,** 1413–1432 (2017).

170. Kossmeier, M., Tran, U. S. & Voracek, M. Visual inference for the funnel plot in meta-analysis. *Zeitschrift für Psychologie* (2019).

171. Cremers, J. & Klugkist, I. One direction? A tutorial for circular data analysis using R with examples in cognitive psychology. *Frontiers in psychology* **9,** 396689 (2018).

172. Bartsch, U., Simpkin, A. J., Demanuele, C., Wamsley, E., Marston, H. M. & Jones, M. W. Distributed slow-wave dynamics during sleep predict memory consolidation and its impairment in schizophrenia. *NPJ Schizophr.* **5,** 18 (2019).

173. Dehnavi, F., Koo-Poeggel, P. C., Ghorbani, M. & Marshall, L. Spontaneous slow oscillation–slow spindle features predict induced overnight memory retention. *Sleep* **44**, zsab127 (2021).

174. Kruschke, J. K. & Liddell, T. M. The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychon. Bull. Rev.* **25,** 178–206 (2018).

175. Roebber, J. K., Lewis, P. A., Crunelli, V., Navarrete, M. & Hamandi, K. Effects of anti-seizure medication on sleep spindles and slow waves in drug-resistant epilepsy. *Brain Sci.* **12,** 1288 (2022).

176. Cox, R. & Fell, J. Analyzing human sleep EEG: A methodological primer with code implementation. *Sleep Med. Rev.* **54,** 101353 (2020).