Accepted Manuscript (post-peer-review; accepted for publication in Neurobiology of Learning and Memory).

Citation: Itamar Lerner, & Sarah C. Hamm (in press). Sleep promotes illusory word compositions, a distinct form of false memory. *Neurobiology of Learning and Memory*.

https://doi.org/10.1016/j.nlm.2025.108129

Sleep Promotes Illusory Word Compositions, a Distinct Form of False Memory

Authors: Itamar Lerner[1], Sarah C. Hamm[1]

[1]Department of Psychology, The University of Texas at San Antonio, San Antonio, TX 78249 USA

Corresponding author:
Itamar Lerner
One UTSA Circle
San Antonio, Texas, 78249 USA
Email: Itamar.lerner@utsa.edu

Abstract

Extensive evidence supports the beneficial effects of sleep on memory and learning, including the consolidation and reorganization of memories and the extraction of regularities from encoded experiences. Nevertheless, some studies suggest that sleep may also increase false memories, potentially as a byproduct of regularities extraction. Physiologically, time-compressed memory replay in the hippocampus during non-rapid-eye-movement (nREM) sleep is believed to contribute to the consolidation process, although the functional significance of time compression remains elusive. Recently, we proposed that compressed replay might allow associating events that happened at disparate times, thus supporting the extraction of regularities with a temporal nature. This model predicted that sleep might also facilitate a distinct kind of false memories, in which two separate events occurring consecutively are encoded as a single composite event. Here, we tested this prediction by exposing male and female adults to separate word pairs (e.g., car, pet) that could form a new composite word if combined (carpet). We then tested their memory for composite words following a period of sleep or wake. Confirming our main prediction, we found that sleep actively facilitated false composite memories. Furthermore, EEG recordings indicated the involvement of nREM sleep in the process, albeit in a nuanced manner: While some slow-wave or spindle-related parameters predicted increase in false memories, others were associated with fewer false memories and a decline in veridical memories. The latter result resembles previous findings from non-composite false memory studies and could suggest a competitive mechanism between semantic and episodic consolidation during sleep.

Keywords: Slow Wave Sleep, Sleep and Memory, Memory Consolidation, False Memories, Sleep Spindles, Slow Oscillations, DRM

Highlights

- Sleep is promoting the false recognition of words composed of two previously seen word components.
- False composite memories are independent of the order of presentation of their components but rely on temporal proximity.
- Slow wave sleep markers are predictive of false composite memories, but only when their components were presented in the forward direction.

1. Introduction

Substantial evidence suggests that sleep plays a significant role in memory consolidation, from the simple facilitation of declarative memories to the reorganization of memories into schema-like forms (Rasch & Born, 2013; Lewis & Durrant, 2011). Memory reorganization manifests in various way, including the implicit detection of patterns within encoded experiences, the explicit recognition of rules governing a set of encoded events, the extraction of gist from partially overlapping stimuli and the generalization of specific episodes to novel circumstances – all of which have been shown to benefit from sleep (e.g., Durrant et al., 2011; Ellenbogen et al., 2007; Friedrich et al., 2015; Graveline & Wamsley, 2017; Lerner & Gluck, 2018; Wilhelm et al., 2013). Perhaps the most striking example of memory reorganization is exemplified by studies demonstrating that sleep supports the insightful discovery of hidden regularities. In these studies, subjects are presented with a sequence of stimuli and asked to respond to each stimulus by following a simple rule (Fischer et al., 2006; Wagner et al., 2004). Unbeknownst to subjects, a hidden structure governs the series of presentations such that, if discovered, it could improve performance significantly. Results show that following sleep, subjects are far more likely to explicitly discover the hidden regularities compared to subjects who stayed awake. Another example of a sleep-dependent effect often attributed to memory reorganization is demonstrated using the Deese-Roediger-McDermott (DRM) false memory paradigm (Roediger & McDermott, 1995). In this task, participants memorize a list of words with a semantically related theme (e.g., *hospital*, *nurse*, *patient*) and are then tested on whether they falsely remember seeing the theme word (e.g., *doctor*). While results have not always been consistent, sleep in between study and test was occasionally shown to increase erroneous recall of the theme words, particularly when the list of studied words is short (Newbury & Monaghan, 2019).

Physiological evidence from human and rodent studies shows that non-rapid-eye-movement (nREM) sleep, particularly slow wave sleep (SWS), may be key to these effects (Wilhelm et al., 2013; Yordanova et al., 2012). During SWS, recently encoded memories are replayed in the hippocampus in an accelerated, “time-compressed” pace as part of a hippocampal-cortical dialogue (Diba & Buzsaki, 2007; Ji & Wilson, 2007). Theoretical frameworks such as the Active System Consolidation hypothesis (Diekelmann & Born, 2010) suggest that replay may contribute to the strengthening of common features in newly encoded memories while blunting their idiosyncratic elements, effectively leading to the extraction of "gist" and the integration of those memories within the general knowledge structure residing in the cortex (McClelland et al., 1995; Lewis & Durrant, 2011). Nevertheless, the precise mechanism and the relevance of the time-compression attribute remain largely unknown (Abel et al., 2013).

Recently, we have proposed a 'temporal scaffolding' model of SWS to explain the role of accelerated replay in insightful processes (Lerner & Gluck, 2019; 2022). Based on a review of the literature, we showed that explicit detection of hidden rules following sleep is mostly evident when the rules are temporal; that is, rules in which a stimulus happening at one timepoint predicts another stimulus happening a few seconds later. Such rules are not easily detected in real time when unexpected; however, if the relevant sequences are encoded in the hippocampus and then replayed during SWS at an accelerated pace, events that happened seconds apart are brought into a much shorter timescale (50-100ms; August & Levy, 1999), sufficient for associative Hebbian mechanisms to form the crucial temporal associations and allow recognition of the hidden regularities.

Although our model was developed to explain how sleep contributes to the detection of temporal patterns, its mechanism – relying on time-compressed replay – gives rise to another prediction: The possible formation of a distinct type of sleep-dependent false memory that is not directly based on semantic relationships. Specifically, if several unrelated but temporally proximal events are encoded and then replayed in a compressed timescale one after the other, they could potentially be consolidated into a single, integrated memory. This prediction is based on the well-documented contribution of sleep to the consolidation of associative links between declarative memories. For example, in the classic paired-associates paradigm, subjects memorize word pairs like *dog – leaf* and later asked to recall the second word of each pair after being presented with the first. Sleep in between exposure and testing, particularly SWS, has been repeatedly shown to strengthen this association, especially when the two words are related in meaning (e.g., *dog – bone*), reflecting a tendency to retain semantic themes (Ellenbogen et al., 2006; Payne et al., 2012; Plihal & Born, 1997; Walker & Stickgold, 2009). Moreover, similar effects can also emerge when stimuli pairs are encountered incidentally, without a deliberate memorization attempt (e.g., Schmieding et al., 2024). By the same token, we project that during sleep, separately encoded word pairs may be brought "together" through time-compressed replay and form associations such that they end up being consolidated as a single memory, particularly if that memory has a meaning of its own. For example, if subjects are presented with two consecutive words (e.g., *car* and *pet*) that could be combined into a composite word (*carpet*), the two distinct memories of *car* and *pet* might be integrated following sleep to become one false memory of seeing *carpet* – a memory reorganization of sorts, but one that builds on temporal proximity rather than just semantics.

The current study aimed to test this prediction. We speculated that compared to wake, an afternoon nap would encourage the false recognition of composite words whose components were separately encountered prior to napping, and that SWS-related measures associated with memory replay would correlate with this effect. Confirming these hypotheses would suggest a potential functional role for time-compression in memory consolidation and, consequently, in the extraction of temporal regularities.

2. Methods

2.1. Overview of Study Design and Predictions

We examined the effects of sleep on the formation of false composite memories by comparing how an afternoon nap, compared to wake, affected false recognition of composite words. Participants were first exposed to components of the composite words (e.g., *for* and *mat*), presented either sequentially within the same trial or split into distant trials, and later underwent a surprise memory recognition test for the composite words (e.g., *format*). Inclusion of the "split" condition served as control to the main sequential presentation: Both cases yield identical exposure to the components, but given the long time gap and numerous intervening items in between matching components of the split condition, only the sequential condition is expected to be influenced by time-compressed replay (i.e., the rapid reactivation of encoded sequences hypothesized to drive the predicted effects in the sleep group). The memory test also included old non-composite words presented earlier during exposure, as well as novel words never presented before. The ability to differentiate old words from (a) composite words presented sequentially, (b) composite words split across trials, and (c) totally novel words, was used as the outcome variable. Word type was manipulated within-subject, whereas wake versus sleep was manipulated between subjects, with one wake group tested soon after finishing the exposure phase ('Wake-Immediate'), another wake group tested 2 hours later while remaining awake throughout ('Wake-Delayed'), and the sleep group tested 2 hours later while sleeping in between, thus forming a 3 x 3 mixed design. We predicted that sleep, compared to either wake groups, would facilitate the false recognition of composite words compared to novel words, but only when their components were presented sequentially during exposure (thus more likely to be replayed in a sequence) rather than when they were split into different trials. Moreover, we predicted that all participants who stayed awake, regardless of the interval length between exposure and test, would show similar outcomes, thus supporting the assertion that it is sleep that actively facilitates false composite memories rather the passive passage of time. See Fig. 2A for a summary of these behavioral predictions. We further predicted that various physiological measures of sleep that are known to be associated with replay and memory consolidation, such as the percent of time spent in SWS, the magnitude of slow oscillations (SO) and sleep spindles, and the degree of SO-spindle coupling, would be positively correlated with the degree of false composite memory.

An additional subdivision of conditions was used to investigate a secondary prediction of the model. Specifically, exposure trials with sequentially presented composite words could have the components displayed in the forward direction (*for* –> *mat*) or in the backward direction (*mat* –> *for*). This added distinction was based on findings showing memory replay during SWS tends to occur more often in a forward than a backward manner – that is, an encoded sequence of events is more likely to be replayed in the same order as the original experience rather than in reverse (Diba & Buzsaki, 2007; Wikenheiser et al., 2013). We therefore predicted that sleep, while affecting memories for both directions of presentation, would nevertheless have a stronger effect in facilitating false memories of composite words whose components were presented sequentially in the forward direction compared to those presented backwards (Fig. 2B).

2.2. Participants

One hundred and six participants (N = 63 Females) were recruited for this study from the undergraduate student population at The University of Texas at San Antonio via campus ads and research participant pools and received either course credit or monetary compensation for their participation. The number of participants was chosen based on power calculations (Gpower 3.1.9.2) informed by a previous pilot study (Lerner et al., 2019), to detect within-between interactions in a repeated measures ANOVA with 3 groups (sleep and two wake controls) and 3 repeated measures, assuming a medium-small effect size of partial η^2 = 0.04, α = 0.05, power (1-β) = 0.8, with sphericity assumed and no correlation between repeated measures, yielding a minimum sample size of 99. Exclusion criteria included a history of sleep deficiencies, visual impairment consisting of uncorrected vision, a history of major neurological or psychiatric disorders, or head injury leading to unconsciousness. Participants were assigned by stratified randomization to either the experimental group (Sleep) or one of the two control groups (Wake-Delayed, Wake-Immediate; See Experimental Procedure), matching the groups on age, education, gender, and sleep habits (Table 1). Several additional participants were removed from the study due to their responses indicating a failure to follow instructions (clicking indiscriminately on the same response again and again; 4 participants), failure to fall asleep (2 participants) or missing behavioral data (1 participant).

2.3 Behavioral Task

2.3.1. Task Design

Our false composite words task was partly inspired by previous studies demonstrating memory conjunction errors (e.g., Underwood & Zimmerman, 1973; but see Discussion for important differences between the two paradigms) and included two sessions: exposure and testing. In each trial of the exposure session, participants were presented with a pair of words in succession and asked to indicate whether one of the words or both/neither included the letter *e* by pressing one of two buttons. The task was chosen to ensure sufficient attention was given to the words without highlighting their meaning, with equal division between the two correct answers. The stimuli consisted of 3- and 6-letter words, with all four combinations of word lengths (3 and 3, 3 and 6, 6 and 3, 6 and 6) appearing across trials in equal numbers. Critically, unbeknownst to participants, the 3-3 trials were composed of words that could be combined to form a six-letter composite word (e.g., *carpet*, composed of *car* and *pet*). In these trials, the words could be shown either in the direction consistent with the composite word (Forward condition: *car* followed by *pet*), or in the opposite direction (Backward condition: *pet* followed by *car*). In other trials, where 3-letter words were paired with 6-letter words (3-6 or 6-3), half of the 3-letter words could be combined to create composite words across trials (Split condition: e.g., the composite word *carpet* being split so that in one trial, *car* would be followed by *doctor,* and in another trial *pet* would be followed by *beyond*). None of the 3-letter words in either the Forward, Backward or Split condition were semantically related to the 6-letter composite word they were part of. The remaining 3-6 and 6-3 trials, as well as all 6-6 trials, contained words that were neither part of nor containing any composites.

In the testing session, participants received a surprise memory test. They were presented with a list of 6-letter words one-by-one and asked to indicate for each, on a scale of 1 to 6, how confident they felt that those words appeared in the first session (with a score of 1 indicating they were confident that the word did not appear previously and a score of 6 indicating they were confident that the word did appear previously). Half of the words in the list were 6-letter words previously presented in the first phase ("Old" condition). The other half were comprised of either completely novel words ("Novel" condition) or of composite words made from the two 3-letter words that appeared during the first experimental session, equally distributed between the Forward, Backward, and Split conditions (see Fig. 1A). The order of presentation of word pairs in the exposure session and words in the testing session was random.

2.3.2. Stimuli preparation

We compiled a list of eighteen 6-letter composite words, each representing the combination of two 3-letter words (e.g., *car–pet; dam–pen; for–mat*). The 18 words were organized into three groups of six words each, with similar average frequency (M = 15.56, 15.41, 15.48 occurrences/million for the three groups, $F(2,15) = 0.0006$, $p > 0.99$. See full list of words in Supplementary Table S1; frequency values were based on data from Davies, 2008-). The three word groups were used for the composite word conditions in the experiment (Forward, Backward, Split), with each 3-letter component serving as a 3-letter word in the exposure session and the full 6-letter composite word appearing in the testing session. For counterbalancing purposes, six versions of the exposure session stimuli were created, with each having a different composite word group used for a different condition. The six versions were cycled through all participants based on participant number.

Other than the 18 composite words, we compiled an additional twelve 3-letter words and forty-eight 6-letter words to serve as the remaining stimuli in the 3-6, 6-3 and 6-6 trials of the exposure session for a total of 48 word-pair trials. The testing session included the 18 composite words, twenty-four of the 6-letter words that appeared in the exposure session, and additional six 6-letter words serving as the Novel condition for a total of 48 trials. Finally, there were also 16 practice trials prior to the exposure session, divided equally between the 3-3, 3-6, 6-3 and 6-6 combinations. No 3-letter strings (either 3-letter words or the first or second halves of 6-letter words) were repeated across the stimuli used in the task.

2.4. Experimental Procedure

Participants arrived to the lab in the afternoon and underwent the exposure and testing sessions with an interval in between (Fig. 1B). In the Sleep group, the interval totaled 120 minutes during which the participants were given the opportunity to sleep while attached to a Polysomnography device (PSG). The PSG was removed prior to the beginning of the testing session. In the Wake-Delayed group, the interval also totaled 120 minutes, and participants were not allowed to sleep. Instead, they engaged in non-stimulating activities (watching National Geographic episodes). They were also equipped with the PSG during this interval. In the Wake-Immediate group, the interval totaled 15 minutes during which participants were not allowed to sleep, with no PSG monitoring. In this group, participants sat quietly in the lab and did not watch any videos or use any electronic device.

All experimental sessions were conducted in a quiet room equipped with a bed, nightstand, desktop computer and some home-like decorations. Participant first received written instructions explaining the exposure session and then underwent a short practice. In each trial of the practice, the first word appeared in the middle of the screen for 500ms and was immediately followed by the second word, also appearing for 500ms. The screen then remained blank until participants gave their response by pressing either the ‘L’ or the ‘A’ button on the keyboard. They then received feedback (smiley face for correct answers, sad face for incorrect answers, appearing for 1 second), followed by an inter-trial interval (ITI) of 1.5 seconds before the next trial began. Following practice, participants were informed on screen that the actual experiment was about to begin. They were then presented with the experimental trials, which followed the same presentation schedule as practice but did not include feedback (Fig.1C, left). Instead, a crosshair appeared for 500ms after each response, followed by an ITI of 1 second before the next trial began. At the end of the exposure session, participants had a group-dependent interval as explained above, after which they received the testing session. They were first presented with instructions informing them of a surprise memory test on the words they have seen in the previous session; then, they were presented with one word at a time, appearing in the middle of the screen, with a Likert scale below. The scale had 6 levels with a message indicating "Confidently New" and "Confidently Old" appearing to the left and right of it, respectively. Participants responded by moving a mouse curser to the level that indicated how positive they were that the word appeared in the previous session (1 representing new and 6 representing old) and pressed the mouse button to proceed to the next trial following a 1 second ITI (Fig.1C, right). Finally, at the conclusion of the testing session, participants were given a post-experimental questionnaire to determine if they became aware of the hidden composite nature of some of the test stimuli. They were asked whether any unexpected patterns connecting the words in the exposure and testing sessions were observed, and if so, to describe the nature of the connection and provide examples.

2.5. Sleep Monitoring and Data Extraction

Sleep data was collected using the LiveAmp EEG system (Brain Vision LLC). The LiveAmp device was set up following the standard 10-20 system. Due to logistical or equipment errors, data from 6 participants in the Sleep group were not collected. All physiological sleep analysis was performed on the remaining 29 participants.

From the PSG recordings, we extracted parameters that are commonly considered to signal memory consolidation during nREM sleep (Rasch & Born, 2013; Kumral et al., 2023), including the percent of time spent in SWS, SO amplitude (peak to peak), spindle density, spindle power, and SO-spindle coupling phase and strength (see Supplementary Materials for similar results considering additional measures). A low-pass filter of 0.1Hz was applied to the raw EEG data collected with the 10-20 system to eliminate baseline drifts, and derivations were calculated for EEG channels C4-M1, C3-M2, F4-M1, F3-O2, O2-M1, O1-M2 and EOG channels E1-M2, E2-M1, in addition to the EMG channel. Sleep scoring was conducted by a trained sleep technician to determine the sleep stage in epochs of 30-second length and served to compute the total time spent in each sleep stage. SO and spindle parameters were extracted from the central derivations (C4-M1, C3-M2; results did not change markedly when using forward derivations instead) by a well-validated automatic algorithm using the default settings (YASA; Vallat & Walker, 2021). Results were then averaged over the two central derivations, except for when one derivation was substantially nosier than the other, in which case only the value for the less noisy derivation was used.

Measures of SO-spindle coupling were extracted following standard procedures (e.g. Schreiner et al., 2021; Hahn et al., 2020). First, artifact removal was applied on each channel derivation of each participant using EEGLAB's Independent Component Analysis (ICA) function. SOs for each participant were identified by filtering the data of each channel derivation between 0.16–2 Hz and detecting zero-crossings for all sections previously determined to belong to sleep stages N2 and N3. Potential SO events were considered for segments between each two consecutive positive-to-negative zero-crossings that met standard SO duration criteria (segment length between 0.8 to 2 seconds). Out of these segments, only those among the top 25% of peak-to-peak amplitude were identified as SOs ('SO amplitude criteria'). Five-second long epochs (±2.5 seconds centered on the trough of the SO segment) were extracted for each SO out of the raw signal. For spindle-detection, the raw data of each channel derivation of each participant was filtered between 12-16 Hz and a Hilbert transform was applied to retrieve the instantaneous amplitude. Then, for all sections previously determined to belong to sleep stages N2 and N3, we calculated the root mean square (RMS) using a moving average of 200 ms and set a 75% percentile of RMS values as the spindle amplitude threshold. A spindle event was detected for each instance where the RMS values exceeded the threshold for at least 0.5 seconds but no longer than 3 seconds. Five-second long epochs (±2.5 seconds centered on the peak of the spindle segment) were extracted for each spindle out of the raw signal. SO-spindle events were identified as those spindle-centered epochs in which the spindle peak occurred within 1.5 seconds following an SO trough (Schreiner et al., 2021).

For SO-spindle coupling analysis, we normalized the SO-spindle events (using z-score with mean and standard deviations obtained for each participant and channel; Ladenbauer et al., 2021), filtered them in the 0.16 - 2 Hz band and applied the Hilbert transform. The same procedure was repeated for the spindle band (12 – 16 Hz). To avoid filter edge artifacts, we only considered the time range within −2 to 2 seconds. We then extracted, for each spindle peak of each SO-spindle event, the instantaneous SO phase angle, and the resulting distribution of phases across SO-spindle events was tested against uniformity using the Rayleigh test. Finally, to measure the degree of coupling, we calculated the mean SO phase angle and the corresponding resultant vector length for each participant in each of the two central channel derivations. Matlab's circular toolbox was used to extract phase angles, vector lengths, and to produce the resulting figures.

2.6. Statistical Analyses

Based on the confidence levels of each participant in each testing condition (Composite, Split, Novel), Receiver Operating Characteristic (ROC) curves were created by comparing the False Positive ("False Alarm") rate of each condition to the True Positive ("Hit") Rate of the Old words condition. The area under the curve (AUC) was then calculated using the extrapolation technique (Stanislaw & Todorov, 1999) as a measure of participants' ability to differentiate between old and new words of each condition (both the Composite and Split conditions are considered "new" because participants did not see any of the composite words as a single word before). The same analysis was conducted for the Forward Composite and the Backward Composite conditions separately. In cases the extrapolation technique yielded a bad fit (less than 7% of the cases, roughly equally distributed among the groups), we calculated the AUC by the simple Trapezoidal rule (Yeh, 2002).

For statistical analysis, we ran a marginal linear model with AUC as the dependent variable and Condition and Group as within- and between-subject factors, respectively. Analysis was performed in SPSS 27.0 using the Mixed models procedure employing robust covariances estimation. First, we compared the Sleep, Wake-Immediate and Wake-Delayed groups in the three critical word conditions: Composite, Split and Novel. The model thus included a main factor of Group with 3 levels and a main factor of Condition with 3 levels, as well as their interaction, using an unstructured covariance matrix. In addition, to control for the different versions of the exposure stimuli used to counterbalance between word groups, the model also included the stimuli version as a block factor, together with its interactions with all other factors (Pollastek & Well, 1995). Follow-up tests included pairwise comparisons between conditions for each group and group comparisons for each condition, corrected for multiple comparisons using the Holm-Sidak method. Finally, we ran a similar marginal linear model analysis to compare the two composite word conditions (Forward vs. Backward).

To examine the associations with sleep physiology, we conducted a multiple regression analysis for the Sleep group participants, with AUC in each of the three critical conditions as the predicted variable and the physiological parameters of interests as predictors. A follow-up regression analysis was conducted separately for the Forward and Backward composite conditions to test whether each showed significant predictor effects on its own. Additionally, to compare the Forward and Backward regression models directly, we ran a multivariate regression (seemingly unrelated regression – SUR; Zellner, 1962) followed by a Wald test of parameter equality to assess whether specific sets of predictors differed in their effects across the two conditions. Finally, to further clarify the sources of some of the effects, we ran a second set of regression models with the same predictors but using raw confidence scores as the outcome variables, as detailed in Results. The Holm-Sidak method was used to correct for multiple comparisons across word conditions in each level of the analysis. Additional analyses of the associations between physiological sleep parameters and task performance are described in the Supplementary Materials.

Lastly, we also examined whether participants gained insight into the composite nature of the words by analyzing the post-experimental questionnaire. Participants were considered to have gained insight if, when asked whether any unexpected patterns connecting the words in the exposure and testing sessions were observed, they indicated that they noticed some words from the exposure session were combined in the testing session. Fisher's exact test was conducted to examine if the number of participants gaining insight differed between groups.

3. Results

3.1. Task Performance

We first verified that the groups did not differ in overall accuracy when performing the exposure session. A one-way analysis of variance showed that all groups performed the task with high accuracy ($M = 0.92$, $M = 0.91$, $M = 0.87$, for the Sleep, Wake-Delayed and Wake-Immediate groups, respectively) with no statistically significant differences between them ($F(2,103) = 1.14$, $p = 0.32$).

Next, we examined the expression of false composite memories during testing. Mean raw confidence levels for all groups and conditions are presented in Table 2. To analyze the effect of sleep, we compared participants' sensitivity to old versus new words by building ROC curves for the main categories of new words (Composite, Split and Novel; see Supplementary Fig. S1A) and calculating the AUC as the sensitivity measure. Means and standard errors of the AUC are presented in Fig. 2C. A marginal linear model comparing the groups showed a significant main effect of Condition ($F(2,264) = 8.51$, $p < 0.001$) and a significant interaction between Group and Condition ($F(4,264) = 2.60$, $p = 0.036$). Following the significant interaction, Sidak-holm corrected pairwise comparisons within each group showed that for the Sleep group, the Composite condition yielded a significantly lower sensitivity than either the Split ($t(264) = 4.21$, $p < 0.001$) or the Novel ($t(264) = 4.46$, $p < 0.001$) condition, whereas there was no difference between the Split and the Novel conditions ($p = 0.29$). In contrast, none of the pairwise comparisons were significant for any of the wake groups (all *p*'s > 0.21). Comparisons of the groups within each condition showed that for the Composite condition, the Sleep group displayed a significantly lower sensitivity than the Wake-Delayed group ($t(264) = 2.83$, $p = 0.015$) and a similar trend when compared to the Wake-Immediate group ($t(264) = 2.01$, $p = 0.09$), whereas the two wake groups did not differ from each other ($p = 0.231$). In contrast, there were no significant group differences for either the Split or the Novel conditions (all *p*'s > 0.58). Since the two wake groups did not differ in the Composite condition, we followed up with an independent t-test comparing the Sleep group in that condition to the two wake groups combined. The analysis showed a significant difference, with the Sleep group exhibiting smaller sensitivity than the combined wake group ($t(104) = 2.59$, $p = 0.011$).

Having established the first predicted effect, we turned to examine whether the Forward and Backward Composite conditions differed between sleep and wake. We ran a second marginal model, identical to the first except that only the Forward and Backward Composite conditions were included. Means and standard errors are presented in Fig. 2D. The analysis showed a significant main effect of Group ($F(2,176) = 3.39$, $p = 0.036$), but neither Condition nor the interaction between Condition and Group were significant (both *p*'s > 0.67). Pairwise comparisons showed that across the two composite conditions, the Sleep group had a significantly lower sensitivity that the Wake-Delayed group ($t(264) = 2.54$, $p = 0.035$), though the difference between the Sleep and the Wake-Immediate groups did not reach statistical significance (($t(264) = 1.86$, $p = 0.124$), and neither did the difference between the two wake groups ($p = 0.309$). Since the wake groups did not differ and only the main effect of Group was significant, we followed up by comparing the Sleep group with the two wake groups combined, across the two composite conditions. An independent t-test confirmed that the Sleep group exhibited significantly smaller sensitivity than the combined wake group ($t(104) = 2.34$, $p = 0.021$). To summarize, while sleep lowered the sensitivity to composite words overall, there was no significant difference between the forward and backward composite words for any group.

Finally, since the AUC, our measure of sensitivity to old versus new words, was calculated by combining data from the Old condition with the other conditions, it did not provide a baseline memory measure of how each of the groups performed. To examine whether the groups differed on that respect, we compared their raw recognition confidence scores in the Old condition. Neither a one-way ANOVA comparing the three groups nor an independent t-test comparing the Sleep group to the combined wake control group yielded a significant effect (both p's > 0.36).

3.2. Analysis of post-experimental questionnaires

We compared how many participants in each group became aware of the presence of composite words in the testing session. We found that only 3 participants in the Sleep group and 4 in each of the Wake control groups (8.6%, 11.1%, and 11.4% for the Sleep, Wake-Delayed, and the Wake-Immediate groups, respectively) gained such insight, with nearly all of them also able to give at least one example of a composite word. Fischer's exact test showed that the groups did not differ on this aspect ($p = 1.0$)

3.3. Associations between Sleep Physiology and Task Performance

Sleep parameters of interest were extracted as described in Methods. These included the percent of time spent in SWS, SO amplitude, sleep spindle density and power, and SO-Spindle coupling phase and strength. For the SO-spindle analysis, we ran the Rayleigh test for each participant to examine whether the distribution of instantaneous SO phases when peak spindle amplitudes occur is significantly different from uniformity. We found that for 27 out of the 29 participants with available data, the distribution was significantly different from uniform ($p < 0.05$), replicating conclusions from previous studies about the existence of a mechanism at play that maintains coupling precision between spindles and SOs (Schreiner et al., 2021; Helfrich et al., 2018; Staresina et al., 2015). Fig. 3A displays the grand average EEG signal across all SO-spindle events, time-locked to the spindle peak amplitude (black line), together with the corresponding grand average of the same segment filtered in the SO band (blue line). Mean preferred phase was -69.41° ± 27.94°, with all preferred angles falling within the [0 180] range, and the mean vector length corresponding to the mean preferred angle was 0.88 (Fig. 3B).

3.3.1. Analysis of Sleep Parameters and AUC Scores

To examine the relations between sleep physiology and behavioral performance in our experiment, we first calculated the pairwise correlations among all sleep parameters of interest and confirmed that none were highly correlated with each other (max|r| = 0.38; Fig. 3C), indicating that multicollinearity is unlikely to substantially affect the results (cf. Dormann et al., 2013). We then ran multiple regression models for the AUC scores of participants in the Sleep group in each main experimental condition (Composite, Split and Novel; significance corrected for 3 multiple comparisons using Holm-Sidak) with the sleep parameters as predictors. Fig. 4 (upper three rows) and Table 3 display the results.

The regression model was significant for the Composite condition, ($F(6, 21) = 4.45$, uncorrected $p = 0.005$), with effects driven by a negative correlation with the percent of time spent in SWS ($\beta = -0.391$, $p = 0.002$), a marginally significant negative correlation with SO amplitude ($\beta = -0.004$, $p = 0.067$), and a positive correlation with spindle power ($\beta = 0.343$, $p = 0.003$). For the Novel condition, the model was significant as well ($F(6, 21) = 3.79$, uncorrected $p = 0.01$), with effects driven by a negative correlation with SO amplitude ($\beta = -0.005$, $p = 0.017$) and a marginally significant negative correlation with SWS percent ($\beta = -0.221$, $p = 0.077$). The model did not reach significance for the Split condition (uncorrected $p = 0.056$), though inspection of the individual factors again showed negative associations with SWS and SO amplitude (Table 3). To follow up on the significant effect for the composite words, we reran the model separately for the AUC scores of the Forward and Backward Composite conditions (significance corrected for 2 multiple comparisons). Table 4 presents the results. We found that for the Forward condition, the model was significant ($F(6, 21) = 3.39$, uncorrected $p = 0.017$), driven, as in the overall Composite condition, by a negative correlation with SWS percent ($\beta = -0.47$, $p = 0.006$) and a marginally significant negative correlation with SO amplitude ($\beta = -0.005$, $p = 0.075$), as well as a positive correlation with spindle power ($\beta = 0.509$, $p = 0.002$). For the backward condition, the model was not significant (uncorrected $p = 0.135$). To follow up on the different results between the Forward and Backward conditions, we then compared the two directly. To that end, we conducted a joint Wald test for the 2 significant predictors found in the Forward model (SWS percent, spindle power). Results showed a marginal effect ($\chi^2(2) = 5.10$, $p = 0.078$), indicating a potentially different influence of the two predictors for the Forward and Backward conditions.

To summarize, the regression models partially supported our hypothesis by showing that greater percentage of time spent in SWS and higher amplitude of slow oscillations predict reduced differentiability between real and false memories. In contrast, spindle power unexpectedly showed the opposite effect, with higher power predicting better differentiability. These correlations were evident for the Forward composite but not the Backward composite condition when examined separately. In addition, contrary to our prediction, the SO amplitude effect – and, to a lesser degree, the SWS effect – were also observed for non-composite words.

3.3.2. Analysis of Sleep Parameters and Raw Confidence Scores

Our previous analysis demonstrated the existence of various correlations between AUC scores and SWS-related physiological measures, some supporting our hypothesis and some contradicting it. However, since AUC scores represent the ability to distinguish between encountered and unencountered words, it remains unclear if these correlations reflected associations with increased false memories, a decline in real memories, or both. Moreover, when the same correlation emerges across both composite and non-composite conditions - as was the case with the negative correlation of AUC and SO amplitude – it could be a byproduct of a single measure used to compute the AUC for all conditions, namely, the raw recognition confidence scores for old words. To further clarify the sources of our effects, we reran the regression models using the raw confidence scores of each word type (Forward, Backward, Split, Novel and Old, correcting for 5 multiple comparisons) as the outcome variables. Results are present in Table 5 and Fig. 4 (bottom three rows).

We found that for the Old condition, the model was significant ($F(6, 21) = 4.96$, uncorrected $p = 0.003$), with effects driven by a negative correlation with SO amplitude ($\beta = -0.029$, $p = 0.005$). The model was also significant for the Forward composite condition ($F(6, 21) = 3.68$, uncorrected $p = 0.012$), with effects driven by a positive correlation with SWS percent ($\beta = 1.979$, $p = 0.045$) and a negative correlation with spindle power ($\beta = -2.658$, $p = 0.006$), as well as a negative correlation with SO-Spindle coupling phase ($\beta = -1.161$, $p = 0.001$). There were no significant effects for the Backward, Split or Novel conditions (all uncorrected p's > 0.75). Comparing the Forward and Backward regression models directly, a joint Wald test for the 3 significant predictors (SWS percent, spindle power, SO-spindle coupling phase) yielded a significant effect ($\chi^2(3) = 8.24$, $p = 0.041$), suggesting they influenced the two conditions differently.

To summarize, the results suggest that the correlation between AUC scores and SO amplitude across the composite and non-composite conditions was likely driven by responses to the old words. In contrast, the correlations with SWS and spindle power in the Forward condition, observed both using the AUC scores and the raw confidence scores, were driven by the responses to the forward composite words (notice that the direction of correlations with confidence scores of composite words is consistent with correlations in the opposite direction with the corresponding AUC scores). In addition, the analysis of the raw confidence scores revealed a contribution of the SO-Spindle coupling phase: The closer the peak spindle amplitude got to alignment with the peak SO phase, the less participants erroneously recognized forward composite words as words they have seen before (see color illustration for the Forward Composite condition in Fig. 3B)

4. Discussion

Our findings in this study can be summarized as follows: (a) an afternoon nap contributes to the formation of false composite memories; sleep, compared to wake, enhances the false recognition of composite words when they are comprised of shorter word components presented in temporal proximity prior to sleep; (b) the order of presentation of the word components prior to sleep does not modulate the effect; (c) the percentage of time spent in SWS, the power of sleep spindles and the coupling phase between spindles and slow waves during the nap mitigate the effect, particularly when the word components are presented in the forward direction that fits their composite presentation after sleep. Higher proportion of SWS predicts more false memories, whereas greater spindle power and better SO-spindle alignment predict fewer false memories; (d) SO amplitude is associated with the weakening of real memories of words presented prior to sleep. In the following we discuss each of these results and suggest future directions to explore.

4.1. Behavioral Effects

Our behavioral results are mostly consistent with the predictions of the temporal scaffolding hypothesis (Lerner & Gluck, 2019; Lerner, 2017a; 2017b), which suggests that temporally compressed replay of an encoded memory sequence during SWS could result in the composition of the sequence into a unitary memory if that memory can be seen as a single entity (as in the case of two words forming together a third, new word). Importantly, there was no difference between the two wake control groups tested either before or after the intermission, supporting the interpretation that sleep was actively changing memory representations rather than simply preventing a change that would have occurred anyway with time. The hypothesis further predicted a stronger effect for a forward presentation of the sequence compared to a backward presentation due to a bias towards forward replay during SWS (Wikenheiser et al., 2013). This effect was not observed behaviorally but was supported at the physiological level, where composite memories in the Forward condition displayed stronger associations with the sleep metrics compared to the Backward condition. Our results allude to several potential explanations for the discrepancy between the behavioral and physiological findings. One possibility is that our experimental manipulation was not strong enough to elicit a detectable behavioral difference between the Forward and Backward conditions. This interpretation is supported by a non-significant but numerical difference between the two conditions, evident in the raw confidence scores of the sleep group (see Table 2). Another possibility is that several sleep mechanisms involved in the consolidation of forward composite words counteracted one another: Whereas SWS amount appeared to shift AUC scores of the Forward condition in one direction, spindle amplitude shifted them in the opposite direction (Table 3). The combination of these effects may therefore have resulted in the Forward and Backward conditions exhibiting similar behavioral scores. Nevertheless, both explanations leave open the question of why AUC scores for the Backward condition were lower than for the Split and Novel conditions in the sleep group. Ultimately, additional experiments may be necessary to clarify this effect, potentially reflecting a sleep-dependent mechanism that is yet unidentified.

4.2. Slow Wave Sleep Metrics as Predictors of Behavior

4.2.1. Slow Wave Metrics and False Memories

The associations between our behavioral findings and the physiological measures of sleep confirmed the involvement of replay-related EEG markers in the process, although in a more nuanced way than originally hypothesized. Whereas the predicted positive correlation between false composite memories and slow wave-related parameters was identified for the percentage of time spent in SWS, spindle power and spindle-SO coupling exhibited the opposite effect. These relationships suggest a more complex mechanism at play, where slow waves and spindles may not always act in concert when contributing to memory consolidation. One possible interpretation of this process is that spindles may have been involved in a form of compensatory mechanism, whereby a tendency to consolidate illusory compositions during slow wave sleep was countered by stronger spindle activity, particularly when they were coupled with slow waves. Spindles and spindle-SO coupling are often thought to reflect a hippocampal-cortical dialogue during which hippocampal memories are replayed, reorganized and transferred to the cortex for long-term storage (Rasch & Born, 2013). However, some models suggest that early forms of reorganization and regularities extraction are already present in the hippocampus itself (Gluck & Myers, 1993; Sucevic & Schapiro, 2023). Considering these models, our results could be explained by assuming: (a) the tendency to form false composite memories occurs within the hippocampus during SWS replay; and (b) stronger spindle activity reflects the disentanglement of those composite memories into their original separate components during the transfer. Such scenario would predict a positive correlation between false memories and the time spent in SWS but a negative correlation with spindle-related metrics. This possibility is strengthened by the fact that the physiological associations were more pronounced in the Forward Composite condition than the Backward condition, further aligning them with the known natural bias towards forward replay during sleep (Wikenheiser et al., 2013). Nevertheless, more evidence is needed to support this hypothesis – for example, by showing a positive correlation between false composite memories and direct indicators of hippocampal replay like sharp wave-ripples (Roumis & Frank, 2015), alongside a negative correlation with cortical-based measures like spindles. Potentially, this prediction could be tested in future experiments that allow measuring hippocampal activity directly, such as animal studies using an adapted version of the current task.

4.2.2. Slow Wave Metrics and Real Memories

Our second finding concerning the association between sleep physiology and behavioral measurements was that raw recognition confidence scores for old words were negatively correlated with the amplitude of SO, signifying a degradation in true memory as the amplitude increased in magnitude. This finding may seem surprising given that no significant difference was found between the sleep and wake groups in the raw recognition scores of old words, and, in addition, most previous sleep studies suggest SWS contributes to the consolidation of memorized words rather than to their weakening (Plihal & Born, 1997; Rasch & Born, 2013). However, a very similar finding to ours was previously reported for false memories using the DRM paradigm. To reiterate, in the DRM task participants are exposed to semantically related words and are later tested on how well they remember these words as well as on whether they falsely remember the unseen theme word linking them. Utilizing this paradigm in a study involving sleep in between exposure and testing, Payne and colleagues (2009) found: (a) a negative correlation between the time spent in SWS during a nap and performance on studied ("veridical") words; (b) no difference in the overall performance on veridical words between the sleep and wake groups; and (c) an increased rate of false memories in the sleep group (see also Newbury & Monaghan, 2019, for a meta-analysis showing no overall difference in veridical memory performance between sleep and wake in the DRM paradigm). These findings were then replicated by the same group, this time also detecting a negative correlation between SWS and false memories (Pardilla-Delgado & Payne, 2017) that echoed previous results in older adults (Lo et al., 2014).

According to Payne and Colleagues, such a negative correlation with SWS may reflect the involvement of the semantic system in memorization during sleep as it tries to efficiently consolidate a list of words while facing the possibility to extract gist information (Payne et al., 2009). Payne and Colleagues suggested that in such circumstances, the system may be encouraged to encode words based on their semantic relatedness rather than the more contextual, episodic-based memorization that SWS is known to enhance. Therefore, participants engaging in more SWS would, unfavorably, over-rely on the less efficient, non-semantic way to encode the DRM stimuli, leading to poorer recollection of veridical and thematic memories alike. Indeed, a tendency of sleep to prioritize abstraction processes at the expense of simple memory consolidation and vice versa has been documented in several previous studies using a variety of different behavioral paradigms (Alger & Payne, 2016; Davidson et al., 2018; Gomez et al., 2006; Lerner et al., 2021). Since our study, like the DRM paradigm, uses word stimuli to test memory, the current findings of SO Amplitude being negatively correlated with memory performance for old words and spindle parameters being negatively correlated with false composite words may reflect a similar process, by which the system invests resources in SWS-dependent episodic memorization at the expense of semantic encoding. Payne and colleagues did not examine any associations between their behavioral measures and more particular SWS metrics such as sleep spindles, SO, or spindle-SO coupling, let alone their combination; therefore, it is not clear if false memories in the DRM paradigm would also reveal a positive relation to some SWS-related parameters if a more robust analysis is conducted. This question would be interesting to examine in future studies, with potential theoretical implications.

4.3. Relation to Other False Memory Paradigms

Finally, our composite words task bears some resemblance to another false memory paradigm, demonstrating "conjunction" errors (Underwood & Zimmerman, 1973). In these studies, participants are first exposed to words like *heartburn* and *drumbeat* and then undergo a memory test where they tend to falsely recognize conjunction words that contain parts of the exposed items, like *heartbeat.* There are, however, some important differences between this paradigm and ours, hinting that different mechanisms are at play. First, semantic relationships between exposed and tested items increase conjunction errors (Leding et al., 2007), whereas no such semantic relations existed in our stimuli. Second, theoretical accounts of conjunction errors point to the contribution of familiarity (Jones et al., 2001), while familiarity per se cannot explain our results given that we found no difference between the Split and Novel conditions. Finally, our results highlight the importance of temporal proximity during exposure to component words whereas no such proximity is required for conjunction errors. Nevertheless, future studies might examine whether sleep also affects the formation of conjunction errors and compare it to the effects found here and in the DRM paradigm.

5. Conclusions

The current study brought evidence for a new form of false memories enhanced by sleep. Unlike previous studies using the DRM paradigm that showed sleep may increase false memories with semantic relations to studied material, the type of false memories demonstrated here were linked to previously presented stimuli only through a low-level temporal association, and their sleep-dependent facilitation is predicted by the temporal scaffolding hypothesis based on the presumed time-compression of memory replay during non-REM sleep. Indeed, as predicted, we found evidence linking the new effect to slow wave sleep, particularly when items were presented in the forward direction, consistent with the known bias towards forward memory replay during sleep. Nevertheless, the linkage between sleep physiology and behavioral effects was more nuanced than predicted, with additional effects showing negative relations between sleep parameters and both false composite and veridical memories, potentially reflecting the involvement of semantic encoding in the process not unlike those present in the DRM paradigm. The full mechanism contributing to these effects remains to be further elucidated.

CRediT authorship contribution statement

Itamar Lerner: Conceptualization, Resources, Methodology, Supervision, Formal Analysis, Writing – Original Draft, Visualization. **Sarah Hamm:** Data Curation, Methodology, Formal Analysis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was partly supported by the National Science Foundation (NSF/EFRI 2317706).

References

Abel, T., Havekes, R., Saletin, J. M., & Walker, M. P. (2013). Sleep, plasticity and memory from molecules to whole-brain networks. *Current Biology, 23*(17), 774–788.

Alger, S. E., & Payne, J. D. (2016). The differential effects of emotional salience on direct associative and relational memory during a nap. *Cognitive, Affective, & Behavioral Neuroscience, 16*(6), 1150–1163.

August, D. A., & Levy, W. B. (1999). Temporal sequence compression by an integrate-and-fire model of hippocampal area CA3. *Journal of Computational Neuroscience, 6*(1), 71–90.

Davidson, P., Carlsson, I., Jönsson, P., & Johansson, M. (2018). A more generalized fear response after a daytime nap. *Neurobiology of Learning and Memory, 151*, 18–27.

Davies, M. (2008-). Word frequency data from The Corpus of Contemporary American English (COCA). Data available online at https://www.wordfrequency.info.

Diba, K., & Buzsaki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. *Nature Neuroscience, 10*(10), 1241–1242.

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience, 11*(2), 114-126.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography, 36*(1), 27–46.

Durrant, S. J., Taylor, C., Cairney, S., & Lewis, P. A. (2011). Sleep-dependent consolidation of statistical learning. *Neuropsychologia, 49*(5), 1322–1331.

Ellenbogen, J. M., Hulbert, J. C., Stickgold, R., Dinges, D. F., & Thompson-Schill, S. L. (2006). Interfering with theories of sleep and memory: sleep, declarative memory, and associative interference. *Current Biology*, *16*(13), 1290-1294.

Fischer, S., Drosopoulos, S., Tsen, J., & Born, J. (2006). Implicit learning–explicit knowing: A role for sleep in memory system interaction. *Journal of Cognitive Neuroscience, 18*(3), 311–319

Friedrich, M., Wilhelm, I., Born, J., & Friederici, A. D. (2015). Generalization of word meanings during infant sleep. *Nature Communications, 6,* 6004.

Gluck, M. A., & Myers, C. E. (1993). Hippocampal mediation of stimulus representation: A computational theory. *Hippocampus*, *3*(4), 491-516.

Gomez, R. L., Bootzin, R. R., & Nadel, L. (2006). Naps promote abstraction in language-learning infants. *Psychological Science, 17*(8), 670–674.

Graveline, Y. M., & Wamsley, E. J. (2017). The impact of sleep on novel concept learning. *Neurobiology of Learning and Memory, 141,* 19–26.

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

Helfrich, R. F., Mander, B. A., Jagust, W. J., Knight, R. T., & Walker, M. P. (2018). Old brains come uncoupled in sleep: slow wave-spindle synchrony, brain atrophy, and forgetting. *Neuron*, *97*(1), 221-230.

Ji, D., & Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience, 10*(1), 100–107.

Jones, T. C., Jacoby, L. L., & Gellis, L. A. (2001). Cross-modal feature and conjunction errors in recognition memory. *Journal of Memory and Language, 44*(1), 131–152.

Kumral, D., Matzerath, A., Leonhart, R., & Schönauer, M. (2023). Spindle-dependent memory consolidation in healthy adults: A meta-analysis. *Neuropsychologia, 189*, 108661.

Lacaux, C., Andrillon, T., Bastoul, C., Idir, Y., Fonteix-Galet, A., Arnulf, I., & Oudiette, D. (2021). Sleep onset is the creative sweet spot. *Science Advances, 7*(8), eabj586.

Ladenbauer, J., Ladenbauer, J., Külzow, N., & Flöel, A. (2021). Memory-relevant nap sleep physiology in healthy and pathological aging. *Sleep*, *44*(7), zsab002.

Leding, J. K., Lampinen, J. M., Edwards, N. W., & Odegard, T. N. (2007). The memory conjunction error paradigm: Normative data for conjunction triplets. *Behavior Research Methods, 39*(4), 920–925.

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Sciences, 15*(8), 343–351.

Lerner, I. (2017a). Sleep is for the brain: Contemporary computational approaches in the study of sleep and memory and a Novel 'Temporal Scaffolding' Hypothesis. In A. Moustafa (Ed.), *Computational Models of Brain and Behavior* (pp. 241-261). Hoboken, NJ: Wiley.

Lerner, I. (2017b). Unsupervised temporal learning during sleep supports insight. *Proceedings of the Conference on Cognitive Computational Neuroscience* (CCN), 2017.

Lerner, I., & Gluck, M. A. (2018). Individual differences in slow-wave sleep predict acquisition of full cognitive maps. *Frontiers in Human Neuroscience, 12,* 404.

Lerner, I., & Gluck, M. A. (2019). Sleep and the extraction of hidden regularities: A systematic review and the importance of temporal rules. *Sleep Medicine Reviews, 47*, 39–50.

Lerner, I., & Gluck, M. A. (2022). Sleep facilitates extraction of temporal regularities with varying timescales. *Frontiers in Behavioral Neuroscience, 16*, 847083.

Lerner, I., Kerbaj, T. P., & Gluck, M. A. (2019). When sleep-dependent gist extraction goes awry: False composite memories are facilitated by slow wave sleep. In *Proceedings of the 41st Annual Meeting of the Cognitive Science Society (CogSci 2019): Creativity+ Cognition+ Computation* (pp. 2119–2124). The Cognitive Science Society.

Lerner, I., Lupkin, S. M., Tsai, A., Khawaja, A., & Gluck, M. A. (2021). Sleep to remember, sleep to forget: Rapid eye movement sleep can have inverse effects on recall and generalization of fear memories. *Neurobiology of Learning and Memory, 180*, 107413.

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in cognitive sciences*, *15*(8), 343-351

Lo, J. C., Sim, S. K., & Chee, M. W. (2014). Sleep reduces false memory in healthy older adults. *Sleep*, *37*(4), 665-671.

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

Newbury, C. R., & Monaghan, P. (2019). When does sleep affect veridical and false memory consolidation? A meta-analysis. *Psychonomic Bulletin & Review, 26*(2), 387–400.

Pardilla-Delgado, E., & Payne, J. D. (2017). The impact of sleep on true and false memory across long delays. *Neurobiology of learning and memory*, *137*, 123-133.

Payne, J. D., Schacter, D. L., Propper, R. E., Huang, L.-W., Wamsley, E. J., & Tucker, M. A. (2009). The role of sleep in false memory formation. *Neurobiology of Learning and Memory, 92*(3), 327–334.

Payne, J. D., Tucker, M. A., Ellenbogen, J. M., Wamsley, E. J., Walker, M. P., Schacter, D. L., & Stickgold, R. (2012). Memory for semantically related and unrelated declarative information: the benefit of sleep, the cost of wake. *PloS one*, *7*(3), e33079.

Plihal, W., & Born, J. (1997). Effects of early and late nocturnal sleep on declarative and procedural memory. *Journal of Cognitive Neuroscience, 9*(4), 534–547.

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(3), 785–794.

Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological Reviews, 93*(2), 681–766.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803–814.

Roumis, D. K., & Frank, L. M. (2015). Hippocampal sharp-wave ripples in waking and sleeping states. *Current Opinion in Neurobiology, 35*, 6–12.

Schmidig, F. J., Geva-Sagiv, M., Falach, R., Yakim, S., Gat, Y., Sharon, O., ... & Nir, Y. (2024). A visual paired associate learning (vPAL) paradigm to study memory consolidation during sleep. *Journal of Sleep Research*, *33*(5), e14151.

Schreiner, T., Petzka, M., Staudigl, T., & Staresina, B. P. (2021). Endogenous memory reactivation during sleep in humans is clocked by slow oscillation-spindle complexes. *Nature communications*, *12*(1), 3112.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31*(1), 137–149.

Staresina, B. P., Bergmann, T. O., Bonnefond, M., Van Der Meij, R., Jensen, O., Deuker, L., ... & Fell, J. (2015). Hierarchical nesting of slow oscillations, spindles and ripples in the human hippocampus during sleep. *Nature neuroscience*, *18*(11), 1679-1686.

Sučević, J., & Schapiro, A. C. (2023). A neural network model of hippocampal contributions to category learning. *Elife*, *12*, e77185.

Underwood, B. J., & Zimmerman, J. (1973). The syllable as a source of error in multisyllable word recognition. *Journal of Verbal Learning and Verbal Behavior, 12*(6), 701–706.

Vallat, R., & Walker, M. P. (2021). An open-source, high-performance tool for automated sleep staging. *eLife, 10*, e70092.

Vickrey, B., & Lerner, I. (2023). Overnight exposure to pink noise could jeopardize sleep-dependent insight and pattern detection. *Frontiers in Human Neuroscience, 17*, 1302836.

Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature, 427*(6972), 352–355.

Walker, M. P., & Stickgold, R. (2006). Sleep, memory, and plasticity. *Annual Review of Psychology, 57,* 139–166.

Wikenheiser, A. M., & Redish, A. D. (2013). The balance of forward and backward hippocampal sequences shifts across behavioral states. *Hippocampus, 23*(1), 22–29.

Wilhelm, I., Rose, M., Imhof, K. I., Rasch, B., Büchel, C., & Born, J. (2013). The sleeping child outplays the adult's capacity to convert implicit into explicit knowledge. *Nature Neuroscience, 16*(3), 391–393

Yeh, S. T. (2002). Using trapezoidal rule for the area under a curve calculation. *Proceedings of the 27th Annual SAS® User Group International (SUGI'02)*, (pp. 1-5).

Yordanova, J., Kolev, V., & Kirov, R. (2012). Brain oscillations and predictive processing. *Frontiers in Psychology, 3*, 416.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American statistical Association*, *57*(298), 348-368

Tables

Table 1. Demographic and Sleep Data of Participants (N = 106)

Variable	Sleep Group	Wake-Delayed Group	Wake-Immediate Group
Number of Subjects	N = 35	N = 36	N = 35
Gender (M / F)	15 / 20	14 / 22	15 / 20
Age	19.54 ± 1.93	19.56 ± 1.98	19.86 ± 1.44
Years of Education	14.11 ± 1.19	13.94 ± 1.50	14.50 ± 1.43
TST (min)	80.88 ± 25.67	-	-
N1 (min)	9.65 ± 6.24	-	-
% N1	14.28 ± 12.18	-	-
N2 (min)	42.62 ± 16.12	-	-
% N2	54.66 ± 15.61	-	-
N3 (min)	16.57 ± 17.36	-	-
% N3	17.50 ± 16.93	-	-
REM (min)	12.05 ± 10.97	-	-
% REM	13.56 ± 12.21	-	-

Demographic data of participants based on group. Sleep parameters presented only for the Sleep group. Numbers above represent Mean ± Standard Deviation. TST = total sleep time. N1 = minutes spent in stage 1 sleep. % N1 = percentage of time spent in stage 1 sleep out of total sleep time. N2 = minutes spent in stage 2 sleep. % N2 = percentage of time spent in stage 2 sleep out of total sleep time. N3 = minutes spent in SWS. % N3 = percentage of time spent in SWS out of total sleep time. REM = minutes spent in Rapid Eye Movement sleep. % REM = percentage of time spent in REM sleep out of total sleep time.

Table 2. Means and standard deviations of the confidence levels for each group and condition

	Sleep	Wake-Delayed	Wake-Immediate
Composite	3.55 (0.61)	3.38 (0.69)	3.29 (0.48)
Composite - Forward	3.62 (0.86)	3.32 (0.84)	3.34 (0.61)
Composite - Backward	3.49 (0.83)	3.44 (0.93)	3.25 (0.68)
Split	3.28 (0.87)	3.42 (0.79)	3.40 (1.06)
Novel	3.06 (0.79)	3.07 (0.92)	3.11 (0.97)
Old	4.03 (0.58)	4.22 (0.64)	4.07 (0.54)

Higher values represent higher confidence that a word has been seen before.

Table 3. Multiple regression for AUC scores in each condition using the sleep predictors

	Composite		Split		Novel	
Model fit	F(6,21) = 4.4468 p = 0.005, R^2 = 0.4337		F(6,21) = 2.4820 p = 0.056, R^2 = 0.2477		F(6,21) = 3.7933 p = 0.010, R^2 = 0.3830	
	β	p	β	p	β	p
Intercept	0.3043	0.2756	0.8472	0.0228	0.9270	0.0055
% SWS	-0.3909	0.0016	-0.2735	0.0583	-0.2211	0.0767
SO Amplitude	-0.0037	0.0674	-0.0004	0.0862	-0.0054	0.0170
Spindle Density	-0.0044	0.7966	-0.0180	0.4071	-0.0039	0.8369
Spindle Power	0.3426	0.0026	0.1179	0.3655	0.1360	0.2335
Coupling Phase	0.0106	0.7704	-0.0218	0.6396	-0.0489	0.2299
Coupling Strength	-0.0006	0.9637	0.1327	0.4138	0.0135	0.9233

p values of the overall model fit are before correction for multiple comparisons. Significant or marginally significant effects of predictors (not including intercept) are marked in grey. Adjusted R^2 values are presented. % SWS = Percent of time spent in slow wave sleep; SO = Slow Oscillations.

Table 4. Multiple regression for AUC scores in the Composite Forward and Backward conditions

	Forward		Backward	
Model fit	F(6,21) = 3.3884 p = 0.017, R^2 = 0.3467		F(6,21) = 1.7044 p = 0.169, R^2 = 0.1353	
	β	p	β	p
Intercept	0.1934	0.6235	0.7020	0.0846
% SWS	-0.4748	0.0056	-0.2080	0.1899
SO Amplitude	-0.0051	0.0746	-0.0034	0.2217
Spindle Density	-0.0166	0.4959	-0.0020	0.9330
Spindle Power	0.5092	0.0019	0.1113	0.4460
Coupling Phase	0.0865	0.1057	-0.0670	0.2042
Coupling Strength	0.0738	0.6846	-0.1132	0.5339

p values of the overall model fit are presented before correction for multiple comparisons. Significant or marginally significant effects of predictors (not including intercept) are marked in grey. Adjusted R^2 values are presented. % SWS = Percent of time spent in slow wave sleep; SO = Slow Oscillations.

Table 5. Multiple regression for raw recognition scores using SO-Spindle coupling as predictors

	Forward		Backward		Split		Novel		Old	
Model Fit	F(6,21) = 3.68 p = 0.012 R^2 = 0.374		F(6,21) = 0.09 p = 0.997 R^2 = -0.254		F(6,21) = 0.56 p = 0.751 R^2 = -0.106		F(6,21) = 0.13 p = 0.991 R^2 = -0.240		F(6,21) = 4.96 p = 0.003 R^2 = 0.468	
	β	p	β	p	β	p	β	p	β	p
Intercept	7.30	0.005	4.13	0.190	3.19	0.291	3.17	0.273	6.02	<0.001
% SWS	1.98	0.045	0.53	0.662	0.87	0.464	0.529	0.640	-0.90	0.101
SO amp	0.01	0.719	-0.00	0.871	-0.00	0.876	0.00	0.844	-0.03	0.005
Sp. dns	0.12	0.400	0.12	0.545	0.19	0.303	0.10	0.564	-0.03	0.810
Sp. pow	-2.66	0.005	-0.29	0.801	0.04	0.969	-0.30	0.774	0.54	0.285
Cpl. phs	-1.16	0.001	-0.06	0.884	-0.35	0.379	-0.12	0.743	-0.28	0.123
Cpl. str	-0.95	0.390	0.32	0.820	-1.12	0.419	-0.40	0.762	-0.05	0.933

p values of the overall model fit are presented before correction for multiple comparisons. Significant effects of predictors (not including intercept) are marked in grey. Adjusted R^2 values are presented. % SWS = Percent of time spent in slow wave sleep; SO amp = Slow Oscillations amplitude; Sp. dns = spindle density; Sp. Pow = spindle power; Cpl. Phs = Spindle-SO coupling phase; Cpl. Str = Spindle-SO coupling strength.

Figures

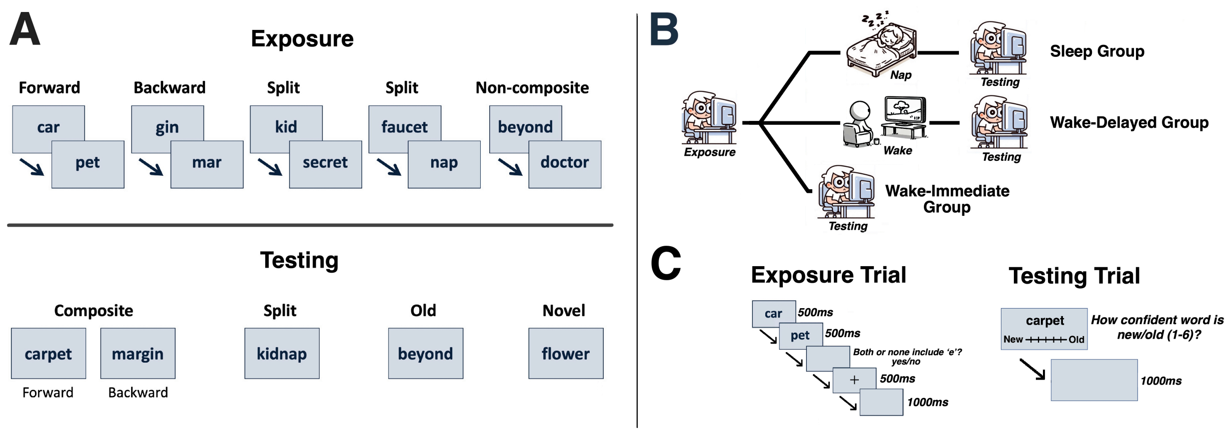


Fig. 1. Experimental design. A) Types of Trials in the exposure and testing Sessions**.** Top row shows the four types of trials administered to participants during the exposure session: Forward, Backward, Split, and Novel. Forward, Backward and Split conditions were counterbalanced across participants. Bottom image shows the four types of trials administered during the testing session: Composite (consisting of word compositions from either the Forward or Backward trials in the exposure session) Split (consisting of word compositions from the Split trials in the exposure session), Old (any non-composite words appearing during exposure), and Novel (words not seen during exposure). B) Example of trial progression for the Exposure (left) and Testing (right) sessions. During exposure, participants indicated if one or both words of a presented pair contained or did not contain the letter ‘e’. During testing, participants indicated how confident they are that a presented word was seen earlier during the exposure session. C) Experimental procedure. The Sleep group performed the exposure session during early afternoon, then had an opportunity to nap for 2 hours, followed by the testing session. The Wake-Delayed group had a similar schedule, only they were not allowed to nap during the 2-hour interval. The Wake-Immediate group performed the exposure session at a similar time, and the testing session 15 minutes afterwards without a long interval in between.

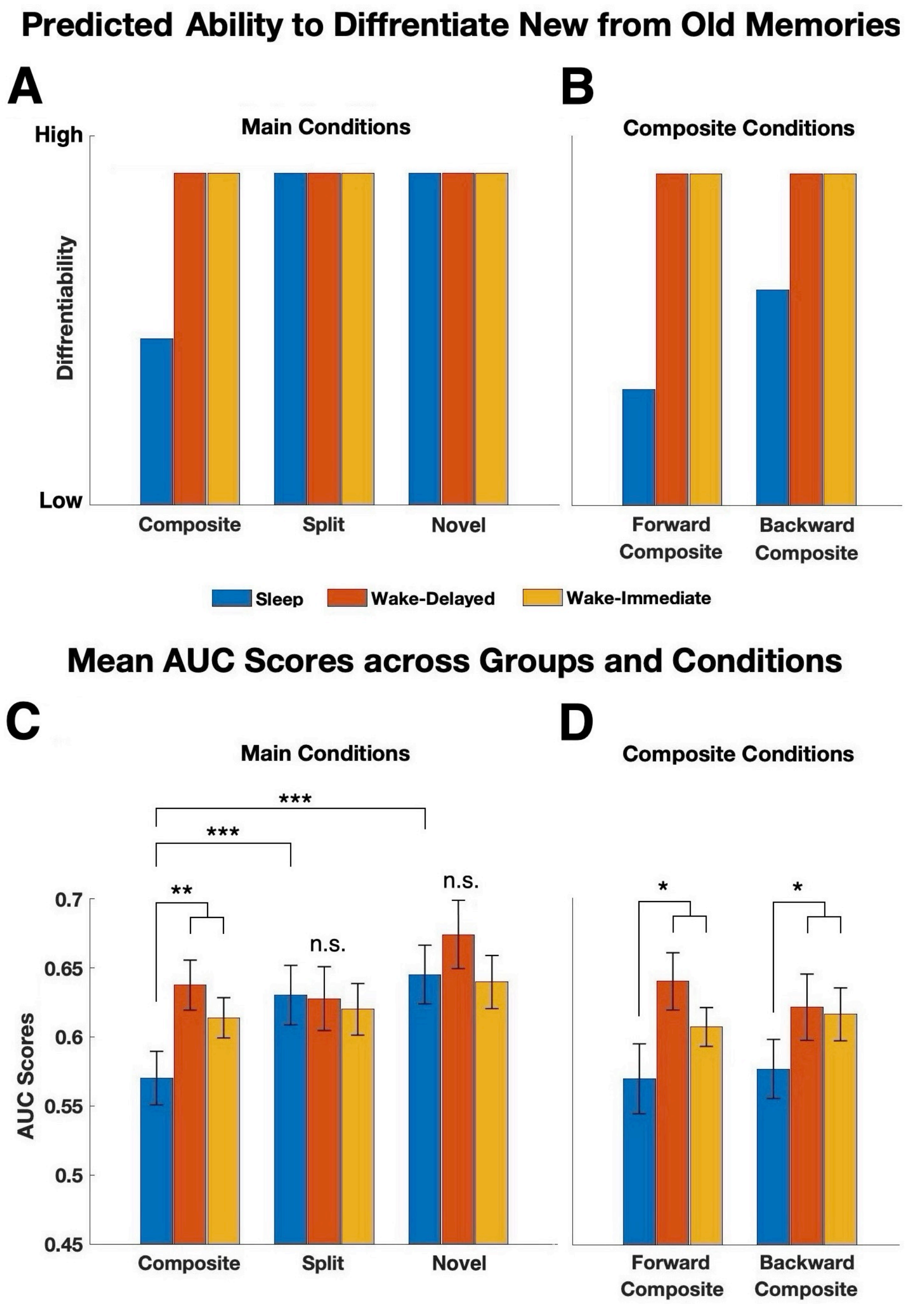


Fig. 2. Main behavioral results of the experiment compared to predicted results**.** A) Illustration of the predicted differences in the ability to differentiate new from old words between the sleep and wake control groups for each of the main conditions (arbitrary units). B) Same as A, for the Forward versus Backward Composite conditions. C) Empirical AUC scores of all groups for the main experimental conditions. D) Same as C, for the Forward versus Backward Composite conditions. *** $p < 0.001$; ** $p = 0.01$; * $p = 0.02$; n.s., not significant.

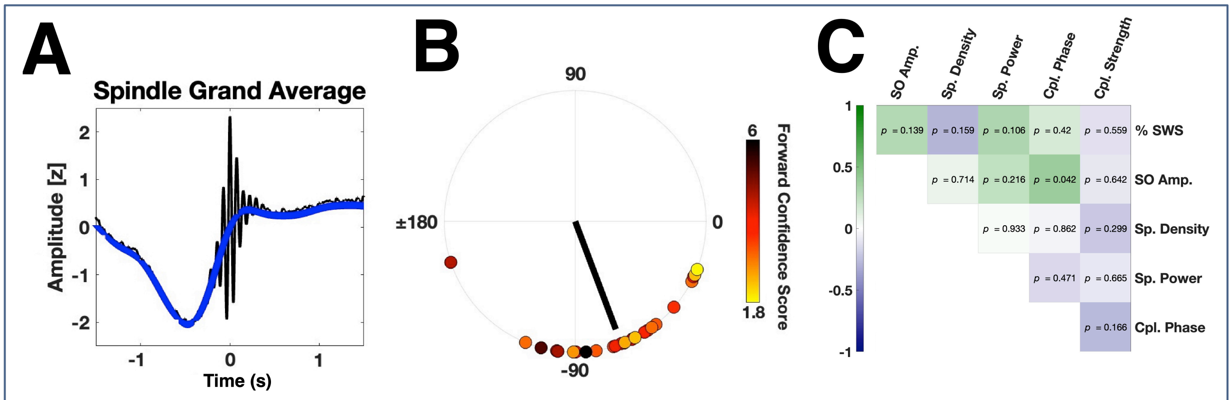


Fig. 3. A) Grand average of EEG central channels across all SO-spindle events and participants, time-locked to spindle peak amplitude (black), together with the same segments filtered in the SO band (blue). B) Preferred SO phase at peak spindle amplitude across participants. Color represents the average raw confidence score in the Forward Composite condition per participant, from low (yellow) to high (dark red). C) Pairwise correlations between the physiological sleep predictors. Blue/green Intensity of each cell represents the value of the correlation (1 – Green, -1 – blue) with the corresponding significance level indicated at the center. % SWS = Percent of time spent in slow wave sleep; SO Amp. = Slow oscillations amplitude; Sp. Density = spindle density; Sp. Power = spindle power; Cpl. Phase = SO-spindle coupling phase; Cpl. Strength = SO-spindle coupling strength.

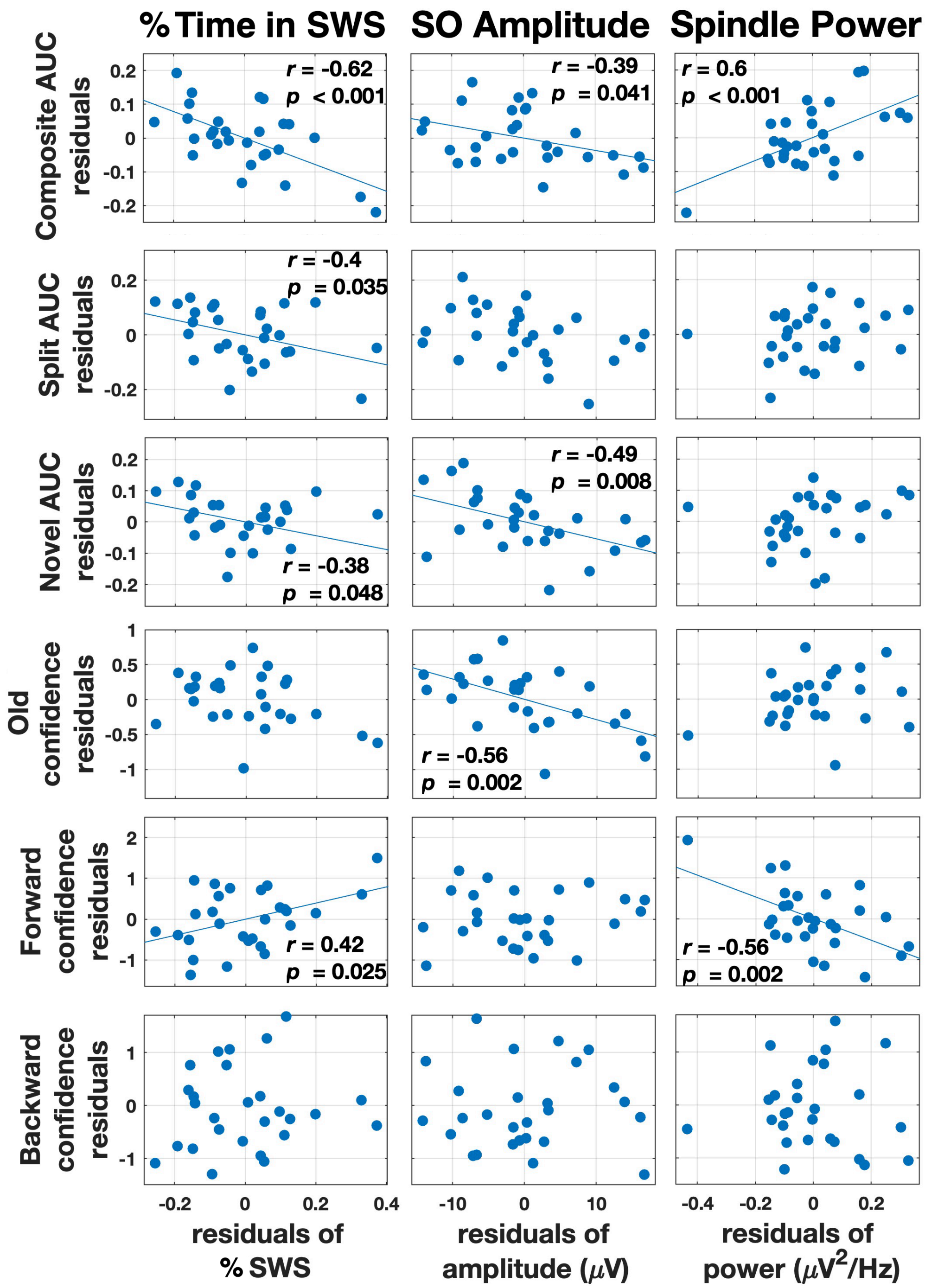


Fig. 4. Partial correlations between main sleep parameters and AUC scores/raw recognition confidence scores for the critical conditions. *r* and *p* values for the residual correlations and the corresponding slope are displayed for significant correlations.

Supplementary Materials

Sleep Promotes Illusory Word Compositions, a Distinct Form of False Memory

Itamar Lerner[1] & Sarah C. Hamm[1]

[1] Department of Psychology, The University of Texas at San Antonio

Corresponding author:
Itamar Lerner
One UTSA Circle
San Antonio, Texas, 78249 USA
Email: Itamar.lerner@utsa.edu

Supplementary Methods

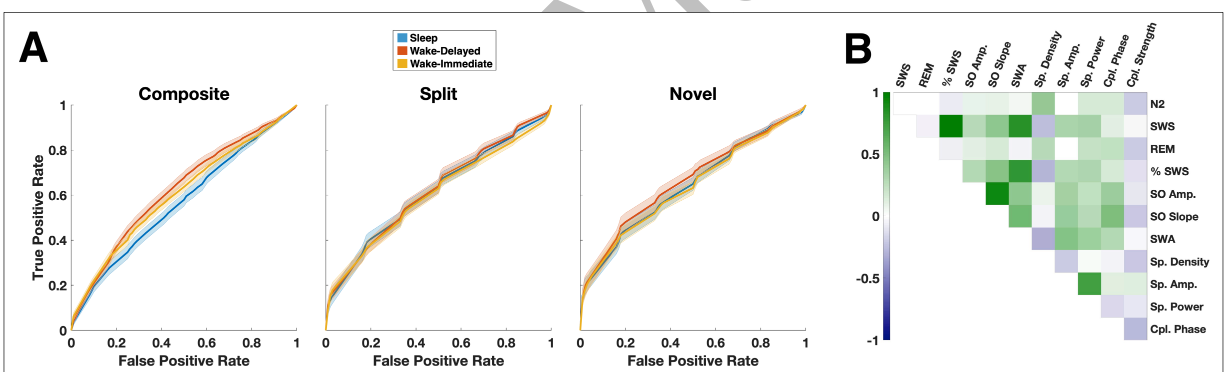
<u>Composite words used in the study</u>

Supplementary Table S1. Composite words used in the study, organized by list.

List 1		List 2		List 3	
Word	*Frequency*	*Word*	*Frequency*	*Word*	*Frequency*
ashram	0.3	campus	40.74	convex	0.57
canvas	13.75	carpet	15.36	dampen	2.38
hamper	4.38	curfew	2.82	donkey	3.81
haptic	0.25	format	25.25	kidnap	9.47
pallet	1.56	parrot	3.96	margin	22.57
refuse	73.09	warden	4.33	profit	54.06
Mean:	15.56		15.41		15.48

Frequency in occurrences/million.

Supplementary Results



Supplementary Fig. S1. A) Average ROC curves for each condition and group. B Pairwise correlations between extended set of sleep parameters. SWS = Slow Wave Sleep. REM = Rapid Eye Movement; SO = Slow Oscillations; SWA = Slow Wave Activity; Amp. = Amplitude; Sp. = Spindles. Cpl = Coupling.

Stepwise Regression Across Extended Set of Sleep Parameters

In our main regression analysis of sleep physiology and behavior, predictors were chosen for their well-known role as indices of memory consolidation during slow wave sleep. However, there are other sleep parameters that have been used in the literature in relation with memory consolidation (e.g., Kumral et al., 2023). Many of these parameters are highly correlated with each other (e.g., $|r|>0.7$; See Fig S1B), preventing their inclusion in one single model due to multicollinearity concerns (Dormann et al., 2013). Nevertheless, it is important to verify that the main results are not dependent on a particular set of selected parameters. To that end, we complemented our multiple regression analyses by employing an alternative analytic approach to identify which predictors out of a larger set of potentially relevant parameters should be preferred when accounting for the behavioral data.

Methods: We ran a stepwise regression analysis for each word condition presented in the main text to select factors that produce the best fit from a model selection perspective. Each of these analyses started with a default intercept-only model. At every step, variables were selected based on p-values of the F-statistics, with p <.05 being the criterion to add a term to the model and p >.10 being the criterion to remove an included term from the model. The set of potential parameters included all the original variables analyzed in the main text (%SWS, SO amplitude, spindle density, spindle power and SO-spindle coupling phase and strength), as well as the raw time spent in the main sleep stages (N2, SWS, REM), additional slow-wave parameters (slow-wave activity (SWA), SO slope), and additional spindle parameters (spindle count, spindle amplitude). Time in each sleep stage was based on sleep scoring, and SO slope and spindle count and amplitude were extracted using the same YASA algorithm described in the main text. SWA was calculated following the method presented by Wilhelm and colleagues (2011): The raw data from the EEG derivations of each participant underwent high-pass (0.1Hz) and low-pass (30Hz) filtering and cleaned using Independent Component Analysis. Then, for all time periods spent during either N2 or SWS, the data were segmented into 10min consecutive time bin (we used 10 instead of 20 minutes as in the original paper because we measured naps rather than overnight sleep, resulting in substantially shorter recordings) and for each bin we extracted the overall power in the SO and delta band (0.5-4Hz). We then calculated the median value over these bins for each derivation, and then the median over derivations from the central and frontal areas (C4-

M1, C3-M2, F1-M1, F2-M2; we used medians rather than means to restrict occasional distortions resulting from channels with high non-biological noise, for example due to movement that was not sufficiently cleaned). This value served as the SWA for each participant. All processing related to SWA was executed with Matlab's EEGLAB and the signal processing toolbox, and the stepwise procedure was accomplished with the function *stepwiselm* in Matlab 2024b.

Results:

For the main AUC scores, results were as follows:

- Composite condition: The model was significant ($F(3, 24) = 4.74$, $p = 0.039$, Adjusted $R^2 = 0.500$) with %SWS and SO amplitude retained as predictors with a negative association and spindle power retained as a predictor with a positive association.
- Split condition: The model was significant ($F(1, 26) = 7.79$, $p = 0.009$, Adjusted $R^2 = 0.201$) with SO amplitude retained as a sole predictor with a negative association.
- Novel condition: The model was significant ($F(1, 26) = 13.87$, $p < 0.001$, Adjusted $R^2 = 0.323$) with SO amplitude retained as a sole predictor with a negative association.

For the Forward and Backward composite AUC scores, results were as follows:

- Forward condition: The model was significant ($F(2, 25) = 6.87$, $p = 0.004$, Adjusted $R^2 = 0.303$) with %SWS and SO amplitude retained as predictors with a negative association and spindle power retained as a predictor with a positive association.
- Backward condition: The model was significant ($F(1, 26) = 6.05$, $p = 0.021$, Adjusted $R^2 = 0.158$) with SO slope retained as a sole predictor with a negative association.

For the raw confidence sores, results were as follows:

- Forward condition: The model was significant ($F(3, 24) = 10.4$, $p < 0.001$, Adjusted $R^2 = 0.510$) with SWS retained as a predictor with a positive association and spindle amplitude and SO-spindle coupling phase retained as a predictors with a negative association.
- Backward condition: no predictors added, default intercept-only model retained.
- Split condition: no predictors added, default intercept-only model retained.
- Novel condition: no predictors added, default intercept-only model retained.
- Old condition: The model was significant ($F(1, 26) = 22.6$, $p < 0.001$, Adjusted $R^2 = 0.444$) with SO slope retained as a sole predictor with a negative association.

Overall, results from the stepwise regression analysis closely mirrored findings from the multiple regression models (compare with Tables 2–4 in the main text) with differences mostly limited to substitutions among conceptually similar and highly correlated parameters (SWS and %SWS; SO amplitude and SO slope; spindle amplitude and spindle power; see Fig S1B), confirming the robustness of our main findings.

References:

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography, 36*(1), 27–46.

Kumral, D., Matzerath, A., Leonhart, R., & Schönauer, M. (2023). Spindle-dependent memory consolidation in healthy adults: A meta-analysis. *Neuropsychologia, 189*, 108661.

Wilhelm, I., Diekelmann, S., Molzow, I., Ayoub, A., Mölle, M., & Born, J. (2011). Sleep selectively enhances memory expected to be of future relevance. *Journal of Neuroscience, 31*(5), 1563–1569.