Deconstructing Procedural Memory: Different Learning Trajectories and Consolidation of Sequence and Statistical Learning

SIMOR, Peter[1,2*], ZAVECZ, Zsofia[3,4,5*], HORVÁTH, Kata[3,4,6], ÉLTETŐ, Noémi[5], TÖRÖK, Csenge[5], PESTHY, Orsolya[5], GOMBOS, Ferenc[7,8], NEMETH, Dezso[4,5**], JANACSEK, Karolina[4,5**]

[1] Institute of Psychology, Eötvös Loránd University, Budapest, Hungary

[2] Nyírő Gyula Hospital/National Institute of Psychiatry and Addictions, Budapest, Hungary

[3] Doctoral School of Psychology, Eötvös Loránd University, Budapest, Hungary

[4] Brain, Memory and Language Research Group, Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary

[5] Department of Clinical Psychology and Addiction, Institute of Psychology, Eötvös Loránd University, Budapest, Hungary

[6] Department of Cognitive Psychology, Institute of Psychology, Eötvös Loránd University, Budapest, Hungary

[7] Department of General Psychology, Pázmány Péter Catholic University, Budapest, Hungary

[8] MTA-PPKE Adolescent Development Research Group, Hungarian Academy of Sciences, Budapest, Hungary

* These authors contributed equally to this work.

** These authors contributed equally to this work.

Corresponding authors:

Dezso Nemeth, PhD and Karolina Janacsek, PhD, Department of Clinical Psychology and Addiction, Institute of Psychology, Eötvös Loránd University, Damjanich utca 41-43, H-1071 Budapest, Hungary

Email address: nemethd@gmail.com, janacsekkarolina@gmail.com, petersimor@gmail.com

Abstract

Procedural learning is a fundamental cognitive function that facilitates efficient processing of and automatic responses to environmental stimuli. Here, we examined training-dependent and off-line changes of two sub-processes of procedural learning: namely, sequence learning (acquisition of order-based associations) and statistical learning (acquisition of frequency-based associations). Healthy young adults completed a procedural learning task, and were retested after a delay containing either active wakefulness, quiet rest, or daytime sleep. Performance in Sequence Learning increased gradually during training and during additional practice after the delay, while Statistical Learning plateaued early. Although, on a behavioral level, Sequence and Statistical Learning were similar across groups after the delay, cortical oscillations were associated with performance within the sleep group only. Moreover, sleep spindle parameters showed differential associations with Sequence and Statistical Learning. Our findings can contribute to a deeper understanding of the dynamic changes of parallel learning and consolidation processes that underlie procedural memory formation.

Keywords: procedural learning, sequence learning, statistical learning, sleep, EEG, consolidation

Introduction

Procedural learning, the development of perceptual and motor skills through extensive practice is a crucial ability that facilitates efficient processing of and automatic responses to complex environmental stimuli. Procedural learning is evidenced by enhanced performance as well as functional changes in the neural network underlying behavior (Howard et al., 2004; Fletcher et al., 2005). Learning performance does not only depend on training during acquisition but also on the post-learning period (Karni et al., 1998; Doyon et al., 2009; Durrant et al., 2011). Nevertheless, there are intensive debates questioning whether the acquired memories are stabilized or enhanced during post-learning, off-line periods (Maquet et al., 2000; Peigneux et al., 2006; Nemeth et al., 2010; Pan and Rickard, 2015). Mixed findings emerging in this field suggest that different processes *within* the procedural learning domain may show different trajectories during learning and off-line periods. At least two processes underlying procedural learning can be distinguished: sequence learning and statistical learning (Howard et. al., 1997, Nemeth et al., 2013). Sequence learning refers to the acquisition of a series of (usually 5-12) stimuli that repeatedly occur in the same *order* (with no embedded noise in deterministic sequences, or with some embedded noise in probabilistic sequences). In contrast, statistical learning refers to the acquisition of shorter-range relationships among stimuli that is primarily based on *frequency* information (i.e., differentiating between more frequent and less frequent runs (e.g., triplets) of stimuli (Armstrong et al. 2017; Perruchet and Pacton, 2006; Thiessen et al., 2013; Siegelman et al., 2017). Recent studies revealed that these two parallel learning processes show distinct electrophysiological characteristics (Kóbor et al., under revision). Previous research has not directly contrasted the consolidation of these two processes. Here, we show - using a visuomotor probabilistic sequence learning task - that performance in sequence learning

compared to statistical learning (acquisition of order vs. frequency information) show marked practice-dependent improvements before and after off-line periods.

Studies on sequence learning showed enhanced behavioral performance after an off-line period spent asleep compared to an equivalent period spent awake, especially if individuals acquired an explicit, abstract or complex representation of the sequence (Robertson et al., 2004; Spencer et al., 2006; King et al., 2017). On the other hand, learning probabilistic sequences (Song et al., 2007a, Nemeth et al., 2010) does not seem to benefit from post-learning sleep on the behavioral level, while on a neural level, it has been shown that post-learning sleep is involved in the reprocessing and optimization of the acquired probabilistic sequential information (Peigneux et al., 2003). Importantly, in these probabilistic sequence learning studies the behavioral index of learning encompassed the acquisition of both order- and frequency-based information, thus, the consolidation of sequence learning and statistical learning was not examined separately (Song et al., 2007a, Nemeth et al., 2010). There is limited evidence that statistical learning in the auditory domain benefits from sleep (Durrant et al., 2011, 2013), however to date, no study has investigated the consolidation, and more specifically, the role of sleep in statistical learning in the visuomotor domain.

Although sequence learning and statistical learning seem to require different cognitive mechanisms (Howard and Howard, 1997; Perruchet and Pacton, 2006; Nemeth et al., 2013; Thiessen et al., 2013; Siegelman et al., 2017), in everyday learning scenarios, humans might rely simultaneously on both forms of learning. Previous studies investigated the consolidation of these processes in separate task conditions. Therefore, the first aim of our study was to examine the consolidation of sequence learning and statistical learning simultaneously, in the same experimental context. Previous studies suggest that sequence learning may, whereas statistical learning may not benefit from post-learning sleep and sleep-specific oscillations (slow wave activity and spindles); however, these studies applied awake control groups

engaged in daytime activities during the off-line periods (King et al., 2017). Several studies indicate that even a daytime nap might facilitate the consolidation of procedural memories compared to an equivalent time spent awake. Nevertheless, in these studies the amount of interference within the awake control groups was not controlled either, because participants of the wake group can usually do their daily routines before going back to the laboratory (Mednick et al., 2011). As the amount of interference might influence off-line memory processing (Mednick et al., 2011), our second aim was to examine the off-line change of sequence learning and statistical learning after three different post-learning conditions: active wakefulness, quiet rest, and daytime sleep. We hypothesized that sequence learning would be enhanced after daytime sleep and quiet rest (i.e., due to low interference) compared to active wakefulness, whereas off-line change in statistical learning would be independent from the post-learning condition. Finally, in light of previous literature on the positive influence of slow oscillatory and sleep spindling activity on neural plasticity and memory consolidation (Rasch and Born, 2013) we hypothesized that enhanced sequence learning after sleep would be associated with these sleep-specific oscillations.

Materials and methods

Participants

Participants (all native Hungarians) were selected from a large pool of undergraduate students from the Eötvös Loránd University in Budapest. The first step of the selection procedure consisted of the completion of an online questionnaire assessing sleep quality and mental health status. Sleep-related questionnaires included the Pittsburgh Sleep Quality Index (PSQI, Buysse et al., 1989; Takács et al., 2016), and Athens Insomnia Scale (AIS, Soldatos et al., 2003; Novák, 2004). Participants that showed poor sleep quality based on previous normative measurements were not included. The Hungarian version of the short (nine item)

Beck Depression Inventory (BDI, Rózsa et al., 2001) was used to exclude participants with signs of mild to moderate/severe depression, therefore, participants only with a score less than 10 were included. Respondents reporting current or prior chronic somatic, psychiatric or neurological disorders, or the regular consumption of pills other than contraceptives were also excluded. In addition, individuals reporting the occurrence of any kind of extreme life event (e.g., accident) during the last three months that might have had an impact on their mood, affect and daily rhythms were not included in the study. At the first encounter with the assistant, participants were instructed to follow their usual sleep-wake schedules during the week prior to the experiment and to refrain from consuming alcohol and all kinds of stimulants 24 hours before the day of the experiment. Sleep schedules were monitored by sleep agendas, as well as by the adapted version of the Groningen Sleep Quality Scale (Simor et al., 2009) in order to assess individuals' sleep quality the night before the experiment. The data of participants reporting poor sleep quality the night before the experiment (> 7 points) were not considered in the analyses.

After the above selection procedure, 96 right-handed (28 males, M_{age} = 21.66±1.98) participants with normal or corrected-to-normal vision were included in the study. Participants were randomly assigned to one of three groups: an Active Wake, a Quiet Rest, or a Nap group. Individuals unable to fall asleep in the Nap group (n = 10) as well as those falling asleep in the awake groups (n = 5) were excluded from the final analyses. Furthermore, 3 additional participants were excluded due to the absence of learning in the training session. Therefore, the final behavioral analyses were based on the data of 78 participants (20 males, M_{age} = 21.71±1.97), with 25, 26, and 27 participants in the Active Wake, Quiet Rest, and Nap group, respectively. In case of the EEG analyses, the data of 12 participants was excluded due to technical artifacts rendering EEG recordings less reliable. Therefore, physiological analyses were restricted to EEG data with sufficient quality (Active

Wake, N = 20; Quiet Rest, N = 21, Nap = 25). All participants provided written informed consent before enrollment and received course credits for taking part in the experiment. The study was approved by the research ethics committee of the Eötvös Loránd University, Budapest, Hungary (201410). The study was conducted in accordance with the Declaration of Helsinki.

Task

Behavioral performance was measured by the explicit version of the Alternating Serial Reaction Time (ASRT) task (Fig. 1, Nemeth et al., 2013). In this task, a stimulus (a dog's head, or a penguin) appeared in one of four horizontally arranged empty circles on the screen, and participants had to press the corresponding button (of a response box) when it occurred. Participants were instructed to respond as fast and accurate as they could. The task was presented in blocks with 85 stimuli. A block started with five random stimuli for practice purposes, followed by an 8-element alternating sequence that was repeated ten times. The alternating sequence was composed of fixed sequence (pattern) and random elements (e.g., 2-R-4-R-3-R-1-R, where each number represents one of the four circles on the screen and "R" represents a randomly selected circle out of the four possible ones). The response to stimulus interval was set to 120 ms (Song et al., 2007a; Nemeth et al., 2010). In the explicit ASRT task participants are informed about the underlying structure of the sequence, and their attention is drawn to the alternation of sequence and random elements by different visual cues. In our case, a dog always corresponded to sequence elements, and a picture of a penguin indicated random elements (Figure 1A). Participants were informed that penguin targets had randomly chosen locations whereas dog targets always followed a predetermined pattern. They were instructed to find the hidden pattern defined by the dog in order to improve their performance. For each participant, one of the six unique permutations of the four possible ASRT sequence

stimuli was selected in a pseudo-random manner, so that the six different sequences were used equally often across participants (Howard and Howard, 1997; Nemeth et al., 2010).

The task consisted of a total of 40 blocks. Participants completed 25 blocks during the *training phase*. This was followed by a short (3 minutes long) break in order to minimize the fatigue effect due to massed practice (Rickard et al., 2008; Rieth et al., 2010). After the break, participants were tested on the task for 5 more blocks that constituted the *testing phase*. Subsequently, participants spent an approximately one-hour long off-line period in one of the three conditions (Active Wake, Quiet Rest, and Nap). Finally, they completed a *retesting phase*: 10 more blocks of the same task. The training phase lasted approximately 30 minutes, the testing phase 5 minutes, and the retesting phase 10 minutes. Awareness of the sequence (pattern elements) was measured after each block. Participants had to type in the regularities they noticed during the task using the same response buttons they used during the ASRT blocks. This method allowed us to determine the duration (in terms of the number of blocks) participants needed to learn the sequence correctly as defined by consistently reporting the same sequence from that point on in the remaining blocks.

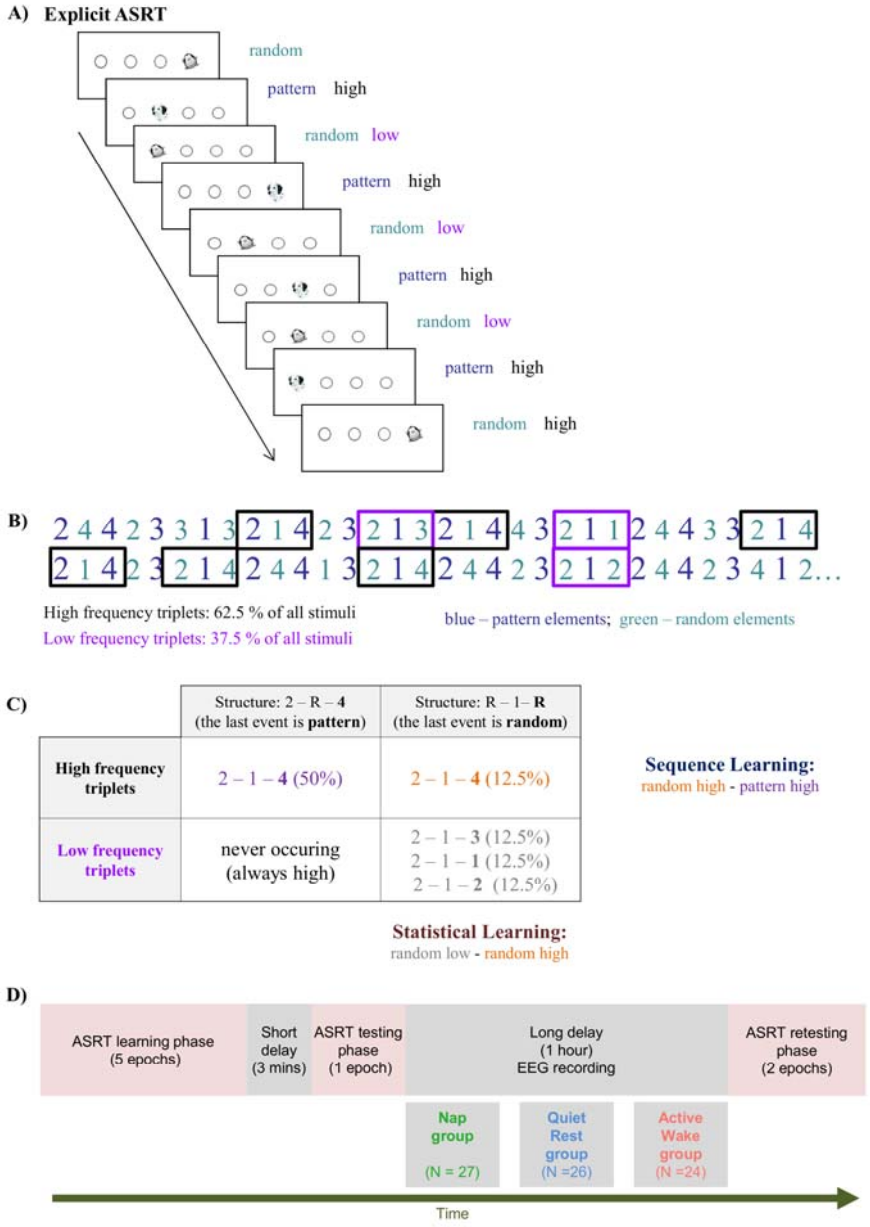

Figure 1. The modified Alternating Serial Reaction Time (ASRT) task. **A)** Pattern and random trials are presented in an alternating fashion. Pattern trials are marked with a picture of a dog, random ones with that of a penguin. Pattern trials always appear in a given location with high probability. Random trials include trials that appear in a given location with high probability and trials that appear in a given location with low probability. **B)** As the ASRT task contains an alternating sequence structure (e.g., 2R4R3R1R, where numbers correspond to the four locations on the screen and the letter R represents randomly chosen locations), some runs of three consecutive elements (called triplets) occur more frequently than others. For subsequent analyses, we determined for each stimulus whether it was the last element of a high-frequency triplet (black frames) or the last element of

a low-frequency triplet (purple frames). **C)** We assessed *Statistical Learning* by comparing the responses for those random elements that were the last elements of a high frequency triplet, opposite to those that were the last of a low frequency triplet. In contrast, *Sequence Learning* was quantified as the difference between responses for pattern elements (which were always high frequency triplets) vs. random-high frequency triplet elements. **D)** Study Design. The training phase consisted of five epochs (25 blocks). The testing and retesting phases comprised one and two (that is, 5 and 10 blocks), respectively.

Trial types and learning indices

The alternating sequence of the ASRT task forms a sequence structure in which some of the runs of three successive elements (henceforth referred to as triplets) appear more frequently than others. In the above example, triplets such as 2X4, 4X3, 3X1, and 1X2 (X indicates the middle element of the triplet) occur frequently since the first and the third elements can either be pattern or random stimuli. However, 3X2 and 4X2 occur less frequently since the first and the third elements can only be random stimuli. Figure 1B and 1C illustrate this phenomenon with the triplet 2-1-4 occurring more often than other triplets such as 2-1-3, 2-1-1, and 2-1-2. The former triplet types are labeled as *high-frequency* triplets whereas the latter types are termed as *low-frequency* triplets (see Figure 1C and Nemeth et al., 2013).

The third element of a high-frequency triplet is highly predictable (with 62.5 % probability) from the first element of the triplet. In contrast, in low-frequency triplets the predictability of the third element is much lower (based on a probability of 12.5 %). According to this principle, each stimulus was categorized as either the third element of a high- or a low-frequency triplet. Moreover, trials are differentiated by the cues (dog and penguin) indicating whether the stimulus belongs to the pattern or the random elements. In case of pattern trials, participants can use their explicit knowledge of the sequence to predict the trial, thus we differentiate high-frequency triplets with the last element being a pattern

from those triplets in which the last one is a random element. This way, the task consists of three trial types: 1) elements that belong to the explicit sequence and at the same time appear as the last element of a high-frequency triplet are called *pattern* trials; 2) random elements that appear as the last element of a high-frequency triplet are called *random high* trials; and 3) random elements that appear as the last element of a low-frequency triplet are termed *random low* trials (see the example in Figure 1C).

To disentangle the two key learning processes underlying performance on the explicit ASRT task, we differentiate *Sequence Learning* and *Statistical Learning* (Figure 1C). *Sequence Learning* is measured by the difference in reaction times (RT) between random high and pattern elements (the average RT for random high elements minus the average RT for pattern elements). These elements share the same statistical properties (both correspond to the third element of high-frequency triplets), but have different sequence properties (i.e., pattern vs. random elements). Thus, greater Sequence Learning is determined as faster responses to pattern in contrast to random high trials. *Statistical Learning* is assessed by comparing the responses for those random elements that were the last elements of a high-frequency triplet, opposite to those that were the last of a low-frequency triplet (the average RT for random low elements minus the average RT for random high elements). These elements share the same sequence properties (both are random) but differ in statistical properties (i.e., they correspond to the third element of a high or a low-frequency triplet). Hence, faster responses to random high compared to random low trials yields greater Statistical Learning. In sum, Sequence Learning quantifies the advantage (in terms of RT) due to the awareness of the sequential pattern, whereas Statistical Learning captures purely frequency-based learning (Nemeth et al., 2013).

Procedure

One to two weeks prior the experiment, participants were invited to the laboratory in order to familiarize them with the environment, and to assess their working memory and executive functions based on the Wisconsin Card Sorting Test (PEBL's Berg Card Sorting Test; Fox et al., 2013) and the Digit Span (Racsmány et al., 2005) and Counting Span (Conway et al., 2005) tasks, respectively. Participants were instructed to complete sleep agendas reporting the schedules, duration and subjective quality of their sleep. At the day of the experiment, participants arrived at the laboratory at 10.00 AM. They completed the GSQS assessing previous nights' sleep quality. Additionally, their subjective stress levels scored on a 10-point Likert scale ("On a scale from 0-10 how stressed are you feeling now?"), as well as the Karolinska Sleepiness Scale (KSS, Akerstedt and Gillberg, 1990) were administered. Subsequently, EEG caps with 64 electrodes were fitted by two assistants. Testing started at 11.30 AM and took place in a quiet room equipped with a large computer screen, a response box and EEG recording device. After listening to the instructions, participants had the opportunity to practice the task in order to get familiar with the stimuli and the response box; however, all stimuli appeared in a random fashion during the practice session. This was followed by the explicit ASRT task composed of the *training phase*, *testing phase*, *off-line period*, and *retesting phase* (Figure 1D). A 3-min long break was inserted between the learning and the testing phases during which the fitting of the EEG caps were monitored and impedances were reset under 10 kΩ. The off-line period extended from 12.30 to 13.30. Participants assigned to the Active Wake group were instructed to watch an approximately one-hour long documentary. (They were allowed to select from documentaries of different topics such as natural sciences, nature or history). Participants of the Quiet Rest group were asked to sit quietly with eyes closed in a comfortable chair. They were instructed by the assistant to open their eyes for 1 minute, every 5 minutes or in case the EEG recording showed any sign of sleep onset (slow eye movements, attenuation of alpha waves and

presence of theta oscillations). Participants in the Nap group had the opportunity to spend a daytime nap in the laboratory. The off-line period took place (in all groups) at the same room in which learning, testing and retesting occurred, and was monitored by EEG. Before the retesting phase, participants were asked to complete again the KSS and the scale assessing the level of stress.

EEG recording

EEG activity was measured by using a 64-channel recording system (BrainAmp amplifier and BrainVision Recorder software, BrainProducts GmbH, Gilching, Germany). The Ag/AgCl sintered ring electrodes were mounted in an electrode cap (EasyCap GmbH, Herrsching, Germany) on the scalp according to the 10% equidistant system. During acquisition, electrodes were referenced to a scalp electrode placed between Fz and Cz electrodes. Horizontal and vertical eye movements were monitored by EOG channels. Three EMG electrodes to record muscle activity, and one ECG electrode to record cardiac activity were placed on the chin and the chest, respectively. All electrode contact impedances were kept below 10 kΩ. EEG data was recorded with a sampling rate of 500 Hz, band pass filtered between (0.3 and 70 Hz).

In order to remove muscle and eye movement related artifact from the awake EEG data (Active Wake and Quiet Rest groups), EEG preprocessing was performed using the Fully Automated Statistical Thresholding for EEG artifact Rejection (FASTER) toolbox (http://sourceforge.net/projects/faster, Nolan et al., 2010) implemented in EEGLAB (Delorme and Makeig, 2004) under Matlab (The Mathworks). The data was first re-referenced to the Fz electrode, notch filtered at 50 Hz, and band-pass filtered between 0.5 – 45 Hz. Using a predefined *z*-score threshold of ±3 for each parameter, artifacts were detected and corrected regarding single channels, epochs, and independent components (based on the infomax

algorithm Bell and Sejnowski, 1995). This way, data was cleared from eye-movement, muscle and heartbeat artifacts. The data was then re-referenced to the average of the mastoid electrodes (M1 and M2). Remaining epochs containing artifacts were removed after visual inspection on a 4-seconds long basis. In case of the sleep recordings (Nap group), data was re-referenced to the average of the mastoid electrodes, and sleep stages as well as conventional parameters of sleep macrostructure were scored according to standardized criteria (Berry et al., 2012) by two experienced sleep researchers. Periods of NREM sleep (Stage 2 and SWS) were considered for subsequent analyses. Epochs containing artifacts were visually inspected and removed on a 4-seconds basis.

Spectral power and sleep spindle analyses of artifact-free segments were performed by a custom made software tool for EEG analysis (FerciosEEGPlus, © Ferenc Gombos 2008-2017). Overlapping (50%), artifact-free, four-second-epochs of all EEG derivations were Hanning-tapered and Fourier transformed by using the FFT (Fast Fourier Transformation) algorithm in order to calculate the average power spectral densities. The analyzed frequencies spanned between 0.75-31 Hz in the Nap group, and between 1.5 - 25 Hz in the awake groups. Low frequencies (0.75-1.5 Hz) were not considered in the awake conditions due to the negligible and unreliable contribution of measurable cortical activity at this frequency range during wakefulness. In addition, frequencies above 25 Hz were unreliable in the awake data due to technical and movement-related artifacts. We summed up frequency bins to generate five frequency bands for the wake groups: delta (1.5-4 Hz), theta (4.25-8), alpha (8.25-13), sigma (13.25-16), and beta (16.25-25 Hz) frequency bands, and five frequency domains for the sleep group: delta (0.75-4 Hz), theta (4.25-8), alpha (8.25-13), sigma (13.25-16), and beta (16.25-31 Hz) frequency ranges. In order to reduce the number of parameters, we averaged bandwise spectral power measures of Frontal (frontal, frontocentral and frontotemporal),

Central (central and centrotemporal), and Posterior (parietal, parietotemporal and occipital) electrode derivations.

We quantified sleep spindling activity by the Individual Adjustment Method (IAM, (Bódizs et al., 2009; Ujma et al., 2015) that considers individual spectral peaks to detect spindles in each participant. This method defines frequency boundaries for slow and fast spindle based on the spectral power of NREM sleep. These individualized boundaries are used as frequency limits for slow and fast spindle bandpass filtering (FFT-based, Gaussian filter, 16 s windows) of the EEGs. Thresholding of the envelopes of the band-pass filtered recordings are performed by individual and derivation-specific amplitude criteria (See the description of the method in more detail in Bódizs et al., 2009; Ujma et al., 2015). We used spindle density (spindles/min) and the average amplitude (μV) of slow and fasts spindles as different measures of spindling activity. To reduce the number of statistical comparisons, we averaged spindle measures of Frontal, Central, and Posterior electrode derivations similarly to spectral power measures.

Statistical analyses

Statistical analyses were carried out with the Statistical Package for the Social Sciences version 22.0 (SPSS, IBM) and R (Team, 2014). The blocks of the explicit ASRT task were collapsed into epochs of five blocks to facilitate data processing and to reduce intra-individual variability. The first epoch contained blocks 1–5, the second epoch contained blocks 6–10, etc. We calculated median reaction times (RTs) for all correct responses, separately for pattern, random high and random low trials for each epoch. Note that for each response (n), we defined whether it was the last element of a high- or a low-frequency triplet. Two kinds of low-frequency triplets were eliminated: repetitions (e.g., 222, 333) and trills (e.g., 212, 343). Repetitions and trills corresponded to low frequency triplets for all

participants and individuals often show pre-existing response tendencies to such triplets (Howard et al., 2004). By eliminating these triplets, we attempted to ensure that differences between high vs. low-frequency triplet elements emerged due to learning and not to pre-existing response tendencies.

To show the performance trajectories of RTs for different trial types, and to explore their differences, we performed a mixed design analyses of variance (ANOVA) with EPOCH (1-8) and TRIAL TYPE (pattern, random high, random low) as within-subject factors, and GROUP (Active Wake, Quiet Rest, Nap) as a between-subject factor. To evaluate the effect of epoch and trial type we performed post-hoc comparisons (Fisher's LSD).

In order to examine the changes in Statistical and Sequence Learning that occur during the training phase, we applied a mixed-design ANOVA with EPOCH (1 -5) and LEARNING TYPE (Statistical Learning, Sequence Learning) as within-subject factors, and GROUP (Active Wake, Quiet Rest and Nap) as a between-subject factor. Post-hoc comparisons were applied to evaluate changes in performance during the training phase in case of Sequence and Statistical Learning.

To examine offline changes occurring between testing and retesting sessions we used a similar mixed-design ANOVA with EPOCH (6-8) and LEARNING TYPE (Statistical Learning, Sequence earning) as within-subject factors, and GROUP (Active Wake, Quiet Rest and Nap) as a between-subject factor. Post-hoc comparisons were run to contrast performances of the testing phase (6^{th} epoch) and the retesting phases (7^{th} and 8^{th} epochs).

Greenhouse-Geisser epsilon (ε) correction was used if necessary. Original *df* values and corrected *p* values (if applicable) are reported together with partial eta-squared (η_p^2) as a measure of effect size.

Finally, we aimed to examine the associations between EEG spectral power measured during the off-line period and change in learning performance across the testing and retesting

phase, in each group separately. Off-line changes in Sequence and Statistical Learning were defined as the difference between the learning scores of the first retesting (7th epoch) session and the testing session (6th epoch). Thus, a positive value indicated improvement in learning performance after the off-line period. Furthermore, we aimed to examine whether EEG spectral power measured during off-line periods predicted additional performance change after longer re-learning, therefore, we calculated a secondary off-line change score contrasting learning scores of the 8th (2nd half of the retesting session) with those of the 6th epoch (testing session).

The associations between sleep spindles and off-line changes of the above measures were also examined (within the sleep group only). Pearson correlation coefficients or (if normality was violated) Spearman rank correlations were run between spectral power values (of each region and band) and off-line changes in learning scores. The issue of multiple comparisons was addressed by the False Discovery Rate correcting for type 1 error (Benjamini and Hochberg, 1995).

Results

Group characteristics

Groups were matched in age, gender, working memory, executive function, and initial sleepiness and stress level (Table 1). However, after the one hour long offline period, the groups differed in sleepiness ($F_{2,75}$ = 3.19, p = 0.05), post-hoc test showed that the Nap group scored significantly higher on the KSS (indicating lower sleepiness) than the Active Wake group (p = 0.02), however the difference was not significant after FDR correction.

Table 1. Descriptive characteristics of groups

	Active Wake	Quiet Rest	Nap group

Variable	group Mean (SD)	group Mean (SD)	Mean (SD)	p value
Age (years)	22.08 (2.04)	22.00 (1.94)	21.15 (1.83)	p = 0.16
Gender (male, %)	28%	22%	27%	p = 0.88
GSQS	1.96 (1.72)	2.31(2.13)	2.33 (1.96)	p = 0.75
Stress scale 1	2.65 (2.09)	2.55 (1.43)	3.33 (1.98)	p = 0.35
Stress scale 2	2.59 (1.28)	2.00 (1.33)	1.77 (1.41)	p = 0.17
KSS 1	6.44 (1.26)	6.81 (1.13)	6.19 (1.52)	p = 0.24
KSS 2	5.64 (1.19)	5.96 (1.70)	6.62 (1.30)	p = 0.05
Digit span	6.32 (1.31)	5.88 (1.14)	6.26 (1.06)	p = 0.36
Counting span	3.91 (1.50)	3.59 (0.72)	3.48 (0.81)	p = 0.33
WCST – number of perseverative errors	15.67 (9.23)	14.31 (3.23)	13.19(5.86)	p = 0.40

Note GSQS – Groningen Sleep Quality Scale, KSS - Karolinska Sleepiness Scale, WCST - Wisconsin Card Sorting Test

Sleep parameters of the Nap group are listed in Table 2. In the Nap group, only 1 participant reached REM phase during sleep, thus we only report the characteristics of Non-REM sleep.

Table 2. Descriptive characteristics of sleep parameters in the Nap group

Variable	Mean (SD)
Sleep duration (min)	41.16 (12.35)
Sleep efficiency (%)	70.28 (16.27)

Wake duration (min)	16.53 (7.77)
S1 duration (min)	6.02 (3.62)
S2 duration (min)	17.93 (6.59)
SWS duration (min)	16.89 (12.82)

Note S1 – Stage 1, S2 – Stage 2, SWS – Slow Wave Sleep

Awareness of the sequence in the groups

By the end of the task all but 4 participants reported the correct sequence. The RTs and learning indices of these 4 participants were within 1 standard deviation from the mean, thus their learning did not differ from those who gained explicit knowledge of the sequence. There were no differences across the groups in the time needed for gaining explicit knowledge about the sequence during the experiment defined as the number of the block in which the participant could report the sequence structure ($F_{2,73} = 1.68$, $p = 0.19$). On average, participants gained explicit knowledge of the sequence after the 3rd block ($M = 3.46$, $SD = 4.44$). 87% of participants reported the sequence during the first epoch (first 5 blocks), and 95% reported it in the first 2 epochs (first ten blocks).

Are performance trajectories of responses to different trial types different between groups?

Overall, participants in the different groups responded with similar RTs (main effect of GROUP: $F_{2,75} = 0.80$, $p = 0.46$, partial $\eta^2 = 0.02$). Irrespectively of trial types, RTs significantly decreased across epochs (main effect of EPOCH: $F_{7,525} = 175.26$, $p < 0.0001$, partial $\eta^2 = 0.70$), indicating general skill improvements due to practice (Figure 2). The GROUP x EPOCH interaction was not significant ($F_{14,525} = 1.18$ $p = 0.32$, partial $\eta^2 = 0.03$), suggesting that general skill improvements were similar in the groups. Furthermore, participants showed significant Sequence and Statistical Learning (main effect of TRIAL TYPE: $F_{2,150} = 52.04$, $p < 0.0001$, partial $\eta^2 = 0.41$): they responded faster to pattern than

random high trials ($p < 0.0001$), and faster to random high compared to random low trials ($p < 0.0001$). The GROUP x TRIAL TYPE interaction was not significant ($F_{4,150} = 0.80$, $p = 0.46$, partial $\eta^2 = 0.02$) indicating that there was no difference between the groups in performance for different trial types. In addition to that, the EPOCH x TRIAL TYPE interaction was significant ($F_{14,1050} = 11.93$, $p < 0.0001$, partial $\eta^2 = 0.14$), indicating different learning trajectories in case of the three trial types (see Figure 2). Although participants became faster for all trial types during the course of the task, responses to pattern trials showed greater gains in comparison to both random trials: Average reaction times of pattern trials decreased from 357.89 to 257.56 ms ($p < 0.0001$), of random high trials from 370.98 to 326.14 ms ($p < 0.0001$), and of random low trials from 388.26 to 349.65 ms ($p < 0.0001$). Practice-dependent improvement in response to pattern trials was significantly higher than the improvement in case of random high ($t_{77} = 4.81$, $p < 0.0001$) and random low ($t_{77} = 5.45$, $p < 0.0001$) trials. The improvement in responses to random high and random low trials was only marginally different ($t_{77} = 1.84$, $p = 0.07$). The GROUP x EPOCH x TRIAL TYPE interaction was not significant ($F_{28,1050} = 0.66$, $p = 0.68$, partial $\eta^2 = 0.02$), suggesting that performance trajectories to the different trial types were similar among the groups.

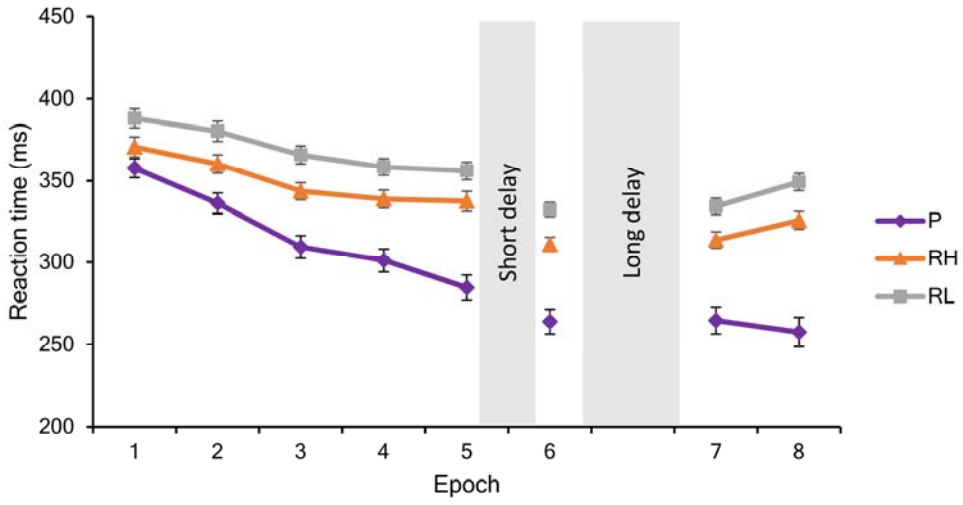

Figure 2. Performance during the training, testing and retesting sessions. Mean reaction times and standard errors are visualized in response to pattern (P), random high (RH), and random low (RH) trials during each epoch.

Do Sequence and Statistical Learning during training differ between groups?

Sequence and Statistical Learning during the training phase were similar across the groups (main effect of GROUP: $F_{2,75} = 1.10$, $p = 0.34$, partial $\eta^2 = 0.03$). Irrespectively of learning type, performance improved across epochs of training (main effect of EPOCH: $F_{4,300} = 10.92$, $p < 0.0001$, partial $\eta^2 = 0.13$). The GROUP x EPOCH interaction was not significant ($F_{8,300} = 0.59$, $p = 0.68$, partial $\eta^2 = 0.02$), suggesting that improvement during training was similar between the groups. In addition, the main effect of LEARNING TYPE was significant ($F_{1,75} = 3.93$, $p = 0.05$, partial $\eta^2 = 0.05$): participants showed greater Sequence Learning compared to Statistical Learning ($M = 32.50$ vs. $M = 19.64$, $p < 0.0001$). The GROUP x LEARNING TYPE interaction was not significant ($F_{2,75} = 0.81$, $p = 0.45$, partial $\eta^2 = 0.02$), suggesting that the difference between Sequence and Statistical Learning were similar among the groups. Furthermore, a significant interaction between EPOCH and LEARNING TYPE emerged ($F_{4,300} = 5.52$, $p = 0.002$, partial $\eta^2 = 0.07$): as illustrated in Figure 3, participants exhibited a steep increase in Sequence Learning during the training phase (the average learning score increased from 13.09 to 53.31 from the 1st epoch to the 5th, ($p < 0.001$), whereas Statistical learning occurred in the beginning of the task and remained unchanged by the end of the training phase (the average learning score increased from 17.28 to 18.64 from the 1st epoch to the 5th, $p = 0.68$). The GROUP x EPOCH x LEARNING TYPE interaction was not significant ($F_{8,300} = 0.58$, $p = 0.72$, partial $\eta^2 = 0.02$), suggesting that training-dependent patterns of Sequence Learning and Statistical Learning were similar across the groups.

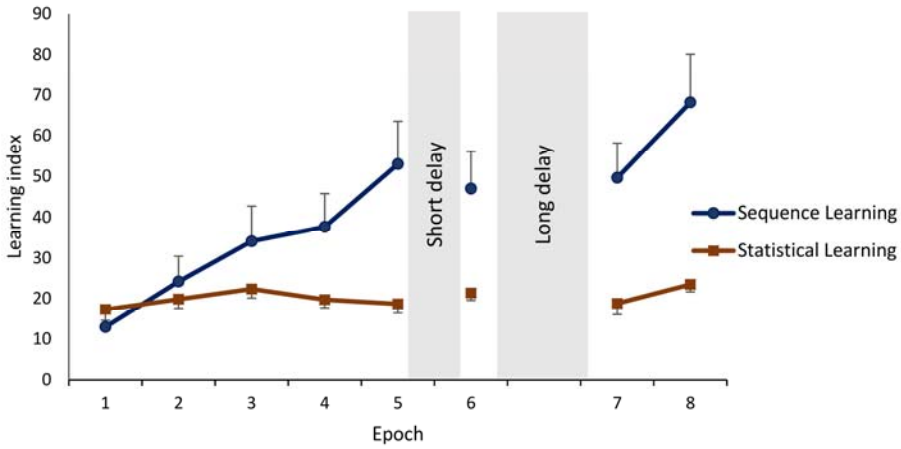

Figure 3. Learning and off-line changes in Sequence and Statistical Learning. Means and standard errors of Sequence Learning and Statistical Learning during each epoch. Sequence Learning exhibited a steep increase during training and additional practice after the off-line periods, whereas Statistical Learning remained unchanged throughout the sessions.

Are off-line changes in Sequence and Statistical Learning different across the groups?

The three groups did not show different patterns of Sequence and Statistical Learning from the testing to the retesting sessions, as neither the main effect of GROUP ($F_{2,75} = 0.65$, $p = 0.53$, partial $\eta^2 = 0.02$), nor the interactions GROUP x EPOCH ($F_{4,150} = 0.52$, $p = 0.67$, partial $\eta^2 = 0.01$), GROUP x LEARNING TYPE ($F_{2,75} = 0.65$, $p = 0.53$, partial $\eta^2 = 0.02$), and GROUP x EPOCH x LEARNING TYPE ($F_{4,150} = 0.73$, $p = 0.55$, partial $\eta^2 = 0.02$) emerged as significant predictors. The lack of a group effect is shown in Figure 4 that illustrates off-line changes (6^{th} vs the 7^{th} epoch) in Sequence and Statistical Learning separately for each group. Similarly to the training phase, participants exhibited higher scores in Sequence Learning than in Statistical Learning (main effect of LEARNING TYPE: $F_{1,75} = 10.72$, $p = 0.002$, partial $\eta^2 = 0.13$). Moreover, learning indices produced robust changes across epochs as indicated by a significant main effect EPOCH ($F_{2,150} = 18.99$, $p < 0.0001$, partial $\eta^2 = 0.20$). More specifically, overall performances (regardless of learning type) were unchanged from the

testing phase (6^{th} epoch) to the first retesting epoch (7^{th}) ($p = 0.86$), but improved ($p < 0.0001$) from the testing phase to the end of the retesting session (8^{th} epoch), and from the first retesting epoch to the second (7^{th} epoch vs 8^{th} epoch) ($p < 0.0001$). Furthermore, Sequence Learning and Statistical Learning scores showed different patterns after the off-line period (see Epoch 7 and 8 in Figure 3), as indicated by the significant EPOCH x LEARNING TYPE interaction ($F_{2,150} = 5.31$, $p = 0.009$, partial $\eta^2 = 0.07$). Neither Sequence Learning nor Statistical Learning seemed to show immediate (early) gains after the off-line period. Sequence Learning scores did not significantly change from the testing phase to the first epoch of retesting (6^{th} epoch, $M = 47.02$ vs. 7^{th} epoch, $M = 47.69$, $p = 0.85$). Similarly, Statistical Learning remained unchanged from testing to the first retesting (6^{th} epoch, $M = 21.39$ vs. 7^{th} epoch, $M=19.96$, $p = 0.56$). Nevertheless, additional practice produced robust changes in Sequence Learning, that increased significantly from the testing phase to the second epoch of the retesting phase (8^{th} epoch, $M = 68.19$, $p = 0.001$), whereas Statistical Learning did not show any significant changes by the end of the retesting phase (8^{th} epoch: $M = 23.51$, $p = 0.41$).

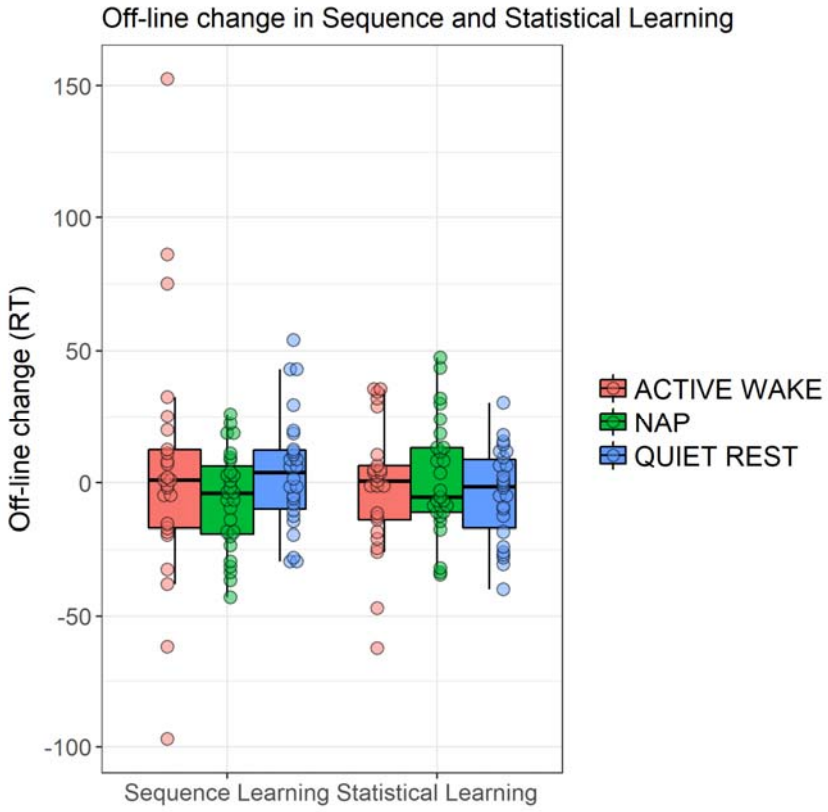

Figure 4. Off-line changes in learning indices within the three groups. Dots show individual data points, the vertical line within the boxes show the medians, boxes represent the first and third quartiles, whiskers indicate the interquartile range of 1.5.

Associations between EEG spectra and off-line changes

Off-line changes in Sequence and Statistical Learning as indexed by the difference scores between the 7th (first half of retesting phase) and the 6th epochs' (testing phase) scores were not associated with spectral EEG power measures in any of the three groups. Additional off-line-changes in Sequence Learning as indexed by the difference scores between the 8th (second half of retesting phase) and the 6th epochs' (testing phase), however, showed a positive association with frontal theta power ($r = 0.52$, $p= 0.008$) within the nap group. Nevertheless, the correlation did not reach statistical significance after FDR correction of multiple comparisons. Since region-wise averaging of electrodes might not capture

associations between behavioral measures and spectral power of a more local nature, we examined (on an exploratory level) the associations between theta activity and off-line change (8^{th} vs 6^{th} epoch) in Sequence Learning within the nap group. As shown in Figure 5**,** associations with theta band power were prominent at frontal electrode sites, peaking at left frontopolar locations. Finally, we examined the associations between off-line (8^{th} vs 6^{th} epoch) change in Sequence Learning and bin-wise EEG spectral power averaged across all electrodes (within the Nap group). Post-sleep improvement in Sequence Learning correlated only with slow frequency activity between 2-7.75 Hz (all bins $p < 0.01$).

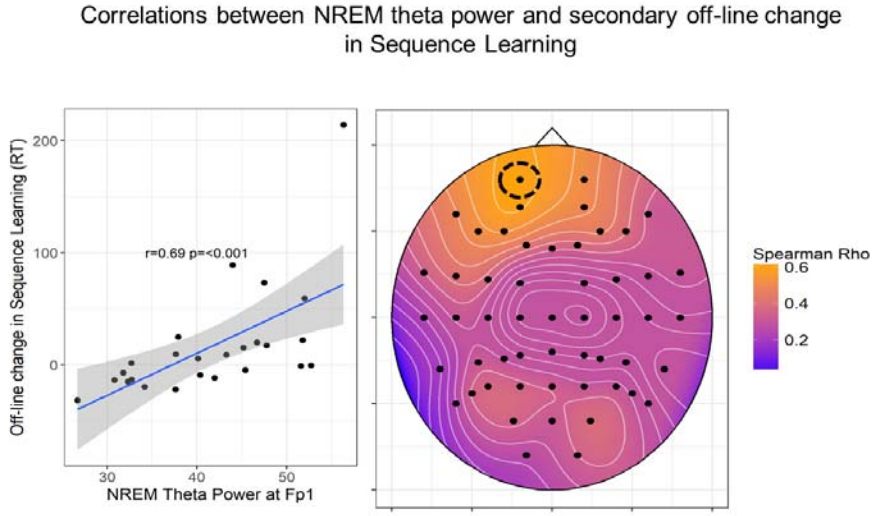

Figure 5. Association between NREM theta power and further gains (8^{th} vs. 6^{th} epoch) in off-line change in case of Sequence Learning. The heat plot on the right indicates the magnitude of Spearman Rho correlation coefficients, the scatterplot on the left shows the association in a prominent (left frontal) electrode site. Correlation remained significant ($r = 0.64$, $p < 0.001$) after the exclusion of the outlier.

Off-line change (8^{th} vs 6^{th} epoch) in Statistical Learning was not associated with spectral power measures within the nap group, and no other associations emerged within the Quiet Rest and Active Wake groups.

In sum, individual differences in off-line changes in Sequence and Statistical Learning assessed immediately after the long delay (6^{th} vs. 7^{th} epoch) were not associated with spectral EEG power. After extended practice, subsequent (6^{th} vs. 8^{th} epoch) changes in Statistical Learning were not associated with spectral EEG power measures, but further improvements in Sequence Learning were predicted by high delta and theta activity during sleep within the Nap group.

Associations between sleep spindles and off-line changes

Off-line change (7^{th} vs 6^{th} epoch) in Sequence Learning showed a negative correlation with slow spindle density in Frontal ($r = -0.52$, $p = 0.008$), Central ($r = -0.54$, $p = 0.006$) and Posterior ($r = -0.53$, $p = 0.006$) derivations. Slow spindle amplitude, fast spindle density and amplitude were not associated with the off-line change in Sequence Learning. Negative correlations between slow spindle density and off-line change in Sequence Learning remained significant after FDR correction. As Figure 6A indicates negative correlations of similar magnitude ($r \approx -0.35 – -0.6$) emerged in all electrode sites, except right temporal regions.

Off-line change in Statistical Learning was negatively correlated with fast spindle amplitude (Frontal: $r = -0.43$, $p = 0.03$; Central: $r = -0.47$, $p = 0.02$; Posterior: $r = -0.44$, $p = 0.03$), but was not related either to fast spindle density or slow spindle density/amplitude. Correlations between fast spindle amplitude and off-line change in Statistical Learning were not significant after FDR correction. Regarding the topographical distribution of the correlations between fast spindle amplitude and off-line change in Statistical Learning, coefficients showed similar magnitudes ($r \approx -0.3 – -0.56$), and peaked at left centro-posterior electrode sites (Figure 6B).

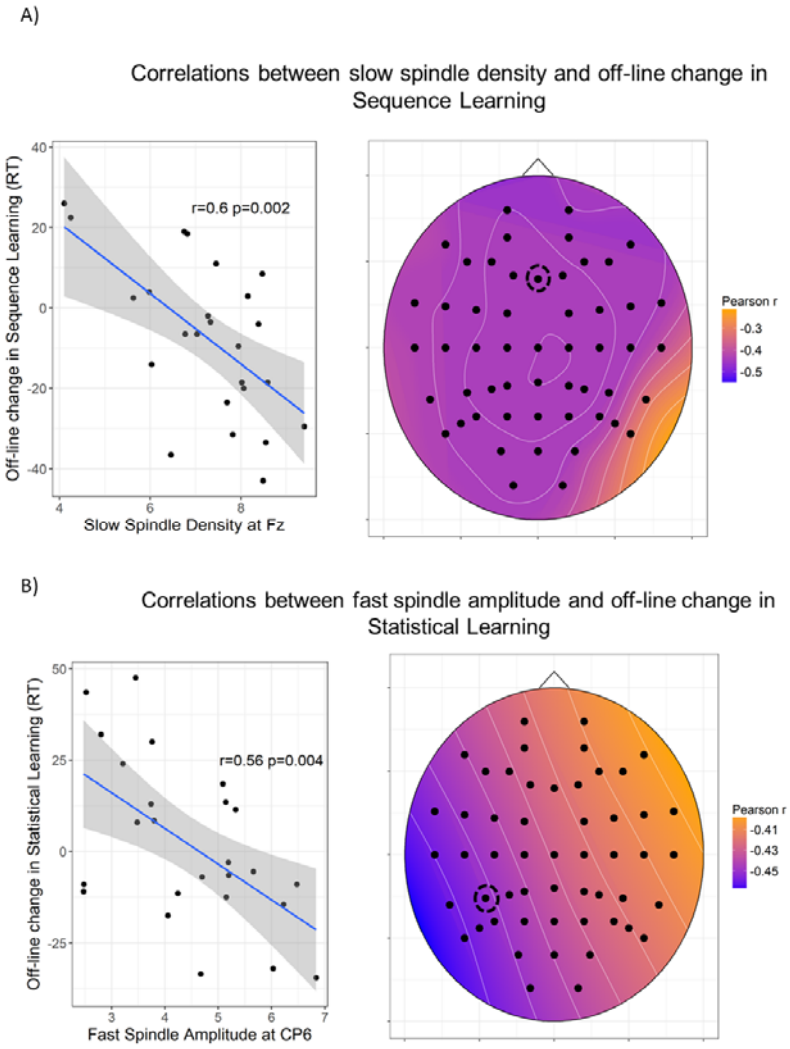

Figure 6. Associations between spindle parameters and off-line changes in learning indices. Scatterplots of prominent electrode sites and heat plots visualizing the magnitude of correlation coefficients are shown. Slow spindle density and fast spindle amplitude were negative correlates of off-line changes in Sequence Learning **(A)** and Statistical Learning **(B)**, respectively.

To examine whether the negative correlation between off-line changes in performance and spindle parameters were linked to overall Sequence/Statistical Learning ability, we applied partial correlations with learning performance of the training phase as a covariate. Learning performance here was computed as the differences in Sequence and Statistical learning between the 5th and the 1th epochs of the training phase. Slow spindle density

remained a negative correlate of off-line change in Sequence Learning even after controlling for this initial Sequence Learning performance (Frontal: r = -0.5, p = 0.02; Central: r = -0.52, p = 0.009; Posterior: r = -0.51, p = 0.01).

Similarly, partial correlations were computed between fast spindle amplitude and off-line change in Statistical Learning with Statistical Learning performance as a covariate. The correlations remained significant even after partialing out this initial Statistical Learning performance (Frontal: r = -0.37, p = 0.07; Central: r = -0.43, p = 0.03; Posterior: r = -0.36, p = 0.08).

Additional off-line-changes in Sequence and Statistical Learning as indexed by the difference scores between the 8th (second half of retesting phase) and the 6th epochs' (testing phase) were not associated to any of the extracted spindle parameters.

Discussion

Our aim was to investigate performance trajectories in Sequence and Statistical Learning during extensive practice and after off-line periods spent in different vigilance states. In order to examine these processes in the same experimental context, we applied a paradigm that simultaneously measured sequence and statistical Learning by delineating order and frequency-based information. Our findings indicate that Sequence and Statistical Learning follow markedly different learning curves. Whereas performance in Sequence Learning exhibited a gradual increase during training, Statistical Learning was rapidly acquired but remained unchanged throughout training. After the off-line period, both forms of learning were preserved as no significant changes emerged in either Sequence or Statistical Learning. Nevertheless, Sequence Learning improved after additional practice, whereas Statistical Learning remained stable regardless of further training compared to the testing phase. Performance trajectories were similar across the groups: Performance during training

and after off-line periods did not differ between the Active Wake, Quiet Rest and Nap groups. EEG spectral power assessed during the off-line periods was not associated with off-line changes in Sequence and Statistical Learning in either group. Within the Nap group, slow spindle density was negatively associated with post-sleep improvement in Sequence Learning, and fast spindle amplitude was negatively associated with post-sleep improvement in Statistical Learning. Furthermore, within the Nap group, slow frequency oscillations (high delta and theta power) predicted further improvements in Sequence Learning after additional practice.

From a theoretical perspective it is important to note that sequence learning can also be viewed as a type of statistical learning (for overview see Thiessen et al., 2013, Peruchet and Pacton 2006). For example, in our experimental design sequence learning constitutes of learning regularities where second order transitional probability is one; thus, the next sequence element is predictable based on the previous sequence element with 100% certainty, meaning that all sequence elements will always appear in the same *order* over the entire course of the task. This is in contrast to statistical learning where the second order transitional probability is less than one: namely, high frequency triplets can be predicted with approximately 62.5% certainty compared to the 12.5% of low frequency triplets, meaning that these triplets (i.e., a given high frequency triplet after a given low frequency triplet, etc.) do not always occur in the same order during the task. Importantly, sequence learning and statistical learning as defined in our study show different developmental trajectories (Nemeth et al., 2013) and different neurophysiological background (Kóbor et al., under revision). Our results also indicate that sequence and statistical learning are markedly different sub-processes of procedural learning. The acquisition of order-based information seems to be a gradual process, and shows a steep improvement during training. In contrast, frequency-based information is acquired rapidly and remains stable irrespective of further training. This

finding corroborates earlier data (Nemeth et al., 2013) that showed different developmental trajectories of sequence and statistical learning between 11 and 40 years of age but did not analyze the time course of these learning types. Here, we extended this previous study by investigating the time course of learning from the acquisition period to the post-consolidation period with inserting a napping, a quiet rest or a wakeful, alert vigilance condition in-between. The levels that participants achieved in Sequence and Statistical Learning after extensive practice were unchanged following the one-hour long off-line period, regardless of state of vigilance.

We had a special focus on the off-line change and the effect of sleep on Sequence Learning and Statistical Learning. In order to differentiate between the specific effects of sleep and from the indirect effect of reduced interference during off-line periods, we included a quiet rest control group into the design. On the behavioral level, we found no sleep-dependent consolidation neither in Sequence Learning nor in Statistical Learning. The lack of evidence for the beneficial influence of sleep on statistical learning is in line with previous studies that used probabilistic sequence learning tasks (Peigneux et al., 2003, 2006; Song et al., 2007a; Nemeth et al., 2010; Hallgató et al., 2013), however, we should note that these studies did not differentiate between order-based and frequency-based learning mechanisms. Here, we aimed to investigate the influence of sleep on pure (frequency-based) statistical learning in the perceptual-motor domain. Other studies examined sleep-dependent consolidation on statistical learning in the auditory domain (Durrant et al., 2011, 2013) and contrary to our results, found improved performance after sleep compared to wakefulness. Discrepancies between these studies and our findings might stem from methodological differences (overnight sleep and longer daytime naps in Durrant and colleagues' study) as well as the examined modality (auditory system vs. perceptual-motor system). Nevertheless, it

is important to highlight that Durrant and colleagues (Durrant et al., 2011) did not include a quiet rest condition that might be favorable in napping studies.

Interestingly, and contrary to our expectations sleep did not facilitate off-line improvement in Sequence Learning either. In case of perceptual-motor sequence learning, Robertson and colleagues (Robertson et al., 2004) reported sleep-dependent consolidation in the explicit version of the Serial Reaction Time task using deterministic sequences. Discrepant findings between the present and Robertson and colleagues' study can be the result of different sequence structures applied in the SRT and ASRT task. In addition, other confounding factors, such as the effects of fatigue or reactive inhibition (Török et al., 2017) might have a different impact on these tasks. For instance, effects of fatigue are typical to occur in learning tasks (Rickard et al., 2008; Brawn et al., 2010; Pan and Rickard, 2015), however, ASRT learning scores seem to be relatively immune against the influence of fatigue (Török et al., 2017). Furthermore, recent studies raised concerns about the reliability of the deterministic SRT task (Stark-Inbar et al., 2017; West et al., 2017) while the ASRT proved to be a more reliable measure of sequence learning (Stark-Inbar et al., 2017).

Performance in Sequence and Statistical Learning did not show off-line improvements immediately after the long delay period; however, performance in Sequence Learning exhibited further gains after additional practice. Interestingly, the extent of further, training-dependent improvement was associated with slow oscillatory activity within the Nap group. This finding indicates that not sleep *per se*, but sleep-specific oscillations are associated with late performance gains after sleep and additional practice. This finding corroborates a recent study (Maier et al., 2017) showing that daytime sleep does not contribute to immediate improvements after sleep in a finger tapping task, but facilitates subsequent gains after additional training. Here, we extend these considerations by suggesting that slower oscillatory activity including the (high) delta and the theta frequency ranges (from 2 to 7.75 Hz) during

daytime sleep are specifically predictive of further, (late) improvements in Sequence Learning. Slow frequency oscillations peaking at anterior locations and spanning between 1 Hz to 8 Hz reflect the homeostatic and restorative capacity of sleep as power in these frequencies is increased after prolonged wakefulness (Borbély et al., 1981; Marzano et al., 2010) in fronto-central derivations and predict improved performance in cognitive tasks (Mander et al., 2010). Furthermore, the homeostatic increase in spectral power between 2-7 Hz is state-independent (Marzano et al., 2010) making these oscillations likely candidates to reflect restorative processes during a daytime nap, with lower homeostatic pressure. Whether the association between slow frequency activity and further improvement in Sequence Learning reflects processes of sleep-related memory consolidation, or a non-specific effect of restorative sleep facilitating performance remains a question of further research.

Sleep spindle parameters within the Nap group were negatively associated with off-line changes in performance: slow spindle density and fast spindle amplitude showed negative associations with early off-line changes in Sequence Learning and Statistical Learning, respectively. These findings are hard to interpret as they are at odds with the majority of previous findings that reported a positive association between spindle parameters, general cognitive abilities, and off-line gains in performance in a variety of declarative and procedural learning tasks (see Rasch and Born, 2013 for a comprehensive review). Still, negative correlations were also reported to some extent although in samples including children (Chatburn et al., 2013), and psychiatric patients (Nishida et al., 2016). In our study, associations between spindle parameters and off-line changes in performance might not simply stem from trait-like effects, as associations were unchanged if we controlled for the confounding effects of training-dependent learning performance. Nevertheless, given the lack of baseline EEG measurements, we cannot fully discern trait- and state-like effects in the present study. Moreover, only the association between slow spindle density and the off-line

change in Sequence Learning remained significant after the correction for multiple comparisons, whereas previous studies mainly linked sleep-dependent cognitive benefits to fast spindle activity. In sum, off-line changes in Sequence Learning and Statistical Learning were associated with different spindle parameters, nevertheless, the relevance of these associations should be examined in further studies, including baseline sleep measurement without pre-sleep learning experience.

To conclude, here we were able to assess the time-course of two fundamental learning processes, namely Sequence Learning and Statistical Learning separately and showed that Statistical Learning is acquired rapidly and remains unchanged even after extended practice, whereas Sequence Learning is gradually developed. On the behavioral level, both sequence and statistical knowledge were retained and were independent of whether the offline period included sleep or not. On the neural level, however, measures of cortical oscillations (spindling and slow frequency power) assessed during the off-line period were associated with individual differences in performance gains within the sleep group only. Moreover, off-line changes in Sequence Learning and Statistical Learning showed different neural correlates. These findings suggest that sleep has not an all-in-one-effect on memory consolidation, and future studies should focus on mapping systematically which learning and memory mechanisms might and might not benefit from sleep and sleep-specific oscillations. Learning and memory should be assessed on a process level (such as Sequence Learning and Statistical Learning in the current study) in order to characterize the time-course of these processes on the behavioral level as well as their neural correlates more precisely.

Acknowledgements

This research was supported by the Research and Technology Innovation Fund, Hungarian Brain Research Program (KTIA NAP 13-2-2015-0002); Hungarian Scientific Research Fund

(NKFI PD 124148, PI: K.J., NKFI PD 115432, PI: P.S.); and Janos Bolyai Research Fellowship of the Hungarian Academy of Sciences (to K. J. and P.S.). The authors declare no competing financial interests.

References

\AAkerstedt T, Gillberg M (1990) Subjective and objective sleepiness in the active individual. Int J Neurosci 52:29–37.

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.

Armstrong, B. C., Frost, R., & Christiansen, M. H. (2017). The long road of statistical learning research: past, present and future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711).

Bell AJ, Sejnowski TJ (1995) An information-maximization approach to blind separation and blind deconvolution. Neural Comput 7:1129–1159.

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol 57:289–300.

Berry RB, Brooks R, Gamaldo CE, Harding SM, Marcus CL, Vaughn BV (2012) The AASM manual for the scoring of sleep and associated events. Rules Terminol Tech Specif Darien Ill Am Acad Sleep Med Available at: http://www.aasmnet.org/resources/pdf/scoring-manual-preface.pdf [Accessed May 15, 2017].

Bódizs R, Körmendi J, Rigó P, Lázár AS (2009) The individual adjustment method of sleep spindle analysis: methodological improvements and roots in the fingerprint paradigm. J Neurosci Methods 178:205–213.

Borbély AA, Baumann F, Brandeis D, Strauch I, Lehmann D (1981) Sleep deprivation: effect on sleep stages and EEG power density in man. Electroencephalogr Clin Neurophysiol 51:483–495.

Brashers-Krug T, Shadmehr R, Bizzi E (1996) Consolidation in human motor memory. Nature 382:252–255.

Brawn TP, Fenn KM, Nusbaum HC, Margoliash D (2010) Consolidating the effects of waking and sleep on motor-sequence learning. J Neurosci Off J Soc Neurosci 30:13977–13982.

Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ (1989) The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 28:193–213.

Chatburn A, Coussens S, Lushington K, Kennedy D, Baumert M, Kohler M (2013) Sleep Spindle Activity and Cognitive Performance in Healthy Children. Sleep 36:237–243.

Conway ARA, Kane MJ, Bunting MF, Hambrick DZ, Wilhelm O, Engle RW (2005) Working memory span tasks: A methodological review and user's guide. Psychon Bull Rev 12:769–786.

Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21.

Doyon J, Korman M, Morin A, Dostie V, Hadj Tahar A, Benali H, Karni A, Ungerleider LG, 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 195:15–26.

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 N Y N 1991 23:2467–2478.

Durrant SJ, Taylor C, Cairney S, Lewis PA (2011) Sleep-dependent consolidation of statistical learning. Neuropsychologia 49:1322–1331.

Fletcher PC, Zafiris O, Frith CD, Honey R a. E, Corlett PR, Zilles K, Fink GR (2005) On the Benefits of not Trying: Brain Activity and Connectivity Reflecting the Interactions of Explicit and Implicit Sequence Learning. Cereb Cortex 15:1002–1015.

Fox CJ, Mueller ST, Gray HM, Raber J, Piper BJ (2013) Evaluation of a Short-Form of the Berg Card Sorting Test. PLOS ONE 8:e63885.

Hallgató E, Győri-Dani D, Pekár J, Janacsek K, Nemeth D (2013) The differential consolidation of perceptual and motor learning in skill acquisition. Cortex J Devoted Study Nerv Syst Behav 49:1073–1081.

Howard DV, Howard Jr JH, Japikse K, DiYanni C, Thompson A, Somberg R (2004) Implicit sequence learning: effects of level of structure, adult age, and extended practice. Psychol Aging 19:79.

Howard JH, Howard DV (1997) Age differences in implicit learning of higher order dependencies in serial patterns. Psychol Aging 12:634–656.

Karni A, Meyer G, Rey-Hipolito C, Jezzard P, Adams MM, Turner R, Ungerleider LG (1998) The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. Proc Natl Acad Sci U S A 95:861–868.

King BR, Hoedlmoser K, Hirschauer F, Dolfen N, Albouy G (2017) Sleeping on the motor engram: The multifaceted nature of sleep-related motor memory consolidation. Neurosci Biobehav Rev 80:1–22.

Kóbor, A., Takács, A., Kardos, Z., Janacsek, K., Horváth, K., Csépe, V., & Nemeth, D. (under revision). ERPs differentiate the sensitivity to statistical probabilities and the learning of sequential structures during procedural learning.

Korman M, Raz N, Flash T, Karni A (2003) Multiple shifts in the representation of a motor sequence during the acquisition of skilled performance. Proc Natl Acad Sci U S A 100:12492–12497.

Krakauer JW, Shadmehr R (2006) Consolidation of motor memory. Trends Neurosci 29:58–64.

Maier JG, Piosczyk H, Holz J, Landmann N, Deschler C, Frase L, Kuhn M, Klöppel S, Spiegelhalder K, Sterr A, Riemann D, Feige B, Voderholzer U, Nissen C (2017) Brief periods of NREM sleep do not promote early offline gains but subsequent on-task performance in motor skill learning. Neurobiol Learn Mem 145:18–27.

Mander BA, Reid KJ, Baron KG, Tjoa T, Parrish TB, Paller KA, Gitelman DR, Zee PC (2010) EEG measures index neural and cognitive recovery from sleep deprivation. J Neurosci Off J Soc Neurosci 30:2686–2693.

Maquet P, Laureys S, Peigneux P, Fuchs S, Petiau C, Phillips C, Aerts J, Del Fiore G, Degueldre C, Meulemans T, Luxen A, Franck G, Van Der Linden M, Smith C, Cleeremans A (2000) Experience-dependent changes in cerebral activation during human REM sleep. Nat Neurosci 3:831–836.

Marzano C, Ferrara M, Curcio G, De Gennaro L (2010) The effects of sleep deprivation in humans: topographical electroencephalogram changes in non-rapid eye movement (NREM) sleep versus REM sleep. J Sleep Res 19:260–268.

McGaugh JL (2000) Memory--a century of consolidation. Science 287:248–251.

Mednick SC, Cai DJ, Shuman T, Anagnostaras S, Wixted J (2011) An opportunistic theory of cellular and systems consolidation. Trends Neurosci 34:504–514.

Nemeth D, Janacsek K, Fiser J (2013) Age-dependent and coordinated shift in performance between implicit and explicit skill learning. Front Comput Neurosci 7:147.

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.

Nishida M, Nakashima Y, Nishikawa T (2016) Slow sleep spindle and procedural memory consolidation in patients with major depressive disorder. Nat Sci Sleep 8:63–72.

Nishida M, Walker MP (2007) Daytime naps, motor memory consolidation and regionally specific sleep spindles. PloS One 2:e341.

Nolan H, Whelan R, Reilly RB (2010) FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection. J Neurosci Methods 192:152–162.

Novák M (2004) Alvászavarok és életmin\Hoség. Mentális Egészségtudományok Magatartástudomány Program.

Pan SC, Rickard TC (2015) Sleep and motor learning: Is there room for consolidation? Psychol Bull 141:812–834.

Peigneux P, Laureys S, Fuchs S, Destrebecqz A, Collette F, Delbeuck X, Phillips C, Aerts J, Del Fiore G, Degueldre C, Luxen A, Cleeremans A, Maquet P (2003) Learned material content and acquisition level modulate cerebral reactivation during posttraining rapid-eye-movements sleep. NeuroImage 20:125–134.

Peigneux P, Orban P, Balteau E, Degueldre C, Luxen A, Laureys S, Maquet P (2006) Offline Persistence of Memory-Related Cerebral Activity during Active Wakefulness. PLoS Biol 4 Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1413571/.

Perruchet P, Pacton S (2006) Implicit learning and statistical learning: one phenomenon, two approaches. Trends Cogn Sci 10:233–238.

Racsmány M, Lukács Á, Németh D, Pléh C (2005) A verbális munkamemória magyar nyelvű vizsgálóeljárásai. Magy Pszichol Szle 60:479–506.

Rasch B, Born J (2013) About sleep's role in memory. Physiol Rev 93:681–766.

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.

Rieth CA, Cai DJ, McDevitt EA, Mednick SC (2010) The role of sleep and practice in implicit and explicit motor learning. Behav Brain Res 214:470–474.

Robertson EM, Pascual-Leone A, Press DZ (2004) Awareness modifies the skill-learning benefits of sleep. Curr Biol CB 14:208–212.

Rózsa S, Szádóczky E, Furedi J (2001) Psychometric properties of the Hungarian version of the shortened Beck Depression Inventory. Psychiatr Hung 16:384–402.

Siegelman N, Bogaerts L, Christiansen MH, Frost R (2017) Towards a theory of individual differences in statistical learning. Philos Trans R Soc Lond B Biol Sci 372.

Simor P, Köteles F, Bódizs R, Bárdos G (2009) A questionnaire based study of subjective sleep quality: the psychometric evaluation of the Hungarian version of the Groningen Sleep Quality Scale. Mentálhig És Pszichoszomatika 10:249–261.

Soldatos CR, Dikeos DG, Paparrigopoulos TJ (2003) The diagnostic validity of the Athens Insomnia Scale. J Psychosom Res 55:263–267.

Song S, Howard JH, Howard DV (2007a) Sleep does not benefit probabilistic motor sequence learning. J Neurosci Off J Soc Neurosci 27:12475–12483.

Song S, Howard JH, Howard DV (2007b) Implicit probabilistic sequence learning is independent of explicit awareness. Learn Mem Cold Spring Harb N 14:167–176.

Spencer RMC, Sunm M, Ivry RB (2006) Sleep-dependent consolidation of contextual learning. Curr Biol CB 16:1001–1005.

Stark-Inbar A, Raza M, Taylor JA, Ivry RB (2017) Individual differences in implicit motor learning: task specificity in sensorimotor adaptation and sequence learning. J Neurophysiol 117:412–428.

Takács J, Bódizs R, Ujma PP, Horváth K, Rajna P, Harmat L (2016) Reliability and validity of the Hungarian version of the Pittsburgh Sleep Quality Index (PSQI-HUN): comparing psychiatric patients with control subjects. Sleep Breath Schlaf Atm 20:1045–1051.

Team RC (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013.

Thiessen ED, Kronstein AT, Hufnagle DG (2013) The extraction and integration framework: a two-process account of statistical learning. Psychol Bull 139:792–814.

Török B, Janacsek K, Nagy DG, Orbán G, Nemeth D (2017) Measuring and filtering reactive inhibition is essential for assessing serial decision making and learning. J Exp Psychol Gen 146:529–542.

Ujma PP, Gombos F, Genzel L, Konrad BN, Simor P, Steiger A, Dresler M, Bódizs R (2015) A comparison of two sleep spindle detection methods based on all night averages: individually adjusted vs. fixed frequencies. Front Hum Neurosci 9:52.

West G, Vadillo MA, Shanks DR, Hulme C (2017) The procedural learning deficit hypothesis of language learning disorders: we see some problems. Dev Sci.