Sleep Strengthens Predictive Sequence Coding

Nicolas D. Lutz,^{1,2} Ines Wolf,^{1*} Stefanie Hübner,^{1*} Jan Born,^{1,3} and Karsten Rauss¹

¹Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, 72076 Tübingen, Germany, ²Graduate Training Centre of Neuroscience/IMPRS for Cognitive and Systems Neuroscience, University of Tübingen, 72074 Tübingen, Germany, and ³Werner Reichardt Centre for Integrative Neuroscience, University of Tübingen, 72076 Tübingen, Germany

Predictive-coding theories assume that perception and action are based on internal models derived from previous experience. Such internal models require selection and consolidation to be stored over time. Sleep is known to support memory consolidation. We hypothesized that sleep supports both consolidation and abstraction of an internal task model that is subsequently used to predict upcoming stimuli. Human subjects (of either sex) were trained on deterministic visual sequences and tested with interleaved deviant stimuli after retention intervals of sleep or wakefulness. Adopting a predictive-coding approach, we found increased prediction strength after sleep, as expressed by increased error rates to deviant stimuli, but fewer errors for the immediately following standard stimuli. Sleep likewise enhanced the formation of an abstract sequence model, independent of the temporal context during training. Moreover, sleep increased confidence for sequence knowledge, reflecting enhanced metacognitive access to the model. Our results suggest that sleep supports the formation of internal models which can be used to predict upcoming events in different contexts.

Key words: abstraction; consolidation; predictive coding; sequence learning; sleep

Significance Statement

To efficiently interact with the ever-changing world, we predict upcoming events based on similar previous experiences. Sleep is known to benefit memory consolidation. However, it is not clear whether sleep specifically supports the transformation of past experience into predictions of future events. Here, we find that, when human subjects sleep after learning a sequence of predictable visual events, they make better predictions about upcoming events compared with subjects who stayed awake for an equivalent period of time. In addition, sleep supports the transfer of such knowledge between different temporal contexts (i.e., when sequences unfold at different speeds). Thus, sleep supports perception and action by enhancing the predictive utility of previous experiences.

Introduction

In a world of constant change, internal models are essential to efficiently interact with our environment. This is the central notion in current theories of hierarchical predictive coding, which postulate top-down prediction signals that “explain away” predictable (model-congruent) information, and bottom-up prediction error signals of unexplained (model-incongruent) events that are used to update predictions of future sensory inputs (Gregory, 1980; Mumford, 1992; Rao and Ballard, 1999; Friston,

2012; Rauss and Pourtois, 2013; Rauss and Born, 2017). Thus, adequate internal models should facilitate the detection of unexpected events (Vinken et al., 2017) while at the same time renormalizing behavior as soon as the environment again unfolds as predicted. There is ample evidence that our brains implement principles of predictive coding (Summerfield et al., 2006; Alink et al., 2010; Smith and Muckli, 2010; Vinken et al., 2017; Kaposvari et al., 2018). However, it remains largely unclear how predictions are distilled from previously encoded information and how they are maintained over extended periods (Hinton et al., 1995).

In the present study, we addressed to what extent sleep contributes to the formation, consolidation, and abstraction of simple task models obtained from sequences of visual stimuli (Nissen and Bullemer, 1987). From a predictive-coding perspective, successful formation of sequence knowledge involves the acquisition of a stable internal model representing the rules that govern the sequence. This should yield informed behavioral responses (Fiser et al., 2010), that is, correct predictions of upcoming stimuli, as expressed in reduced error rates and reaction times; conversely, unexpected stimuli would induce prediction errors, as expressed

Received May 24, 2018; revised Aug. 23, 2018; accepted Aug. 24, 2018.

Author contributions: N.D.L., J.B., and K.R. designed research; N.D.L., I.W., and S.H. performed research; N.D.L. and K.R. analyzed data; N.D.L., J.B., and K.R. wrote the paper.

This work was supported by Deutsche Forschungsgemeinschaft Grant TR-SFB 654 (Plasticity and Sleep). We thank Susanne Diekelmann for help with the conception of the study; and Marco Rütt for technical assistance.

The authors declare no competing financial interests.

*I.W. and S.H. contributed equally to this work.

Correspondence should be addressed to Dr. Karsten Rauss, Institute of Medical Psychology and Behavioral Neurobiology, Otfrid-Müller-Straße 25, 72076 Tübingen, Germany. E-mail: karsten.rauss@uni-tuebingen.de.

DOI:10.1523/JNEUROSCI.1352-18.2018

Copyright © 2018 the authors 0270-6474/18/388989-12\$15.00/0

in higher error rates and/or increased reaction times (Stadler, 1995; Peigneux et al., 2000; Robertson, 2007). Given the known benefits of sleep for memory consolidation (Rasch and Born, 2013), we hypothesized that sleep after sequence learning promotes the formation of internal sequence models, thereby strengthening error detection as well as renormalizing behavior after unexpected events.

We trained a group of healthy humans on deterministic sequences of visual stimuli. We used long sequences to ensure learning would remain implicit (Robertson et al., 2004; Song et al., 2008; Song and Cohen, 2014; Rosenthal et al., 2016), and short response-to-stimulus intervals (RSIs; 200 ± 50 ms) to ensure that stimuli are clearly perceived as a sequence. Given our predictive-coding approach, we focused the analyses on error rates for deviant stimuli, and those immediately following deviants. Such so-called “follower standards” have been shown to initially elicit behavior that is more similar to deviant than standard stimuli (Kaposvari et al., 2018). If sleep supports consolidation of an internal sequence model, we reasoned that sleep compared with wakefulness should increase the number of errors in response to model-incongruent deviants, reflecting stronger prediction errors in the presence of a more consolidated sequence model. Conversely, sleep subjects should display reduced error rates in response to follower standards, reflecting a more adequate sequence model that allows them to bridge the gap between two standard stimuli separated by a random deviant. Our results confirm this hypothesis.

Additionally, we examined whether sleep also supports the emergence of a more abstract task model that is independent of the original training context (Witt et al., 2010). Focusing on temporal aspects, we find that sleep boosts prediction strength when training was performed with longer intervals between stimuli than testing.

Materials and Methods

Participants. In total, 128 young healthy adults (64 females) participated in the experiment (mean age: 25 years; range: 20–37 years). Four different groups ($n_{RSI} = 32$ each) were tested, each consisting of two subgroups ($n_{sleep/wake} = 16$): one subgroup was tested before and after normal nighttime sleep (Sleep groups); the other subgroup was tested once in the morning and once in the evening of the same day (Wake groups) (Fig. 1). Subjects did not take any medication, did not report any neurological or psychological disorders, had a normal sleep–wake cycle (i.e., they usually went to bed between 2200 and 2400 h, and got up between 0700 and 0900 h), were right-handed, and had normal or corrected-to-normal visual acuity. They were not allowed to ingest caf-

Figure 1. Task and experimental design. *a*, Subjects sat in front of a computer screen at a distance of 57 cm while visual stimuli were presented in a sequential order. Their task was to indicate as fast and accurately as possible at which of six possible locations each stimulus was presented, by pressing corresponding buttons on a computer keyboard. Subjects were instructed to fixate a cross in the middle of the screen at all times during the measurements. Stimulus sequences consisted of 12-items and contained 10 predictable standard stimuli (shown as white circles) and one unpredictable deviant stimulus (deviating from the standard location sequence; shown as a black circle). Stimuli immediately following deviants are different from normal standards and called follower standards (shown as a gray circle). Shown is an example deviant sequence containing all three stimulus types. *b*, The stimulus layout consisted of six locations, oriented in a circle around the center of the screen. Grayscale Gabor gratings of 5° of visual angle were displayed one after another in two different stimulus orientations (tilted at an angle of 45° or 135° , respectively). *c*, Example standard and deviant sequences. In the latter, one stimulus deviates from the standard location sequence and is immediately followed by a follower standard (for more information, see Materials and Methods). *d*, Subjects were trained on only the standard stimuli either in the evening (Sleep group) or in the morning (Wake group) and tested after 30 min (pre-retention test). Following sleep versus wakefulness, subjects came back in the morning of the following day (Sleep group) or the evening of the same day (Wake group) and were tested again (post-retention test). Individual deviant stimuli that deviated from the sequence of standard stimuli were pseudo-randomly shown during both pre-retention and post-retention tests.

feine or alcohol during the days of the experimental sessions. Musicians or professional typists were excluded from the study. Eleven subjects were excluded from analysis due to them being outliers on tests for explicitness of the trained regularity (i.e., >2 SDs above the group mean for free recall, both liberal and conservative, or triplet completion, meaning that they displayed high explicit knowledge). Thus, 117 subjects were included in the final analyses (31 in the Short-to-short group, 29 in the

Short-to-long group, 28 in the Long-to-short group, and 29 in the Long-to-Long group; see Experimental design and procedure). However, results remain largely equivalent when taking into account all subjects disregarding their explicitness scores. All subjects gave written informed consent and were paid for participation. The experiment was approved by the ethics committee of the Medical Faculty at the University of Tübingen and conducted in accordance with the approved guidelines.

Stimuli and task. We used visual stimuli, with the layout consisting of a central fixation cross and six peripheral stimulus locations. Subjects saw sequences of grayscale Gabor gratings tilted either at an angle of 45° or 135° (stimulus size: 5° of visual angle, eccentricity: 7.5° of visual angle [span: 5°–10° of visual angle], spatial frequency: 1.5 cycles per degree) (Fig. 1*b*). We chose an RSI of either 200 ± 50 ms (short RSI) or 2000 ± 500 ms (long RSI), depending on group and session (before vs after sleep/wake retention).

Subjects saw Gabor gratings displayed at six peripheral locations on a computer screen. Their task was to indicate the position of each Gabor grating with an appropriate button press while fixating the cross in the middle of the screen at all times. Specifically, subjects used the left/right outermost keys on the lower row of a standard keyboard (Left-CTRL/Right-CTRL) with their ring fingers to respond to the uppermost left/right stimulus locations; the next keys toward the center (Left-WIN/Task keys) with their middle fingers for stimuli positioned at the horizontal meridian on the left/right; and the next keys toward the center (Alt/Right-WIN) to respond to the lower left/right stimuli with their index fingers (Fig. 1*a*). We encouraged subjects to respond as fast and as accurately as possible. Importantly, stimuli were presented until subjects made a correct button press (i.e., their reaction time depended largely on their error rate). Subjects were not informed about the fact that stimuli were ordered regularly in terms of a repeating 12-item sequence. Each subject was only trained on one single, individual sequence. This sequence fulfilled the following conditions: (1) stimuli occurred equally often at each of the six locations; (2) stimuli at a given location were always followed by at least two stimuli at other locations before the first location was stimulated again (no triplets; e.g., 1-4-2-1, not 1-4-1-2); (3) stimuli did not occur at more than two neighboring locations in a row (no runs; e.g., not 1-2-3 or 6-5-4); and (4) stimuli at a given location did not predict stimuli at any of the other locations more than once within the sequence (no double predictions). Thus, stimuli at a given location always predicted two other locations within the sequence. In this context, another requirement was that (5) if a stimulus at a given location predicted stimuli at two other locations, these other locations never predicted the first one (i.e., one-directional predictions). This goes beyond the no-triplet requirement described above because each location was stimulated twice within the sequence. As a consequence, our location sequence followed a second-order conditional rule (Rosenthal et al., 2016): that is, a stimulus at a given location can fully and only be predicted by a combination of the two preceding stimulus locations. To reinforce learning, stimulus orientation was linked to the position of the stimulus within the sequence, with the constraints that (1) the same orientation must not be linked to the same location more than once; and (2) the same orientation must not occur more than twice in a row. When including these orientation rules, the sequences can be considered first-order deterministic, in that the combination of a stimulus' location and orientation is sufficient to predict the next stimulus. Sequences were selected randomly for each subject, but the same set of sequences was used in all Sleep and Wake subgroups. Thus, from a pool of 228 sequences fulfilling the above conditions (grammar: a-c-d-f-b-e-d-b-a-e-c-f; of which each letter [a-f] can be substituted by a screen location [1–6]; Figure 1*c*), 16 were randomly drawn and then allocated to one subject per subgroup (Sleep vs Wake groups), such that each sequence was reused eight times. Each experimental block consisted of 108 trials (i.e., eight repetitions of the 12-item sequence plus two randomly split pieces of one sequence at the beginning and the end of each block, to minimize the risk of subjects becoming consciously aware of the sequence). During training and test phases, there were 30 s breaks between blocks. In the test phases, irregularities were introduced in terms of stimuli that deviated from the original location sequence (i.e., deviants) but did not differ in any other regard. There was one deviant per sequence

repetition (i.e., eight deviants per block). Standard stimuli immediately following deviants are different from normal standards and, here, referred to as “follower standards” (Kaposvari et al., 2018). Like deviants, these stimuli occurred only once per sequence repetition (i.e., eight follower standards per block). We constructed two possible deviant sequence combinations (A and B) of six deviant sequences each, differing in four sequence positions. These two combinations emerge through four substitution possibilities that cancel each other out (i.e., “b” substituted by “c” and “c” substituted by “b”). Half of the subjects were assigned to deviant sequence combination A, and the other half were assigned to combination B. Given that we had eight deviant sequences per block and three blocks per test, we additionally assigned two of the six possible deviant sequences to each block such that, in sum, each deviant sequence was repeated four times during each test phase. Importantly, the introduction of deviants did not violate any of the above constraints. Stimulus presentation was performed using Presentation (Neurobehavioral Systems; RRID:SCR_002521).

Experimental design and procedure. To test our hypotheses, we used a mixed design, with between-subjects factor Sleep/wake, and within-subjects factors Pre/post (pre-retention vs post-retention test) and Stimulus type (standards vs deviants vs follower standards). Each experimental session consisted of blocks that were presented in either a short temporal context (i.e., with an RSI of 200 ± 50 ms) or in a long temporal context (i.e., with an RSI of 2000 ± 500 ms). In the “Short-to-short” and “Long-to-long” groups, subjects were trained and tested in the same temporal context (i.e., with short and long RSIs, respectively); in the “Long-to-short” and “Short-to-long” groups, subjects were trained and tested in different temporal contexts.

Subjects came to the laboratory at 2100 h (Sleep groups) or 0900 h (Wake groups). Before the start of the experiment, we tested their visual acuity (Freiburg Vision Test, FrACT; RRID:SCR_016439) (Bach, 2007), ocular dominance (Dolman method) (Chaurasia and Mathur, 1976; Cheng et al., 2004), vigilance (Diekelmann et al., 2013), and sleepiness (Stanford Sleepiness Scale) (Hoddes et al., 1972). The training phase started with a short warmup of 20 random trials to familiarize subjects with the experimental task. This was followed by five structured blocks comprising only regular standard stimuli. After a break of 30 min, during which subjects listened to relaxing music, sleepiness was tested again, followed by a first test of three blocks (pre-retention test), containing regular standard stimuli, deviants, and follower standards. After the first experimental session, subjects in the Sleep groups went home to sleep in their familiar environment. This was controlled by actigraphy (Actiwatch 2, Philips Respironics). The following morning, subjects from the Sleep groups completed a questionnaire on sleep quality (SF-A/R). Subjects in the Wake groups left the laboratory to engage in their everyday wake activities. They were instructed not to take any daytime naps during the retention interval, which was also controlled by actigraphy. They were further encouraged to refrain from strenuous activities as far as possible and especially not to engage in any other learning activities. Additionally, their daytime activities were recorded via a questionnaire. Subjects then came back to the laboratory at 0900 h (Sleep groups) or 2100 h (Wake groups). First, they completed a series of initial tests. Specifically, we assessed their chronotype (Morningness-Eveningness Questionnaire Self-Assessment Version) (Horne and Ostberg, 1976), handedness (Edinburgh Handedness Inventory) (Oldfield, 1971), as well as their vigilance and sleepiness. After another short warming-up of 20 random trials, subjects were tested in a post-retention test of three blocks, again including standards, deviants, and follower standards. The post-retention test took place either in the same (Short-to-short and Long-to-long groups) or in the other temporal context (Short-to-long and Long-to-short groups). Afterward, subjects were asked increasingly specific questions regarding their knowledge of any regularity in the task, including a paper-and-pencil-based free recall task where they should write down any sequential regularity they noticed. If they had profound knowledge of the sequential pattern, they were marked as potentially explicit. All subjects were then informed about the presence of a 12-item sequence and subsequently performed three computerized tests in which their knowledge about the sequence was further investigated: first, in a free recall task, subjects were asked to reconstruct the sequence

twice with 12 inputs each by using the same buttons as in the main task (this free recall task was repeated in the end of the session). Second, in a triplet-recognition task, subjects watched sequences of three stimuli that were either part of the sequence (correct) or not (foils) (12 triplets each). They were then asked to indicate for each sequence whether they found it familiar or not. After each response, they were additionally asked to indicate how confident they were of their choice, on a 4-point scale from 1 (very unsure) to 4 (very sure). Finally, in a triplet-completion task, subjects were asked to watch sequences of two stimuli (all of which were correct) and then complete the subsequent stimulus location. All possible triplets had to be completed twice (24 triplets in total). After each input, subjects were again asked to indicate confidence about their response. We chose a fixed sequence for the sequence knowledge tests to minimize the risk of subjects noting any regularity in tests providing relatively more sequence information (e.g., the triplet-completion task, given that it contains only correct sequences of two stimuli) that may have improved subjects' performance on tests providing relatively less sequence information (e.g., the free recall task, given that it provides no sequence information whatsoever). We further note that, given this fixed sequence, we are not able to tease apart potential order effects and thus cannot exclude transfer between the sequence knowledge tests. After debriefing, all subjects performed a last test in the originally trained temporal context. In the end, we tested the subjects' visuo-spatial memory span as well as their supraspan. These tests were adopted from Corsi (1973) and implemented in The Psychology Experiment Building Language (RRID:SCR_014794). During all training and test phases, subjects were given feedback of their mean reaction time per block to enhance motivation. Subjects did not have any prior knowledge about the number of blocks and the number of trials within blocks, to reduce effects of anticipatory self-motivation.

Statistical analysis. Predictive sequence coding was analyzed based on subjects' error rates and reaction times to standards versus deviants versus follower standards. A trial was defined as erroneous when subjects made a wrong button press (i.e., a button press corresponding to a location where the current stimulus was not presented). The error rate was then calculated as the amount of error trials divided by the total number of trials per block. For the control analyses on reaction times, we included all trials, no matter if they were correct or erroneous. Reaction times were measured from stimulus onset until correct button press. For the main analysis, we performed mixed ANOVAs with between-subjects factor Sleep/wake, and within-subjects factors Pre/post (pre-retention vs post-retention test) and Stimulus type (standards vs deviants vs follower standards). For control analyses on individual stimulus types, we conducted separate Sleep/wake $\times$ Pre/post ANOVAs for each stimulus type. Additionally, we calculated a prediction-strength index based on our hypothesis that better consolidation of an internal sequence model should lead to high error rates for model-incongruent deviants, and low error rates for model-congruent follower standards. The prediction-strength index was defined as a simple subtraction of error rates for deviants minus error rates for follower standards. To control for possible effects of sequence repetition on prediction strength, we further performed mixed ANOVAs, including the within-subjects factor Block.

For analysis of prediction-informed errors, we calculated joint probabilities for committing errors in accordance with the trained sequence. This was done separately for pre-retention and post-retention test sessions as well as Sleep and Wake groups, according to the following formula:

$$P(A_i \cap B_i) = P(B_i|A_i)P(A_i)$$

where $P(A_i)$ is the probability of errors following deviants; $P(B_i|A_i)$ is the conditional probability of a response corresponding to the predicted location given that a deviant error has been committed; and i denotes session (1 = pre-retention, 2 = post-retention). To further test whether the number of prediction-informed errors exceeded chance level, we calculated the joint probability as follows:

$$P(A_1 \cap B_c) = P(B_c|A_1)P(A_1)$$

where $P(A_1)$ is the probability of errors following deviants in the pre-retention test (i.e., the baseline error rate following deviants before the

sleep/wake manipulation), and $P(B_c|A_1)$ is the conditional chance probability of making a prediction-informed button press given that a deviant error has been made (i.e., 1 of 5 possible erroneous button presses; $P(B_c|A_1) = 1/5$).

For the paper-and-pencil-based free recall task, we evaluated the rate of correct continuous items. For the computerized free recall task, we calculated both "conservative" and "liberal" measures, referring to the maximum number of correct continuous items from the first button press or from any given button press, respectively. For analysis of sleep effects in the triplet-recognition task, we focused on the recognition index d' for high-confidence trials (i.e., trials for which subjects indicated to remember them with a confidence rating of 3 or 4 on a 4-point scale), as previously established (Wagner et al., 1998; Henson et al., 2000; Macrae et al., 2004; Rimmele et al., 2011, 2012; Sterpenich et al., 2014; Lutz et al., 2017). To test for sleep effects on confidence levels in the triplet-completion task, we performed a mixed ANOVA with between-subjects factor Sleep/wake and within-subjects factor Trial type (correct vs incorrect).

Because there was only one deviant and one follower standard per repetition of the 12-item sequence in the test sessions, we wanted to rule out effects due to differences in trial number sizes and associated differences in variances. We thus reanalyzed our data in R (R Project for Statistical Computing, RRID:SCR_001905) using 1000 iterations of random standard samples that matched the number of deviants and follower standards in each block (92 standards vs eight deviants and eight follower standards per block in the main analysis). Samples were taken from eight bins of standard stimuli comprising 11 or 12 stimuli each (i.e., one standard stimulus was taken per bin), to ensure equal stimulus distribution over the course of each block.

We further collected eye tracker data throughout the measurements (The Eye Tribe Tracker, The Eye Tribe ApS). For the eye tracking analysis, data were separated into single events and then split into eye blinks and fixations. ROIs were defined as the area around the fixation cross (central fixation) as well as the respective stimulus location (stimulus gazing). Percentage of gaze time was then calculated as the number of fixations within each ROI divided by the number of eye tracking data points per event. Finally, these data were averaged for each subject across events and measurements.

To assess whether and how subjects slept during the experimental night (Sleep groups) or day (Wake groups), we collected actigraphy data and used additional questionnaires (for more details, see Experimental design and procedure). Subjects were asked to press a marker on the Actiwatch when switching off the lights in the evening and getting up in the morning. In the subjective questionnaires, subjects should indicate the time they switched off the lights, the time when they fell asleep, how often and how long they were awake during the night, as well as the time when they got up. Time in bed (TiB) was calculated as the time between lights off and getting up; sleep period was calculated as the time between falling asleep and waking up; sleep onset latency (SOL) was calculated as the difference between lights off and the time falling asleep; and wake after sleep onset (WASO) was calculated as the time being awake after falling asleep. Congruence between actigraphic and questionnaire data was calculated using correlations as well as simple ratios for each parameter. For raw data analysis of the actigraphy data, we used a standard algorithm as implemented in Respiroics Actiware 5 (RRID:SCR_016440), with a medium threshold for the estimation of wake phases (40 activity counts/epoch for epoch lengths of 15 s) and a detection algorithm of sleep intervals based on minutes of inactivity (10 min of inactivity was set as threshold for the estimation of start and end phases of a sleep interval).

Two-tailed tests were chosen for all statistical analyses. The level of significance was set to $p = 0.05$. Mixed ANOVAs as implemented in the GLM module of SPSS 24 (IBM, RRID:SCR_002865) were used, in combination with follow-up t tests. Greenhouse–Geisser correction of degrees of freedom was applied when the assumption of sphericity appeared to be violated. We report original degrees of freedom and corrected p values in these cases. Analysis of covariance was used to exclude differences in vigilance as a potential confound. ANCOVA results are reported when the inclusion of vigilance values after retention led to significant improvements of the statistical model. This was only the case

Figure 2. Sleep improves predictive sequence coding. *a*, After sleep, subjects showed high error rates for unpredictable deviants and renormalization of performance for predictable follower standards. In contrast, after wakefulness, error rates were low for deviants and not significantly different from follower standards. *b*, These results are further supported by analysis of a prediction-strength index (deviants — follower standards), which was significantly higher after retention periods containing sleep compared with wakefulness. Importantly, this prediction-strength index was >0 only in the Sleep group, but not in the Wake group after retention. Data are mean $\pm$ SEM. $**p < 0.01$, $*p < 0.05$, $(*)p < 0.1$. $N = 15$ and $N = 16$ in the Sleep group and the Wake group, respectively.

for the post-retention test in the Long-to-short group, for both error rates ($p = 0.042$) and reaction times ($p = 0.038$).

Results

Sleep strengthens predictive sequence coding

Subjects ($n = 31$) were trained on a deterministic 12-item sequence of visual stimuli with short RSIs (200 ± 50 ms). The subjects' task was to react as fast and as accurately as possible by pressing keys corresponding to the locations of the stimuli (Nissen and Bullemer, 1987). The next stimulus only appeared after correct user input (i.e., subjects had to correct their input until they made the right choice). In this context, reaction times for correct(ed) responses heavily depend on error rates. Accordingly, the latter reflect a more sensitive measure of prediction, and will be used as our main variable of interest (for reaction time results, see Control analyses).

All subjects included in the analysis had normal nighttime sleep (Sleep group, $n = 15$) or did not sleep during daytime (Wake group, $n = 16$), as monitored by actigraphy and subjective questionnaires (for a comprehensive analysis, see Control analyses). Whereas training contained only standard stimuli, pre-retention and post-retention tests also contained deviant stimuli that differed from the standard location sequence (Fig. 1*d*). Given that subjects first encountered deviant stimuli in the pre-retention test session, sleep should not only stabilize the associations between standard stimuli, but also immunize performance against interruptions by irregular deviants. To test this idea, we included standard stimuli immediately following deviants (i.e., follower standards) as an additional factor level into our analyses (Kaposvari et al., 2018).

An initial analysis across both groups (Sleep/wake), pre-retention and post-retention tests (Pre/post), and the three stimulus types (standards, deviants, and follower standards) indicated that the different stimulus types induced different numbers of errors (Stimulus type, $F_{(2,58)} = 20.07$, $\eta_p^2 = 0.41$, $p < 0.001$, mixed ANOVA; Fig. 2*a*). Importantly, this effect depended on sleep (vs wakefulness) as indicated by a significant Sleep/wake $\times$ Pre/post $\times$ Stimulus type interaction effect ($F_{(2,58)} = 4.22$, $\eta_p^2 = 0.13$, $p = 0.026$, mixed ANOVA). To unravel this interaction, we cal-

culated separate analyses for the pre-retention and post-retention tests. Error rates during the pre-retention test followed a specific linear pattern that was independent of group assignment (standards $<$ follower standards $<$ deviants, $F_{(1,29)} = 22.21$, $\eta_p^2 = 0.43$, $p < 0.001$, linear contrast; Sleep/wake $\times$ Stimulus type, $F_{(2,58)} = 0.47$, $\eta_p^2 = 0.02$, $p = 0.580$, mixed ANOVA). In contrast, during the post-retention test, this linear pattern across groups was overlaid by group-specific processing of the different stimuli (Sleep/wake $\times$ Stimulus type, $F_{(2,58)} = 6.78$, $\eta_p^2 = 0.19$, $p = 0.004$, mixed ANOVA). Specifically, sleep subjects committed more errors in response to deviants than wake subjects (Sleep/wake, $t_{(29)} = 2.44$, Cohen's $d = 0.87$, $p = 0.021$, independent-samples t test). *Post hoc* tests indicated that, after sleep, the error rates in response to follower standards and regular standards were highly similar (indicating similar processing of these two stimulus types) ($t_{(14)} = 0.48$, Cohen's $d = 0.09$, $p = 0.640$, paired t test), whereas after wakefulness, subjects committed more errors in response to follower standards than standards ($t_{(15)} = 2.189$, Cohen's $d = 0.65$, $p = 0.045$, paired t test). The inverse dissociation was seen in the comparison of deviants and follower standards, with fewer errors in response to follower standards committed after sleep ($t_{(14)} = 3.69$, Cohen's $d = 0.95$, $p = 0.002$, paired t test), but not after wakefulness ($t_{(15)} = 0.45$, Cohen's $d = 0.11$, $p = 0.657$, paired t test).

We calculated a "prediction-strength index" (deviants — follower standards; Fig. 2*b*) to summarize how well the data fit our hypothesis of high error rates for deviants and low error rates for follower standards. As suggested by the preceding analyses, this prediction-strength index evolved differently in the two groups (Sleep/wake $\times$ Pre/post interaction, $F_{(1,29)} = 6.08$, $\eta_p^2 = 0.17$, $p = 0.020$, mixed ANOVA). *Post hoc* tests showed that the prediction-strength index was significantly enhanced if the retention period contained sleep compared with wakefulness ($t_{(29)} = 3.27$, Cohen's $d = 1.16$, $p = 0.003$, independent-samples t test), whereas there was no such difference in the pre-retention test ($t_{(29)} = 0.12$, Cohen's $d = 0.04$, $p = 0.909$, independent-samples t test). Additionally, the prediction-strength index was significantly different from zero after sleep ($t_{(14)} = 3.69$, Cohen's $d = 0.95$, $p =$

Figure 3. Higher prediction-informed error rate after sleep. After sleep, subjects displayed significantly higher rates of prediction-informed errors than after wakefulness (i.e., they were more likely to press the button corresponding to the predicted stimulus on a deviant trial). This difference was further reflected in performance compared with chance level (i.e., only the Sleep group performed above chance in the post-retention test) and could not be seen in the pre-retention test. Data are mean ± SEM. ** $p < 0.01$, * $p < 0.05$, (*) $p < 0.1$. $N = 15$ and $N = 16$ in the Sleep group and the Wake group, respectively.

0.002, one-sample t test), but not after wakefulness ($t_{(15)} = 0.45$, Cohen's $d = 0.11$, $p = 0.657$, one-sample t test; these tests are equivalent to the comparison of deviants vs follower standards above), whereas it was not significant before the retention interval in either group (Sleep: $t_{(14)} = 1.45$, Cohen's $d = 0.37$, $p = 0.169$; Wake: $t_{(15)} = 2.00$, Cohen's $d = 0.50$, $p = 0.064$; one-sample t tests).

If sequence knowledge is used to actively predict upcoming stimuli, errors following deviants should not be random but reflect a bias toward the standard stimuli, which would have occurred in the trained sequence (Fiser et al., 2010). To analyze such prediction-informed errors (i.e., responses to deviant stimuli with button presses corresponding to the predicted stimuli), we calculated joint probabilities for both groups before and after the retention interval (Fig. 3; for a detailed description of analysis, see Materials and Methods). We found that the number of prediction-informed errors depended on both consolidation and group assignment (Sleep/wake $\times$ Pre/post interaction ($F_{(1,29)} = 4.54$, $\eta_p^2 = 0.14$, $p = 0.042$, mixed ANOVA). There were no main effects of Sleep/wake ($F_{(1,29)} = 0.93$, $\eta_p^2 = 0.03$, $p = 0.342$, mixed ANOVA) or Pre/post ($F_{(1,29)} = 0.77$, $\eta_p^2 = 0.03$, $p = 0.387$, mixed ANOVA). *Post hoc* tests indicated a higher number of prediction-informed errors after sleep compared with wakefulness in the post-retention test ($t_{(29)} = 2.13$, Cohen's $d = 0.76$, $p = 0.042$, independent-samples t test), with no such difference in the pre-retention test ($t_{(29)} = 0.34$, Cohen's $d = 0.12$, $p = 0.736$, independent-samples t test). Additionally, while prediction-informed errors showed at least a trend for being above chance before the retention interval (Sleep, $t_{(14)} = 1.96$, Cohen's $d = 0.44$, $p = 0.071$; Wake, $t_{(15)} = 3.14$, Cohen's $d = 0.85$, $p = 0.007$, paired t tests comparing prediction-informed error rates with individual baseline error rates, as described in Statistical analysis), this effect was conserved only in the Sleep group at the post-retention test ($t_{(14)} = 2.40$, Cohen's $d = 0.75$, $p = 0.031$, paired t test), whereas it disappeared in the Wake group ($t_{(15)} = 0.35$, Cohen's $d = 0.12$, $p = 0.731$, paired t test). Together, these findings show that deterministic associations between stimuli were encoded equally well in both groups. However, sleep-associated consolidation processes are required to transform these associations into lasting memory traces, which are then used to predict future stimuli and prepare appropriate motor responses.

We did not observe significant differences between Sleep and Wake groups on either error rates for standard stimuli during initial training (Sleep/wake main effect, $F_{(1,29)} = 1.03$, $\eta_p^2 = 0.03$, $p = 0.318$, or Sleep/wake $\times$ Block interaction, $F_{(4,116)} = 1.00$, $\eta_p^2 = 0.03$, $p = 0.397$, mixed ANOVA) or error rates for the three stimulus types during the pre-retention test (see analyses above). Moreover, there were no differences between Sleep and Wake groups on any of the control tests during both pre-retention and post-retention sessions (see Control analyses). Thus, it is unlikely that our results can be explained by circadian effects during either training or test sessions.

Sleep improves predictive sequence coding across temporal contexts

We investigated whether sleep contributes to the emergence of an abstracted internal task model, which can be used in modified contexts (Witt et al., 2010). For this, we manipulated the interval between stimuli (i.e., the RSI between a response and the presentation of the next stimulus) between pre-retention and post-retention sessions. In the standard group (reported above), RSIs were short (200 ± 50 ms) both during training (including the pre-retention test) and post-retention test ("Short-to-short" group). In the "Long-to-long" group ($n_{\text{sleep}} = 15$, $n_{\text{wake}} = 14$), RSIs were increased by a factor of 10 (2000 ± 500 ms) during training and post-retention test. Correspondingly, in two additional groups ("Long-to-short," $n_{\text{sleep}} = 14$, $n_{\text{wake}} = 14$, and "Short-to-long," $n_{\text{sleep}} = 16$, $n_{\text{wake}} = 13$), RSIs differed at training and post-retention test.

We found that, with long RSIs, the prediction-strength index was distinctly lower than with short RSIs ($F_{(1,232)} = 8.87$, $\eta_p^2 = 0.04$, $p = 0.003$, for an ANOVA across all groups and retention tests; for reaction-time results, see Control Analyses). Indeed, the prediction-strength index was not significantly different from zero when using long RSIs. This was true both in the pre-retention test ($t_{(56)} = 1.45$, Cohen's $d = 0.19$, $p = 0.151$, one-sample t test) and in the post-retention test ($t_{(57)} = 0.27$, Cohen's $d = 0.04$, $p = 0.786$, one-sample t test). On the other hand, the prediction-strength index was positive for short RSIs in both sessions (pre-retention, $t_{(59)} = 2.69$, Cohen's $d = 0.35$, $p = 0.009$; post-retention, $t_{(58)} = 3.63$, Cohen's $d = 0.47$, $p = 0.001$, one-sample t tests). To examine whether sleep fosters the formation of an abstracted internal model, we thus restricted our analyses to the prediction-strength index in the Long-to-short group. Testing the difference between Sleep and Wake groups at the post-retention test indicated higher prediction-strength index values after sleep compared with wakefulness ($F_{(1,25)} = 5.32$, $\eta_p^2 = 0.18$, $p = 0.030$, ANCOVA; Fig. 4). The prediction-strength index was again significantly different from zero after sleep ($t_{(13)} = 3.14$, Cohen's $d = 0.84$, $p = 0.008$, one-sample t test), but not after wakefulness ($t_{(13)} = 0.24$, Cohen's $d = 0.06$, $p = 0.812$, one-sample t test). Finally, we also found prediction-informed errors to be significantly above chance only after sleep ($t_{(13)} = 2.32$, Cohen's $d = 0.62$, $p = 0.037$, paired t test), but not after wakefulness ($t_{(13)} = 1.52$, Cohen's $d = 0.41$, $p = 0.153$, paired t test). However, in the Long-to-short group, we failed to find a significant difference for prediction-informed errors after sleep compared with wakefulness ($t_{(26)} = 1.15$, Cohen's $d = 0.44$, $p = 0.260$, independent-samples t test).

Comparisons across Short-to-short and Long-to-short groups additionally indicated that prediction strength in both groups was significantly enhanced after the retention interval for subjects who slept (main effect of Sleep/wake, $F_{(1,55)} = 11.54$,

Figure 4. Sleep improves predictive sequence coding across temporal contexts. **a**, Even though subjects were trained with long RSIs (2000 ± 500 ms), when tested with short RSIs (200 ± 50 ms), they showed high error rates for unpredictable deviants and renormalization of performance for predictable follower standards. In contrast, after wakefulness, error rates were low for deviants and not significantly different from follower standards. **b**, The prediction-strength index (deviants — follower standards) was significantly higher after retention periods containing sleep compared with wakefulness when subjects were tested with short RSIs. Importantly, the prediction-strength index was different from zero only in the Sleep group, after retention. Data are mean $\pm$ SEM. $**p < 0.01$, $*p < 0.05$, $(*)p < 0.1$. $N = 14$ in both Sleep and Wake groups.

Figure 5. Higher confidence in triplet recognition after sleep. **a**, In a triplet-recognition task, we found a significantly higher recognition performance (in terms of d') for high-confidence trials (i.e., trials for which subjects indicated to remember them with a confidence rating of 3 or 4 on a 4-point scale) after sleep compared with wakefulness. These results are in line with those of a triplet-completion task (**b**), where subjects gave significantly higher confidence ratings for correct trials than for incorrect trials after they slept, whereas there was no difference between the two trial types after an equivalent period of wakefulness. Data are mean $\pm$ SEM. $**p < 0.01$, $*p < 0.05$. $N = 15$ and $N = 16$ in the Sleep group and the Wake group, respectively.

$\eta_p^2 = 0.17$, $p = 0.001$, ANOVA), and this effect was independent of training RSI (RSI $\times$ Sleep/wake, $F_{(1,55)} = 1.58$, $\eta_p^2 = 0.03$, $p = 0.214$, ANOVA). Moreover, the prediction-strength index was again significantly different from zero after sleep ($t_{(28)} = 4.11$, Cohen's $d = 0.76$, $p < 0.001$, one-sample t test), but not after wakefulness ($t_{(29)} = 0.62$, Cohen's $d = 0.11$, $p = 0.541$, one-sample t test). When testing across both groups, prediction-informed errors were significantly higher after sleep compared with wakefulness ($F_{(1,55)} = 5.26$, $\eta_p^2 = 0.09$, $p = 0.026$, ANOVA), and this effect was again independent of training RSI (RSI $\times$ Sleep/wake, $F_{(1,55)} = 0.36$, $\eta_p^2 = 0.01$, $p = 0.549$, ANOVA). Additionally, prediction-informed errors were again above chance only after sleep ($t_{(28)} = 3.40$, Cohen's $d = 0.63$, $p = 0.002$, paired t test), but not after wakefulness ($t_{(29)} = 1.12$, Cohen's $d = 0.20$, $p = 0.272$, paired t test).

Higher confidence in triplet recognition after sleep

Importantly, none of the subjects included in our analyses reported full explicit knowledge of the trained sequence. To probe for explicit knowledge of the sequence model after retention, we had subjects perform a paper-and-pencil-based free recall task, a computerized free recall task, where they should reconstruct the 12-item sequence; a triplet-recognition task, where they were asked to indicate for triplets of stimuli (either correct or foil) whether they were part of the sequence or not, and a triplet-completion task, where they were asked to complete stimulus triplets by indicating the position of the last stimulus given information about the two previous ones. In the latter two tasks, subjects additionally rated their confidence after each trial, on a 4-point scale from 1 (very unsure) to 4 (very sure). Given that we had subjects perform the computerized free recall task twice (once after the post-retention test and once in the very end of the post-retention session) and analyzed this task in a liberal as well as a conservative way (see Materials and Methods), we had seven measures of sequence knowledge per group. There were no significant differences between Sleep and Wake subgroups on any of these measures ($p \geq 0.088$, for 28 independent-samples t tests across all groups). Interestingly, however, on the triplet-recognition task, we found a significantly higher recognition index d' , specifically for high-confidence trials (confidence rating ≥ 3) after sleep compared with wakefulness in the Short-to-short group ($t_{(29)} = 2.09$, Cohen's $d = 0.38$, $p = 0.045$, independent-samples t test; Fig. 5a). Furthermore, on the triplet-completion task, the same subjects showed higher confidence ratings for correct trials compared with incorrect trials after sleep ($t_{(14)} = 3.12$, Cohen's $d = 0.81$,

$p = 0.008$, paired t test), but not after wakefulness ($t_{(15)} = 0.47$, Cohen's $d = 0.12$, $p = 0.644$, paired t test; Sleep/wake $\times$ Trial type interaction, $F_{(1,29)} = 5.41$, $\eta_p^2 = 0.16$, $p = 0.027$, mixed ANOVA; Figure 5b). Assuming that confidence ratings are a more sensitive measure to detect conscious access to implicit sequence knowledge (Rosenthal et al., 2016), these findings suggest that sleep promotes metacognitive access to implicitly learned regularities, without leading to full conscious awareness.

Control analyses

Control for different numbers of trials

As there was only one deviant and one follower standard per repetition of the 12-item sequence in the test sessions, the relationship between standards, deviants, and follower standards was 10:1:1. To rule out effects due to the difference in trial numbers, we reanalyzed the main results (see 'Sleep strengthens predictive sequence coding') using 1000 iterations of random samples of standard stimuli that matched the number of deviants and follower standards. The pattern of results for these control analyses is virtually identical to those reported in the main text: we found significantly different stimulus-type patterns after sleep compared with wakefulness in the post-retention test (95% CI: $p = [0.004; 0.004]$, mixed ANOVA), but not in the pre-retention test ($p = [0.659; 0.687]$, mixed ANOVA; Sleep/wake $\times$ Pre/post $\times$ Stimulus type interaction, $p = [0.035; 0.037]$, mixed ANOVA). Specifically, error rates for follower standards did not differ from regular standards after subjects had slept ($p = [0.524; 0.558]$, paired t test) compared with the Wake group ($p = [0.027; 0.033]$, paired t test). Results for comparisons between deviants and follower standards, which constitute the prediction-strength index, remain unchanged, as random samples were only taken from the pool of standard stimuli.

Individual analysis of stimulus types

Individual analyses for each stimulus type across sessions in the Short-to-short group further showed a trend for group-specific processing of deviants (Sleep/wake $\times$ Pre/post, $F_{(1,29)} = 3.31$, $\eta_p^2 = 0.10$, $p = 0.079$, mixed ANOVA), which was driven by a numerical increase in error rates for deviants over sleep, and a numerical decrease over wakefulness (both changes not significant, $p \geq 0.146$). This interaction was not significant for follower standards ($F_{(1,29)} = 1.68$, $\eta_p^2 = 0.06$, $p = 0.206$, mixed ANOVA). However, we found a significant group-specific change for the standard stimuli ($F_{(1,29)} = 6.86$, $\eta_p^2 = 0.19$, $p = 0.014$), which was driven by a decrease in errors committed after wakefulness ($F_{(1,15)} = 6.34$, $\eta_p^2 = 0.30$, $p = 0.024$, mixed ANOVA), and no change over sleep ($F_{(1,14)} = 1.42$, $\eta_p^2 = 0.09$, $p = 0.254$, mixed ANOVA). These findings are in accordance with the notion that sleep benefits sequence-specific learning, whereas an equivalent period of daytime wakefulness may lead to an improvement of general skill knowledge (i.e., an improved visuo-motor and motor-motor coordination) (Cohen et al., 2005; Pace-Schott and Spencer, 2013).

Analysis across sessions and blocks

Because subjects saw each deviant sequence four times during each test session, we tested for possible effects of sequence repetition on prediction strength in the Short-to-short group. A $2 \times 2 \times 3$ mixed ANOVA with between-subjects factor Sleep/wake and within-subjects factors Pre/post and Block revealed no evidence that the prediction-strength index evolves differently across Pre/post sessions (Pre/post, $F_{(1,29)} = 0.38$, $\eta_p^2 = 0.01$, $p = 0.578$; Pre/post $\times$ Block, $F_{(2,58)} = 2.03$, $\eta_p^2 = 0.07$, $p = 0.141$;

Block, $F_{(2,58)} = 0.03$, $\eta_p^2 = 0.04$, $p = 0.334$; linear and quadratic contrasts for Block, $p \geq 0.172$; mixed ANOVA) or within sessions (pre-retention: Block, $F_{(2,58)} = 1.27$, $\eta_p^2 = 0.04$, $p = 0.288$; linear and quadratic contrasts, $p \geq 0.210$; post-retention: Block, $F_{(2,58)} = 2.02$, $\eta_p^2 = 0.07$, $p = 0.142$; linear and quadratic contrasts, $p \geq 0.123$; mixed ANOVAs). This suggests that there was no effect of sequence repetition on prediction strength.

Reaction times

Subjects in the Short-to-short group were successfully trained on the sequence, as also seen by a significant decrease in reaction times during training (main effect of Block, $F_{(4,116)} = 69.11$, $\eta_p^2 = 0.70$, $p < 0.001$; linear contrast, $F_{(1,29)} = 116.68$, $\eta_p^2 = 0.85$, $p < 0.001$; quadratic contrast, $F_{(1,29)} = 6.53$, $\eta_p^2 = 0.18$, $p = 0.016$; mixed ANOVA), and different reaction times for the three stimulus types ($F_{(2,58)} = 44.97$, $\eta_p^2 = 0.61$, $p < 0.001$, mixed ANOVA) (Fig. 6a). In contrast to the error rates, for reaction times we did not find a significant Sleep/wake $\times$ Pre/post $\times$ Stimulus type interaction ($F_{(2,58)} = 1.02$, $\eta_p^2 = 0.03$, $p = 0.365$, mixed ANOVA). Hypothesis-driven *post hoc* analyses, however, revealed significantly different reaction time patterns for the stimulus types between the Sleep and the Wake groups in the post-retention test (Sleep/wake $\times$ Stimulus type interaction, $F_{(2,58)} = 3.81$, $\eta_p^2 = 0.12$, $p = 0.036$, mixed ANOVA), but not in the pre-retention test ($F_{(2,58)} = 0.52$, $\eta_p^2 = 0.02$, $p = 0.598$, mixed ANOVA), which replicates our findings on error rates in the Short-to-short group (compare Sleep strengthens predictive sequence coding). Specifically, we found significantly longer reaction times for deviants compared with follower standards after sleep ($t_{(14)} = 3.41$, Cohen's $d = 0.88$, $p = 0.004$, paired t test), but not after wakefulness ($t_{(15)} = 0.60$, Cohen's $d = 0.15$, $p = 0.559$, paired t test). This was further reflected by a significantly higher reaction time-based prediction-strength index after sleep compared with wakefulness in the post-retention test ($t_{(29)} = 2.30$, Cohen's $d = 0.42$, $p = 0.029$, independent-samples t test), but not in the pre-retention test ($t_{(29)} = 0.88$, Cohen's $d = 0.16$, $p = 0.385$, independent-samples t test) (Fig. 6b). Additionally, this reaction time-based prediction-strength index was significantly different from zero after sleep ($t_{(14)} = 3.41$, Cohen's $d = 0.88$, $p = 0.004$, one-sample t test), but not after wakefulness ($t_{(14)} = 0.60$, Cohen's $d = 0.15$, $p = 0.559$, one-sample t test), whereas this difference was not significant before the retention interval (Sleep: $t_{(14)} = 2.00$, Cohen's $d = 0.52$, $p = 0.065$; Wake: $t_{(14)} = 0.95$, Cohen's $d = 0.25$, $p = 0.358$, one-sample t tests). However, the respective Sleep/wake $\times$ Pre/post interaction term failed to reach significance ($F_{(1,29)} = 1.66$, $\eta_p^2 = 0.05$, $p = 0.208$, mixed ANOVA).

In addition to the results on short (200 ± 50 ms) versus long RSIs (2000 ± 500 ms) for the error-based prediction-strength index (see 'Sleep improves predictive sequence coding across temporal contexts'), we further found the reaction time-based prediction-strength index to be distinctly lower with long RSIs compared with short RSIs ($F_{(1,232)} = 14.54$, $\eta_p^2 = 0.06$, $p < 0.001$, for an ANOVA across all groups and retention tests). The prediction-strength index was again not significantly different from zero when using long RSIs, either in the pre-retention test ($t_{(56)} = 0.32$, Cohen's $d = 0.04$, $p = 0.747$, one-sample t test) or in the post-retention test ($t_{(57)} = 0.08$, Cohen's $d = 0.01$, $p = 0.939$, one-sample t test), whereas it was positive for short RSIs (pre-retention, $t_{(59)} = 2.84$, Cohen's $d = 0.37$, $p = 0.006$; post-retention, $t_{(58)} = 4.54$, Cohen's $d = 0.59$, $p < 0.001$; one-sample t tests).

Figure 6. Sleep improves predictive sequence coding: analysis of reaction times. *a*, Subjects were successfully trained on the sequence, as seen by a significant decrease in reaction times during training and different reaction times for the three stimulus types. After sleep, subjects showed longer reaction times for deviants compared with follower standards. In contrast, after wakefulness, reaction times for deviants were not significantly different from follower standards. *b*, These results are further supported by analysis of a reaction time-based prediction-strength index (deviants – follower standards), which was significantly higher after retention periods containing sleep compared with wakefulness. Data are mean $\pm$ SEM. Importantly, this prediction-strength index was >0 only in the Sleep group, but not in the Wake group after retention. Data are mean $\pm$ SEM. $^{**}p < 0.01$, $^{*}p < 0.05$, $^{(*)}p < 0.1$. $N = 15$ and $N = 16$ in the Sleep group and the Wake group, respectively.

Speed-accuracy tradeoff

Interestingly, subjects responded more slowly to stimuli in long RSI conditions compared with short RSI conditions ($F_{(1,232)} = 6.56$, $\eta_p^2 = 0.03$, $p = 0.011$, for an ANOVA across all groups and retention tests), while at the same time, they committed fewer errors in long versus short RSI conditions ($F_{(1,232)} = 4.62$, $\eta_p^2 = 0.02$, $p = 0.033$, for an ANOVA across all groups and retention tests). Hence, our results support the notion of a speed-accuracy tradeoff. Slower responses to stimuli in the long RSI conditions, together with lower error rates, might be the result of longer but less precisely timed preparation for responses to upcoming stimuli. Such effects are likely to play a role in stimulus type-specific processing, as discussed by Willingham et al. (1997).

Sleep parameters

Sleep data were collected via actigraphy and subjective sleep questionnaires. Specifically, TiB, sleep period, SOL and WASO were collected. Actigraphic data have previously been demonstrated to show high sensitivity and high accuracy, but low specificity when compared with EEG-based polysomnography (i.e., they are good at scoring sleep as sleep but bad at scoring wake as wake) (Marino et al., 2013). Here, congruence between actigraphic and subjective sleep data was very high (mean $\pm$ SEM: $94 \pm 3\%$ for TiB, $94 \pm 3\%$ for sleep period, and $82 \pm 7\%$ for sleep onset latency), as also shown by significant correlations between subjective and objective measures for TiB ($r = 0.979$, $p < 0.001$) and sleep period ($r = 0.969$, $p < 0.001$). We also found a trend for a correlation between subjective and objective SOL ($r = 0.312$, $p = 0.082$). Congruence between actigraphic and subjective data was very low for WASO (mean $\pm$ SEM: $27 \pm 3\%$), reflecting the low specificity of actigraphic sleep recordings. However, we still found a significant correlation between these measures ($r = 0.469$, $p < 0.001$).

The Sleep groups showed normal sleep patterns with a TiB of (mean $\pm$ SEM) 439.55 ± 10.35 min, a sleep period of 397.46 ± 10.09 min, and a sleep onset latency of 14.96 ± 2.26 min. Sleep subjects reported rare wake phases during the night (mean $\pm$ SEM: 7.13 ± 1.51 min), which was also confirmed by the actigraphy data (WASO = 27.06 ± 2.10 min).

Careful analysis of daytime activity questionnaires and actigraphy data obtained in the Wake groups was performed to rule out that any of the subjects slept during the retention interval. Importantly, none of the Wake subjects reported daytime naps. This was largely confirmed by the actigraphy data, with some reported inactivity phases (e.g., watching TV) estimated as sleep in 4 subjects (mean $\pm$ SEM for sleep period: 6.82 ± 4.23 min).

Eye tracking

To ensure that our results are not influenced by oculomotor learning (Albouy et al., 2006), subjects were asked to fixate the center of the screen at all times during the measurements (Fig. 1a). This was controlled with eye tracking. Central fixation during stimulus presentation was high with $77 \pm 2\%$ at 5° of visual angle around the fixation cross and only $6 \pm 1\%$ direct stimulus fixation.

Control tests

Before training as well as before the post-retention test, all subjects were tested on a vigilance task (Diekelmann et al., 2013) and a subjective sleepiness questionnaire (Stanford Sleepiness Scale) (Hoddes et al., 1972). They further performed tests of their visuo-spatial memory span and supra span (Corsi, 1973), as well as their chronotype (Horne and Ostberg, 1976). In the Short-to-short group, there were no significant differences between Sleep and Wake subgroups for any of the control variables before retention (all $p > 0.2$, independent-samples t test) or after retention (all $p > 0.3$, independent-samples t tests) (Table 1). In the Long-to-short group, the Sleep subgroup had a greater block span ($t_{(26)} = 3.73$, Cohen's $d = 0.72$, $p = 0.001$, independent-samples t test) and supra span ($t_{(26)} = 2.21$, Cohen's $d = 0.43$, $p = 0.036$, independent-samples t test), compared with the Wake subgroup. In the Short-to-long group, in the post-retention session, subjects who slept reported to be less sleepy than subjects who stayed awake ($t_{(27)} = 2.16$, Cohen's $d = 0.41$, $p = 0.040$, independent-samples t test), which was confirmed by objective vigilance values ($t_{(27)} = 2.41$, Cohen's $d = 0.46$, $p = 0.023$, independent-samples t test).

Table 1. Control tests^a

	Short-to-short group		Long-to-short group		Short-to-long group		Long-to-long group	
	Sleep group	Wake group	Sleep group	Wake group	Sleep group	Wake group	Sleep group	Wake group
Pre-retention session								
Vigilance	458.73 ± 14.56	471.00 ± 9.81	452.14 ± 9.07	445.07 ± 9.23	431.63 ± 12.23	459.69 ± 12.85	442.90 ± 13.01	446.93 ± 10.48
SSS	2.82 ± 0.20	2.52 ± 0.13	2.93 ± 0.23	2.54 ± 0.13	2.67 ± 0.19	2.35 ± 0.20	3.27 ± 0.26	2.43 ± 0.15*
Post-retention session								
Vigilance	444.33 ± 13.84	460.38 ± 10.64	439.14 ± 5.55	423.43 ± 5.69 ^(*)	425.47 ± 11.15	461.15 ± 12.18*	433.30 ± 8.39	437.64 ± 10.48
SSS	2.63 ± 0.21	2.78 ± 0.28	2.64 ± 0.20	2.86 ± 0.23	2.28 ± 0.19	2.96 ± 0.26*	2.73 ± 0.21	2.79 ± 0.35
Chronotype	3.00 ± 0.14	2.88 ± 0.15	2.93 ± 0.13	2.86 ± 0.10	3.13 ± 0.20	3.00 ± 0.30	2.93 ± 0.12	2.93 ± 0.07
Corsi (block span)	6.47 ± 0.32	6.56 ± 0.30	7.21 ± 0.30	5.86 ± 0.21**	6.73 ± 0.32	6.58 ± 0.40	6.00 ± 0.37	6.69 ± 0.36
Corsi (supra span)	3.73 ± 0.71	4.81 ± 0.91	5.21 ± 0.73	3.14 ± 0.59*	4.19 ± 0.90	5.15 ± 0.96	3.80 ± 0.76	3.64 ± 0.80

^aData are mean ± SEM. Vigilance is measured as reaction time (ms). Chronotype is measured based on the MEQ and ranges from 1 (definitely evening type) to 5 (definitely morning type). The Corsi block span is a measure for the number of visual items that can be remembered in sequence; the supra span test examines the number of repetitions needed until a sequence of visual items is learned implicitly. SSS, Stanford Sleepiness Scale (7-point scale on sleepiness). *N* = 31 for the Short-to-short group, *N* = 28 for the Long-to-short group, *N* = 29 for the Short-to-long group, and *N* = 29 for the Long-to-long group.

***p* < 0.01; **p* < 0.05; (*)*p* < 0.1; significant differences between the Sleep group and the Wake group.

t test). Finally, in the Long-to-long group, subjects in the Sleep subgroup were more sleepy than subjects who stayed awake before the retention interval ($t_{(27)} = 2.74$, Cohen's $d = 0.52$, $p = 0.011$, independent-samples *t* test). However, vigilance performance did not differ ($t_{(27)} = 0.19$, Cohen's $d = 0.04$, $p = 0.850$, independent-samples *t* test). Importantly, we did not find a significant difference between Sleep and Wake subgroups regarding their chronotype in any group (all $p > 0.5$, independent-samples *t* tests), excluding substantial chronological effects on our results.

Discussion

Our results demonstrate that sleep improves predictive sequence coding. While sleep-dependent consolidation of sequence knowledge increased the number of errors in response to deviants, it improved performance for stimuli immediately following deviants. Thus, sleep after acquisition of sequence knowledge helps put performance back on track after unpredictable disturbances. Moreover, when training and testing occurred in different temporal contexts, sleep supported the transfer of sequence knowledge to the new context.

The present study was inspired by theories of hierarchical predictive coding. In this framework, our findings imply that, during training, the strong statistical regularities in our stimulus sequences are extracted from bottom-up input and lead to the formation of an internal model of the trained sequence (Gregory, 1980; Friston, 2012). The following process of memory consolidation during sleep strengthens this model, rendering it more powerful for predicting how the sequence unfolds when it is later reencountered. Subjects who slept after training are therefore able to provide better predictions about the sequence based on top-down memory templates derived from the sequence model. This leads to characteristic changes in behavioral responses to different stimulus types: a boost of prediction errors for unpredictable, model-incongruent deviants, in combination with quick recovery for predictable, model-congruent follower standards (Mumford, 1992).

While an increase in errors to deviant stimuli (i.e., a decrease in performance) after sleep may appear counterintuitive, it indicates that sleep consolidated sequence knowledge, with decremental consequences on post-retention performance if the number of errors is considered in isolation. A more fine-grained analysis based on different error types supports this interpretation: following sleep, subjects were more prone to respond to a deviant stimulus with a button press corresponding to the stimulus that should have occurred. This increase in prediction-informed errors corroborates that sleep enhanced implicit

knowledge of the trained sequence. At the same time, performance improved for a different set of stimuli: following irregular events, subjects who slept were quicker to catch up with the trained sequence, as shown by decreased error rates for follower standards. Together, this pattern indicates that sleep improved knowledge of the sequence model as a whole, rather than simply reinforcing the links between individual items within the sequence. Because deviants were introduced before the retention interval, our findings can be explained by a sleep-induced enhancement of the knowledge that deviants occur from time to time, but that the underlying sequence is never disrupted for more than one item.

Recent research indicates that sleep does not simply reinforce memories but rather supports their abstraction from the concrete learning context (Payne et al., 2009; Diekmann et al., 2010; McKeon et al., 2012; Lutz et al., 2017; Pardilla-Delgado and Payne, 2017). Witt et al. (2010) demonstrated a beneficial effect of sleep in generalizing extrinsic sequence information from one hand to the other in a finger tapping task, suggesting sleep-dependent consolidation of effector-independent representations. To test whether sleep also supports temporal abstraction of sequence knowledge, we manipulated the temporal distance between sequence items. We found that, with long intervals between stimuli (RSIs), neither error rates nor reaction times could be used to assess the presence of a predictive sequence model. This agrees with previous findings (Willingham et al., 1997) and probably reflects masking of predictive behavior when subjects are given enough time to prepare their responses during the RSI. Nevertheless, when subjects were trained with long intervals between stimuli, they showed enhanced prediction strength when tested with short intervals after retention. Crucially, this temporal abstraction was sleep-dependent: we found significant differences in prediction strength after sleep compared with wakefulness, and prediction strength was different from chance only after sleep. This suggests that a functionally equivalent model emerges when sequence knowledge is acquired in a different temporal context. This aspect is particularly important given the variability of stimulus parameters in natural environments (Fiser et al., 2010).

In the context of sequence learning, our study helps to clarify the conditions under which sleep benefits may occur. Previous findings have remained largely inconclusive concerning the role of sleep, especially in implicit sequence learning (Robertson et al., 2004; Fischer et al., 2006, 2007; Spencer et al., 2006; Song et al., 2007a; Nemeth et al., 2010; Song and Cohen, 2014), possibly because tasks and analyses were not focused on the predictive

aspects of sequence learning. We used a modified version of the serial reaction time task, and find that sleep does affect sequence learning. Our use of strongly predictable deterministic sequences may have added to the emergence of robust sleep effects. Training on deterministic perceptual sequences likely results in stronger hippocampal involvement than training on probabilistic sequences (Curran, 1997; Poldrack et al., 2001; Schendan et al., 2003; Song et al., 2007a; Turk-Browne et al., 2010) or purely motor sequences (Rose et al., 2011). While sleep benefits have been demonstrated also for procedural tasks with little hippocampal engagement, hippocampal involvement might increase this benefit. A recent review by King et al. (2017) suggests that, even for procedural tasks, sleep benefits may depend on hippocampal engagement. Hippocampal processing may further have been fostered by the use of a spatially demanding 2D stimulus layout, in contrast to the horizontal layout used in most SRT studies (Nissen and Bullemer, 1987; Giesbrecht et al., 2013). We also note that interleaving frequent standard stimuli and rare deviants may have improved sensitivity for detecting differences between predictable and unpredictable events compared with block-based designs (Willingham et al., 1997).

Finally, in a series of debriefing tests assessing explicit sequence knowledge, we observed higher confidence levels for correct choices after sleep compared with wakefulness. This may indicate that strengthening of newly formed sequence models during sleep promotes metacognitive access to learned stimulus associations (Rosenthal et al., 2016), in agreement with findings of sleep-dependent conversion of implicitly learned regularities to explicit knowledge (Fischer et al., 2006; Yordanova et al., 2017). Although sequence learning can operate outside conscious awareness (Song et al., 2007b), clear-cut sleep effects have only been demonstrated following explicit sequence learning (Robertson et al., 2004; Spencer et al., 2006; Song et al., 2007a; Janacek and Nemeth, 2012). A presumed increase of hippocampal involvement in our protocol, as discussed above, may have prompted improved metacognitive access after sleep (Marshall and Born, 2007).

In conclusion, we show that sleep improves predictive coding of implicitly learned visual sequences. An increase in prediction-informed errors and quicker recovery from deviants suggest that sleep supports the emergence of a more holistic internal sequence model. This in turn provides better prediction of upcoming stimuli, and improved transfer of sequence knowledge across temporal contexts. Sleep also increased confidence for sequence knowledge compared with wakefulness, presumably reflecting enhanced metacognitive access to the processes supporting decision-making during explicitness tests, rather than explicit sequence knowledge per se.

References

- Albouy G, Ruby P, Phillips C, Luxen A, Peigneux P, Maquet P (2006) Implicit oculomotor sequence learning in humans: time course of offline processing. *Brain Res* 1090:163–171. [CrossRef Medline](#)
- Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L (2010) Stimulus predictability reduces responses in primary visual cortex. *J Neurosci* 30:2960–2966. [CrossRef Medline](#)
- Bach M (2007) The Freiburg Visual Acuity Test: variability unchanged by post-hoc re-analysis. *Graefes Arch Clin Exp Ophthalmol* 245:965–971. [CrossRef Medline](#)
- Chaurasia BD, Mathur BB (1976) Eyedness. *Acta Anat (Basel)* 96:301–305. [CrossRef Medline](#)
- Cheng CY, Yen MY, Lin HY, Hsia WW, Hsu WM (2004) Association of ocular dominance and anisometropic myopia. *Invest Ophthalmol Vis Sci* 45:2856–2860. [CrossRef Medline](#)
- Cohen DA, Pascual-Leone A, Press DZ, Robertson EM (2005) Off-line learning of motor skill memory: a double dissociation of goal and movement. *Proc Natl Acad Sci U S A* 102:18237–18241. [CrossRef Medline](#)
- Corsi PM (1973) Human memory and the medial temporal region of the brain. *Diss Abstr Int* 34:891.
- Curran T (1997) Higher-order associative learning in amnesia: evidence from the serial reaction time task. *J Cogn Neurosci* 9:522–533. [CrossRef Medline](#)
- Diekelmann S, Born J, Wagner U (2010) Sleep enhances false memories depending on general memory performance. *Behav Brain Res* 208:425–429. [CrossRef Medline](#)
- Diekelmann S, Wilhelm I, Wagner U, Born J (2013) Sleep to implement an intention. *Sleep* 36:149–153. [CrossRef Medline](#)
- Fischer S, Drosopoulos S, Tsien J, Born J (2006) Implicit learning—explicit knowing: a role for sleep in memory system interaction. *J Cogn Neurosci* 18:311–319. [CrossRef Medline](#)
- Fischer S, Wilhelm I, Born J (2007) Developmental differences in sleep's role for implicit off-line learning: comparing children with adults. *J Cogn Neurosci* 19:214–227. [CrossRef Medline](#)
- Fiser J, Berkes P, Orbán G, Lengyel M (2010) Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn Sci* 14:119–130. [CrossRef Medline](#)
- Friston K (2012) Predictive coding, precision and synchrony. *Cogn Neurosci* 3:238–239. [CrossRef Medline](#)
- Giesbrecht B, Sy JL, Guerin SA (2013) Both memory and attention systems contribute to visual search for targets cued by implicitly learned context. *Vision Res* 85:80–89. [CrossRef Medline](#)
- Gregory RL (1980) Perceptions as hypotheses. *Philos Trans R Soc Lond B Biol Sci* 290:181–197. [CrossRef Medline](#)
- Henson RN, Rugg MD, Shallice T, Dolan RJ (2000) Confidence in recognition memory for words: dissociating right prefrontal roles in episodic retrieval. *J Cogn Neurosci* 12:913–923. [CrossRef Medline](#)
- Hinton GE, Dayan P, Frey BJ, Neal RM (1995) The “wake-sleep” algorithm for unsupervised neural networks. *Science* 268:1158–1161. [CrossRef Medline](#)
- Hoddes E, Dement WC, Zarcone V (1972) The development and use of the Stanford Sleepiness Scale. *Psychophysiology* 9:150.
- Horne JA, Ostberg O (1976) A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *Int J Chronobiol* 4:97–110. [CrossRef Medline](#)
- Janacek K, Nemeth D (2012) Predicting the future: from implicit learning to consolidation. *Int J Psychophysiol* 83:213–221. [CrossRef Medline](#)
- Kaposvari P, Kumar S, Vogels R (2018) Statistical learning signals in macaque inferior temporal cortex. *Cereb Cortex* 28:250–266. [CrossRef Medline](#)
- King BR, Hoedlmoser K, Hirschauer F, Dolfin N, Albouy G (2017) Sleeping on the motor engram: the multifaceted nature of sleep-related motor memory consolidation. *Neurosci Biobehav Rev* 80:1–22. [CrossRef Medline](#)
- Lutz ND, Diekelmann S, Hinse-Stern P, Born J, Rauss K (2017) Sleep supports the slow abstraction of gist from visual perceptual memories. *Sci Rep* 7:42950. [CrossRef Medline](#)
- Macrae CN, Moran JM, Heatherton TF, Banfield JF, Kelley WM (2004) Medial prefrontal activity predicts memory for self. *Cereb Cortex* 14:647–654. [CrossRef Medline](#)
- Marino M, Li Y, Rueschman MN, Winkelman JW, Ellenbogen JM, Solet JM, Dulin H, Berkman LF, Buxton OM (2013) Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep* 36:1747–1755. [CrossRef Medline](#)
- Marshall L, Born J (2007) The contribution of sleep to hippocampus-dependent memory consolidation. *Trends Cogn Sci* 11:442–450. [CrossRef Medline](#)
- McKeon S, Pace-Schott EF, Spencer RM (2012) Interaction of sleep and emotional content on the production of false memories. *PLoS One* 7:e49353. [CrossRef Medline](#)
- Mumford D (1992) On the computational architecture of the neocortex: II. The role of cortico-cortical loops. *Biol Cybern* 251:241–251. [CrossRef Medline](#)
- Nemeth D, Janacek K, Londe Z, Ullman MT, Howard DV, Howard JH Jr (2010) Sleep has no critical role in implicit motor sequence learning in young and old adults. *Exp Brain Res* 201:351–358. [CrossRef Medline](#)
- Nissen MJ, Bullemer P (1987) Attentional requirements of learning: evidence from performance measures. *Cogn Psychol* 19:1–32. [CrossRef Medline](#)

- Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh Inventory. *Neuropsychologia* 9:97–113. [CrossRef Medline](#)
- Pace-Schott EF, Spencer RM (2013) Age-related changes in consolidation of perceptual and muscle-based learning of motor skills. *Front Aging Neurosci* 5:1–7. [CrossRef Medline](#)
- Pardilla-Delgado E, Payne JD (2017) The impact of sleep on true and false memory across long delays. *Neurobiol Learn Mem* 137:123–133. [CrossRef Medline](#)
- Payne JD, Schacter DL, Propper RE, Huang LW, Wamsley EJ, Tucker MA, Walker MP, Stickgold R (2009) The role of sleep in false memory formation. *Neurobiol Learn Mem* 92:327–334. [CrossRef Medline](#)
- Peigneux P, Maquet P, Meulemans T, Destrebecqz A, Laureys S, Degueldre C, Delfiore G, Aerts J, Luxen A, Franck G, Van der Linden M, Cleeremans A (2000) Striatum forever, despite sequence learning variability: a random effect analysis of PET data. *Hum Brain Mapp* 10:179–194. [CrossRef Medline](#)
- Poldrack RA, Clark J, Paré-Blagoev EJ, Shohamy D, Crespo Moyano J, Myers C, Gluck MA (2001) Interactive memory systems in the human brain. *Nature* 414:546–550. [CrossRef Medline](#)
- Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 2:79–87. [CrossRef Medline](#)
- Rasch B, Born J (2013) About sleep's role in memory. *Physiol Rev* 93:681–766. [CrossRef Medline](#)
- Raass K, Born J (2017) A role of sleep in forming predictive code. In: *Cognitive neuroscience of memory consolidation* (Axmacher N, Rasch B, eds), pp 117–132. New York: Springer.
- Raass K, Pourtois G (2013) What is bottom-up and what is top-down in predictive coding? *Front Psychol* 4:276. [CrossRef Medline](#)
- Rimmele U, Davachi L, Petrov R, Dougal S, Phelps EA (2011) Emotion enhances the subjective feeling of remembering, despite lower accuracy for contextual details. *Emotion* 11:553–562. [CrossRef Medline](#)
- Rimmele U, Davachi L, Phelps EA (2012) Memory for time and place contributes to enhanced confidence in memories for emotional events. *Emotion* 12:834–846. [CrossRef Medline](#)
- Robertson EM (2007) The serial reaction time task: implicit motor skill learning? *J Neurosci* 27:10073–10075. [CrossRef Medline](#)
- Robertson EM, Pascual-Leone A, Press DZ (2004) Awareness modifies the skill-learning benefits of sleep. *Curr Biol* 14:208–212. [CrossRef Medline](#)
- Rose M, Haider H, Salari N, Büchel C (2011) Functional dissociation of hippocampal mechanism during implicit learning based on the domain of associations. *J Neurosci* 31:13739–13745. [CrossRef Medline](#)
- Rosenthal CR, Andrews SK, Antoniadis CA, Kennard C, Soto D (2016) Learning and recognition of a non-conscious sequence of events in human primary visual cortex. *Curr Biol* 26:834–841. [CrossRef Medline](#)
- Schendan HE, Searl MM, Melrose RJ, Stern CE (2003) An fMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron* 37:1013–1025. [CrossRef Medline](#)
- Smith FW, Muckli L (2010) Nonstimulated early visual areas carry information about surrounding context. *Proc Natl Acad Sci U S A* 107:20099–20103. [CrossRef Medline](#)
- Song S, Cohen LG (2014) Practice and sleep form different aspects of skill. *Nat Commun* 5:3407. [CrossRef Medline](#)
- Song S, Howard JH Jr, Howard DV (2007a) Sleep does not benefit probabilistic motor sequence learning. *J Neurosci* 27:12475–12483. [CrossRef Medline](#)
- Song S, Howard JH Jr, Howard DV (2007b) Implicit probabilistic sequence learning is independent of explicit awareness. *Learn Mem* 14:167–176. [CrossRef Medline](#)
- Song S, Howard JH Jr, Howard DV (2008) Perceptual sequence learning in a serial reaction time task. *Exp Brain Res* 189:145–158. [CrossRef Medline](#)
- Spencer RM, Sunm M, Ivry RB (2006) Sleep-dependent consolidation of contextual learning. *Curr Biol* 16:1001–1005. [CrossRef Medline](#)
- Stadler MA (1995) Role of attention in implicit learning. *J Exp Psychol Learn Mem Cogn* 21:674–685. [CrossRef](#)
- Sterpenich V, Schmidt C, Albouy G, Matarazzo L, Vanhaudenhuyse A, Boveroux P, Degueldre C, Leclercq Y, Baeteu E, Collette F, Luxen A, Phillips C, Maquet P (2014) Memory reactivation during rapid eye movement sleep promotes its generalization and integration in cortical stores. *Sleep* 37:1061–1075. [CrossRef Medline](#)
- Summerfield C, Egner T, Greene M, Koechlin E, Mangels J, Hirsch J (2006) Predictive codes for forthcoming perception in the frontal cortex. *Science* 314:1311–1314. [CrossRef Medline](#)
- Turk-Browne NB, Scholl BJ, Johnson MK, Chun MM (2010) Implicit perceptual anticipation triggered by statistical learning. *J Neurosci* 30:11177–11187. [CrossRef Medline](#)
- Vinken K, Vogels R, Op de Beeck H (2017) Recent visual experience shapes visual processing in rats through stimulus-specific adaptation and response enhancement. *Curr Biol* 27:914–919. [CrossRef Medline](#)
- Wagner AD, Schacter DL, Rotte M, Koutstaal W, Maril A, Dale AM, Rosen BR, Buckner RL (1998) Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science* 281:1188–1191. [CrossRef Medline](#)
- Willingham DB, Greenberg AR, Thomas RC (1997) Response-to-stimulus interval does not affect implicit motor sequence learning, but does affect performance. *Mem Cognit* 25:534–542. [CrossRef Medline](#)
- Witt K, Margraf N, Bieber C, Born J, Deuschl G (2010) Sleep consolidates the effector-independent representation of a motor skill. *Neuroscience* 171:227–234. [CrossRef Medline](#)
- Yordanova J, Kolev V, Bruns E, Kirov R, Verleger R (2017) Sleep spindles in the right hemisphere support awareness of regularities and reflect pre-sleep activations. *Sleep* 40:11. [CrossRef Medline](#)