

Sleep-independent off-line enhancement and time of the day effects in three forms of skill learning

Ferenc Kemény^{1,2,3} · Ágnes Lukács²

Received: 16 June 2015 / Accepted: 7 January 2016 / Published online: 9 February 2016
© Marta Olivetti Belardinelli and Springer-Verlag Berlin Heidelberg 2016

Abstract The role of sleep in memory and skill-learning processes is an important and widely debated issue. The current study explores the nature of the relationship between sleep and off-line improvement in three tasks for measuring different aspects of skill learning: the serial reaction time (SRT) task, which is a motor sequence learning task; the artificial grammar learning (AGL) task, testing abstract verbal sequence learning; and the weather prediction (WP) task, which is a non-sequential categorization task. Each participant was tested on one of the three tasks twice, either in a Wake condition (with a 12-h off-line period without sleep), or in a Sleep condition (with sleep). Results showed no sleep-related off-line improvement throughout the three tasks in a two-session re-learning design, but a sleep-independent time-based effect was found on the SRT task. No performance boost was observed in the WP and AGL tasks. Performance on the SRT showed a time of the day effect: the Sleep group outperforming the Wake group; however, this effect was restricted to overall response latencies. Taken together, no

evidence was found in favor of sleep-dependent off-line enhancement in skill learning, but methodological concerns warrant further investigations.

Keywords Off-line enhancement · Sleep · Skill learning · Time of day · Consolidation

Introduction

Sleep has a critical role in learning and memory formation, but the specific contribution of sleep to different forms of learning and memory is still a topic of extensive research. Studies usually differentiate two off-line effects related to consolidation: the lack of forgetting and off-line improvement between testing sessions (Lechner et al. 1999; McGaugh 2000). In this paper, we will focus on sleep-related off-line improvement and refer to the lack of forgetting as retention. Off-line enhancement seems to be sleep-dependent for several visual skills such as figure-ground segmentation (Karni et al. 1994), visual discrimination (Stickgold and Walker 2007), contour integration (Gerván and Kovács 2007), and visuomotor saccade learning (Gais et al. 2008). So far, results for the role of sleep in the off-line improvement of visuomotor skill learning are controversial: Several studies provide evidence in favor of sleep-dependent retention processes (Durrant et al. 2011; Kuriyama et al. 2004; Stickgold et al. 2000; Walker et al. 2002, 2003) while there is also a set of results arguing against the critical role of sleep in retention of this form of learning (Cai and Rickard 2009; Hallgató et al. 2013; Nemeth et al. 2010; Pan and Rickard 2015; Rickard et al. 2008; Robertson et al. 2004; Song et al. 2007a; Wilson et al. 2012). The current study focuses on how an off-line period with or without sleep affects

Handling Editor: Paul Verhaeghen, Georgia Institute of Technology.

Reviewers: Darlene Howard, Georgetown University; Timothy Rickard, UC San Diego.

✉ Ferenc Kemény
fkemeny@cogsci.bme.hu

¹ Division of Developmental Psychology, Institute for Psychology, University of Graz, Universitätsplatz 2, 8010 Graz, Austria

² Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary

³ Institute for Psychology, University of Bern, Bern, Switzerland

performance in skill learning, one of the central aspects of human behavior. Skill learning is generally understood as a heterogeneous phenomenon present in many different areas of behavior varying in the engagement of different cognitive functions. In this paper, we focus on sleep-dependent off-line enhancement in three different aspects of learning such as learning of non-abstract motor sequences (serial reaction time task), statistical learning of abstract non-motor sequences (artificial grammar learning), and probabilistic classification learning (weather prediction task).

The majority of results in consolidation of skill learning tested motor sequence learning on a finger-tapping task (Doyon et al. 2009; Kuriyama et al. 2004; Walker et al. 2002, 2003). These results argue that a performance boost is only present when the off-line period involves sleep. Recently, the critical role of sleep in motor skill learning has been challenged by studies using some form of the serial reaction time (SRT) task (see below, Nissen and Bullemer 1987). Several studies by Nemeth et al. (2010; Hallgató et al. 2013) found a retention of sequence-specific knowledge and an off-line improvement of general motor skills that are independent of sleep. They used a 12-h off-line period and a modified version of the SRT task, the alternating serial reaction time (ASRT) task (Howard and Howard 1997) where the appearance of every second element follows a deterministic sequence, while the rest is random.

A recent study by Meier and Cock (2014) tested how the length of the off-line period and the complexity of the sequences affect learning. They compared the effect of 24-h and 1-week off-line periods on sequence learning in both deterministic and probabilistic sequence learning paradigms. As the off-line intervals were 24 and 168 h (both necessarily involving sleep), the study did not compare off-line enhancement with and without sleep. Similar to the previous studies, Meier and Cock (2014) found retention of sequence-related knowledge and strong off-line improvement for general skills.

Sleep-dependent consolidation of sequence-specific skill learning outside the motor skill domain has been studied less extensively. Durrant et al. (2011) used a motor-free statistical sequence learning paradigm to test the effect of sleep on learning. In their task, participants were exposed to a stream of tones with different pitch. Certain sequences were more frequent than others. In the testing sessions, participants had to choose between structured and unstructured exposures. Half of the participants was first tested in the morning (no sleep in the off-line period), and the other half was tested in the evening (sleep in the off-line period). Results showed a significant performance increase from the immediate recall phase to the delayed recall phase only in the Sleep group; no off-line improvement effect was observed in the Wake group. The

experiment was replicated with a 4-h off-line period between 12 p.m. and 4 p.m. either involving or not involving a nap in between. Results showed a similar pattern to Experiment 1: There was a significant sleep-related off-line improvement. These findings are paralleled by evidence for sleep-dependent retention in the abstraction of non-adjacent dependencies in artificial language learning in infants (Gomez et al. 2006; Hupbach et al. 2009). Taken together, results in this domain are interpreted as demonstrating the important role of sleep in the off-line improvement of statistical information.

Skill learning is not limited to the acquisition of sequences. Sleep-related learning was also tested in probabilistic category learning using the weather prediction (WP) task (Djonlagic et al. 2009). In this task, participants are exposed to 1, 2, or 3 out of four possible cues and have to guess whether it would be sunshine or rain. Cues and outcomes have a probabilistic relationship, and participants get feedback after each trial. Participants are expected to improve in prediction performance throughout the task. Similar to the previous studies, Djonlagic et al. (2009) tested participants either in the morning or in the evening and then again after a 12-h off-line period with or without sleep. Results showed a significant off-line improvement in the Observation condition, where cues and outcomes were presented simultaneously. Off-line improvement appeared for both the Sleep and Wake groups, but it was significantly higher in the Sleep group than in the Wake group. Only the Sleep group showed a significant improvement effect when there was only a short period of learning with feedback. In the case of longer learning with feedback, no groups showed a significant increase from the pre-consolidation to the post-consolidation phase; however, the mean performance difference between the two phases was marginally higher in the Sleep group than in the Wake group. Results were interpreted as evidence in favor of sleep-related off-line improvement in probabilistic categorization.

While the current study focuses on sleep-dependent memory consolidation in skill learning, consolidation designs are suitable for testing time of the day effects as well. Yet, very few skill-learning studies address circadian effects, despite the fact that this issue has a long history in other domains of research on memory and language (dating back to Jenkins and Dallenbach 1924). May et al. (1993) tested young and older adults on a standardized Morningness–Eveningness Questionnaire, a self-assessment questionnaire that requires participants to answer questions like how they would feel if they went to bed at 11 p.m., or how they would plan a day with hard physical work. May et al. (1993) found that younger adults are mostly Evening or Neutral types, whereas older adults are mainly Morning types. The authors also tested recognition memory and showed that learning in the evening results in much higher

performance for young adults than in older adults, while there was no group difference in the morning session. That is, recognition memory is strongly affected by the time of the day, especially by the synchrony between the optimal period of learning and the time of testing. With a similar design, May et al. (2005) tested younger adults in peak or off-peak sessions on implicit and explicit stem completion and implicit category generation. Results showed that implicit performance was better off-peak, while explicit performance was better in the personal optimal period. According to the authors, results argue for different circadian schedules for explicit and implicit functioning. These results are paralleled by findings of Wieth and Zacks (2011) who report a similar pattern for analytic versus insight problem-solving: the latter benefited from testing at non-optimal time of day.

As reviewed above, previous studies focused on consolidation effects in skill learning through one specific paradigm, yielding controversial findings and suggesting that the contribution of sleep to skill learning might be task-dependent. In the current study, we use three different tasks to assess the effects of sleep-dependent and sleep-independent 12-h off-line periods. Since skill learning generally involves multiple mechanisms relying on several underlying neural systems (Lukács and Kemény 2014, 2015), our aim was to test and compare the off-line enhancement effects observed in different aspects of skill learning. For this reason, we tested participants on one of the following three tasks: the serial reaction time task (Nissen and Bullemer 1987), measuring the motor-based sequence learning aspect of skill learning, an artificial grammar learning task (Saffran 2002), measuring the motor-free abstract statistical sequence aspect of skill learning, or the weather prediction task (Knowlton et al. 1996), assessing the motor-free non-sequential statistical aspect of skill learning. Relying on three different tasks gives us a unique opportunity to explore whether sleep affects different forms and aspects of skill learning differentially. If sleep enhances skill learning in general, we would see sleep-related enhancement on all three tasks. If the effect is selective to sequence learning, that would result in enhancement in the SRT task and artificial grammar learning. If statistical learning is affected by sleep-dependent enhancement, that would entail increased post-consolidation performance in artificial grammar learning and the weather prediction task. The last possibility is that sleep-related enhancement is different depending on whether the task requires abstraction over the environmental input. Artificial grammar learning and the weather prediction task require participants to abstract and generalize over the input stimuli, manipulate abstract representations, while the serial reaction time task only requires response to surface features.

Methods

Participants

Altogether 130 people (mean age: 20.65; SD: 1.42; range: 18.83–29.08; 53 female/77 male) participated in the study. All participants provided a written informed consent in accordance with the principles set out in the Declaration of Helsinki and the stipulations of the local Institutional Review Board. Participants with known neurological or cognitive deficits were not included in the study. All participants were recruited from the Budapest University of Technology and Economics and participated voluntarily for credit points.

Participants were randomly assigned into six groups along two variables: Task (SRT vs. AGL vs. WP) and Sleep condition (Sleep vs. Wake). Forty-one participants were tested twice on the SRT task—22 starting in the evening (Sleep group) and 19 in the morning (Wake group); 45 participants were tested twice on the AGL task—23 in the Sleep group and 22 in the Wake group, whereas 44 students were engaged in the WP task—21 in the Sleep and 23 in the Wake group.

Stimuli and procedure

The current study focuses on sleep-dependent enhancement over a 12-h off-line period. The first session of testing of the Wake groups took place in the morning, between 7 a.m. and 9 a.m., and retesting occurred 12 h later. We tested the Sleep groups first in the evening between 7 p.m. and 9 p.m. and then 12 h later. All participants were assigned to only one task (taking approximately 20 min), and to either the Sleep or the Wake group. Contrary to previous studies, we tested off-line enhancement with a repeated learning paradigm,¹ that is, participants faced the exactly same task twice with a 12-h off-line period. All experimental paradigms were computerized. Stimulus presentation and data collection was done by E-prime 1.2 (Psychology Software Tools Inc., Pittsburgh, PA).

Serial reaction time (SRT) task

The task was identical to the one we used in previous studies (Lukács and Kemény, 2014, 2015), which was an

¹ Previous studies using the SRT task applied this same design, that is, the same session repeated twice (e.g., Meier and Cock 2014; Nemeth et al. 2010). On the other hand, earlier studies of the AGL and WP tasks employed only a test phase in Session 2: For the cited WP task, the test session was a short session without feedback (Djonlagic et al. 2009), while for AGL, it was a familiarity decision for statistically constrained and unconstrained sequences (Durrant et al. 2011). That is, no further learning could have occurred in the testing phase in these cases.

adaptation of the task used by Nissen and Bullemer (1987). Four circles (diameter = 55 pixel) appeared in the horizontal centerline of a 640 × 480 screen with equal distances between the circles. One of the circles was black (target stimulus), while the other three were white. The aim of the participants was to press the button corresponding to the target stimulus. The response buttons were Y, C, B, M, which are located in the bottom row of Hungarian QWERTZ keyboards with one button in between each pair of response buttons. The target item was on screen until one of the response buttons was pressed. If the answer was incorrect, a short, 560-ms tone signalled the error. The response-to-stimulus interval was set to 250 ms.

There were 12 blocks of 60 button-presses. Unknown to the participants, there was a 12-element-long second-order conditional deterministic sequence present in Blocks 1 to 11 (yielding 5 presentations of the sequence in each block). Block 12 was the transfer block with the target stimulus appearing in a pseudorandom order (no two neighboring stimuli were identical). The sequence was 121423413243, in which number 1 stands for the black circle (and the response button) in the leftmost position. The increasing numbers reflect the shifting of the target stimulus (and required response) to the right.

Participants were asked to keep their fingers on the response keys throughout the task, and to respond as quickly and as accurately as possible. Participants were not informed about the presence of the sequence or the change in the conditions between Blocks 11 and 12.

Artificial grammar learning (AGL) task

The AGL paradigm was adapted from the P language of Saffran (2002) and was identical to the version we used in previous skill-learning studies (Lukács and Kemény 2014, 2015). The task had a training phase and a test phase. During the training phase of approximately 12 m, participants heard 58 sequences repeated twice. Sequences in the training phase were generated by rules of a small phrase structure grammar. The grammar contained four rules (1). The set of rules was applied to a small lexicon (2): each category of the grammar (A, C, D, F, G) could be manifested in one of two or four different forms. (3) provides example sentences.

In the test phase, participants were presented with two sequences and had to decide which sequence was more similar to the sequences heard in the training phase. There were 24 pairs of sequences, each pair consisted of a novel grammatical and an ungrammatical sequence. The order of the pairs and the order of the sequences within the pairs were randomized. Participants were asked to press “1” if they considered the first sentence similar to the earlier sentences and “2” if the second.

(1) Rules of the phrase structure grammar

$$S \rightarrow AP + BP + (CP)$$

$$AP \rightarrow A + (D)$$

$$BP \rightarrow CP + F$$

$$CP \rightarrow C + (G)$$

(2) Lexicon:

A: bif, hep, mib, rud

C: kav, lam, neb, szig

D: lor, gal

F: dup, dók, rász, vot

G: tez, péf

(3) Example sentences:

Bif lor szig péf dók

Hep gal lam péf dók

Mib lam péf vot

Rud gal neb dup

Weather prediction (WP) task

The WP task was an adaptation of the task used by Gluck and colleagues (2002) and identical to the task we used in previous skill-learning studies (Kemény and Lukács 2010; Lukács and Kemény 2014, 2015). In this task, participants faced 1, 2, or 3 out of four possible cues. Cue1 was a square, Cue2 was a triangle, Cue3 was a pentagon, while Cue4 was a rhombus. Participants had to decide whether the outcome would be sunshine or rain. A feedback slide revealed the correct outcome after each decision. Unknown to the participants, each cue had a predefined predictive value with which it predicted sunshine. Cue1 predicted sunshine in 85.7 % of cases, Cue2 in 70 %, Cue3 in 30 %, and Cue4 in 14.3 %. Note that in all other cases, the cues were associated with rain.

Each cue appeared in a 125 × 125 pixel square 144 pixels from the top of the 640 × 480 screen. If there was only one cue present, the cue was located in the horizontal centerline. In the case of two simultaneous cues, the cues appeared on the two sides of the centerline, while if there were three cues present, the central cue appeared in the centerline and the two other cues on each side. After each response, the cues remained on screen with an 83 × 86 pixel icon appearing in the horizontal centerline 343 pixels from the top of the screen. The icon was either a drawing of the sun, or a drawing of a cloud with rain. The feedback was on screen for 1500 ms, then it disappeared, and the new set of items appeared for prediction.

On the appearance of the cues, participants were asked to press ENTER for sunshine and SPACE for rain. There were four blocks of 50 items. Each block consisted of 50 trials. The order of the trials was preset, no two consecutive

Table 1 Types and occurrences of cues or cue combinations per blocks of 50 trials

Cues	Frequency	<i>p</i> (SUN)
A	8	0.875
B	4	0.75
C	4	0.25
D	8	0.125
AB	8	0.875
AC	1	1
BC	2	0.5
BD	1	0
CD	8	0.125
ABC	2	1
ABD	1	1
ACD	1	0
BCD	2	0

The first column (cues) shows which cues are present in a given combination: A is Cue1, B is Cue2, C is Cue3, and D is Cue4. Frequency is the number of appearances within a block of 50 trials. The third column provides the probability at which the given cue or combination is associated with sunshine. Note that the overall predictive value of Cue1 is the weighted average of the patterns that involve Cue1

items presented the same combination of cues. Table 1 summarizes the design of the weather prediction task.

Data analysis

In the SRT task, we analyzed two separate factors: general motor skills and sequence-specific knowledge. After providing an analysis of reaction times by Session by Block and by Group, the next analysis tests off-line improvement by the comparison of the last sequence block of Session 1 (Block 11) and the first sequenced block of Session 2 (Block 13). This comparison reveals how the interplay of general skills and sequence-specific knowledge consolidates. In the case of general motor skills, we compared reaction times of the random block in the two sessions (i.e., Block 12 and Block 24). Comparing the two random blocks is expected to give information on how sequence-free general skills consolidate. In the case of sequence-specific learning, we contrasted the difference between the random block and the last sequenced block between the two sessions (Block 12 RT – Block 11 RT vs. Block 24 RT – Block 23 RT). This analysis shows how re-learning the sequence affects sequence-specific knowledge. For all participants, we calculated median reaction times per Block, and the mean of median reaction times was compared by Group.

In the AGL task, participants' performance was measured by the rate of correct responses to the 24 test pairs.

Participants faced the same task twice, performance was compared by session and by group, to assess the effect of sleep as well as the effect of the 12-h off-line period.

In the WP task, in concert with previous studies (Kemény 2014; Kemény and Lukács 2013a, b; Knowlton et al. 1994), we used a chance maximization strategy in identifying correct responses on the WP task. A response was coded as correct if the participant gave an answer that matched the predictive values of the cues—regardless of the final outcome. That is, if cues 1, 2, and 3 were present, we expected a SUN answer, as the combined predictive value of the three cues was $(85.7 + 70 + 30)/3 = 61.9\%$, which is above chance level (50 %) in predicting SUN. In this case, the correct response was SUN, even if the actual outcome turned out to be RAIN. Off-line enhancement was analyzed with the comparison of Block 4 (last block of Session 1) and Block 5 (first block of Session 2). To obtain data on gradual learning, we also compared improvement from Block 3 to Block 4 (online improvement: improvement due to learning) and improvement from Block 4 to Block 5 (off-line + online improvement: improvement due to learning and off-line effect). In the SRT and AGL tasks, we tested the effect of repeated learning sessions, in which performance was assessed at the end of each session. To test consolidation effects on repeated learning in the WP task, we compared the performance increase between the last block of Session 1 and the last block of Session 2 by the two groups.

After analyzing effects on consolidation, we also tested time of the day effects in each task comparing performance measures of the Sleep and Wake groups at Session 1. We expected this analysis to show potential differences due to the time of assessment: The Sleep group was first tested in the evening and the Wake group in the morning. The analysis focused on the sequence learning score in the SRT task (Block 12 RT – Block 11 RT), on mean performance in the AGL task, and on Block 4 accuracy in the WP task.

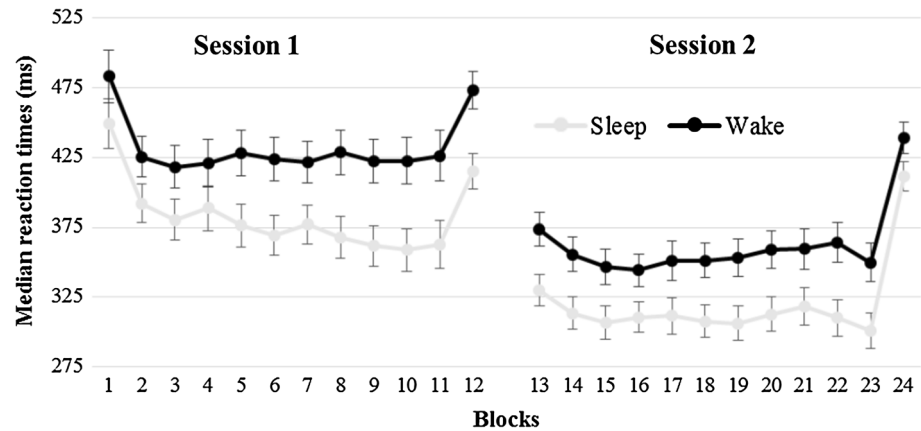
Results

Serial reaction time task

First, we present the analysis on overall reaction times. A $2 \times 2 \times 12$ mixed ANOVA was conducted with Session (Session 1 vs. Session 2) and Block (1 to 12) as within-subject variables, and Group (Sleep vs. Wake) as between-subject variable.² Figure 1 provides mean RTs for both groups on all 24 blocks. Results showed significantly shorter mean RTs for the Sleep group than for the Wake

² Due to sphericity violations, Huynh–Feldt corrections were used for Block and Session \times Block effects.

Fig. 1 Reaction times by Session (Session 1 and Session 2), by Block (1–24), and by Group (Sleep vs. Wake) on the SRT task. Error bars indicate standard errors of mean (SEM)



group, $F(1, 39) = 7.527$, $p < 0.01$, $\eta_p^2 = 0.162$. RTs were also generally shorter for Session 2 than for Session 1, $F(1, 39) = 135.192$, $p < 0.001$, $\eta_p^2 = 0.776$. There was also a significant Block effect, $F(3.282, 127.991) = 38.262$, $p < 0.001$, $\eta_p^2 = 0.495$, and a significant Session \times Block interaction, $F(5.550, 216.444) = 11.326$, $p < 0.001$, $\eta_p^2 = 0.225$. No other effects were significant (all $ps > 0.322$). Results show a general RT difference between the groups in favor of the Sleep group, and between sessions, with shorter RTs on the second session.

The next analysis compares the last sequence block of Session 1 with the first sequence block of Session 2. A 2×2 mixed ANOVA was conducted with Block (Block 11 vs. Block 13) as within-subject variable and Group (Sleep vs. Wake) as between-subject variable. The ANOVA revealed that reaction times in Block 11 were significantly longer than Block 13 RTs, confirmed by a significant main effect of Block, $F(1, 39) = 22.837$, $p < 0.001$, $\eta_p^2 = 0.369$. Also, the Sleep group performed significantly faster, as revealed by a significant main effect of Group, $F(1, 39) = 7.767$, $p < 0.01$, $\eta_p^2 = 0.166$. The Block \times Group interaction was not significant, $p = 0.269$. Results show shorter response latencies in the second session (Block 13), with the Sleep group performing significantly faster.

To test sequence-free general skills, we compared the reaction times of the random blocks in the two sessions by Group. A 2×2 mixed ANOVA was conducted with Session (Block 12 RT vs. Block 24 RT) as within-subject variable and Group (Sleep vs. Wake) as between-subject variable. The ANOVA revealed that RTs for the Session 2 random block were significantly shorter than Session 1 random block RTs, as confirmed by a significant main effect of Session, $F(1, 39) = 7.455$, $p < 0.01$, $\eta_p^2 = 0.160$. RTs in general were smaller for the Sleep group, as revealed by a significant main effect of Group, $F(1, 39) = 7.711$, $p < 0.01$, $\eta_p^2 = 0.165$. There was also a significant Session \times Group interaction, $F(1, 39) = 4.810$, $p < 0.05$, $\eta_p^2 = 0.110$. To further analyze the Session \times Group

interaction, a separate analysis was conducted for each Group comparing Block 12 versus Block 24 RTs. RTs for the two blocks did not differ in the Sleep group, $F(1, 21) = 0.297$, $p = 0.592$, $\eta_p^2 = 0.014$, while Block 24 RTs were significantly shorter than the Block 12 RTs for the Wake group, $F(1, 18) = 7.266$, $p < 0.05$, $\eta_p^2 = 0.288$. Results again show that the Sleep group was generally faster than the Wake group, and that random RTs in the Wake group shortened from Block 12 to Block 24, while RTs of the Sleep group were not affected.

To test further sequence learning with an off-line period involving or not involving sleep, we conducted a 2×2 mixed ANOVA with Session (Block 12 RT – Block 11 RT vs. Block 24 RT – Block 23 RT) as within-subject variable and Group (Sleep vs. Wake) as between-subject variable. The ANOVA revealed a significant main effect of Session, $F(1, 39) = 28.591$, $p < 0.001$, $\eta_p^2 = 0.423$: the Block 12 – Block 11 RT difference was smaller than the Block 24 – Block 23 RT difference. No other effects were significant (both $ps > 0.186$).

Artificial grammar learning

A 2×2 mixed ANOVA was conducted with Session (Session 1 vs. Session 2) as within-subject variable and Group (Sleep vs. Wake) as between-subject variable. The ANOVA only revealed a significant main effect of Group, showing higher performance for the Sleep group in general, $F(1, 43) = 5.133$, $p < 0.05$, $\eta_p^2 = 0.107$. This Group effect appeared to be stable with time, as the Session \times Group interaction was not significant, $p = 0.601$. The Session main effect was not significant either ($p = 0.185$), showing no difference between Sessions 1 and 2. The Sleep group achieved 60.7 % accuracy (Standard Error = 2.4 %) in differentiating grammatical versus ungrammatical strings in Session 1, and 65 % (SE = 2.6 %) in Session 2, while the Wake group performed at 55.7 % (SE = 2.5 %) in Session 1 and 57.6 % (SE = 2.7 %) in Session 2.

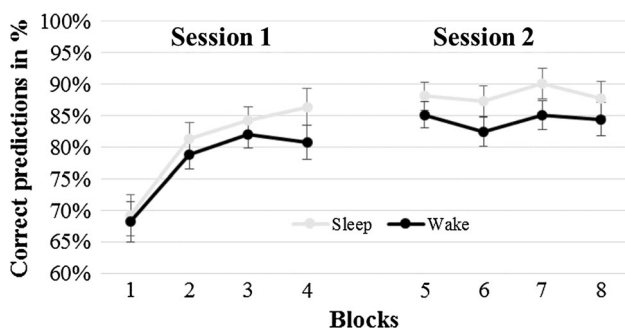


Fig. 2 Performance (% correct) on the weather prediction task by Session (Session 1 and Session 2), by Block (1–8), and by Group (Sleep vs. Wake). Error bars indicate standard errors of means (SEM)

Weather prediction task

We used a 2×2 mixed ANOVA to test the within-subject effect of Session (Block 4 vs. Block 5) and the between-subject effect of Group (Sleep vs. Wake). A significantly higher performance in the post-consolidation block (Block 5) was revealed by a significant main effect of Session, $F(1, 42) = 4.914$, $p < 0.05$, $\eta_p^2 = 0.105$. No other effects were significant (all $ps > 0.198$). See Fig. 2 for performance measures by Group, Block, and Session.

To test whether the above improvement from Block 4 to Block 5 was simply due to further learning, or learning and off-line enhancement as well, we compared the improvement from Block 3 to Block 4 with the improvement from Block 4 to Block 5. The previous takes place within Session 1 and can only be due to online learning, while the latter is intersessional and is due to the interaction of online learning and off-line enhancement. A 2×2 mixed ANOVA with Improvement (Intrasession vs. Intersession) as within-subject variable and Group (Sleep vs. Wake) as between-subject variable showed no significant effects, that is, the improvement from Block 3 to Block 4 is not different from the improvement from Block 4 to Block 5, and this effect does not change with Group (all $ps > 0.212$). That is, learning progresses in the post-consolidation block in the same pace as it progressed in the pre-consolidation block, no further performance increase was observed.

To test consolidation effects on repeated learning, we conducted a 2×2 mixed ANOVA with Session (Last block of Session 1 vs. Last block of Session 2) as within-subject and Group (Sleep vs. Wake) as between-subject variable. The ANOVA revealed no significant effects (all $ps > 0.135$).

Time of the day effects

To assess the time of the day effects in the SRT task, we conducted a 2×2 mixed ANOVA with Block (Sequence vs. Random) as within-subject and Group (Sleep vs. Wake)

as between-subject variable. Results revealed a significant main effect of Block, indicating sequence learning, $F(1, 39) = 41.292$, $p < 0.001$, $\eta_p^2 = 0.514$. The Sleep group showed a general RT advantage, as revealed by a significant main effect of Group, $F(1, 39) = 8.741$, $p < 0.01$, $\eta_p^2 = 0.183$. The Block \times Group interaction was not significant ($p = 0.725$).

Time of the day effects for the AGL task was tested by a univariate ANOVA with accuracy as the dependent variable and Group (Sleep vs. Wake) as independent variable. The ANOVA revealed no significant group-based differences, $F(1, 43) = 2.108$, $p = 0.154$, $\eta_p^2 = 0.047$. Block 4 accuracy was the dependent measure in the WP task with Group (Sleep vs. Wake) as between-subject variable. The ANOVA showed no significant difference between the groups, $F(1, 42) = 1.969$, $p = 0.168$, $\eta_p^2 = 0.045$.

Discussion

The central aim of the current study was to test the effect of sleep over a 12-h period with three different skill-learning paradigms. We found that there was a significant decrease in reaction times in the sequence blocks of the SRT task in both the Sleep and the Wake groups, while for random block RTs, we only observed a decrease in reaction times in the Wake group. Results of the AGL task showed significantly better performance in the Sleep group which was unrelated to sleep, as it was present at the initial testing already. There were no signs of performance increase from the first session to the second. The Sleep and Wake groups showed similar performance throughout the two sessions of the WP tasks. In both groups, performance at the beginning of Session 2 was significantly better than performance at the end of Session 1, but the size of performance change from Block 4 to Block 5 was not different from online learning between Block 3 and Block 4 in the first session. We found no evidence in favor of sleep-related differences in the off-line enhancement in the learning of deterministic motor sequences, in the extraction of regularities from auditory sequences and in probabilistic category learning in a non-sequential task. On the other hand, several aspects of sleep-independent enhancement were observed.

In the serial reaction time task, we observed that reaction times in general became shorter after an off-line period. This effect was only present in the sequence blocks, suggesting that the general decrease in reaction times is in fact due to sequence-related enhancement, and not to the consolidation of general motor skills (although there was an enhancement of general motor skills in the Wake group). Although these results are in line with findings by Robertson et al. (2004), they are not fully compatible with all aspects of other earlier results. As reviewed above,

several studies (Meier and Cock 2014; Nemeth et al. 2010) found off-line improvement of general motor skills but no improvement (and also no decline) for sequence-specific knowledge using the same off-line periods. A potential reason for differences in results from Nemeth et al.'s is that they used a different, probabilistic version of the task, the alternating serial reaction time task. A possible explanation of the difference in results is that sequence-specific off-line improvement only appears for deterministic sequences and not probabilistic information. This account is supported by a lack of improvement for statistical learning in our results for AGL and WP, but contradicted by earlier findings of sleep-dependent off-line improvement of statistical information on the similar tasks (Djonlagic et al. 2009; Durrant et al. 2011). Another possible explanation is that sequence learning measures are very low in the ASRT task, which can cause insensitivity to consolidation effects: Participants of six out of eight conditions showed a sequence learning effect below 10 ms in Hallgató et al. (2013), and all learning scores reported by Meier and Cock (2014) were 10 ms or below.

On the other hand, Meier and Cock found evidence of off-line improvement of general motor skill learning using deterministic sequences too, with a 24-h off-line interval. It is possible that there is an initial enhancement for sequence-specific knowledge in deterministic (but not probabilistic) sequences over a period of 12-h. This effect is independent of sleep, as both groups in our study showed a similar pattern. There is also a small wakefulness-related enhancement of general skills: We showed a decrease in RTs for the random blocks only in the Wake group (results in line with Song et al. 2007b). Integrating the results with Meier and Cock's (2014) findings suggests that the sequence-related off-line improvement effect disappears in the second 12 h and gives way to further wake-related off-line enhancement of general motor skills.

The results show that off-line enhancement on the serial reaction time task is different from the other two tasks. One of the crucial differences is that the SRT task has a motor component, while the other two tasks do not have one. Although this finding could suggest that a motor component is necessary for off-line memory enhancement, such a conclusion is contradicted by a large set of evidence for the existence of motor-unrelated off-line enhancement (off-line enhancement after in visual discrimination: Karni et al. 1994; Stickgold et al. 2000, or in the perceptual learning of spoken language Fenn et al. 2003).

Another crucial difference between the serial reaction time task versus AGL and WP is that focusing on surface features can be sufficient for learning in the serial reaction time task, while participants have to abstract away from specific stimuli in the weather prediction and artificial grammar learning tasks. There is considerable debate on

the nature of learning in the SRT task (Deroost et al. 2006; Kemény and Lukács 2011; Remillard 2003; Willingham et al. 2000); in one view, learning is explained by integrating surface stimuli into a single sequence representation. This also applies to the classical design of artificial grammar learning using finite-state grammars. The current experiment however employs a phrase structure grammar where rules are defined over categories instead of elements, which precludes mapping surface sequences of syllables as a sufficient strategy for above chance learning performance. In this design, even chunking requires categorization, as there are a number of items that can appear in each position. Similarly, solving the weather prediction task requires the participants to separate cues, cumulate their predictive values, combine the predictive values, and make a decision based on the combinations (Reber et al. 1996).

If off-line enhancement only appears in skill-learning tasks that operate on surface properties of stimuli, then it is possible to integrate the current results with previous studies of statistical learning. In previous studies of statistical learning showing sleep-related enhancement (Durrant et al. 2011, 2013), the task required participants to identify the transitional probabilities between triplets, i.e., to process and map statistical information on surface elements. The AGL paradigm used in the current study on the other hand required the categorization of elements and the application of rules to the categories themselves, and not to surface stimuli. That is, the smaller the involvement of abstraction, the more likely that off-line enhancement takes place. Note, however, that this hypothesis makes predictions for off-line enhancement regardless of sleep. Another possible explanation for the lack of sleep-dependent off-line enhancement in our study is ceiling effect. Our data show that performance only minimally increases even after the same amount of training in the second session. Hence, it is possible that participants maximally extract what they can by the end of Session 1. However, the fact that the Sleep group shows better performance than the Wake group argues against this account: At least for participants in the Wake group, further improvement is possible, but it does not take place from Session 1 to Session 2 either.

The current findings are not in line with previous results on the weather prediction task (Djonlagic et al. 2009). Djonlagic et al. (2009) showed a sleep-related enhancement for the weather prediction task, but only with a reduced training phase of 100 items and not with the generally used 200-item training. Their conclusion was that learning did not take place after the longer training due to a ceiling effect in learning. Using a repeated learning design, we provided evidence that participants do not reach the highest possible performance even after 200 items of training in Session 1, as further online learning in the second session increases performance. This suggests that

the lack of sleep-related enhancement is not an artifact of at ceiling performance. On the other hand, since there is only a small overall increase in performance in Session 2, we cannot rule out ceiling effects. Further research with shorter training sessions is required to disentangle ceiling effects and lack of sleep-related enhancement in probabilistic categorization. It is also an important methodological difference that we tested the re-learning of statistical information in both the AGL and the WP tasks, while the cited studies only employed a test phase in Session 2 with no feedback.

Our hypothesis that an off-line period is only beneficial in non-abstract learning is seemingly difficult to integrate with infant studies. These studies show that an off-line period triggers generalization. Studies though are not conclusive whether this generalization is sleep- (Friedrich et al. 2015) or wakefulness-dependent (Werchan and Gómez 2014). A possible integration of these studies and the current results concerns the role of off-line enhancement. Previous theories highlight the role of off-line consolidation in the decontextualization and generalization of memories (Gómez 2011). If the role of consolidation is truly to decontextualize and generalize representation, it would entail a difference between representations of different levels of abstractness: Surface features can be easily generalized, while features that are already interpreted are only decontextualized. This would explain that only uninterpreted stimuli benefit from the off-line period, as they yet lack the first step in generalization. This would also entail however that pre- and post-consolidation representation differs in terms of abstractness. No skill consolidation studies addressed this question previously. It is also important to note that results from adult studies are not easy to generalize to infants and vice versa, due to developmental changes.

Note, however, that the current experiment was not designed to test this hypothesis, but the hypothesis was deduced from the results. It is also important to note that while the tasks share a number of characteristics, there are also a number of yet unexplained differences. Hence, further experiments are required to identify whether and to what extent can the lack of abstraction contribute to consolidation processes.

Another possible interpretation of the results, as pointed out by the reviewers of the manuscript, is that the specific effect on the SRT task might be due to the release from fatigue. Pan and Rickard (2015) argue that lengthy motor learning tasks may trigger reactive inhibition that can block improvement, but this inhibition decreases with the introduction of short pauses (Rickard et al. 2008). Because of this effect, it is unclear whether the time-based consolidation effect obtained in the study is confounded by the effect of reactive inhibition. This effect might account for

the effect observed in the SRT task and not the other two tasks. The current study used self-paced breaks between blocks, and as a result, the length of the breaks was not controlled. In this case, our results from the SRT task argue for the lack of consolidation effects. Further studies are required to control break length and clarify this problem.

Time of the day effect

Throughout all three tasks, we observed a time of the day effect, at least to some extent. It is important to highlight that it is not a random group effect, as we used a between-subject design: Each task was tested with a different set of participants. In the SRT task, we obtained significantly shorter reaction times for the Sleep group (evening first advantage), but no difference in the disruption from sequence to random block. That is, the advantage was only present in general skills and not sequence-specific learning (cf. Janacek and Nemeth 2013). There were also no sleep-related effects, but a stable phenomenon, preserved throughout all later stages in both sessions: Session 2 reaction times of the Wake group were in the same range as Session 1 reaction times of the Sleep group (see Fig. 1). That is, a first session in the morning resulted in much longer RTs than a first session in the evening. The second session could have levelled out the RT differences, if Wake group participants had shown better performance in their second session in the evening, reaching the same level as those having the second session in the morning. This is not what the results show: The time of the day effect is carried on to the second session, and the initial group difference is preserved.

A similar pattern was observed in artificial grammar learning. The Sleep group showed a general advantage: They outperformed the Wake group in both sessions. However, despite the high numerical difference, there was no significant statistical difference between the groups when only Session 1 performances were compared. The time of the day effect was statistically not significant in the weather prediction task. However, as shown in Fig. 2, performance of the Wake group is generally below the Sleep group's performance. Note, however, that in both the AGL and the WP tasks, the lack of statistical significance could be due to the lack of statistical power.

As the current study employed six groups with three different tasks with two different times of initial testing, results on the time of the day effect seem to be convincing. As described above, previous studies showed that recognition memory (May et al. 1993) as well as explicit stem completion and implicit category generation (May et al. 2005) and insight problem-solving (Wieth and Zacks 2011) are affected by the time of testing in young adults. The latter studies also involved assessments of morningness or

eveningness and suggested enhanced implicit performance in the suboptimal time windows, which was mainly morning for young adults.

Contrary to previous observations, our results show that participants performed better when they were first tested in the evening, i.e., in their peak time (which is suggested to be their peak- and so non-optimal time for implicit learning), and the evening advantage was carried over to the second testing event. That is, our results do not argue for having a different circadian rhythm for implicit learning of skills than for explicit functions. They do not provide clear evidence, though, as the study itself was not designed to directly test this question. Previous consolidation studies have not taken circadian effects into account and mostly reported the difference scores between pre-consolidation and post-consolidation performance. For this reason, we often do not have the results for absolute performance in the evening versus morning groups. A greater overall learning can lead to greater performance difference between the two sessions, which in turn could be interpreted as a sign of off-line enhancement. In this sense, some observations of sleep-dependent processes may only be a side effect of the circadian effect. This is in line with results by Rickard et al. (2008), showing that if time of the day effects is taken into account, sleep-related effects also disappear. Note, however, that this is still the case despite that we only report a significant time of the day effect in the overall reaction times of the SRTT and not for the disruption of the sequence. Contradictory findings together with the lack of studies call for targeted and systematic research on the time-of-day effects in different forms of skill learning.

Study limitations

The current study focused on sleep-dependent effects on learning, an issue that has a long history in cognitive psychological research. Our aim was to not only focus on the off-line enhancement of an acquired representation, but also to test how the consolidation of the previously learned representation can enhance further learning. In this design, the obtained results are affected by both off-line enhancement and further learning, and the effects are not dissociable. While the design may seem unusual, it highlights the possibility that off-line enhancement does not necessarily affect the representation that has been acquired earlier, but may have an effect on the learning process itself. This may or may not lead to better learning in Session 2. Unfortunately, the current design does not allow disentangling the effects of enhanced sensitivity for further improvement and the off-line enhancement of the already existing representation.

Another methodological concern is related to the serial reaction time task. The current design employed one

random block per session, and the location of the random block is at the end of the session. In assessing general skill learning, we compared the random blocks and found no difference in the Sleep group and a lowering of reaction times in the Wake group. One might speculate, though, that RT levels for the random block are a result of sequence-specific learning, and the more a participant is exposed to a sequence, the higher the disruption will be; further studies should address this issue. At the same time, previous studies of the ASRT show that RTs for random elements show a continuous decrease along with RTs for the sequence element (e.g., Howard and Howard 1997). That is, the disruption caused by sequence learning does not grow with amount of learning which argues against the above speculation. Note, however, that random elements have fixed and predictable locations in the ASRT task; hence, the application of the results to the classical SRT task should be done with caution. In sum, we can assume that random blocks are generally not subject to interference, only to general practice; hence, the random block RTs analyzed in the current paper are not contaminated by further learning.

We provided evidence that the effect of a 12-h consolidation period is independent of sleep. However, we have scaling limitations at least on the serial reaction time task and the artificial grammar learning task. Results in the SRT task showed that RTs decrease after the consolidation period to the same extent in the two groups, with a lower average response latency in the Sleep group. It is possible that further RT decrease in the Sleep group is prevented by a floor effect. The same argument applies to the Sleep group in the AGL task, that is, they might have reached the maximum possible performance. Further studies are required to reveal scaling limitations.

Conclusion

The current study tested the role of sleep-dependent off-line enhancement in three different skill-learning tasks and found that sleep has no critical role in off-line enhancement in the three reported forms of skill learning. Comparing the three tasks, we found that regardless of sleep, the 12-h off-line period only had a beneficial effect on the SRT task, a motor-based sequence learning task, while neither motor-free abstract sequence learning (assessed by artificial grammar learning), nor sequence-free statistical learning (measured by the weather prediction task) was subject to performance boost after an off-line period. These results suggest that off-line periods are especially effective in performance improvement in the learning of skills requiring no abstraction. It is still a question how general motor skills and sequence-specific learning are affected by sleep-

independent off-line improvement. Our results show improvement for sequence-specific learning, but further studies are required using non-sequential motor learning tasks to test off-line effects in other areas of motor learning. We also provided evidence for a time of the day effect in skill learning: Being exposed to some skill-learning tasks results in a more advanced performance if learning takes place in the evening. This advantage is preserved even after 12 h. Further studies are required to understand the nature of this time of the day effect and its relation to sleep-related off-line processes. So far, results suggest that performance boost after an off-line period only appears in non-abstract cognitive skill learning.

Acknowledgments We are grateful to Karolina Janacsek and Dezső Németh for their comments on the manuscript. We thank Kata Fazekas for her help in data collection. During the publication process, Ferenc Kemény was a Sciex NMS-CH postdoctoral fellow (14.120, “MUST–Multimodal Sequence Learning”).

References

- Cai DJ, Rickard TC (2009) Reconsidering the role of sleep for motor memory. *Behav Neurosci* 123(6):1153–1157. doi:[10.1037/a0017672](https://doi.org/10.1037/a0017672)
- Deroost N, Zeeuws I, Soetens E (2006) Effector-dependent and response location learning of probabilistic sequences in serial reaction time tasks. *Exp Brain Res* 171:469–480
- Djonlagic I, Rosenfeld A, Shohamy D, Myers C, Gluck M, Stickgold R (2009) Sleep enhances category learning. *Learn Mem* 16:751–755
- Doyon J, Korman M, Morin A, Dostie V, Hadj Tahar A, Benali H, Carrier J (2009) Contribution of night and day sleep vs. simple passage of time to the consolidation of motor sequence and visuomotor adaptation learning. *Exp Brain Res* 195(1):15–26. doi:[10.1007/s00221-009-1748-y](https://doi.org/10.1007/s00221-009-1748-y)
- Durrant SJ, Taylor C, Cairney S, Lewis PA (2011) Sleep-dependent consolidation of statistical learning. *Neuropsychologia* 49(5):1322–1331. doi:[10.1016/j.neuropsychologia.2011.02.015](https://doi.org/10.1016/j.neuropsychologia.2011.02.015)
- Durrant SJ, Cairney SA, Lewis PA (2013) Overnight consolidation aids the transfer of statistical knowledge from the medial temporal lobe to the striatum. *Cereb Cortex (New York, NY: 1991)* 23(10):2467–2478. doi:[10.1093/cercor/bhs244](https://doi.org/10.1093/cercor/bhs244)
- Fenn KM, Nusbaum HC, Margoliash D (2003) Consolidation during sleep of perceptual learning of spoken language. *Nature* 425(6958):614–616. doi:[10.1038/nature01951](https://doi.org/10.1038/nature01951)
- Friedrich M, Wilhelm I, Born J, Friederici AD (2015) Generalization of word meanings during infant sleep. *Nat Commun*. doi:[10.1038/ncomms7004](https://doi.org/10.1038/ncomms7004)
- Gais S, Rasch B, Wagner U, Born J (2008) Visual-procedural memory consolidation during sleep blocked by glutamatergic receptor antagonists. *J Neurosci* 28(21):5513–5518. doi:[10.1523/JNEUROSCI.5374-07.2008](https://doi.org/10.1523/JNEUROSCI.5374-07.2008)
- Gerván P, Kovács I (2007) Sleep dependent learning in contour integration. *J Vis* 7(9):48. doi:[10.1167/7.9.48](https://doi.org/10.1167/7.9.48)
- Gluck MA, Shohamy D, Myers C (2002) How do people solve the “weather prediction” task? Individual variability in strategies for probabilistic category learning. *Learn Mem* 9:408–418
- Gómez RL (2011) Memory, sleep and generalization in language acquisition. In: Aron I, Clark EV (eds) *Experience, variation and generalization: learning a first language*. John Benjamins, Amsterdam, pp 261–276
- Gomez RL, Bootzin RR, Nadel L (2006) Naps promote abstraction in language-learning infants. *Psychol Sci* 17:670–674
- Hallgató E, Györi-Dani D, Pekár J, Janacsek K, Nemeth D (2013) The differential consolidation of perceptual and motor learning in skill acquisition. *Cortex* 49(4):1073–1081. doi:[10.1016/j.cortex.2012.01.002](https://doi.org/10.1016/j.cortex.2012.01.002)
- Howard JH, Howard DV (1997) Age differences in implicit learning of higher order dependencies in serial patterns. *Psychol Aging* 12:634–656
- Hupbach A, Gomez RL, Bootzin RR, Nadel L (2009) Nap-dependent learning in infants. *Dev Sci* 12(6):1007–1012. doi:[10.1111/j.1467-7687.2009.00837.x](https://doi.org/10.1111/j.1467-7687.2009.00837.x)
- Janacsek K, Nemeth D (2013) Implicit sequence learning and working memory: correlated or complicated? *Cortex* 49(8):2001–2006. doi:[10.1016/j.cortex.2013.02.012](https://doi.org/10.1016/j.cortex.2013.02.012)
- Jenkins JG, Dallenbach KM (1924) Obliviscence during Sleep and Waking. *Am J Psychol* 35(4):605–612. doi:[10.2307/1414040](https://doi.org/10.2307/1414040)
- Karni A, Tanne D, Rubenstein BS, Askenasy JJ, Sagi D (1994) Dependence on REM sleep of overnight improvement of a perceptual skill. *Science (New York, NY)* 265(5172):679–682
- Kemény F (2014) Self-insight in probabilistic categorization—not implicit in children either. *Front Psychol* 5:233. doi:[10.3389/fpsyg.2014.00233](https://doi.org/10.3389/fpsyg.2014.00233)
- Kemény F, Lukács Á (2010) Impaired procedural learning in language impairment: results from probabilistic categorization. *J Clin Exp Neuropsychol* 32:249–258
- Kemény F, Lukács Á (2011) Perceptual effect on motor learning in the serial reaction time task. *J Gen Psychol* 138:110–126
- Kemény F, Lukács Á (2013a) Self-insight in probabilistic category learning. *J Gen Psychol* 140(1):57–81
- Kemény F, Lukács Á (2013b) Stimulus dependence in probabilistic category learning. *Acta Psychol* 143(1):58–64. doi:[10.1016/j.actpsy.2013.02.008](https://doi.org/10.1016/j.actpsy.2013.02.008)
- Knowlton BJ, Squire LR, Gluck MA (1994) Probabilistic classification learning in amnesia. *Learn Mem* 1:106–120
- Knowlton BJ, Mangels JA, Squire LR (1996) A neostriatal habit learning system in humans. *Science* 273:1399–1402
- Kuriyama K, Stickgold R, Walker MP (2004) Sleep-dependent learning and motor-skill complexity. *Learn Mem* 11(6):705–713. doi:[10.1101/lm.76304](https://doi.org/10.1101/lm.76304)
- Lechner HA, Squire LR, Byrne JH (1999) 100 years of consolidation—remembering Müller and Pilzecker. *Learn Mem* 6(2):77–87. doi:[10.1101/lm.6.2.77](https://doi.org/10.1101/lm.6.2.77)
- Lukács Á, Kemény F (2014) Domain-general sequence learning deficit in specific language impairment. *Neuropsychology* 28(3):472–483. doi:[10.1037/neu0000052](https://doi.org/10.1037/neu0000052)
- Lukács Á, Kemény F (2015) Development of different mechanisms of skill learning throughout the lifespan. *Cogn Sci* 39:383–404
- May CP, Hasher L, Stoltzfus ER (1993) Optimal time of day and the magnitude of age differences in memory. *Psychol Sci* 4(5):326–330. doi:[10.1111/j.1467-9280.1993.tb00573.x](https://doi.org/10.1111/j.1467-9280.1993.tb00573.x)
- May CP, Hasher L, Foong N (2005) Implicit memory, age, and time of day. *Psychol Sci* 16(2):96–100. doi:[10.1111/j.0956-7976.2005.00788.x](https://doi.org/10.1111/j.0956-7976.2005.00788.x)
- McGaugh JL (2000) Memory—a century of consolidation. *Science (New York, NY)* 287(5451):248–251
- Meier B, Cock J (2014) Offline consolidation in implicit sequence learning. *Cortex* 57:156–166. doi:[10.1016/j.cortex.2014.03.009](https://doi.org/10.1016/j.cortex.2014.03.009)
- Nemeth D, Janacsek K, Londe Z, Ullman MT, Howard DV, Howard JH (2010) Sleep has no critical role in implicit motor sequence learning in young and old adults. *Exp Brain Res* 201:351–358
- Nissen MJ, Bullemer P (1987) Attentional requirements of learning—evidence from performance-measures. *Cogn Psychol* 19:1–32

- Pan SC, Rickard TC (2015) Sleep and motor learning: is there room for consolidation? *Psychol Bull* 141(4):812–834. doi:[10.1037/bul0000009](https://doi.org/10.1037/bul0000009)
- Reber PJ, Knowlton BJ, Squire LR (1996) Dissociable properties of memory systems: differences in the flexibility of declarative and nondeclarative knowledge. *Behav Neurosci* 110:861–871
- Remillard G (2003) Pure perceptual-based sequence learning. *J Exp Psychol Learn Mem Cogn* 29:581–597
- Rickard TC, Cai DJ, Rieth CA, Jones J, Ard MC (2008) Sleep does not enhance motor sequence learning. *J Exp Psychol Learn Mem Cogn* 34:834–842
- Robertson EM, Pascual-Leone A, Press DZ (2004) Awareness modifies the skill-learning benefits of sleep. *Curr Biol* 14:208–212
- Saffran JR (2002) Constraints on statistical language learning. *J Mem Lang* 47:172–196
- Song S, Howard JH, Howard DV (2007a) Sleep does not benefit probabilistic motor sequence learning. *J Neurosci* 27:12475–12483
- Song S, Howard JH, Howard DV (2007b) Sleep does not benefit probabilistic motor sequence learning. *J Neurosci* 27(46):12475–12483. doi:[10.1523/JNEUROSCI.2062-07.2007](https://doi.org/10.1523/JNEUROSCI.2062-07.2007)
- Stickgold R, Walker MP (2007) Sleep-dependent memory consolidation and reconsolidation. *Sleep Med* 8(4):331–343. doi:[10.1016/j.sleep.2007.03.011](https://doi.org/10.1016/j.sleep.2007.03.011)
- Stickgold R, Whidbee D, Schirmer B, Patel V, Hobson JA (2000) Visual discrimination task improvement: a multi-step process occurring during sleep. *J Cogn Neurosci* 12:246–254
- Walker MP, Brakefield T, Morgan A, Hobson JA, Stickgold R (2002) Practice with sleep makes perfect: sleep-dependent motor skill learning. *Neuron* 35(1):205–211
- Walker MP, Brakefield T, Seidman J, Morgan A, Hobson JA, Stickgold R (2003) Sleep and the time course of motor skill learning. *Learn Mem* 10:275–284
- Werchan DM, Gómez RL (2014) Wakefulness (not sleep) promotes generalization of word learning in 2.5-year-old children. *Child Dev* 85(2):429–436. doi:[10.1111/cdev.12149](https://doi.org/10.1111/cdev.12149)
- Wieth MB, Zacks RT (2011) Time of day effects on problem solving: when the non-optimal is optimal. *Think Reason* 17(4):387–401. doi:[10.1080/13546783.2011.625663](https://doi.org/10.1080/13546783.2011.625663)
- Willingham DB, Wells LA, Farrell JM, Stemwedel ME (2000) Implicit motor sequence learning is represented in response locations. *Mem Cogn* 28:366–375
- Wilson JK, Baran B, Pace-Schott EF, Ivry RB, Spencer RMC (2012) Sleep modulates word-pair learning but not motor sequence learning in healthy older adults. *Neurobiol Aging* 33(5):991–1000. doi:[10.1016/j.neurobiolaging.2011.06.029](https://doi.org/10.1016/j.neurobiolaging.2011.06.029)