DOCTORAL (PhD) DISSERTATION

ZSÓFIA ZAVECZ

NEUROCOGNITIVE BACKGROUND OF PROCEDURAL MEMORY:

NEURAL OSCILLATIONS AND SLEEP

2020

EÖTVÖS LORÁND UNIVERSITY
FACULTY OF EDUCATION AND PSYCHOLOGY

Zsófia Zavecz

Neurocognitive background of procedural memory: neural oscillations and sleep

Doctoral School of Psychology
Head of the School: Zsolt Demetrovics, DSc, professor, Eötvös Loránd University

Clinical Psychology and Addiction Program
Head of the Program: Zsolt Demetrovics, DSc, professor, Eötvös Loránd University

Supervisors
Dezső Németh, DSc, professor, Eötvös Loránd University
Karolina Janacsek, PhD, assistant professor, Eötvös Loránd University

Budapest, 2020

EÖTVÖS LORÁND TUDOMÁNYEGYETEM

ADATLAP a doktori értekezés nyilvánosságra hozatalához

I. A doktori értekezés adatai

A szerző neve: Zavecz Zsófia

A doktori értekezés címe és alcíme: Neurocognitive background of procedural memory: neural oscillations and sleep

A doktori iskola neve: Pszichológiai Doktori Iskola

A doktori iskolán belüli doktori program neve: Klinikai pszichológia és addiktológia program

A témavezető neve és tudományos fokozata: Németh Dezső, PhD, DSc és Janacsek Karolina, PhD

A témavezető munkahelye: ELTE PPK Pszichológiai Intézet

MTA Adatbázis-azonosító: 10057879

DOI-azonosító[1]: 10.15476/ELTE.2020.109

II. Nyilatkozatok

1. A doktori értekezés szerzőjeként[2]

a) hozzájárulok, hogy a doktori fokozat megszerzését követően a doktori értekezésem és a tézisek nyilvánosságra kerüljenek az ELTE Digitális Intézményi Tudástárban. Felhatalmazom az ELTE PPK Doktori Iskola hivatalának ügyintézőjét Barna Ildikót, hogy az értekezést és a téziseket feltöltse az ELTE Digitális Intézményi Tudástárba, és ennek során kitöltse a feltöltéshez szükséges nyilatkozatokat.

b) kérem, hogy a mellékelt kérelemben részletezett szabadalmi, illetőleg oltalmi bejelentés közzétételéig a doktori értekezést ne bocsássák nyilvánosságra az Egyetemi Könyvtárban és az ELTE Digitális Intézményi Tudástárban;[3]

[1] A kari hivatal ügyintézője tölti ki.
[2] A megfelelő szöveg aláhúzandó.
[3] A doktori értekezés benyújtásával egyidejűleg be kell adni a tudományági doktori tanácshoz a szabadalmi, illetőleg oltalmi bejelentést tanúsító okiratot és a nyilvánosságra hozatal elhalasztása iránti kérelmet.

c) kérem, hogy a nemzetbiztonsági okból minősített adatot tartalmazó doktori értekezést a minősítés (…….*dátum*)-ig tartó időtartama alatt ne bocsássák nyilvánosságra az Egyetemi Könyvtárban és az ELTE Digitális Intézményi Tudástárban;[4]

d) kérem, hogy a mű kiadására vonatkozó mellékelt kiadó szerződésre tekintettel a doktori értekezést a könyv megjelenéséig ne bocsássák nyilvánosságra az Egyetemi Könyvtárban, és az ELTE Digitális Intézményi Tudástárban csak a könyv bibliográfiai adatait tegyék közzé. Ha a könyv a fokozatszerzést követőn egy évig nem jelenik meg, hozzájárulok, hogy a doktori értekezésem és a tézisek nyilvánosságra kerüljenek az Egyetemi Könyvtárban és az ELTE Digitális Intézményi Tudástárban.[5]

2. A doktori értekezés szerzőjeként kijelentem, hogy
a) a ELTE Digitális Intézményi Tudástárba feltöltendő doktori értekezés és a tézisek saját eredeti, önálló szellemi munkám és legjobb tudomásom szerint nem sértem vele senki szerzői jogait;

b) a doktori értekezés és a tézisek nyomtatott változatai és az elektronikus adathordozón benyújtott tartalmak (szöveg és ábrák) mindenben megegyeznek.

3. A doktori értekezés szerzőjeként hozzájárulok a doktori értekezés és a tézisek szövegének plágiumkereső adatbázisba helyezéséhez és plágiumellenőrző vizsgálatok lefuttatásához.

Kelt: Budapest, 2020. július 11.

a doktori értekezés szerzőjének aláírása

[4] A doktori értekezés benyújtásával egyidejűleg be kell nyújtani a minősített adatra vonatkozó közokiratot.
[5] A doktori értekezés benyújtásával egyidejűleg be kell nyújtani a mű kiadásáról szóló kiadói szerződést.

Acknowledgements

I am grateful for all the people who accompanied me on my scientific journey up to now. Special thanks to the members of my lab, the Brain, Memory and Language Lab who became my scientific family and 'work-friends'. Most importantly, I would like to thank my supervisor and PI of the lab, Dezső Németh for all the opportunities and support that he provided. I am grateful for all the scientific advances and the inspiring conversations but also for all the social events and lab conferencings that he made possible, and the general atmosphere in the office that is full of energy mostly because of his personality. I am also grateful for the other core member of the lab, and my co-supervisor, Karolina Janacsek, whose academic preparedness is outstanding and who's methodological and analytical rigorousness is an inspiration. Besides my supervisors, I am grateful for my fellow labmates, most importantly my office mate, Eszter Tóth-Fáber who turned to me with infinite patience and endless help, saying yes (and laughing) after each and every "Can I ask?" question. Her inputs were crucial also for writing this dissertation. I am also thankful for my other 'work-friends', Csenge Török and Kata Horváth for the work together and the shared memories from conferences and other outings. I am also thankful to know and work with Noémi Éltető, Orsolya Pesthy, Mariann Kiss and Balázs Török from the lab.

I had another very important mentor in my (not so long) research career, Péter Simor. I am forever grateful for his tutoring and that he introduced me to sleep research and set the fundamentals of my knowledge of this field that I can build on ever since. I appreciate his time and energy that he invested in training me (and other young scientists alongside with me). I believe that his contribution to my knowledge was essential for the successes that I had so far in my career. I am also happy that we shared a workspace with his lab, the Sleep and Cognition Lab, and grateful to know the members, especially Richárd Reichardt and Noémi Báthori.

I also appreciate the contribution of Tamás Nagy to my career with his expertise, enthusiasm and suggestions to make research transparent and reproducible. I also learnt a lot from Andrea Kóbor, Ádám Takács and Brigitta Tóth, about EEG, work ethics and researcher paths. I would like to thank Eszter Csábi and Dr. Pálma Benedek for providing

opportunity to work in a sleep clinic. I am also grateful to my opponents, Anett Ragó and Bertalan Polner for their insightful comments and encouragement. I am also happy that I could collaborate with Zsófia Takács and Réka Kassai from the Cognitive Development and Education Research Group, and to share a workspace (and sweets and snacks) and have inspiring discussions with István Tóth-Király, Beáta Bőthe, Mónika Koós, János Salamon and Miriam Jánvári from the Social Interventions Lab.

Another important contribution to my studies was my time in The Netherlands, and I am grateful to get to know the outstanding researchers of the Donders Institute. I am thankful for Mike X Cohen to provide me an understanding of signal processing and EEG that I never thought I could have. He is an exceptional teacher, and I am grateful to have had the opportunity to learn from him. I am also grateful for Martin Dresler, who included me in his lab whole-heartedly. Besides my advisors, I am thankful for Lousie Catheryne Barne for the scientific (and friendly) discussions and for introducing me to the Predictive Brain Lab as well as an inspiring group of academic friends, Karine Roversi, Nina Roth Mota, Mariana Carvalho, Renata Cupertino and João Ferreira.

I am also thankful to work with other students, who have challenged me to understand and explain everything to the fullest. I would like to highlight Bernadett Mikula and Szonja Weigl who grew to be my collaborators over time.

Last but not least, I would like to thank my family and my friends outside of academia who supported me in my pursue of a scientific career. Special thanks to my sister, who was a voluntary participant in many of my research projects. Finally, I would like to thank my love, Rafael Pedrosa for inspiring me with his outstanding research, bringing me different perspectives, being a thinking partner, teaching me work ethics and how to make great figures and ultimately supporting me in any way he can.

Thank you all!

Table of contents

List of figures

List of tables

List of publications included in the dissertation

Zavecz, Z., Horváth, K., Solymosi, P., Janacsek, K., & Nemeth D. (2020) Frontal-midline theta frequency and probabilistic learning: A transcranial alternating current stimulation study. *Behavioural Brain Research,* 393, 112733.

Zavecz, Z., Nagy, T., Galkó, A., Nemeth, D., & Janacsek, K. (2020). The relationship between subjective sleep quality and cognitive performance in healthy young adults: Evidence from three empirical studies. *Scientific Reports,* 10, 4855.

Simor, P.*, **Zavecz, Z.***, Horvath, K., Elteto, N., Török, C., Pesthy, O., ... & Nemeth, D. (2018). Deconstructing procedural memory: Different learning trajectories and consolidation of sequence and statistical learning. *Frontiers in Psychology,* 9, 2708.

Simor, P.*, **Zavecz, Z.***, Csábi, E., Benedek, P., Janacsek, K., Gombos, F., & Németh, D. (2017). Delta and theta activity during slow-wave sleep are associated with declarative but not with non-declarative learning in children with sleep-disordered breathing. *Sleep Spindles & Cortical Up States, 1(1), 55-66.*

* Indicates shared first author positions (equal contribution)

Total impact factor of the published studies: 8.91

Note: Each co-author has granted permission for the given publication to be included in the current dissertation.

List of abbreviations

ACC	Accuracy
AHI	Apnea/Hypopnea Index
AIS	Athens Insomnia Scale
ANOVA	variance analysis
ASRT	Alternating Serial Reaction Time
BDI	Beck's Depression Inventory
BF	Bayes Factors
BIC	Bayesian Information Criterion
BMI	Body Mass Index
CAP	Cyclic Alternating Pattern
EEG	Electroencephalography
EF	Executive function
FDR	False Discovery Rate
FFT	Fast Fourier Transform
fMRI	Functional magnetic resonance imaging
GSQS	Groningen Sleep Quality Scale
IAM	Individual Adjustment Method
KSS	Karolinska Sleepiness Scale
M	Mean
MEG	Magnetoencephalography
MEQ	Morningness-Eveningness Questionnaire
MTL	Medial temporal lobe
NREM	Non-Rapid eye movement sleep
OSA	Obstructive Sleep Apnea
PET	Positron Emission Tomography
PSG	Polysomnography
PSQI	Pittsburgh Sleep Quality Index

REM	Rapid eye movement sleep
RSI	Response to stimulus interval
RT	Reaction time
SD	Standard deviation
SDB	Sleep-disordered breathing
SEM	Standard error of the mean
SRT	Serial Reaction Time
SWS	Slow-wave sleep
tACS	Transcranial Alternating Current Stimulation
WCST	Wisconsin Card Sorting Test
WM	Working memory

General introduction

Growing up we all heard the saying to put a book underneath our pillow to be able to remember its content. This age-old saying seems to have some scientific ground to it. The relationship between sleep and memory has been studied widely in the last century: the beneficial effect of sleep on memory performance has been shown repeatedly (Diekelmann & Born, 2010; Stickgold, 2005; Walker & Stickgold, 2004). However, the precise mechanisms underlying this relationship remain poorly understood, as well as there is no consensus regarding which domains of memory are affected by sleep. The aim of my doctoral studies is to provide a deeper understanding of memory processes (in particular procedural memory processes) by investigating their behavioral characteristics, neural background, and their relationship to sleep. Understanding and improving our memory processes is valuable for the healthy population to boost everyday learning or consolidation performance. Moreover, it could help to treat clinical disorders, such as amnesia, post-traumatic stress disorder or dementia.

Memory systems

Memory is not a unified construct, therefore it does not serve a single function. Within long-term memory, we differentiate multiple memory systems, such as explicit and implicit (Graf & Schacter, 1985) or declarative and non-declarative (Squire, 1992b) traditionally, but for a more recent alternative classification, see Henke (2010). These two traditional distinctions overlap (therefore the terms are often used interchangeably): explicit or declarative memories are consciously accessible and are dependent on the medial temporal lobe (MTL), whereas non-declarative or implicit memories are defined as the lack of consciousness and MTL dependence (P. J. Reber, 2013). These memory systems are further divided into subtypes. Declarative (or explicit) memory consists of episodic and semantic memory, that can be differentiated based on whether the awareness that they require, extends to the self (episodic memory) or is limited to the object of the memory (semantic, Endel Tulving, 1972). Non-declarative (or implicit) memory is an umbrella-term that consists of a heterogeneous group of memory types, such as procedural memory, conditioning, priming, and habituation (Squire & Zola, 1996). In the dissertation, I will mainly focus on procedural memory (but see Study 1).

Procedural memory

Procedural learning is a crucial ability that facilitates efficient processing of and automatic responses to complex environmental stimuli. It underlies the development of perceptual and motor skills and habits through extensive practice (Fiser & Aslin, 2002; Kaufman et al., 2010; Saffran, Aslin, & Newport, 1996; Turk-Browne, Scholl, Johnson, & Chun, 2010; Ullman, 2004). The acquisition of procedural knowledge is often unintentional, requires extended practice and the acquired representations are rigid (Kóbor, Janacsek, Takács, & Nemeth, 2017; A. S. Reber, 1967; Szegedi-Hallgató et al., 2017). Importantly, in complex procedural memory tasks, there could be both explicit and implicit processes present. Even if the instruction and knowledge of regularities are explicit, skill learning still requires practice. It seems that these explicit and implicit processes operate simultaneously in procedural memory tasks (Sanchez & Reber, 2013; Song, Howard, & Howard, 2007a). While the behavioral characteristics are also a topic in the dissertation, the main focus is the neural background of procedural memory.

Neural background

The classical memory taxonomies relied heavily on findings of the neural background of different memory processes. These findings arose from studies of amnesiac patients and human and animal lesions, as well as studies using neuroimaging techniques, such as positron emission tomography (PET) or magnetic resonance imaging (MRI) (Knowlton, Ramus, & Squire, 1992; Nissen, Nissen, Willingham, & Hartman, 1989; Nyberg, Cabeza, & Tulving, 1996; Squire, 1992a; E. Tulving & Markowitsch, 1998; Zola-Morgan & Squire, 1984; Zola-Morgan, Squire, & Amaral, 1986). Based on these techniques, procedural memory has been shown to be dependent on the basal ganglia and the striatum (Graf & Schacter, 1985; Squire & Zola, 1996). These early studies mostly aimed to reveal relevant brain areas underlying memory, with a focus on differentiating memory types and phases. However, this approach has a limited contribution to the understanding of the precise neural mechanisms underlying memory processes.

More recent studies focus on these neural mechanisms applying two main advances in their approach. First, there is an increasing focus on connectivity and pattern analysis

of brain activity, rather than identifying specific brain regions underlying memory processes (Poldrack, 2012). Second, and relatedly, as these patterns and connections are not necessarily captured adequately by neuroimaging techniques, there is increasing tendency to study brain activity underlying memory processes with psychophysiological techniques, such as magnetoencephalography (MEG) and electroencephalography (EEG) (Caplan & Glaholt, 2007; Klimesch, Freunberger, Sauseng, & Gruber, 2008). Psychophysiological techniques capture neurophysiological events directly, in contrast to neuroimaging techniques, where metabolic changes that we measure can be coupled with various underlying neurophysiological events, such as multiple evoked and induced oscillatory effects. This oscillatory synchronization is a key mechanism that integrates anatomically distributed processing and facilitates neuronal communication, thereby supporting synaptic plasticity (Buzsáki & Draguhn, 2004). The spatial scale of oscillatory synchronization can range from local (cortical columns or neighboring neurons) to global, large-scale synchronization that can connect distinct brain areas (Varela, Lachaux, Rodriguez, & Martinerie, 2001). In my doctoral studies I was using EEG to study these oscillatory dynamics underlying different procedural memory processes (Study 1 and 3).

Importantly, inconsistencies seem to occur in the neural background of procedural memory due to differences in the specific paradigms used in different studies. Differences occur in part because the tasks encompass several different functions, such as acquisition of the regularity, sensorimotor integration, model formation, or movement control (Hikosaka, Nakamura, Sakai, & Nakahara, 2002; Janacsek et al., 2020; Penhune & Steele, 2012; Willingham, 1998). Moreover, underlying brain activity in procedural memory tasks involving regularities can differ as a function of these regularities being implicit or explicit (Fletcher et al., 2005; Schendan, Searl, Melrose, & Stern, 2003; Willingham, Salidis, & Gabrieli, 2002). This task-dependent neural background accentuates that procedural learning is also not a unitary process, and instead, we have to differentiate subprocesses of it. The current studies focusing on the neural mechanisms underlying different memory processes can help us to link the characteristics of the tasks to these neural mechanisms. This approach could eventually lead to a novel taxonomy of memory processes focusing on underlying mechanisms rather than underlying brain areas (see for example, Henke, 2010). Furthermore, revealing the precise mechanisms could be the

foundation of targeted enhancement that can be used to boost everyday memory performance or treat memory-related clinical symptoms.

Subprocesses: Sequence and statistical memory

At least two processes underlying procedural learning can be distinguished: sequence and statistical learning (Kóbor et al., 2018; Nemeth, Janacsek, & Fiser, 2013), which is also referred to as rule-based and statistical learning (Maheu, Meyniel, & Dehaene, 2020). Sequence learning refers to the acquisition of a series of (usually 5–12) stimuli that repeatedly occur in the same order. In contrast, statistical learning refers to the acquisition of shorter-range relationships among stimuli that are primarily based on frequency/probability information (i.e., differentiating between more frequent/probable and less frequent/probable runs (e.g., pairs, triplets, etc.) of stimuli). Majority of the state-of-the-art studies (and also Study 1 and 2 in the dissertation) however, do not differentiate between these (or other) subprocesses and quantify procedural learning with a mixed measure of acquiring both frequency/probability and sequential information. In Study 3 and 4, we set out to better characterize these subprocesses and explore their differences, particularly in relation to sleep. For a summary of the types of memory investigated in each study of the dissertation, see Table 1.

Measurement: The Alternating Serial Reaction Time task

To measure procedural learning, we used the Alternating Serial Reaction Time (ASRT) task in all of the four studies (J. H. Howard, Jr. & Howard, 1997; Nemeth, Janacsek, & Fiser, 2013; Nemeth, Janacsek, Londe, et al., 2010). This task has been proven to have good test-retest reliability as well as sensitivity to individual differences in performance (Stark-Inbar, Raza, Taylor, & Ivry, 2016). Furthermore, this task enables us to separate learning processes that rely on improvements in visuo-motor coordination (general skill learning) and those that rely on the extraction of regularities (sequence and statistical learning) (J. H. Howard, Jr. & Howard, 1997; Nemeth, Janacsek, Londe, et al., 2010).

In Study 1 and 2 we used the original version of the ASRT task (J. H. Howard, Jr. & Howard, 1997; Nemeth, Janacsek, Londe, et al., 2010), which measures incidental

learning, incorporating however both sequence and statistical learning. In this perceptual-motor four-choice reaction time task, a stimulus appears in one of four horizontally arranged empty circles on the screen, and participants have to press a corresponding button on a keyboard or response box. The appearance of stimuli follows a predetermined alternating sequence order, such that sequence elements alternate with random ones. Due to the sequence structure in the ASRT task, some runs of three consecutive elements (referred to as triplets in our studies) occur more frequently and consequently are more probable (high-frequency or high-probability triplets) than others (low-frequency or low-probability triplets). The learning is quantified as the difference in reaction times (RTs) and accuracy (ACC) between the high- and low-frequency/probability triplets (note that we use the terms frequency and probability in relation to triplets interchangeably). Importantly, however, the comparison of RTs and ACC of high- vs. low-frequency triplets does not take into account whether the last elements of the high-frequency triplets are sequence or random element, and consequently, provides a mixed measure of sequence and statistical learning. To measure both of these processes within the same behavioral paradigm, we used a modified version of the ASRT task in Study 3 and 4: the cued ASRT (referred to as explicit ASRT in Study 3 and 4).

In contrast to the original (uncued) version of the ASRT task, in the cued version (Nemeth, Janacsek, & Fiser, 2013), sequence and random elements are differentiated on the stimulus level: sequence elements are marked with a picture of a dog, random ones with that of a penguin. Thus, not only high- and low-frequency triplets are well distinguished but also triplets with the last element of a sequence or random stimuli. Together, these characteristics enable us to measure both sequence and statistical learning. Sequence learning is quantified as a difference between responses for triplets ending with sequence vs. random elements that were high-frequency. Statistical learning is quantified as a difference between responses for those random elements that were the last elements of high-frequency triplets vs. those that were the last elements of low-frequency triplets. Importantly, sequence and statistical learning are also distinguishable in the original uncued version of the task, however, it seems to appear only after extended practice, over the time course of multiple sessions over several days (D. V. Howard et al., 2004; J. H. Howard, Jr. & Howard, 1997).

Besides the differentiation of the visual cues, participants are informed about the underlying structure of the sequence in this version, and their attention is drawn to the alternation of sequence and random elements by the different visual cues. Due to this modification, the learning could partially become explicit and intentional in this paradigm. However, it has been shown, that the explicit instruction does not influence the learning performance measured by reaction time and accuracy in this task (Song et al., 2007a), therefore the explicit instruction does not necessarily change the procedural learning itself explicit. For further details of the characteristics of the ASRT tasks, see the task descriptions in the respective studies.

Table 1.1 Summary of memory types investigated in the four studies

Study	**Memory systems**	**Subtypes**	**Task**
Study 1[6]	Declarative	Episodic memory	Story recall
	Non-declarative	Procedural memory (mixed learning index)	ASRT
Study 2[7]	Non-declarative	Procedural memory (mixed learning index)	ASRT
Study 3[8]	Non-declarative	Procedural memory (sequence and statistical learning)	Explicit ASRT
Study 4[9]	Non-declarative	Procedural memory (sequence and statistical learning)	Explicit ASRT

Note: ASRT: Alternating Serial Reaction Time

[6] Simor, P., Zavecz, Z., Csábi, E., Benedek, P., Janacsek, K., Gombos, F., & Németh, D. (2017). Delta and theta activity during slow-wave sleep are associated with declarative but not with non-declarative learning in children with sleep-disordered breathing. Sleep Spindles & Cortical Up States, 1(1), 55-66.

[7] Zavecz, Z., Horváth, K., Solymosi, P., Janacsek, K., & Nemeth, D. (2020). Frontal-midline theta frequency and probabilistic learning: A transcranial Alternating Current Stimulation study. *Behavioural Brain Research*, 112733.

[8] Simor, P., Zavecz, Z., Horváth, K., Éltető, N., Török, C., Pesthy, O., Gombos, F., Janacsek, K., & Nemeth, D. (2019). Deconstructing procedural memory: Different learning trajectories and consolidation of sequence and statistical learning. *Frontiers in Psychology*, *9*, 2708.

[9] Zavecz, Z., Nagy, T., Galkó, A., Nemeth, D., & Janacsek, K. (2020). The relationship between subjective sleep quality and cognitive performance in healthy young adults: Evidence from three empirical studies. *Scientific reports*, *10*(1), 1-12.

Phases of memory: learning, consolidation and retrieval

Investigating the neural background of procedural memory, it is also important to differentiate the memory phases, such as learning, consolidation and retrieval. The initial phase is learning (also referred to as acquisition in procedural memory studies), i.e., encoding of sensory information. Following that, consolidation is the mechanism through which the encoded memory representations get stable, less susceptible to future interferences. Neuropsychological evidence for the existence of consolidation comes from investigating the temporally graded retrograde amnesia (Squire, 2009). Furthermore, pharmacological studies in animals also show evidence for the phenomenon of consolidation: certain manipulations are only effective shortly after encoding, but not after a period of consolidation (McGaugh & Izquierdo, 2000). As a synthesis of studies from different areas, McGaugh (2000) concluded that probably there are more than one processes underlying memory consolidation. He differentiated the cellular stabilization (occurring initially after encoding) and the systems consolidation, which involves large-scale organization of the memory trace. After the successful consolidation, we can retrieve (recall or recollect) the memory when needed. These phases sometimes are referred to as on-line (during practice) or off-line (between practice sessions) memory processes.

The studies in my dissertation investigate learning (on-line memory process) and consolidation (off-line memory process) more closely (Fig. 1): in Study 1, we studied brain activity during consolidation, but both in relation to learning capacity and consolidation; in Study 2, we studied brain activity during learning in relation to the learning performance; in Study 3, we studied brain activity during consolidation in relation to consolidation, but we discuss learning performance on the behavioral level independently of the brain activity; and finally in Study 4, we studied the effects of sleep in general on learning capacity.

Procedural memory and sleep

Sleep has been suggested as a crucial phenomenon for memory consolidation (sleep-dependent memory consolidation, Diekelmann, Wilhelm, & Born, 2009; Walker & Stickgold, 2004). However, it is not clear, whether sleep *per se* is necessary for

consolidation, or it is just an ideal state for that process, due to lack of interference from the environment (Mednick, Cai, Shuman, Anagnostaras, & Wixted, 2011). Researchers argue that a quiet resting period (which also lacks interference) is just as beneficial for consolidation as sleep (Brokaw et al., 2016; Craig & Dewar, 2018). In Study 3, our aim was to test sleep-dependent consolidation of procedural memory, by comparing the memory performance of participants sleeping, quietly resting or actively resting (watching a movie) during the off-line period.

Majority of the studies investigating sleep and memory focus on sleep-dependent memory consolidation. However, it is also important to note that sleep quality in general affects various domains of cognitive performance as well, such as executive functions and attention (Jones & Harrison, 2001) and the learning capacity itself (Walker & Stickgold, 2004). Long-term insufficiency in sleep duration or quality decreases the learning performance (Curcio, Ferrara, & De Gennaro, 2006). Study 1 and 4 target these long-term effects of sleep on memory performance. The studies on sleep and memory have led to mixed findings: again, different memory systems/types seem to have different relations to sleep (see Study 1, and for a review Stickgold, 2005).

Some studies investigating the relationship between sleep and procedural memory have shown associations (Fenn, Nusbaum, & Margoliash, 2003; Fischer, Hallschmid, Elsner, & Born, 2002; Stickgold, James, & Hobson, 2000; Walker, Brakefield, Morgan, Hobson, & Stickgold, 2002), whereas others did not (Csabi et al., 2015; Nemeth, Csábi, Janacsek, Varszegi, & Mari, 2012; Wilhelm, Diekelmann, & Born, 2008). However, these studies not only did not use the same task to measure procedural memory (which could influence this relationship, see Procedural memory section in this Introduction) but also used different methods to study the relationship of sleep with memory.

Methods of sleep and memory research

There are two main types of sleep and memory research: 1) comparing memory performance after a certain amount of time spent either sleeping or awake and 2) investigating the effect of insufficient sleep time or quality on memory (i.e., sleep deprivation, and sleep disorder studies). Both types of memory research can incorporate comparing groups (e.g., participants with or without sleep disorder) and studying associations between sleep parameters (sleep duration, time spent in certain sleep stages,

arousals, etc.) and memory performance. Study 1 in the dissertation falls in the second category by investigating memory performance in a sleep-disordered population in relation to brain activity during sleep. Study 3 falls in the first category, by investigating the effect of alertness state (i.e., wake/quiet rest/active rest) during the off-line period after learning on memory performance.

Objective and subjective sleep parameters

Moreover, we also have to differentiate sleep and memory studies relying on objective (measured by actigraph or EEG) and subjective (self-reported) sleep measurements. Previous studies have shown that subjective and objective sleep parameters could differ (Armitage, Trivedi, Hoffmann, & Rush, 1997; Guedes et al., 2016; Landry, Best, & Liu-Ambrose, 2015a). A widely used sleep parameter that is assessed subjectively (i.e., participants judging their own sleep) is sleep quality. Subjective sleep quality can vary from the objective sleep quality because it is estimated introspectively by a combination of instinctive parameters, including the initiation of sleep, sleep continuity (number of awakenings), and/or depth of sleep. While Study 1 and Study 3 were investigating the relationship between objective sleep measures and procedural memory, Study 4 aims to explore how subjective measures of sleep are associated with procedural memory.

Considering objective sleep parameters, there are several parameters that are commonly assessed. Sleep has different stages, that can be divided into two main classes: non-rapid eye movement (NREM) sleep and rapid eye movement (REM) sleep. NREM sleep can be further divided into three substages (1–3) corresponding in that order to increasing depth of sleep (Berry et al., 2012). The deepest sleep stage, i.e., NREM 3 is also referred to as slow-wave sleep (SWS). The duration and the proportion of these sleep stages (i.e., the macrostructure of sleep) are commonly used to characterize sleep. However, each of these stages has their distinct neural characteristics, with different dominant oscillatory activity and other phenomena, such as spindles and K-complexes in NREM 2 (microstructure of sleep). In both Study 1 and 3 where we assessed objective sleep parameters, we investigated both macro- and microstructure of sleep in relation to memory performance. For an outline of the complex relationship we were aiming to explore, see Figure 1.

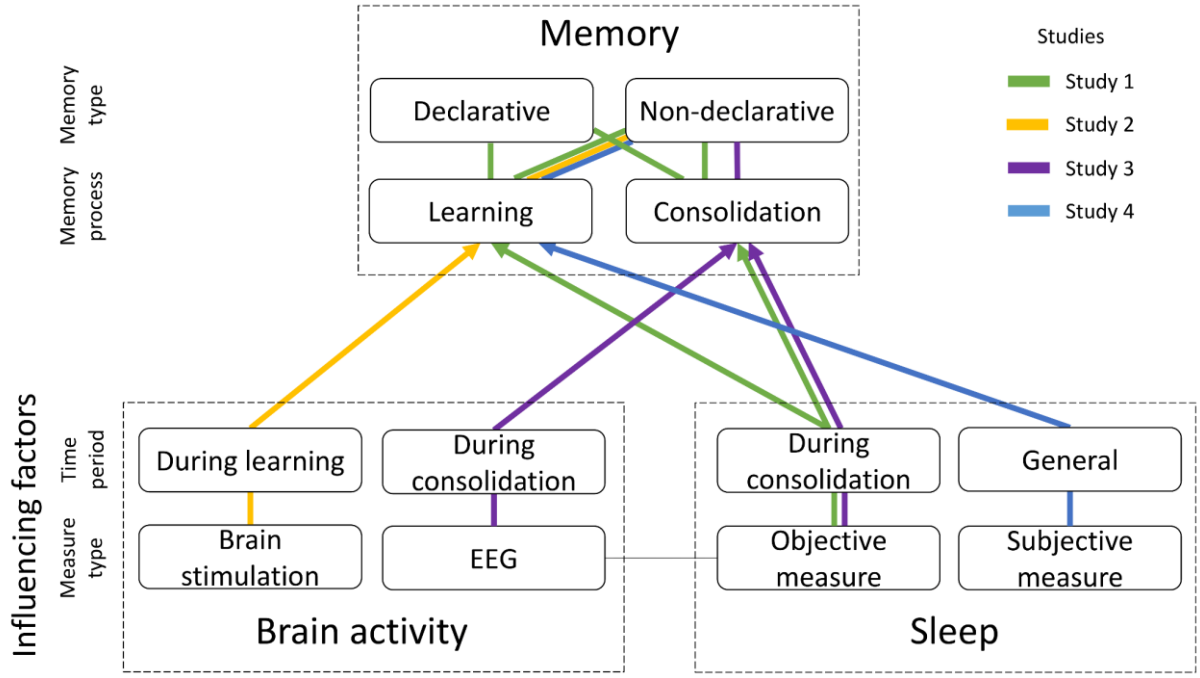

Figure 1.1 Outline of the studied components in the four studies. In all four studies our aim was to provide a deeper understanding of non-declarative (procedural) memory by investigating its neural background, and its relationship to sleep. In the first study (green lines) we investigated the differential association of declarative and non-declarative memory to sleep parameters in a sleep-disordered population. In the second study (yellow lines) we investigated the neural background of non-declarative memory by brain stimulation. In the third study (purple lines) we investigated the effect of sleep vs. wakefulness after learning on non-declarative memory consolidation. In the fourth study (blue lines) we investigated the association of subjective sleep quality and non-declarative memory performance.

Research questions

How declarative and non-declarative memory are related to sleep in a sleep-disordered population?

In Study 1, we investigated the differential association of declarative and non-declarative memory with sleep in a sleep-disordered population. More precisely, we explored how sleep disruptions affect memory performance in pediatric Sleep-disordered breathing (SDB), a prevalent sleep disorder. SDB comprises a broad spectrum of breathing-related sleep problems from primary snoring to the most severe forms of obstructive sleep apnea (Marcus, 2001). Several studies showed that SDB has a

detrimental impact on children's behavior, affect, and cognitive performance, including memory (Beebe & Gozal, 2002; Blunden, Lushington, Lorenzen, Martin, & Kennedy, 2005; Csábi, Benedek, Janacsek, Katona, & Nemeth, 2013; Gottlieb et al., 2004; Gottlieb et al., 2003; Halbower et al., 2006; Kohler et al., 2009; O'Brien et al., 2004). However, the precise mechanism that is leading to cognitive impairment and behavioral problems is unknown. A possible candidate for such disrupted but crucial mechanism is slow-wave sleep. Slow-wave sleep seems to play an important role in memory consolidation (Ferri et al., 2008; Mander et al., 2013; Marshall, Helgadottir, Molle, & Born, 2006; Rasch & Born, 2013) and appears to be altered in case of insufficient sleep (Cajochen, Foy, & Dijk, 1999; Munch et al., 2004). In light of previous studies that reported attenuated SWS-specific slow frequency oscillations in children with SDB (Jussila et al., 2016; Kheirandish-Gozal et al., 2007), our aim was to investigate whether SWS spectral power is associated with learning capacity and overnight memory consolidation within a group of children with SDB. Moreover, we applied both a declarative and a non-declarative memory task in order to further explore the specificity of sleep-related memory impairments in SDB.

Is theta oscillation crucial for procedural memory?

In Study 2, we investigated the neural background of procedural memory by directly manipulating oscillatory activity during learning. Identification of critical brain dynamics and manipulation together give an opportunity to influence memory performance, that can be used in treating clinical symptoms or boost everyday memory performance.

Previous studies showed competition between neural networks related to executive function/working memory vs. procedural learning (Albouy et al., 2015; Albouy et al., 2008; Ashby & O'Brien, 2005; Daw, Niv, & Dayan, 2005; Poldrack et al., 2001). Theta synchronization has been associated with the former (Gevins, Smith, McEvoy, & Yu, 1997; Hsieh & Ranganath, 2014; Jensen & Tesche, 2002; Meyer, Grigutsch, Schmuck, Gaston, & Friederici, 2015; Onton, Delorme, & Makeig, 2005; Scheeringa et al., 2009; Summerfield & Mangels, 2005; Tóth et al., 2014) while desynchronization with the latter (Tóth et al., 2017) in correlational studies. In this study, our aim was to test the causal

relationship between theta synchronization and procedural learning with non-invasive transcranial alternating current (tACS) stimulation.

Is sleep essential for the consolidation of different subprocesses of procedural memory?

In Study 3, we investigated the consolidation of sequence and statistical knowledge in case of sleep, quiet rest or active rest. Our primary goal was to test whether sleep has a beneficial effect on procedural memory consolidation compared to wakefulness. Based on more recent evidence, sleep and rest without interferences can have similar beneficial effects on memory consolidation (Brokaw et al., 2016; Craig & Dewar, 2018; Mednick et al., 2011). Therefore, we compared the consolidation performance of groups of participants who, after learning, either slept, rested quietly or watched a movie. As several studies indicate that not sleep per se, but specific oscillations during sleep facilitate post-sleep improvements in behavioral performance (Rasch and Born, 2013), we also recorded EEG during the consolidation period. We explored associations between the spectral composition of brain activity and the consolidation performance within each off-line activity group. Furthermore, we wanted to test whether different subprocesses of procedural memory, sequence and statistical learning differ in their benefit from sleep. 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). While the sleep-dependent memory consolidation of statistical information is largely unexplored, studies with a mixed measure of sequence and statistical learning did not find benefit of post-learning sleep on consolidation (Nemeth, Janacsek, Londe, et al., 2010; Song, Howard, & Howard, 2007b). Thus, in this study, our aim was to fill this gap by investigating whether sleep in the post-learning period has differential associations with sequence and statistical learning.

Are subjective measures of sleep associated with different subprocesses of procedural memory?

In Study 4, we investigated whether subjective sleep quality is associated with procedural memory. Previous studies have shown that subjective and objective sleep parameters, such as sleep duration or sleep efficiency differ (Armitage et al., 1997; Guedes et al., 2016; Landry, Best, & Liu-Ambrose, 2015b). Extreme deviations can occur between subjective and objective measures in sleep disorders, such as insomnia or sleep-state misperception. According to Zhang and Zhao (2007) and Stepanski et al. (1989), subjective sleep quality of insomniac patients determines both seeking medication and type of effective treatment. Furthermore, one's belief about their own sleep quality induce placebo and nocebo effects both in insomniac patients and healthy individuals (Draganich & Erdal, 2014; Gavriloff et al., 2018). Thus, subjective sleep quality has therapeutic importance, as well as further explanatory value for cognitive performance compared to objective measures. However, scientific evidence on the relationship between subjective sleep quality and cognition is still inconclusive, and memory, in particular, procedural memory has been scarcely investigated in relation to subjective sleep quality. In this study, we aimed to fill this gap by providing an extensive investigation on the relationship between subjective sleep quality and cognitive performance including procedural memory in healthy young adults.

Together, the studies in the dissertation will provide a deeper understanding of procedural memory on two levels: on the 1) behavioral and 2) neural level. On the behavioral level, the aim is to dissect and characterize subprocesses of procedural learning, namely sequence and statistical learning. On the neural level, the aim is to reveal crucial bran dynamics underlying procedural memory and to clarify its relationship with sleep by investigating brain activity both during wake and sleep in relation to both learning and consolidation.

Delta and theta activity during slow-wave sleep are associated with declarative, but not with non-declarative learning in children with sleep-disordered breathing[10]

Abstract

Sleep-disordered breathing (SDB) is a prevalent sleep disorder among young children and is associated with daytime impairments, such as behavioral dysregulation, affective symptoms, and reduced cognitive performance. Microstructural changes of non-rapid eye movement sleep, particularly the reduction of slow frequency oscillations during slow-wave sleep (SWS) might be associated with impaired learning among children with SDB. In this study, we investigated the associations between learning capacity, overnight memory retention, and post-learning, spectral power density of SWS within a clinical sample of children (n = 27) with SDB. Participants performed a declarative (the "War of the Ghosts") and a non-declarative (the "Alternating Serial Reaction Time") memory task at night, before their clinical (night-time polysomnographic) evaluation. Memory retention was assessed in the morning. Overnight changes in performance in the declarative and non-declarative task were not related to relative spectral power measures of SWS. Nevertheless, declarative learning capacity was positively correlated with relative delta (1.25–4 Hz) and negatively with relative theta (4.25–8 Hz) power. Although, statistical learning was not associated with spectral power, general skill learning was positively associated with delta and negatively associated with theta power. Associations in case of declarative learning remained significant beyond the influence of age; however, in case of general skill learning the associations with delta and theta power were explained by age. These findings indicate that among children with SDB, oscillations within the delta and theta band during SWS are associated with declarative learning capacity, but are independent from non-declarative, statistical learning.

Keywords: sleep disordered breathing (SDB); declarative learning; implicit learning; statistical learning; EEG; oscillations

[10] Simor, P., Zavecz, Z., Csábi, E., Benedek, P., Janacsek, K., Gombos, F., & Németh, D. (2017). Delta and theta activity during slow-wave sleep are associated with declarative but not with non-declarative learning in children with sleep-disordered breathing. Sleep Spindles & Cortical Up States, 1(1), 55-66.

Introduction

Sleep-disordered breathing (SDB) is a highly common complaint in prepubertal children with prevalence rates between 7-12 % (Brunetti et al., 2001; Castronovo et al., 2003; Ersu et al., 2004; Ferreira et al., 2000), but in some cases reported up to 34.5 % (Castronovo et al., 2003). SDB comprises a broad spectrum of breathing-related sleep problems from primary snoring to the most severe forms of obstructive sleep apnea (OSA) (Marcus, 2001). Whereas OSA, that is characterized by apnea, hypopnea, transient hypoxia, hypercarbia and related arousals during sleep is diagnosed in 1-3% of children (Ali, Pitson, & Stradling, 1993; Bixler et al., 2009), milder forms of SDB, in which sleep disruptions and impaired gas exchange are not detected, are largely under-diagnosed (Blunden et al., 2005). Here we investigate how sleep disruptions affect cognitive functioning in SDB.

A growing number of studies indicate that moderate to severe OSA has a detrimental impact on children's behavior, affect, and cognitive performance (Beebe & Gozal, 2002; Blunden et al., 2005; Gottlieb et al., 2003; O'Brien et al., 2004). The latter is corroborated by findings linking symptoms of SDB to behavioral dysregulation (Rosen et al., 2004), inattention/hyperactivity (Chervin et al., 2002), as well as to impaired learning, attention, and executive function (Csábi et al., 2013; Gottlieb et al., 2004; Halbower et al., 2006; Kohler et al., 2009). These adverse effects might be driven by disrupted restorative functions of night-time sleep and reduced oxygen delivery resulting in neuronal damage (Beebe & Gozal, 2002; Blunden & Beebe, 2006). Although a recent study (Hunter & Gozal, 2016) involving a large number of pre-school aged children showed that the clinical severity of SDB symptoms [e.g., apnea/hypopnea index (AHI), arousals, oxygen desaturation] is associated with poorer cognitive abilities in a dose-dependent manner (i.e. the more severe the symptoms are, the worse the performance is); converging evidence indicates that compared with healthy, non-snoring controls even milder forms of SDB, such as habitual snoring, are predictive of impaired cognitive and behavioral profile (Archbold, Giordani, Ruzicka, & Chervin, 2004; Bourke et al., 2011a; Csabi et al., 2015). For instance, intellectual abilities and academic functions (Bourke et al., 2011a), declarative memory performance, and executive skills (Gottlieb et al., 2004), as well as parent-rated neurobehavioral functions (Bourke et al., 2011b) were similarly impaired in school-aged children with moderate-to-severe or mild SDB symptoms.

Nevertheless, it is not clear, whether in children with mild SDB, poorer behavioral and cognitive profiles are associated with abnormal nocturnal respiratory patterns, since studies have provided inconclusive results in this regard (Bourke et al., 2011a, 2011b; Hunter & Gozal, 2016). Although sleep fragmentation provoked by apneic events is considered to be another important mechanism that might lead to daytime cognitive impairments (Beebe & Gozal, 2002; Blunden & Beebe, 2006), data regarding the link between disrupted sleep and cognitive performance in mild SDB is scarce. Sleep macrostructure seems to be unaltered in children with SDB, but more subtle indices of homeostatic sleep regulation suggest that abnormal respiration might interfere with cortical, slow frequency oscillations during deep sleep, specifically during slow wave sleep (SWS) (Jussila et al., 2016; Kheirandish-Gozal et al., 2007), albeit findings are not absolutely conclusive (Yang et al., 2010). The lower rate of A1 subtype arousals as quantified by the Cyclic Alternating Pattern (CAP) (Kheirandish-Gozal et al., 2007), and frontally reduced activity in slower frequencies (< 4 Hz) during deep sleep (Jussila et al., 2016) suggest that specific neural oscillations are relatively attenuated in children with SDB.

The frontally localized A1 subtype of CAP (Ferri et al., 2008) such as low-frequency oscillations, indexed by delta (1-4 Hz) power (Cajochen et al., 1999; Munch et al., 2004), reflect the restorative capacity of the brain (Mander et al., 2010) and seem to play an important role in memory consolidation (Ferri et al., 2008; Mander et al., 2013; Marshall et al., 2006; Rasch & Born, 2013). In line with the role of SWS in memory consolidation, (Guo, Igue, Malhotra, Stickgold, & Djonlagic, 2013) that adults suffering from OSA are characterized by diminished SWS and reduced overnight improvement in a verbal-associates task, compared to a healthy control group. Interestingly, the OSA group showed reduced SWS during the experimental night only, when pre-sleep learning occurred. According to the authors, diminished post-training increase in slow wave activity might have contributed to impaired memory consolidation during sleep.

The expression of slow frequency oscillations (more frequently quantified by EEG spectral power) during SWS is strongly dependent on the integrity of the prefrontal cortex (Mander et al., 2013), and seems to be critical for the efficiency of cognitive functions that rely mainly on prefrontal and related (e.g. hippocampus) brain regions (Ferrara & De Gennaro, 2011; Mander et al., 2013; Mander et al., 2010). As a matter of

fact, SDB in children seems to impinge specifically on tasks that involve sustained attention, executive functions or declarative learning (Archbold et al., 2004; Bourke et al., 2011a; Csábi et al., 2013; Gottlieb et al., 2003). On the other hand, in case of an implicit, non-declarative learning task that does not require (or might even benefit from the reduction of) cognitive control functions (Nemeth, Janacsek, Polner, & Kovacs, 2013) children with SDB showed equivalent performance to controls (Csábi et al., 2013; Csabi et al., 2015). More specifically, Csábi and colleagues (Csábi et al., 2013; Csabi et al., 2015) reported impaired declarative learning, but intact non-declarative learning in children with SDB. Furthermore, the patient and the control group showed similar overnight memory retention in both tasks, indicating intact consolidation in children with SDB (Csabi et al., 2015). Nevertheless, in this study the association between task performance and polysomnographic measures was not examined.

In light of previous studies that reported attenuated SWS-specific slow frequency oscillations in children with SDB (Jussila et al., 2016; Kheirandish-Gozal et al., 2007), our aim was to investigate the associations between SWS spectral power, learning performance, and overnight memory retention within a group of children with SDB. To further explore the specificity of sleep-related cognitive impairments in SDB, we applied a declarative, verbal memory task and an implicit, non-declarative statistical learning task. We hypothesize that SWS spectral power is associated with memory retention in SDB. To the best of our knowledge, this is the first study investigating the relationship between SWS, learning performance, and memory retention in children with SDB.

Methods

Participants

Twenty-seven children participated in the experiment. Age, breathing events during sleep, body mass index, and sleep parameters are listed in *Table 1*. All participants were reported to snore by their parents and underwent an overnight polysomnography (PSG) for clinical evaluation at the Sleep Disorders Laboratory of Heim Pál Children's Hospital, Budapest, Hungary. All these patients met the International Classification of Sleep Disorders criteria's (Darien, 2014) for primary snoring (N = 23) or OSA (N = 4). The diagnostic criteria for Primary Snoring are complaint of snoring made by and observer

(e.g., the parent). Polysomnographic monitoring in case of this disease demonstrates inspiratory or expiratory sounds often occurring for prolonged episodes during the total sleep time (this can be measured by snoring index), but no associated abrupt arousals, arterial oxygen desaturation, or cardiac disturbances. The diagnostic criteria of OSA are frequent episodes of obstructed breathing occur during sleep, and complaint of excessive sleepiness or insomnia. Polysomnographic monitoring demonstrates obstructive apneas (this is measured by the AHI), frequent arousals from sleep and arterial oxygen desaturation associated with the apneic episodes.

The snore index of the snoring patients (M = 25.52, SD = 44.16, range: 0–155) significantly differed from zero [$t(22) = 2.77$, $p = .01$]. The AHI of the participants who had been diagnosed with OSA (M = 23.05, SD = 37.60, range: 1–79) did not significantly differ from zero [$t(3) = 1.27$, $p = .31$], probably due to the low number of patients (N = 4). Given that the neurobehavioral deficits characterizing children with primary snoring seem to be similar to those found in children with OSA (Gozal & O'Brien, 2004), we did not intend to examine the OSA and snoring subgroups separately. Nevertheless, apart from the main analyses, we performed a separate analysis for the primary snoring subgroup only (these analyses are presented in the Supplementary Material). The data of one subject was removed from the analyses in relation to the declarative task, and of another subject from the analyses of the non-declarative task, due to lack of motivation to perform the specific task. All SDB patients were untreated prior to and during the experimental night. Informed written parental consent and verbal assent of the children were provided. Participants did not receive any financial compensation for their participation. Ethics approval was obtained by the Ethics Committee at Heim Pál Children's Hospital, Budapest.

Table 2.1 Age, breathing events during sleep, body mass index (BMI), and sleep parameters of participants

Variable	Mean (Std. Deviation)
Age (years)	8.52 (2.12)
Gender (male, %)	59.25%
BMI	18.28 (4.72)
Sleep efficiency (%)	87.44 (6.99)
Relative wake duration (%)	12.22 (6.64)
Relative S1 duration (%)	2.96 (2.34)
Relative S2 duration (%)	42.04 (9.26)
Relative S3 duration (%)	33.41 (9.86)
Relative REM duration (%)	21.74 (5.52)
AHI	3.48 (15.24)
Maximum desaturation (%)	5.29 (7.16)
Desaturation index	6.33 (21.04)
Snore index	25.61 (42.05)

Note: BMI=body-mass index, kg m2, AHI: apnea/hypopnea index, measured as number of events per hour; desaturation index: measured as number of desaturations per hour; snore index: measured as snoring events per hour.

Tasks

Declarative memory task

Declarative memory performance was measured by the classical "The War of the Ghosts" test (Bartlett, 1932; Bergman & Roediger, 1999). This is a story recall test, which is widely used to measure declarative, episodic memory (Andreano & Cahill, 2006; Bartlett, 1932; Bergman & Roediger, 1999; Schwabe et al., 2009). In this test, children are asked to listen and repeat a story which consists of 36 information chunks. Based on the standardized scoring, 1 point is given if an information chunk is correctly recalled,

and 0.5 point is given if it is only partly correct (capturing the gist of the sentences) (Bartlett, 1932; Csábi et al., 2013; Gauld & Stephenson, 1967).

Non-declarative memory task

We used the Alternating Serial Reaction Time (ASRT) Task in order to assess non-declarative learning performance. In this task, a stimulus (a dog's head) appears in one of the four empty circles displayed in the middle of the screen and participants have to press the corresponding button as quickly and accurately as possible (Nemeth, Janacsek, Londe, et al., 2010). The computer used was equipped with a special keyboard with four marked keys (Z, C, B and M on a QWERTY keyboard), each corresponding to one of the horizontally aligned circles. The task consisted of two sessions, the first session (Learning Phase) consisted of 25 blocks, and the second session (Testing Phase) consisted of 5 blocks. Each block consisted 85 key presses — the first 5 stimuli were random for practice purposes, then an eight-element alternating sequence (e.g., 1r4r3r1r, where numbers represent the four places on the screen, and r represents an event randomly selected from the four possible places) repeated 10 times. A different ASRT sequence was selected for each participant based on a permutation rule so that each of the six unique permutations of the four repeating events occurred. Consequently, six different sequences were used across participants. Similarly to earlier studies (Nemeth, Janacsek, Londe, et al., 2010), stimuli were presented 120 ms after the previous response (response-to-stimulus interval, RSI). Each block required about 1.5-2 min and the entire Learning Phase took approximately 40–50 minutes, and the Testing Phase took approximately 10-15 minutes. Between blocks, participants received feedback about their overall reaction time (RT) and accuracy (ACC) on the screen and then rested 10–20 s before starting a new block.

Due to the structure of the sequences in the ASRT task, some triplets or runs of three consecutive events occur more frequently (high-frequency triplets) than others (low-frequency triplets). For example, in the above illustration, 1_**4**, 2_**3**, 3_**1** and 4_**2** (where "_" indicates the middle element of the triplet) would occur often because the third element (bold numbers) could be derived from the sequence or could also be a random element. In contrast, 1_**3** or 4_**1** would occur less frequently because in this case, the third element could only be random. Note that the final event of high-frequency triplets is therefore more predictable from the initial event when compared with the low-frequency

triplets [also known as non-adjacent second-order dependency, (Remillard, 2008)]. Therefore, before analyzing the data we determined whether each item was the last element of a high-frequency or low-frequency triplet. Out of the 64 possible triplets, the 16 high-frequency triplets occurred 62.5% of the time and the 48 low-frequency triplets occurred 37.5% of the time. Note that the final event of high-frequency triplets is therefore more predictable from the initial event compared with the low-frequency triplets.

Previous studies have shown that as people practice the ASRT task, they come to respond more quickly and more accurately to the high-frequency triplets than low-frequency triplets, revealing statistical learning (D. V. Howard et al., 2004; J. H. Howard, Jr. & Howard, 1997; Janacsek, Fiser, & Nemeth, 2012; Nemeth, Janacsek, Londe, et al., 2010; Song et al., 2007b). In addition, general skill learning is revealed in the ASRT task by the overall speed-up due to practice, irrespective of the triplet types. Thus, the ASRT task enables to measure both statistical and general skill learning.

Finally, it is important to note that the task remained implicit for the participants throughout the experiment. According to previous experiments with the ASRT task, even after an extended practice of 10 days, participants are not able to recognize the hidden sequence (D. V. Howard et al., 2004).

Procedure

PSG recordings were performed in the Sleep Disorders Laboratory of Heim Pál Children's Hospital, Budapest, Hungary. All children accomplished first the declarative and then the non-declarative task in two separate sessions, prior to sleep, and after sleep. The order of the tasks was fixed. Memory performance was assessed at 7-9 p.m. in the evening (Learning Phase), and 12 hours later after night-time sleep, at 7-9 a.m. in the morning (Testing Phase). This study was performed within the frames of the clinical evaluation, therefore children spent only one night in the laboratory, and no adaptation night was applied.

Polysomnography

The PSG was performed with the Somnomedics Somnoscreen plus device and software (Randersacker, Germany). PSG was configured to record EEG leads C4, C3 referenced to the mathematically linked mastoids (A2, A1) as well as bipolar EOG, chin EMG, ECG, snoring (by nasal cannula), respiratory effort signals, SpO2, pulse rate, and body position. EEG electrodes (C4, C3, A2, A1) were placed in accordance with the 10–20 electrode placement system (Jasper, 1958). Children were also fitted with two EOG electrodes (left and right EOG channels), monitoring vertical and horizontal eye-movements; two EMG electrodes (bipolar channels) for the chin, bipolar ECG electrodes; in addition to internal body position sensors, a pulse oximeter, a nasal flow thermistor (for measuring snoring), and thoracic and abdominal respiration sensors. Ag/AgCl EEG cup electrodes were fixed with Ten20 EEG conductive paste (Weaver and Company, Aurora, CO, USA). Hardware filters (-6 dB filters) were set between 0.3 Hz (high-pass) and 100 Hz (low-pass), signals were collected and digitized with 256 Hz/channel sampling rate (synchronous) with 8 bits resolution. Impedances were kept below 6 kOhms.

Spectral analyses

Sleep stages and conventional parameters of sleep macrostructure were scored in accordance with standardized criteria (Silber et al., 2007) by two experienced sleep researchers. Spectral analyses were performed by a custom-made software tool for full night sleep EEG analysis (FerciosEEGPlus, © Ferenc Gombos 2008-2016). 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 for whole night SWS [non-rapid eye movement (NREM) Stage3 sleep] between 1 and 25 Hz. Since the absolute power values may be biased due to age-dependent differences of the thickness and conductivity of the skull, (Carrier, Land, Buysse, Kupfer, & Monk, 2001), we applied the relative spectral power values. Relative spectral power values were obtained for each frequency bin (width: 0.25 Hz) by dividing the absolute power of the given frequency bin with the total spectral power (the sum of the absolute power of the whole range of analysis between 1

Hz and 25 Hz). The relative power values reflect the relative contribution of a given frequency range to the total spectrum. To reduce the number of parameters, we summed up frequency bins to generate five frequency band windows: delta (1.25-4 Hz), theta (4.25-8), alpha (8.25-11), sigma (11.25-15), and beta (15.25-25 Hz) frequency bands. We have extracted these measures from SWS, because slow frequency oscillations are predominant during the deepest stage of sleep. Moreover, due to technical artefacts occurring in some participants during the last third of the night (comprising mainly Stage 2 and REM sleep), we have decided to exclude the analyses of Stage 2 periods and focus exclusively on SWS.

Statistical Analysis

Statistical analyses were carried out with the Statistical Package for the Social Sciences version 22.0 (SPSS, IBM) and MATLAB (version 7.10.0.499, R2010a, The MathWorks, Inc., Natick, MA). In case of the declarative learning task, we used three measures: evening score, morning score, and memory consolidation. The latter was obtained by subtracting the evening score from the morning score (higher scores indicating reduced forgetting). In case of the non-declarative learning task, to facilitate data processing, the blocks of ASRT were organized into epochs of five blocks. The first epoch contained blocks 1–5, the second epoch contained blocks 6–10, etc. We calculated mean accuracy scores (ACCs) for all responses and median reaction times (RTs) for correct responses only; separately for high- and low-frequency triplets and for each subject and each epoch. Note that for each response (*n*), we defined whether it was a high- or a low-frequency triplet by considering whether it was more or less predictable from the event *n*-2. For the analyses reported below, as in previous research (J. H. Howard, Jr. & Howard, 1997; Nemeth, Janacsek, Londe, et al., 2010; Song et al., 2007b), two kinds of low-frequency triplets were eliminated: repetitions (e.g., 222 and 333) and trills (e.g., 212 and 343). Repetitions and trills were low frequency for all participants and people often showed pre-existing response tendencies to them (D. V. Howard et al., 2004; J. H. Howard, Jr. & Howard, 1997). By eliminating them we attempted to ensure that any high- vs. low-frequency differences are due to learning and not to pre-existing response tendencies.

For each epoch, a learning score was also calculated as the difference between triplet types in RT (RT for low-probability triplets minus RT for high-probability triplets) and ACCs (ACC for high-probability triplets minus ACC for low-probability triplets). To evaluate performance changes due to statistical learning, we conducted repeated measures analyses of variance (ANOVAs – see detailed description below) separately for accuracy and RT. 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 the measure of effect size. To investigate the offline (overnight) changes of statistical learning, we compared the ACCs and RTs from the last epoch of Session 1 (Epoch 5) and the epoch of Session 2 (Epoch 6 assessed in the morning). These variables were submitted to a repeated measures design ANOVA with TRIPLET (high- vs. low-frequency) and EPOCH (last epoch of Session 1 and the epoch of Session 2) as within-subject factors. Additionally, we subtracted the learning index of last epoch (fifth) of the evening session from the first epoch of the morning session (sixth epoch) (this way, the positive value shows overnight learning, and the negative shows forgetting) indexing memory consolidation in terms of ACC and RT.

Normality of data distribution was verified based on the kurtosis and skewness of the data as well as the Kolmogorov-Smirnov test. To study the associations between learning performance, overnight change (memory consolidation) and SWS spectral power, Pearson's correlation analyses were conducted. Spearman' s correlation coefficient was used when normality was violated. To control for the confounding factor of age (that might affect learning as well as SWS spectral power), we applied a hierarchical linear regression analysis including age as a predictor in our models.

Results

Behavioral data

Declarative memory (story recall)

First, we verified whether immediate recall (at the evening) significantly differed from morning recall. According to the paired sample t-test a significant difference emerged reflecting forgetting from evening to morning (mean evening score = 6.68, *SD* = 4.32; mean morning score = 5.46, *SD* = 4.36; $t(25) = 2.721$, $p = .011$).

Non-declarative memory (ASRT)

We conducted a repeated measure ANOVA on the 5 epochs of the first session with TRIPLET (high- vs. low-frequency) and EPOCH (1-5) as within subject factors and ACCs as the dependent variable. The main effect of TRIPLET was significant [$F(1, 24) = 43.96$, $\eta 2p = .65$, $p < .001$] indicating statistical learning, that is, higher ACCs for the high frequency triplets compared with the low-frequency ones (90.10% vs. 87.40%, respectively). The main effect of EPOCH was also significant [$F(4, 96) = 4.17$, $\eta 2p = .15$, $p = .004$], indicating that accuracy decreased across epochs (*Fig. 1/A*). The TRIPLET x EPOCH interaction showed a trend [$F(4, 96) = 2.17$, $\eta 2p = .08$, $p = .077$]: the ACC for high frequency triplets decreased less, than for low frequency triplets.

Regarding RT, we conducted a similar repeated measure ANOVA on the 5 epochs of the first session with TRIPLET (high- vs. low-frequency) and EPOCH (1-5) as within subject factors and RTs as the dependent variable. The main effect of TRIPLET was significant [$F(1,24) = 61.20$, $\eta 2p = .72$, $p < .001$], indicating statistical learning, that is, shorter RTs for high-frequency triplets compared with the low-frequency ones. The main effect of EPOCH was also significant [$F(1.98,46.78) = 73.04$, $\eta 2p = .75$, $p < .001$], due to reduced RTs across epochs, that reflects general skill learning. The TRIPLET x EPOCH interaction was not significant [$F(2.67,63.53) = 1.93$, $\eta 2p = .07$, $p = .14$], indicating that statistical learning was similar across the epochs(*Fig. 1/B*).

To investigate the offline changes of statistical learning we compared the ACCs of the last epoch of Session 1 (Epoch 5) with the ACCs of the epoch of Session 2 (Epoch 6). These variables were submitted to a repeated measures ANOVA with TRIPLET (high- vs. low-frequency) and EPOCH (last epoch of Session 1 and epoch of Session 2) as within-subject factors. The ANOVA yielded a significant main effect of TRIPLET ($F(1,22) = 56.28$, $\eta 2p = .72$, $p < .001$), indicating that, overall, participants were more accurate on high frequency triplets compared to the low frequency ones. The main effect of EPOCH was also significant ($F(1,22) = 17.80$, $\eta 2p = .45$, $p < .001$), due to more accurate responses in the morning compared to the evening session. The TRIPLET x EPOCH interaction was not significant ($F(1,22) = .12$, $\eta 2p = .005$, $p = .74$) indicating that statistical learning measured by accuracy, remained unchanged from the evening to the morning (*Fig. 1/A*).

Regarding overnight changes in RTs, we compared the RTs from the last epoch of Session 1 (Epoch 5) with the RTs of the epoch of Session 2 (Epoch 6) by a similar repeated measures ANOVA. A significant main effect of TRIPLET ($F(1,22) = 55.08$, $\eta 2p = .72$, $p < .001$) was found, indicating statistical learning, that is, RTs were shorter for high frequency triplets compared to the low frequency ones. The main effect of EPOCH was also significant ($F(1,22) = 17.18$, $\eta 2p = .44$, $p < .001$), such that RTs decreased across epochs. The TRIPLET x EPOCH interaction was not significant ($F(1,22) = 1.72$, $\eta 2p = .07$, $p = .20$), indicating that statistical learning as measured by RT, remained unchanged from the evening to the morning (*Fig. 1/B*).

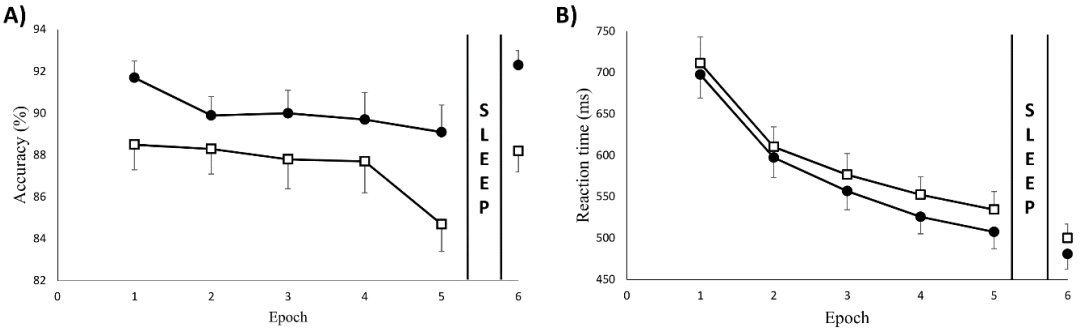

Figure 2.1 The results of statistical learning on accuracy (A) and reaction time (B) measures. Accuracy (A) and RT for correct responses (B) can be seen as a function of epoch (1-6) and trial type (high- vs. low-frequency triplets). Black circles: high-frequency triplets. White squares: low-frequency triplets. The gap between the curves indicates the statistical learning performance. Error bars indicate standard error of mean (SEM).

Associations between behavioral performance and SWS spectral power

Declarative Memory (story recall)

SWS spectral power in the delta range showed a positive correlation with the evening story recall score (r = .59, p = .001, *Fig. 2A*), whereas a negative correlation was found with the theta band (r = -.65, p < .001, *Fig. 2B*). All other frequency bands showed non-significant (*p*s > .68) correlations with the evening score. Similar correlations were found between the morning story recall score and band-wise spectral power measures (delta: r = .472, p = .02, theta: r = -.52, p = .006), all other *p*s > .38. No significant correlations were found between spectral power measures (all *p*s > .59) and overnight memory consolidation (i.e., the change in performance from evening to morning).

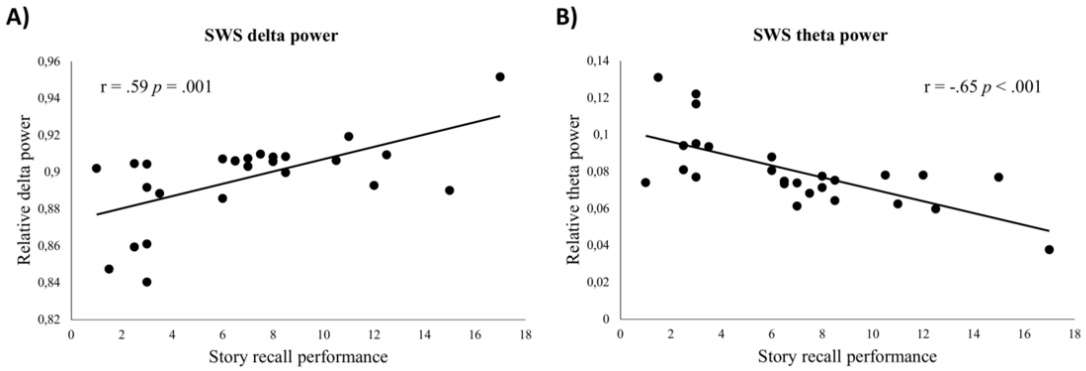

Figure 2.2 Correlation between slow wave sleep delta (A) and theta (B) power spectrum and immediate (evening) story recall performance.

To control for the confounding factor of age that might influence both memory performance and SWS, we conducted a regression analysis with evening (immediate) story recall performance as the dependent factor, and age and SWS delta spectral power as separately entered independent variables. In the first model, performance in story recall was significantly associated with age (Std. beta = .57, p = .002). In the second model where both age and delta spectral power were entered, age (Std. beta = .40, p = .018), and delta power (Std.beta = .46, p = .009) were both significant predictors of immediate story recall. We conducted the same regression analysis with evening story recall performance as dependent variable, and age and SWS theta power as separately entered independent variables. In the final model, age was not significantly associated with story recall performance (Std. beta = .29, p = .11), but theta power remained a significant predictor (Std.beta = .-53, p = .006). Both delta and theta power increased the explained variance of evening recall beyond the explained variance of age. Model parameters are detailed in *Table 2*.

Table 2.2 Linear regression models with evening story recall performance as dependent variable

Entered variables in linear regression models	Std. beta	*t* value	*p* value	Model Summary
Model 1				Adj. R^2 = .30, *p* = .002
Age	.57	3.41	.002	

Model 2				Adj. R^2 = .46, p = .009
Age	.40	2.54	.018	
SWS delta power	.46	2.88	.009	
Model 3				Adj. R^2 = .48, p = .006
Age	.29	1.67	.11	
SWS theta power	-.53	-3.06	.006	

Bin-wise correlations between story recall performance and power spectrum

To explore in more detail the oscillatory activity associated with declarative learning, we performed a post-hoc, bin-wise analyses within the delta and theta range in relation to evening memory performance. As plotted in *Fig. 3*, the spectral power in 1.25-1.5 Hz frequencies was positively, whereas frequency bins between 4-7 Hz were negatively associated to evening recall.

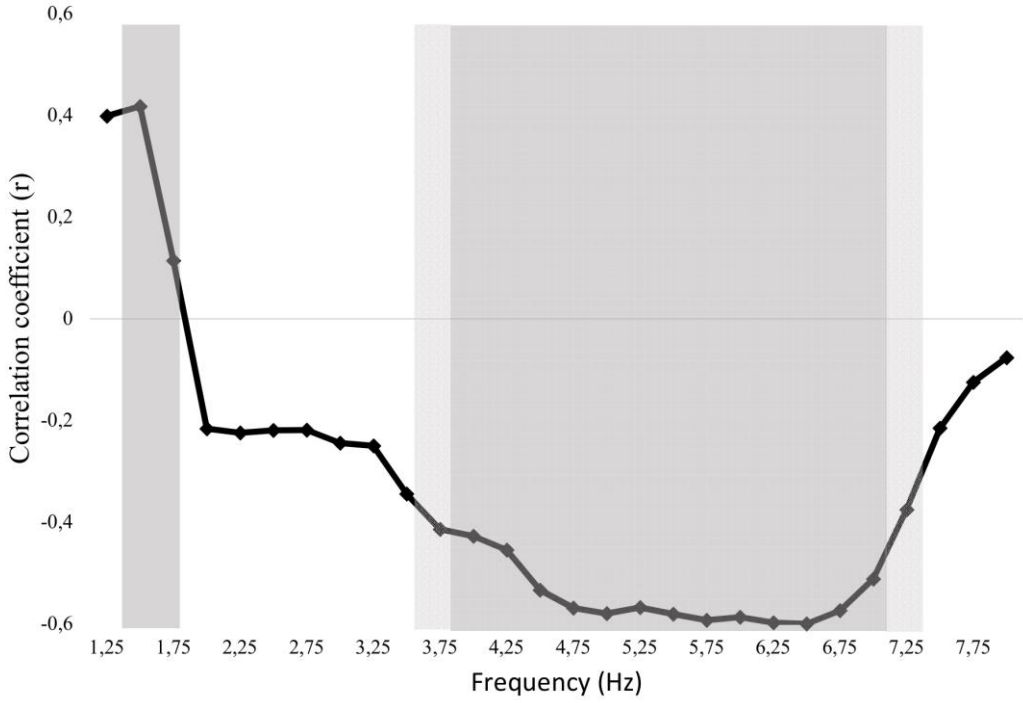

Figure 2.3 Bin-wise correlation coefficients between 1 and 8 Hz spectral power and evening story recall performance. Gray background illustrates statistically significant ($p < .05$) correlations, light gray background illustrates trend ($p < .1$).

Non-declarative memory (ASRT)

SWS spectral power measures were not associated with the statistical learning score in the evening (based on the last, fifth epoch) (all *p*s > .22), or in the morning session (all *p*s > .41) in terms of ACC. Moreover, spectral power measures were not associated with overnight consolidation (all *p*s > .25) of statistical learning (overnight change in ACC). Similarly, no significant correlations emerged between statistical learning performance in the evening (all *p*s > .25), or in the morning session (all *p*s > .11) in terms of RT, and spectral power measures were not associated with overnight consolidation (all *p*s > .28) (overnight change in RT).

Unlike statistical learning, SWS spectral power measures were associated with general skill learning in case of ACCs. Similarly to story recall, SWS spectral power in the delta range showed a positive correlation with the average ACCs (averaged across high- and low-frequency triplets) assessed in the evening (based on the last, fifth epoch, r = .44, p = .028), whereas a negative correlation was found with theta band power (r = -.433, p = .03). All other frequency bands showed non-significant (*p*s > .45) correlations with the average ACCs in the evening. Similarly, although stronger correlations were found between the morning ACCs and band-wise spectral power measures (delta: r = .658, p = .001; theta: r = -.668, p < .001, all other *p*s > .47). No significant correlations were found between spectral power measures (all *p*s > .25) and overnight change in average ACCs (i.e., consolidation of general skill learning).

In case of general skill learning indexed by averaged RTs for high- and low-frequency triplets, no significant correlations emerged between skill learning and spectral power (all *p*s > .10). Neither we found significant correlations between the overnight RTs change and spectral power measures, although theta band power correlated with overnight change on a trend level (r = -.391, p = .07, all other *p*s > .12).

Similarly to story recall, we controlled for the confounding factor of age that might influence both memory performance and SWS. First, we conducted a regression analysis with average evening ACCs as the dependent factor, and age and SWS delta spectral power as separately entered independent variables. In the first model, ACCs was significantly associated with age [Std. beta = .51, p = .009; Adj. R^2 = .23, $F(1,23)$ = 8.24 p = .009]. In the second model, the influence of age remained significant [Std. beta = .38, p = .05], but delta power was not a significant predictor [Std.beta = .29, p = .14] of ACCs.

This model was also significant [Adj. R^2 = .27, $F(2,24) = 5.51$ p = .011], but the R^2 change (.07) was not significant [$F(1,22) = 2.31$, p = .14] indicating that the inclusion of delta power as a predictor did not significantly improve the model. We conducted the same regression analysis with average evening ACCs as dependent variable, and age and SWS theta power as separately entered independent variables. In the third model where both age and theta spectral power were entered, neither age (Std. beta = .36, p = .11), nor theta power (Std.beta = -.26, p = .24) were significant predictors of ACCs. This model was also significant (Adj. R^2 = .25, F (2,24) = 4.93 p = .017), but the R^2 change (.05) was not significant (F (1,22) = 1.46, p = .24), indicating that the inclusion of theta power as a predictor did not significantly improve the model.

Analysis of the primary snoring subjects

To verify whether the above correlations were not produced due to impaired learning specifically within the OSA (n = 4) subgroup, we performed the same analyses based on the data of the primary snoring subgroup only (n = 23). The exclusion of the OSA patients did not modify our results in case of the declarative and the non-declarative learning task. Delta power positively (r = .62, p = .002) and theta power negatively (r = -.67 p = .001) correlated with declarative learning capacity and general skill learning (delta range: r = .59, p = .004; theta range: r = - .47, p = .03). Whereas the associations in case of declarative learning were significant beyond the influence of age, the correlations between SWS spectral power and skill learning were not significant after controlling for age. (see the *Supplemental Material* for a detailed description).

Discussion

The principal aim of this study was to examine the associations between SWS-specific oscillatory activity and memory consolidation within a group of children with SDB. Inter-individual variability of post-learning, night-time SWS spectral power did not predict overnight changes in performance either in case of a declarative or in a non-declarative learning task. Whereas no associations were found between SWS spectral power and indices of memory consolidation, delta and theta power were associated with declarative learning capacity. Delta power during post-learning SWS was positively

associated with short- and long-term memory retention, assessed immediately after encoding, and after a night-time sleep, respectively. On the other hand, faster oscillatory activity, indexed by the theta range was a negative correlate of short-and long-term memory performance. Given that both memory performance and SWS spectral power might be substantially influenced by cortical maturation, we also considered the effects of age. The associations between SWS power and declarative memory performance remained significant and accounted for a large portion of the variance (16 % for delta and 18 % for theta) beyond the effects of age. In contrast, non-declarative statistical learning was not associated with SWS spectral power measures.

Our results indicate that slow frequency activity, in particular oscillations around 1 Hz are associated with better declarative learning capacity, whereas higher frequency activity between 4 and 7 Hz correlate with poorer performance among children with SDB. Two earlier studies (Jussila et al., 2016; Kheirandish-Gozal et al., 2007) reported attenuated slow frequency activity in children with SDB. Abnormal respiratory patterns could result in subtle changes in sleep physiology that might not be revealed by conventional macrostructural measures. Our findings suggest that the predominance of slow frequency (~ 1 Hz) activity, as well as the reduction of faster (4-7 Hz) theta oscillations during SWS reflect better memory performance in children with SDB. Slow frequency activity of NREM sleep, quantified by the CAP A1 was consistently linked to better cognitive outcomes in healthy adults (Arico et al., 2010; Drago et al., 2011; Ferri et al., 2010) and children (Bruni et al., 2012). Given that slow frequency oscillations (with spectral power between 0.25 and 2.5 Hz) are the main contributors of the visually detected CAP A1 subtypes (Ferri, Bruni, Miano, & Terzano, 2005), our findings indicating better declarative memory performance in relation to slower, and worse performance associated with faster frequencies, are in line with the concept of slow oscillations during SWS as sensitive biomarkers of healthy cognition (Tononi & Cirelli, 2006) or even neurodegeneration (Maestri et al., 2015). A large number of studies linked slow oscillations (~ 1 Hz) to sleep-dependent memory consolidation (for an extensive review see (Rasch & Born, 2013); moreover, reduced increase in post-training SWS seems to be associated with impaired declarative memory consolidation in adults with OSA (Guo et al., 2013).

Nevertheless, in our sample SWS spectral power was not associated with overnight changes in performance, but only with general learning capacity. This finding might suggest that the associations between SWS spectral power and declarative learning are driven by trait-dependent variance, instead of state-like effects of sleep on memory reprocessing. Such trait-dependent associations between cognitive measures and sleep-specific oscillations were mainly reported for sleep spindles (Bódizs et al., 2005; Lustenberger, Maric, Durr, Achermann, & Huber, 2012; Ujma et al., 2014) but also for slow oscillations in case of parahippocampal-hippocampal recordings (Bódizs, Békésy, Szűcs, Barsi, & Halász, 2002). Although trait-dependent aspects might account for our findings, associations between SWS power and memory performance could also be driven by learning-induced changes in EEG oscillations, as the expression of nocturnal slow frequency activity is particularly sensitive to previous learning experience (Molle, Marshall, Gais, & Born, 2004; Tononi & Cirelli, 2006). Therefore, state-like and trait-like effects in this study cannot be clearly discerned and should be explored in further investigations.

Whereas declarative learning was related to spectral power measures of SWS, non-declarative statistical learning and overnight change in performance were not associated with SWS-specific oscillations. This finding coheres with earlier studies indicating that non-declarative statistical learning assessed by the ASRT does not benefit from sleep (Nemeth et al., 2012; Nemeth, Janacsek, Londe, et al., 2010). More specifically, statistical learning did not produce off-line improvements in young and old participants and was not influenced by sleep (Nemeth, Janacsek, Londe, et al., 2010). Furthermore, adults diagnosed with OSA (Csabi, Varszegi-Schulz, Janacsek, Malecek, & Nemeth, 2014; Nemeth et al., 2012) as well as children with SDB (Csábi et al., 2013) does not seem to exhibit impaired non-declarative learning, suggesting that statistical learning captured by the ASRT is independent of the influence that sleep might have on cognitive functions. Although others reported sleep-dependent behavioural and neurophysiological effects (sleep-dependent memory consolidation) in case of similar probabilistic learning tasks (Durrant, Cairney, & Lewis, 2013; Durrant, Taylor, Cairney, & Lewis, 2011; Urbain et al., 2013), these tasks differ in their methodology and presumably, also in the underlying neural networks (Durrant et al., 2013; Durrant et al., 2011; Janacsek, Ambrus, Paulus,

Antal, & Nemeth, 2015; Nemeth, Janacsek, Király, et al., 2013; Urbain et al., 2013) that subtend them.

Moreover, statistical learning within the ASRT task is implicit, and occurs without explicit awareness (Nemeth, Janacsek, Londe, et al., 2010; Song et al., 2007b). Several studies indicate, that sleep-related benefits of memory consolidation are restricted to skill-learning paradigms that require attention, intentional learning (Wilhelm et al., 2011), explicit (verbally accessible) representations of the sequence structure (Robertson, Pascual-Leone, & Press, 2004; Song & Cohen, 2014), that are clearly not present in the ASRT task (J. H. Howard, Jr. & Howard, 1997).

General skill learning in terms of accuracy, but not consolidation of skill learning was positively related to delta and negatively to theta power in SWS (*see Supplementary Material*), resembling the association found in case of declarative learning. This finding might be explained by at least partly overlapping cognitive processes underlying declarative learning and ACC performance measures. It has been previously shown that declarative learning is highly reliant on controlled, attention-dependent cognitive processes (Eichenbaum, 2000). Similarly, accuracy performance measures have been suggested to rely on controlled, selective attentional processes to some extent (Prinzmetal, McCool, & Park, 2005). Nevertheless, the association between general skill learning and SWS spectra was explained by age, indicating that both ACC-related processes (Janacsek et al., 2012) and SWS activities (Buchmann et al., 2011) undergo robust age-related changes within this age range.

Some limitations of this study should be considered. First of all, although slow frequency oscillations were associated with declarative learning in our sample, we do not know if this correlation is specific to children with SDB, since we did not have a healthy control group. Given that we performed this study within the frames of a clinical evaluation, due to ethical and technical reasons, we did not include a baseline night without pre-sleep learning experience. Although the associations between delta/theta power and learning capacity suggest a trait-like effect, trait-dependent and state-dependent effects cannot be differentiated since learning experience might also influence oscillatory activity of post-learning SWS. Our analyses focused on spectral power specifically during SWS, due to the predominance such oscillations during that sleep stage. Spectral activity during Stage 2 sleep might have also contributed to our analyses,

however, due to a large number of technical artefacts in some participants during the last third of the night (comprising mainly Stage 2 and REM sleep), we have decided to focus exclusively on SWS sleep.

In spite of these limitations, this study indicates that among children with SDB, slow frequency oscillations within the delta and theta band during SWS are related to declarative learning capacity, but are independent of non-declarative, statistical learning. These preliminary findings emphasize the relevance of oscillatory activity of SWS on specific cognitive processes, and contribute to the characterization of cognitive functions and deficits of children with SDB. Future studies should further characterize which memory systems are specifically affected by fragmented sleep, and disentangle trait-dependent and state-dependent aspects of the interrelations between sleep and cognitive performance.

Frontal-midline theta frequency and probabilistic learning: A transcranial Alternating Current Stimulation study[11]

Abstract

Probabilistic learning is a fundamental cognitive ability that extracts and represents regularities of our environment enabling predictive processing during perception and acquisition of perceptual, motor, cognitive, and social skills. Previous studies show competition between neural networks related to executive function/working memory vs. probabilistic learning. Theta synchronization has been associated with the former while desynchronization with the latter in correlational studies. In the present paper our aim was to test causal relationship between fronto-parietal midline theta synchronization and probabilistic learning with non-invasive transcranial alternating current (tACS) stimulation. We hypothesize that theta synchronization disrupts probabilistic learning performance by modulating the competitive relationship. Twenty-six young adults performed the Alternating Serial Reaction Time (ASRT) task to assess probabilistic learning in two sessions that took place one week apart. Stimulation was applied in a double-blind cross-over within-subject design with an active theta tACS and a sham stimulation in a counter-balanced order between participants. Sinusoidal current was administered with 1 mA peak-to-peak intensity throughout the task (approximately 20 minutes) for the active stimulation and 30 seconds for the sham. We did not find an effect of fronto-parietal midline theta tACS on probabilistic learning comparing performance during active and sham stimulation. To influence probabilistic learning, we suggest applying higher current intensity and stimulation parameters more precisely aligned to endogenous brain activity for future studies.

Keywords: statistical learning, transcranial electric stimulation, procedural learning, neural oscillations, competition

[11] Zavecz, Z., Horváth, K., Solymosi, P., Janacsek, K., & Nemeth, D. (2020). Frontal-midline theta frequency and probabilistic learning: A transcranial Alternating Current Stimulation study. *Behavioural Brain Research*, 112733.

Introduction

Probabilistic learning (often referred to as statistical learning as well) is a fundamental cognitive ability that underlies automatic behaviors and skills, such as motor, linguistic or social skills and habits (C. Conway, Bauernschmidt, Huang, & Pisoni, 2010; Fiser & Aslin, 2001; Frost, Armstrong, & Christiansen, 2019; Lieberman, 2000; Poldrack & Foerde, 2008; Pothos, 2007; Ullman, 2015). It facilitates the extraction of statistical regularities from the environment and enables predictions of environmental events. Several studies discussed the neural background of probabilistic learning using functional magnetic resonance imaging (fMRI) (Schapiro, Kustner, & Turk-Browne, 2012; Stillman et al., 2013; Turk-Browne, Scholl, Chun, & Johnson, 2009), magnetoencephalography (MEG) (Paraskevopoulos, Chalas, & Bamidis, 2017), electroencephalography (EEG) (Kóbor et al., 2018; Tóth et al., 2017) or neuropsychology (Janacsek, Borbély-Ipkovich, Nemeth, & Gonda, 2018; Nemeth, Janacsek, Balogh, et al., 2010; P. J. Reber, 2013; Takács et al., 2018). However, these studies used correlational methods only. In the present paper our aim was to test the causal relationship between brain activity and probabilistic learning by directly manipulating oscillatory activity with non-invasive electric brain stimulation.

Oscillatory synchronization is a fundamental mechanism for information transmission between neural populations and for forming larger networks (Fries, 2005; Salinas & Sejnowski, 2001; Singer, 1993). For instance, theta (4-7 Hz) activity was consistently observed particularly within the fronto-midline areas during working memory and declarative memory tasks (Gevins et al., 1997; Hsieh & Ranganath, 2014; Jensen & Tesche, 2002; Meyer et al., 2015; Onton et al., 2005; Scheeringa et al., 2009; Summerfield & Mangels, 2005; Tóth et al., 2014). Tóth et al. (Tóth et al., 2017) showed in an EEG study that theta activity was correlated with probabilistic learning as well: weaker phase synchronization in theta frequency was associated with better learning performance. Thus, in contrast to declarative and working memory, in theta frequency, desynchronization, and not synchronization seems to be beneficial for probabilistic learning. This is in line with the competition framework in which there is an antagonistic relationship between fronto-hippocampal and striatal networks and related functions such as working and declarative memory vs. probabilistic and sequence learning (Albouy et

al., 2015; Albouy et al., 2008; Ashby & O'Brien, 2005; Daw et al., 2005; Poldrack et al., 2001).

A possible method to test causal relationships between brain networks and cognitive performance is brain stimulation. Transcranial Alternating Current Stimulation (tACS) is a suitable method to influence oscillatory brain activity (Antal et al., 2008; Antal & Paulus, 2013). Based on the above presented evidence for the role of theta frequency in prefrontal-dependent processes (including working memory) and the antagonistic relationship of these processes with probabilistic learning (Ambrus et al., 2019; Filoteo, Lauritzen, & Maddox, 2010; Janacsek et al., 2012; Nemeth, Janacsek, Polner, et al., 2013; Virag et al., 2015), we hypothesized that induced theta synchronization is detrimental for probabilistic learning. Thus, in the present paper, we used a frontal-midline theta frequency tACS stimulation to disrupt probabilistic learning.

Methods

Participants

Twenty-six young adults (19 females) were selected from a large pool of undergraduate students from the Eötvös Loránd University in Budapest (M_{Age} = 21.38 years, SD = 1.52 years; $M_{Years\ of\ education}$ = 14.46 years, SD = 1.45 years). Participants had no previous history of neurological, psychiatric or cardiovascular disorders, brain injuries and they had no metal implants in the head or neck area. They reported not taking any substances that affect the nervous system. All participants completed all sessions: two sessions with different stimulation conditions (sham vs. active stimulation) during the probabilistic learning task and an additional session for other neuropsychological tests. They were naïve regarding the exact purpose of the study and did not know in which session they were assigned to receive active or sham stimulation. Participants gave written and verbal informed consent before participating and received course credits for taking part in the experiment. The experiment was in accordance with the guidelines of the Declaration of Helsinki, and was approved by the ethics committee of the Eötvös Loránd University, Budapest, Hungary (identifier: 2016/120).

Experimental Design

This study utilized a within-subject, cross-over design consisting of two stimulation sessions: 1 mA active tACS stimulation and sham stimulation (Fig. 1). These sessions took place one week apart from each other, starting at the same time of the day to eliminate time-of-day effects. The order of the sessions was counterbalanced across participants, and the stimulation was double-blinded. Therefore, neither the main investigator nor the participant was aware of the current stimulation condition. A second investigator who was not involved in the interaction with participants was responsible for setting the stimulation only. The stimulation was administered simultaneously with the probabilistic learning task (Alternating Serial Reaction Time, ASRT task). In the two sessions, participants learned two different, partly overlapping sequences. The overlap was controlled across participants (see Probabilistic learning section in Tasks for details).

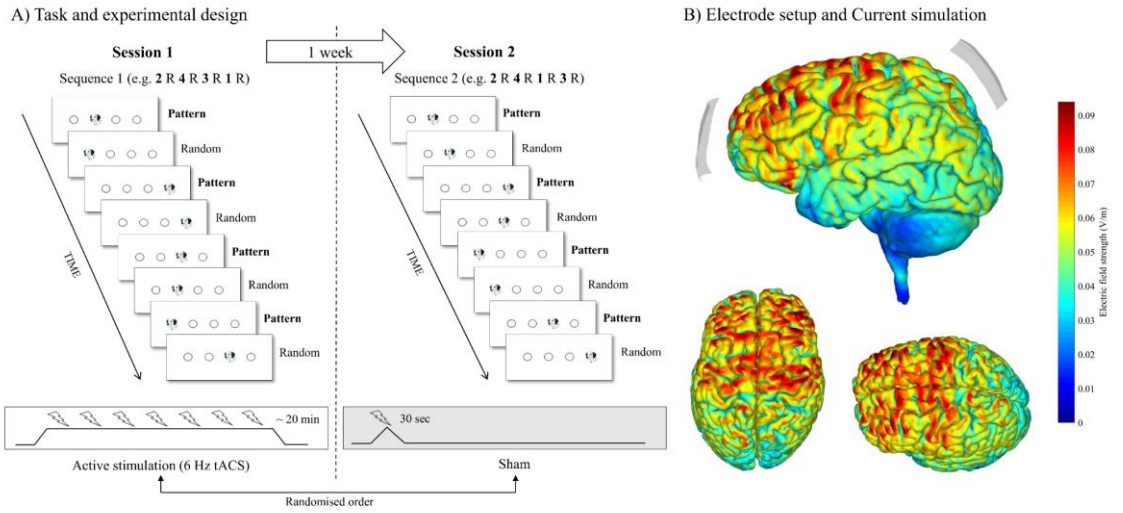

Figure 3.1 Overview of the experimental design and stimulation parameters. A) Task and experimental design. The stimulation was carried out in a double-blind, placebo-controlled crossover design. Healthy young adults participated in two sessions (one week apart) during which they received 1 mA active theta frequency tACS stimulation, or sham stimulation in a counterbalanced order. Active tACS stimulation was administered throughout the task (approximately 20 minutes), while sham stimulation lasted only 30 seconds. In both cases there were 30 seconds ramp up and ramp down periods. Participants completed the Alternating Serial Reaction Time (ASRT) task both times to assess probabilistic learning performance. In this task, pattern elements alternate with random ones, constituting a probabilistic sequence, in which some runs of three consecutive trials ("triplets") occur more frequently than others. We refer to probabilistic learning as a performance difference between high-probability compared to low-probability triplets. Participants learned two different probabilistic sequences during the two sessions. B) Electrode setup and current simulation. A battery driven constant current stimulator delivered a sinusoidal alternating current

stimulation to the participant's scalp via two 5 cm × 5 cm electrodes placed over positions Fpz and Pz according to the international 10-20 system. TACS was applied at a peak-to-peak current intensity of 1 mA oscillating at 6 Hz. To model tACS, we performed a simulation on a template head model by using a free software package called Simulation of Non-invasive Brain Stimulation (for details, see section 'Transcranial Alternating Current Stimulation (tACS)' in the main text). The spatial distribution of the absolute electric field magnitudes in the gray-matter compartment is in mV/mm. We used a robust maximum (99.9th percentile) of the absolute values for the scale limit. Lateral (top), top (bottom left) and superior lateral (bottom right) views are presented. The mean and maximal electric field strength of the robust maximum in the frontal, paracentral (pre- and post-central and central gyri and sulci) and parietal (superior gyri and sulci) regions were 0.088, 0.096, 0.083, 0.093, 0.072, 0.074 V/m respectively.

Tasks

Probabilistic learning - The Alternating Serial Reaction Time (ASRT) task (J. H. Howard, Jr. & Howard, 1997; Nemeth, Janacsek, Londe, et al., 2010) was used to measure probabilistic sequence learning. In this task, a stimulus (a dog's head) appeared in one of the four empty circles on the screen, and participants had to press the corresponding button as fast and as accurately as possible (Fig. 1A). The target remained on the screen until the participant pressed the correct button. The response-to-stimulus interval (RSI) was 120 ms. The computer was equipped with a special keyboard with four marked keys (Z, C, B and M on a QWERTY keyboard), each corresponding to one of the horizontally aligned circles. The ASRT task consisted of 20 blocks, with 85 trials per block. The first five stimuli were random for practice purposes, then an eight-element alternating sequence 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). As one block took 1-1.5 min, the whole task took approximately 20-25 min.

Due to the alternating sequence in the ASRT task, some triplets or runs of three consecutive events are more probable (*high-probability triplets*) than others (*low-probability triplets*). For example, in the abovementioned sequence (2-R-4-R-3-R-1-R), 2-X-4 is a high-probability triplet (where X denotes to any of the four possible positions), since the first and the third elements can either be a pattern or a random stimulus. However, 2-X-1, 2-X-2, and 2-X-3 are low-probability triplets, since the first and the

third elements can only be a random stimulus. Therefore, for analyzing the data we determined whether each trial was the last element of a high-probability or a low-probability triplet. Note that in this way, we determine the probability of each triplet throughout the task in a sliding window manner (i.e., one stimulus is the last element of a triplet, but also the middle and the first element of the consecutive triplets). The high-probability triplets are five times more predictable than the low-probability triplets. Therefore, the last element of a triplet is more predictable in high-probability triplets compared to low-probability ones. Previous studies have shown that as people practice the ASRT task, they come to respond more quickly and more accurately to the high-probability triplets compared to low-probability triplets, revealing probabilistic learning (Howard et al., 2004; Howard and Howard, 1997b; Janacsek, Fiser, and Nemeth, 2012; Nemeth et al., 2010; Song, Howard, and Howard, 2007).

The ASRT task was performed in two sessions during the experiment, with 20 blocks in each session. For this, pairs of sequences were created, where the two sequences shared 2 position orders out of the 4 (e.g., **2**-R-**4**-R-3-R-1-R and **2**-R-**4**-R-1-R-3-R, see Fig. 1A) which results in a 25% overlap in high-probability triplets between the sequences. One of these pairs of sequences was randomly assigned to each participant to keep constant the overlap in the two sequences amongst participants.

Finally, it is important to note that participants were unaware of the underlying alternating sequence structure, thus they acquired the probabilistic regularities incidentally and that knowledge remained implicit throughout the task. This was confirmed using a short questionnaire (Nemeth, Janacsek, Londe, et al., 2010; Song et al., 2007b) after the second stimulation session. The questionnaire included the following two increasingly specific questions: “Have you noticed anything special regarding the task?”, “Have you noticed some regularity in the sequence of stimuli?”. The experimenter rated subjects' answers on a 5-point scale where 1 denoted “Nothing noticed” and 5 denoted “Total awareness”. None of the participants reported noticing regularities in the ASRT task.

Transcranial Alternating Current Stimulation (tACS)

A commercial, battery driven constant current stimulator (DC-Stimulator Plus, NeuroConn, Ilmenau, Germany) delivered a sinusoidal alternating current stimulation to

the participant's scalp via two 5 cm × 5 cm electrodes. The electrodes were covered with a thin layer of electrode gel and were placed over positions Fpz and Pz according to the international 10-20 system (Fig. 1B). This frontal-midline electrode montage choice was based on a previously reported stimulation design (Chander et al., 2016). Impedances were kept below 30 kΩ (average impedance was 8.25±3.83 kΩ). TACS was applied at a peak-to-peak current intensity of 1 mA oscillating at 6 Hz. While recent papers suggest using higher current intensity (Vöröslakos et al., 2018), these intensities can cause intense discomfort. In our study, to ensure that all participants complete both sessions and to maintain blindness of the participants to the stimulation settings, we decided to use a smaller current intensity that was proven successful in previous studies (Ambrus et al., 2015; Polanía, Nitsche, Korman, Batsikadze, & Paulus, 2012; Wischnewski, Zerr, & Schutter, 2016). To avoid possible discomfort during the onset of tACS, the stimulation current was gradually ramped up from 0 to 1.0 mA over a period of 30 s. After the 30 s ramp up, the stimulation intensity was maintained for the length of the task (approximately 20 minutes) in case of the active stimulation condition. To control for tACS-unspecific effects (such as fatigue and beliefs of the participant), there was a sham (placebo) stimulation condition, consisting of 30 s of stimulation following the 30 s ramp up. In both conditions there was a 30 s ramp down period after the stimulation.

To model tACS, we performed a simulation on a template head model by using a free software package called Simulation of Non-invasive Brain Stimulation (SimNIBS; version 2.1.2, Fig. 1B). SimNIBS generates anatomically realistic, multi-compartment head models from structural magnetic resonance imaging by using the finite element method. The head mesh entailed ca. 3,500,000 tetrahedral elements and five compartments. We used standard, isotropic conductivity values for the compartments, all values are expressed in S/m: white matter = 0.126; gray matter = 0.275; cerebrospinal fluid = 1.654; bone = 0.01; scalp = 0.465; eyes = 0.5; silicon rubber electrode = 29.4; conductive medium = 1.0. The physical dimensions of both electrodes were 50 × 50 mm and 4 mm thick. The thickness of the conductive medium was set to 2 mm. The electric field was modeled by using 0.5 mA peak to baseline intensities. To quantify the strength of the induced electric field in particular brain areas, we used the parcellation of human cortical gyri and sulci proposed by Destrieux, Fischl, Dale, and Halgren (2010). We computed the mean and maximal electric field strength of the robust maximum (99.9th

percentile) in the following regions of interest (ROIs): frontal (superior, middle and orbital gyri and sulci), paracentral (pre- and post-central and central gyri and sulci), and parietal (superior gyri and sulci). The electric field strength was $Mean_{max}$ = 0.088 V/m, Max_{max} = 0.096 V/m in the frontal, $Mean_{max}$ = 0.083 V/m, Max_{max} = 0.093 V/m in the paracentral and $Mean_{max}$ = 0.072 V/m, Max_{max} = 0.074 V/m in the parietal regions.

Statistical analysis

Statistical analyses were carried out with the Statistical Package for the Social Sciences version 22.0 (SPSS, IBM) and JASP Version 0.11.1 (Team, 2019). To facilitate data processing, the blocks of ASRT were organized into four epochs of five blocks in each session. The first epoch contained blocks 1–5, the second epoch contained blocks 6–10, etc. We calculated mean accuracy scores (ACCs) for all responses and median reaction times (RTs) for correct responses only, separately for high- and low-probability triplets and for each subject and each epoch. Note that for each trial we defined whether it was the last element of a high- or a low-probability triplet. Two kinds of low-probability triplets were eliminated from the analysis: repetitions (e.g., 222 and 333) and trills (e.g., 212 and 343), as people often showed pre-existing response tendencies to them (Howard and Howard, 1997a; Howard et al., 2004, Howard and Howard, 1997a; Nemeth et al., 2010; Song et al., 2007).

Overall RTs significantly differed between the two sessions (as revealed by the significant main effect of SESSION in the repeated-measures ANOVA on RTs with SESSION (First vs. Second), EPOCH (1-4) and TRIPLET TYPE (High vs. Low) as within-subject factors: $F(1, 25) = 39.510$, $p < .0001$, $\eta^2_P = .612$): participants were faster when completing the task for the second time (M_{RT} = 369.70, SEM = 5.31, M_{RT} = 336.20, SEM = 5.29 for the first and the second session, respectively). Therefore, we calculated z-scores within each subject in each session to eliminate the effects of different baseline speeds when comparing performance between the two sessions. A similar ANOVA computed on accuracy data revealed no significant difference between the two sessions (main effect of SESSION: $F(1, 25) = 0.376$, $p = .545$, $\eta^2_P = .015$).

For each epoch, we calculated learning scores both for RT and ACC data. For RT, the learning score was calculated as the difference between the z-transformed RTs for low-probability triplets minus the z-transformed RTs for high-probability triplets. For

ACC, the learning score was calculated as the raw ACCs for high-probability triplets minus the raw ACCs for low-probability triplets. In both cases higher learning scores indicated better learning. To evaluate changes in probabilistic learning as a function of stimulation, we conducted mixed-design analyses of variance (ANOVAs) separately for the RT and ACC learning scores with STIMULATION (Sham vs. Active) and EPOCH (1-4) as within-subject factors and ORDER (Sham first vs. Stimulation first) as a between-subject factor. We included the ORDER between-subject factor to ensure that the order in which participants received sham and active stimulation did not influence the effects of stimulation. Greenhouse–Geisser epsilon (ε) correction was used when necessary. Original df values and corrected p-values (if applicable) are reported together with partial eta-squared (η^2_P) as the measure of effect size.

Furthermore, as suggested by Biel and Friedrich (2018) we conducted the same mixed-design ANOVAs separately for the RT and ACC learning scores with STIMULATION and EPOCH as within-subject factors and ORDER as a between-subject factor with a Bayesian approach as well. The Bayesian ANOVA contrasts the predictive performance of competing models instead of F-tests of main effects and interactions (Rouder, Engelhardt, McCabe, & Morey, 2016). Models were compared using BF_{10}, which quantifies the evidence in favor of each model relative to the best model in the respective comparison. To summarize the importance of the within-subject factors across all models, we also performed model averaging, which provides us with evidence for inclusion for main effects and interactions ($BF_{inclusion}$). The inclusion Bayes factor quantifies the change from prior inclusion odds to posterior inclusion odds and can be interpreted as the evidence in the data for including a predictor.

To ensure that the partially overlapping sequence in the task between the two sessions did not distort the effects of the stimulation, we recomputed learning scores excluding the responses (RT and ACC) to those triplets that were high-probability in both sessions and ran frequentist and Bayesian repeated-measures ANOVAs on these modified RT and ACC learning scores over time and stimulation (see section '*Does the partial overlap between the sequences practiced during the two stimulation sessions influence the effects of the stimulation?*' and Fig. S1 in the Supplementary results). Importantly, the results after the elimination of the overlapping high-probability triplets are identical to the results without the elimination of these triplets and are not discussed further in the main text.

Lastly, as a post-hoc analysis we investigated the effects of baseline performance on the stimulation. We ran four additional mixed-design ANOVAs (both frequentist and Bayesian) including a between-subject factor for good vs. poor initial/baseline performance in four measures of ASRT (average reaction times, reaction time learning scores, average accuracy, accuracy learning scores) on the learning scores over time and stimulation (see section '*Does baseline performance influence the effects of the stimulation?*' in the Supplementary results). We did not find a differential effect of the stimulation in good vs. poor performers based on initial speed, accuracy, RT or ACC probabilistic learning.

Results

Do RT learning scores differ between stimulation conditions?

The frequentist mixed-design ANOVA on the z-transformed RT learning scores revealed a significant Intercept ($F(1, 24) = 66.277$, $p < .001$, $\eta^2_P = .734$), suggesting that learning occurred in the ASRT task. The main effect of EPOCH was also significant ($F(3, 72) = 6.663$, $p < .001$, $\eta^2_P = .217$), indicating that the learning scores increased throughout the task, independent of the stimulation condition (Fig. 2A). However, we did not find any significant differences between the active stimulation and sham conditions either in overall learning (main effect of STIMULATION: $F(1, 24) = 0.093$, $p = .763$, $\eta^2_P = .004$) or in the time course of learning (STIMULATION * EPOCH interaction: $F(3, 72) = 0.637$, $p = .593$, $\eta^2_P = .026$). The order of the stimulation sessions did not seem to affect the overall learning scores (main effect of ORDER: $F(1, 24) = 2.345$, $p = .139$, $\eta^2_P = .089$), the trajectory of the learning scores (ORDER x EPOCH interaction: $F(3, 72) = 0.048$, $p = .986$, $\eta^2_P = .002$), the effect of stimulation (ORDER x STIMULATION interaction: $F(1, 24) = 0.974$, $p = .333$, $\eta^2_P = .039$) or the trajectory of the learning scores during the two stimulation conditions (ORDER x EPOCH x STIMULATION interaction: $F(3, 72) = 0.627$, $p = .600$, $\eta^2_P = .025$).

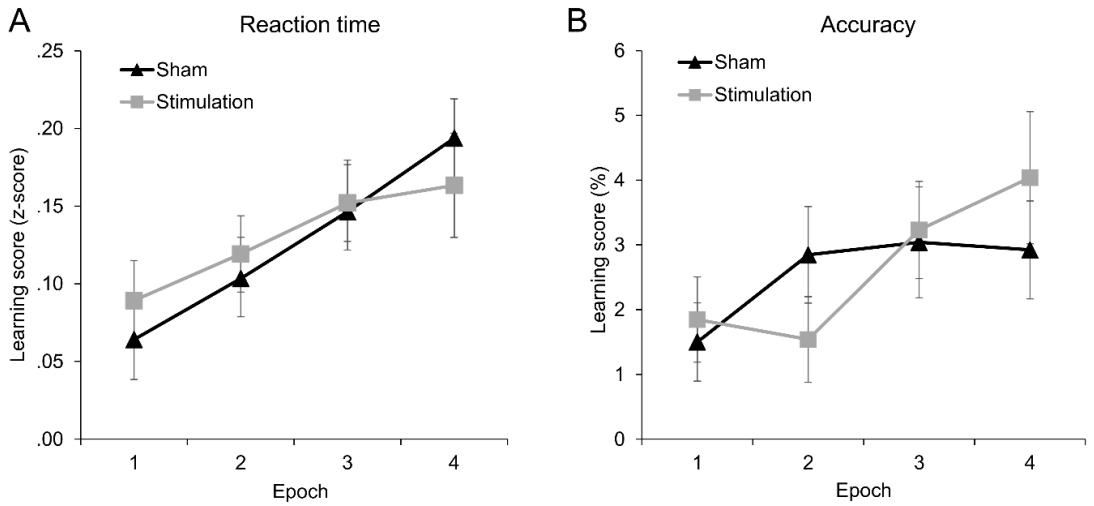

Figure 3.2 Probabilistic learning in terms of reaction times (A) and accuracy (B) in the active stimulation vs. sham conditions across the four epochs of the ASRT task. There was no significant difference between the active stimulation in theta frequency (grey squares) and sham (black triangles) conditions either in overall learning or in the time course of learning. Error bars indicate the Standard Error of Mean (SEM).

The analysis of effects (model-averaged results) of the Bayesian mixed-design ANOVA on the z-transformed RT learning scores showed that the main effect of Epoch should be included in the model ($BF_{inclusion}$ = 74.684), while the effects related to the Stimulation and the Session order should not (all $BF_{inclusion}$ < 1, Table 1). Thus, based on the Bayesian analysis of effects, the learning scores changed throughout the task, but they were independent of the stimulation condition or the order of the stimulation.

Table 3.1 Model-averaged results of Bayesian ANOVA for RT learning scores

Effects	**P(incl)**	**P(incl|data)**	**BF inclusion**
Stimulation	0.737	0.170	0.073
Epoch	0.737	0.995	74.684
Order	0.737	0.437	0.278
Stimulation ∗ Epoch	0.316	0.014	0.030
Stimulation ∗ Order	0.316	0.030	0.067
Epoch ∗ Order	0.316	0.023	0.051
Stimulation ∗ Epoch ∗ Order	0.053	1.564e -5	2.816e -4

Note: The Effects column denotes predictors of interest, the column P(incl) shows the prior inclusion probability, P(incl | D) shows the posterior inclusion probability, and $BF_{Inclusion}$ shows the inclusion Bayes factor.

As our primary interest was the effect of the stimulation on probabilistic learning and the number of models was too high with the ORDER between-subject factor, as well as there was no evidence to include that factor, we recomputed the Bayesian ANOVA with only the STIMULATION and EPOCH as within-subject factors. Based on this Bayesian ANOVA, the best model for our data was with only the main effect of Epoch (Table 2). This model with the main effect of Epoch was ~6.5 times more likely than any model including the effect of the Simulation. Altogether the Bayesian ANOVA for the RT learning scores provides evidence for the model with only the main effect EPOCH to explain best our data. This suggests that while the learning scores changed during the task, this was independent of the stimulation condition and the order of the stimulation condition.

Table 3.2 Bayesian model comparisons for RT learning scores

Models	**P(M)**	**P(M|data)**	**BF $_M$**	**BF $_{10}$**	**error %**
Epoch	0.200	0.853	23.178	1.000	
Stimulation + Epoch	0.200	0.129	0.595	0.152	1.559
Null model	0.200	0.013	0.052	0.015	2.618
Stimulation + Epoch + Stim. ✻ Epoch	0.200	0.004	0.017	0.005	0.610
Stimulation	0.200	0.0006	0.003	0.0008	0.978

Note: All models include Subject. The Model column shows the predictors included in each model, the P(M) column the prior model probability, the P(M | D) column the posterior model probability, the BF_M column the posterior model odds, and the BF_{10} column the Bayes factors of all models compared to the best model. The final column, 'error' is an estimate of the numerical error in the computation of the Bayes factor. All models are compared to the best model and are sorted from highest Bayes factor to lowest.

Do ACC learning scores differ between stimulation conditions?

The frequentist mixed-design ANOVA on the ACC learning scores revealed a significant Intercept ($F(1, 24) = 62.307$, $p < .001$, $\eta^2_P = .722$), suggesting that learning occurred in the ASRT task. The main effect of EPOCH showed a trend ($F(3, 72) = 2.237$, $p = .091$, $\eta^2_P = .085$), indicating that the learning scores increased throughout the task, independent of the stimulation condition (Fig. 2B). We did not find significant differences between the active stimulation and sham conditions either in overall learning (main effect

of STIMULATION: $F(1, 24) = 0.054$, $p = .819$, $\eta^2_P = .002$) or in the time course of learning (STIMULATION * EPOCH interaction: $F(3, 72) = 1.065$, $p = .359$, $\eta^2_P = .042$). The order of the stimulation sessions did not seem to affect the overall learning scores (main effect of ORDER: $F(1, 24) = 1.874$, $p = .184$, $\eta^2_P = .072$), the trajectory of the learning scores (ORDER x EPOCH interaction: $F(3, 72) = 0.249$, $p = .862$, $\eta^2_P = .010$), the stimulation (ORDER x STIMULATION interaction: $F(1, 24) = 1.831$, $p = .189$, $\eta^2_P = .071$) or the trajectory of the learning scores during the two different stimulation condition (ORDER x EPOCH x STIMULATION interaction: $F(3, 72) = 1.731$, $p = .182$, $\eta^2_P = .067$).

The analysis of effects (model-averaged results) of the Bayesian mixed-design ANOVA on the ACC learning scores showed that none of the effects related to Epoch, Stimulation or Session order should be included in the model (all $BF_{inclusion} < 1$, Table 3). Thus, based on the Bayesian analysis of effects, the learning scores were stable throughout the task and they were independent of the stimulation condition or the order of the stimulation.

Table 3.3 Model-averaged results of Bayesian ANOVA for ACC learning scores

Effects	**P(incl)**	**P(incl|data)**	**BF inclusion**
Stimulation	0.737	0.089	0.035
Epoch	0.737	0.220	0.101
Order	0.737	0.618	0.578
Stimulation ∗ Epoch	0.316	0.005	0.011
Stimulation ∗ Order	0.316	0.014	0.031
Epoch ∗ Order	0.316	0.005	0.012
Stimulaltion ∗ Epoch ∗ Order	0.053	2.526e -5	4.547e -4

Note: The Effects column denotes predictors of interest, the column P(incl) shows the prior inclusion probability, P(incl | D) shows the posterior inclusion probability, and $BF_{Inclusion}$ shows the inclusion Bayes factor.

Again, as our primary interest was the effect of the stimulation on probabilistic learning and the number of models was too high with the ORDER between-subject factor, as well as there was no evidence to include that factor, we recomputed the ANOVA with only the STIMULATION and EPOCH within-subject factors. This Bayesian ANOVA showed that the best model for our data is the Null model (Table 4). This Null model is ~6 times more likely than any model including the Stimulation factor. Altogether the

Bayesian ANOVA for the ACC learning scores provides evidence for the Null model to explain best our data. This suggests that learning scores were stable throughout the task and were independent of epochs, the stimulation condition and the order of the stimulation condition.

Table 3. 4 Bayesian model comparisons for ACC learning scores

Models	**P(M)**	**P(M|data)**	**BF $_{M}$**	**BF $_{10}$**	**error %**
Null model	0.200	0.533	4.566	1.000	
Epoch	0.200	0.328	1.956	0.616	0.523
Stimulation	0.200	0.081	0.353	0.152	1.680
Stimulation + Epoch	0.200	0.050	0.210	0.093	2.373
Stimulation + Epoch + Stim. ✻ Epoch	0.200	0.008	0.031	0.014	1.860

Note: All models include Subject. The Model column shows the predictors included in each model, the P(M) column the prior model probability, the P(M | D) column the posterior model probability, the BF_M column the posterior model odds, and the BF_{10} column the Bayes factors of all models compared to the best model. The final column, 'error' is an estimate of the numerical error in the computation of the Bayes factor. All models are compared to the best model and are sorted from highest Bayes factor to lowest.

To reveal possible patterns in the stimulation effects, we visualized individual learning score trajectories for both stimulation conditions separately for RT and ACC learning scores (see section '*Are there any obvious patterns in the stimulation effects for different individuals?*' and Fig. S2-S3 in Supplementary materials). Furthermore, to explore visually whether the order of the conditions influenced the effect of stimulation, we grouped the participants based on whether they completed the sham condition (Fig. S2A and S3A), or the active stimulation condition first (Fig. S2B and S3B). Altogether, the plots did not unravel obvious subgroups based on the difference between the active stimulation and sham conditions either in overall learning or in the time course of learning. Furthermore, the order of the stimulation did not seem to interact with the effects of the stimulation, further supporting the findings reported above.

Discussion

In the current study, our aim was to alter probabilistic learning by applying theta tACS during learning in a double-blinded cross-over within-subject design. We did not find differences either in overall learning performance or the time course of learning between the active stimulation and sham conditions. Moreover, Bayesian model comparisons provided evidence for no effect of stimulation on the learning performance.

Contrary to our expectations, we did not find an effect of the tACS on probabilistic learning. It is possible that the chosen parameters for the tACS stimulation, such as the fronto-parietal midline montage, the relatively weak (1 mA) current intensity, and/or the chosen theta frequency were not appropriate to influence probabilistic learning. Importantly, however, previous studies successfully influenced other cognitive functions (such as short term and working memory, or decision making) with stimulation parameters similar to ours (Chander et al., 2016; Polanía, Moisa, Opitz, Grueschow, & Ruff, 2015; Violante et al., 2017; Vosskuhl, Huster, & Herrmann, 2015), suggesting that these stimulation parameters might be effective for altering some cognitive functions but not others. Specifically, these studies aimed to influence prefrontal-network dependent, expectation/hypothesis-driven (top-down) cognitive processes. It is possible that stimulus-driven, bottom-up processes such as probabilistic learning can be successfully influenced by different frequency and/or electrode positions. Previous studies using similar, bottom-up tasks with deterministic sequential regularities (Serial Reaction Time Task, SRTT) reported alpha and beta frequencies to be successful for stimulation (Antal et al., 2008; Pollok, Boysen, & Krause, 2015). Antal et al. (2008) showed that alpha frequency tACS specifically improved motor sequence learning in contrast to beta or gamma frequencies over the primary motor cortex. Pollok et al. (2015) successfully applied both alpha and beta frequency tACS over the left primary motor cortex to improve motor sequence learning. Note that while these studies tested multiple frequencies to influence sequence learning, neither of them applied theta frequency. Importantly, these tasks were deterministic sequence learning tasks, which potentially rely more on motor representations as opposed to the ASRT task that we used in the current study, therefore we did not rely on these results when determining our stimulation parameters. To the best of our knowledge, our study was the first to test if probabilistic learning can be influenced by tACS and we chose theta frequency stimulation as it has been proven successful in

several studies investigating working memory and it has not been studied in tasks with acquiring regularities of stimuli. Future studies are needed to investigate whether different frequency bands (in particular alpha or beta) or different electrode montages (targeting motor cortex, or frontal or parietal areas selectively) are more suitable to influence probabilistic learning.

It is also possible that desynchronization instead of synchronization with the same parameters would have a bigger impact on probabilistic learning (although opposite effect). In support of this, Alekseichuk, Pabel, Antal, and Paulus (2017) found that fronto-parietal synchronization induced by 0° tACS did not significantly influence brain connectivity (measured via EEG) and working memory performance. In contrast, fronto-parietal desynchronization induced by 180° tACS affected both connectivity and performance. We did not have the appropriate equipment to induce desynchronization in the current study, but based on the finding of Tóth et al. (Tóth et al., 2017), that desynchronization in theta frequency is associated with better probabilistic learning, it would worth testing this stimulation design in case of a probabilistic learning paradigm (see for example the design in (Violante et al., 2017)).

Picking the appropriate stimulation parameters enables electrical stimulation to induce changes in brain activity and, therefore, possibly behavior. Thut, Schyns, and Gross (2011) claim that the entrainment of endogenous brain oscillations by tACS is possible if there is phase-alignment between the stimulation and internal oscillators. For this, an internal oscillator is needed, namely entrainment can occur only if there is a neural population that exhibits oscillations at the stimulation frequency under natural conditions. Moreover, the closer the external rhythm is to the internal one, the smaller the force needed to entrain endogenous oscillations (Pikovsky, Kurths, Rosenblum, & Kurths, 2003). Antal and Herrmann (2016) showed that the electrical current intensity with the standard stimulation strengths of 1-2 mA can be sufficient to induce changes in the brain activity but the induced voltage gradients in the brain are small. Based on our simulation, the induced electric field was up to 0.1 V/m, in particular in frontal and paracentral brain regions in our study. Altogether, tACS with 1 mA stimulation strength (as in our study) will likely influence brain activity only if the chosen stimulation frequency and stimulated brain areas match the patterns of naturally occurring brain activity during the given task. Thut et al. (2017) suggested several approaches to increase the alignment between the

brain stimulation and the ongoing endogenous activity, for example, setting the stimulation parameters by obtaining instantaneous phase or power of oscillatory brain activity from simultaneous EEG/MEG recording, or using EEG/MEG recordings prior to interventions to detect the individual frequency of the oscillation of interest. Further studies with more precise alignment could clarify if fronto-parietal theta entrainment can influence probabilistic learning.

Beyond the stimulation parameters, other factors could also influence the effects of the stimulation. We studied healthy young adults who generally perform well in cognitive tasks (Craik & Bialystok, 2006; Zwart, Vissers, Kessels, & Maes, 2017) and therefore their performance may be less susceptible to the effect of the stimulation. However, this is unlikely the case in our study as we also tested the effects of baseline performance on stimulation (see section '*Does baseline performance influence the effects of the stimulation?*' in Supplementary results) and did not find differential effects of the stimulation in participants performing worse at the beginning of the task. Nevertheless, the effect of theta tACS stimulation on probabilistic learning in a population with poorer cognitive performance remains to be explored.

Limitations

Similarly to most of the previous tACS studies, we did not monitor the brain activity during the stimulation, therefore there is no evidence that the stimulation induced changes in the endogenous activity. Furthermore, offline monitoring of brain activity preceding the stimulation is also lacking. This design would have enabled us to pick an individual theta frequency for each participant. Stimulating with the frequency matching the participant's dominant frequency could promote stronger stimulation effects (Antal & Herrmann, 2016). However, previous studies used similar tACS stimulation successfully to alter behavior. Lastly, as our stimulation parameters relied on previous studies that targeted working memory performance, a working memory control task could have been used to validate these parameters within the current sample. However, as our aim was not replication but to test the effect of simulation on probabilistic learning, we decided not to include other tasks in the stimulation conditions.

Conclusions

To the best of our knowledge, our study was the first to apply tACS to influence probabilistic learning. We did not find statistically significant effects of fronto-parietal midline theta tACS (with ~0.1 V/m electrical field strength) on probabilistic learning comparing behavior during active and sham stimulation. Our results draw attention to possible methodological flaws in electrical stimulation experiments. It is possible that with greater current intensity and/or with stimulation parameters more precisely aligned to endogenous brain activity during probabilistic learning, stimulation effects could be observed.

Deconstructing Procedural Memory: Different Learning Trajectories and Consolidation of Sequence and Statistical Learning[12]

Abstract

Procedural learning is a fundamental cognitive function that facilitates efficient processing of and automatic responses to complex environmental stimuli. Here, we examined training-dependent and off-line changes of two sub-processes of procedural learning: namely, sequence learning and statistical learning. Whereas sequence learning requires the acquisition of order-based relationships between the elements of a sequence, statistical learning is based on the acquisition of probabilistic associations between elements. Seventy-eight healthy young adults (58 females and 20 males) completed the modified version of the Alternating Serial Reaction Time task that was designed to measure Sequence and Statistical Learning simultaneously. After training, participants were randomly assigned to one of three conditions: active wakefulness, quiet rest, or daytime sleep. We examined off-line changes in Sequence and Statistical Learning as well as further improvements after extended practice. Performance in Sequence Learning increased during training, while Statistical Learning plateaued relatively rapidly. After the off-line period, both the acquired sequence and statistical knowledge was preserved, irrespective of the vigilance state (awake, quiet rest or sleep). Sequence Learning further improved during extended practice, while Statistical Learning did not. Moreover, within the sleep group, cortical oscillations and sleep spindle parameters showed differential associations with Sequence and Statistical Learning. Our findings can contribute to a deeper understanding of the dynamic changes of multiple parallel learning and consolidation processes that occur during procedural memory formation.

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

[12] Simor, P., Zavecz, Z., Horváth, K., Éltető, N., Török, C., Pesthy, O., Gombos, F., Janacsek, K., & Nemeth, D. (2019). Deconstructing procedural memory: Different learning trajectories and consolidation of sequence and statistical learning. *Frontiers in Psychology*, *9*, 2708.

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 (Fletcher et al., 2005; D. V. Howard et al., 2004). Learning performance does not only depend on training during acquisition but also on the post-learning period (Doyon et al., 2009; Durrant et al., 2011; Karni et al., 1998). Nevertheless, there are intensive debates questioning whether the acquired memories are stabilized or enhanced during post-learning, off-line periods (Doyon et al., 2009; Krakauer & Shadmehr, 2006; Mantua, 2018; Maquet et al., 2000; Nemeth, Janacsek, Londe, et al., 2010; Pan & Rickard, 2015; Philippe Peigneux et al., 2006; Rickard, Cai, Rieth, Jones, & Ard, 2008). 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 (Nemeth et al., 2013; Kóbor et al., 2018). 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., pairs, triplets, etc.) of stimuli. 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) shows 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 (King, Hoedlmoser, Hirschauer, Dolfen, & Albouy, 2017; Robertson et al., 2004; Spencer, Sunm, & Ivry, 2006). On the other hand, learning probabilistic sequences (Song et al., 2007a, Nemeth et al., 2010), in contrast to deterministic ones, 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, 2007b; Nemeth et al., 2010). There are several studies that investigated the long term retention of statistical learning (Kim, Seitz, Feenstra, & Shams, 2009; Kobor, Janacsek, Takacs, & Nemeth, 2017; Nemeth & Janacsek, 2011), and there is limited evidence that statistical learning in the auditory domain benefits from sleep (Durrant et al., 2011, 2013). Nevertheless, the consolidation, and more specifically, the role of sleep in statistical learning within the visuomotor domain remains largely unexplored.

The Alternating Serial Reaction Time (ASRT) task is a unique tool to investigate statistical and sequence learning within the same experiment (Howard and Howard, 1997; Nemeth et al., 2013). In this perceptual-motor four-choice reaction time (RT) task, participants are required to respond to visual stimuli appearing on the screen. In this task, predetermined sequential (termed as pattern) trials alternate with random ones (e.g., 2R4R3R1R, where numbers correspond to the four locations on the screen presented in the same sequential order during the entire task, and the letter R represents randomly chosen locations) that results in some chunks of stimuli being more frequent than others (see Figure 1) and enables us to measure the acquisition of both order and frequency information. Namely, sequence learning is defined as acquiring order information, in that consecutive elements in the sequence (denoted with numbers in the above example) can be predicted with 100% certainty based on the previous sequence element (i.e., the 2nd order transitional probability for the sequence trials is equal to one), while random trials are unpredictable (random stimuli can occur at any of the four possible locations with the same probability). However, as mentioned above, the alternating stimulus structure also results in some chunks of stimuli (three consecutive trials, called *triplets*) occurring more frequently than others (62.5% vs. 12.5%, respectively). For instance, the triplet 2X4 (where X denotes any location out of the four possible ones) would occur more frequently as its first and third item can originate either from sequential/pattern or random stimuli.

In contrast, the triplet 2X1 would occur less frequently as this combination can originate only from random stimuli (for more details see Figure 1 and the section "Materials and Methods"). Statistical learning is defined as acquiring this frequency information [which also represents a 2nd order regularity, where the transitional probability is less than one; for more detailed explanation see (Kóbor et al., 2018)]. To disentangle sequence and statistical learning in the ASRT task, sequence learning is assessed by contrasting sequential/pattern and random stimuli, while controlling for frequency information (i.e., analyzing only high-frequency trials). In contrast, statistical learning is assessed by contrasting high- vs. low-frequency trials while controlling for order information (i.e., analyzing only the random trials) (Nemeth et al., 2013; Kóbor et al., 2018). The learning trajectories for both sequence and statistical learning can be tracked by how different behavioral indices, such as RT and accuracy, change over the course of the task (Howard et al., 2004; Nemeth et al., 2013). To the best of our knowledge, no study has yet tracked the temporal dynamics of learning sequential structures (order information) as well as statistical probabilities (frequency information) within the same experimental design focusing not only on the learning phase but also on consolidation and on further performance changes in a post-consolidation testing phase.

Although sequence learning and statistical learning seem to require different cognitive mechanisms (Nemeth et al., 2013) in everyday learning scenarios, humans might rely simultaneously on both forms of learning. Nevertheless, 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 or more specific oscillatory activity (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).

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 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.

Although post-learning sleep seem to facilitate learning capacity in different cognitive domains (Feld & Diekelmann, 2015), several studies indicate that not sleep *per se*, but specific oscillations during sleep facilitate post-sleep improvements in behavioral performance (Rasch and Born, 2013). Among these oscillations, slow waves and sleep spindles emerge as important candidates that reflect processes of memory consolidation and synaptic plasticity (Diekelmann & Born, 2010; Fogel & Smith, 2011; Ulrich, 2016). Slow waves around 1 Hz and especially fast sleep spindles (13-16 Hz) are considered as hallmarks of the reactivation and neocortical redistribution of hippocampus-dependent memories (Diekelmann and Born, 2010). In addition, slow frequency oscillations ranging between 1 and 8 Hz were linked to the restorative (homeostatic) function of sleep (Achermann, Dijk, Brunner, & Borbély, 1993; Marzano, Ferrara, Curcio, & Gennaro, 2010). In order to examine the associations between cortical oscillations and behavioral performance, we explored the EEG correlates of off-line changes in sequence and statistical learning. We hypothesized that slow frequency oscillations and fast sleep spindles within the sleep group would be positively associated with the post-sleep gains in sequence learning, but not with those of statistical learning.

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, Novak, Mucsi, Shapiro, Rethelyi, & Kopp, 2004; Soldatos, Dikeos, & Paparrigopoulos, 2000). 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, Szádóczky, & Furedi, 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 3 months that might have had an impact on their mood, affect and daily rhythms were not included in the study. Only right-handed individuals as verified by the Edinburgh handedness inventory (Oldfield, 1971) were invited to the laboratory. 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 h 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, $M_{age} = 21.66 \pm 1.98$) 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, $M_{age} = 21.71 \pm 1.97$), with 25, 26, and 27 participants in the Active Wake, Quiet Rest, and Nap group, respectively (see Table 1). 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, $N = 25$). 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 (2015/279). 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 (Figure 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 10 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*. As the relatively long training phase can introduce fatigue leading to a general decline in performance measures (e.g., slower reaction times at the end of the training phase that do not reflect the acquired knowledge but the effect of fatigue), a retesting session after a long delay (spent asleep or in wakefulness) can result in a spurious increase in performance because of the release from fatigue. This way, the measure of off-line consolidation is confounded by the effect of fatigue (or more specifically, the release from fatigue) (Pan and Rickard, 2015). In order to control for this factor, the training session was followed by a short (3 min long) break in order to minimize the

fatigue effect due to massed practice (Rickard et al., 2008; Rieth, Cai, McDevitt, & Mednick, 2010). After the break, participants were tested on the task for 5 more blocks that constituted the *testing phase*. Subsequently, participants spent an approximately 1-h 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 min, the testing phase 5 min, and the retesting phase 10 min. 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.

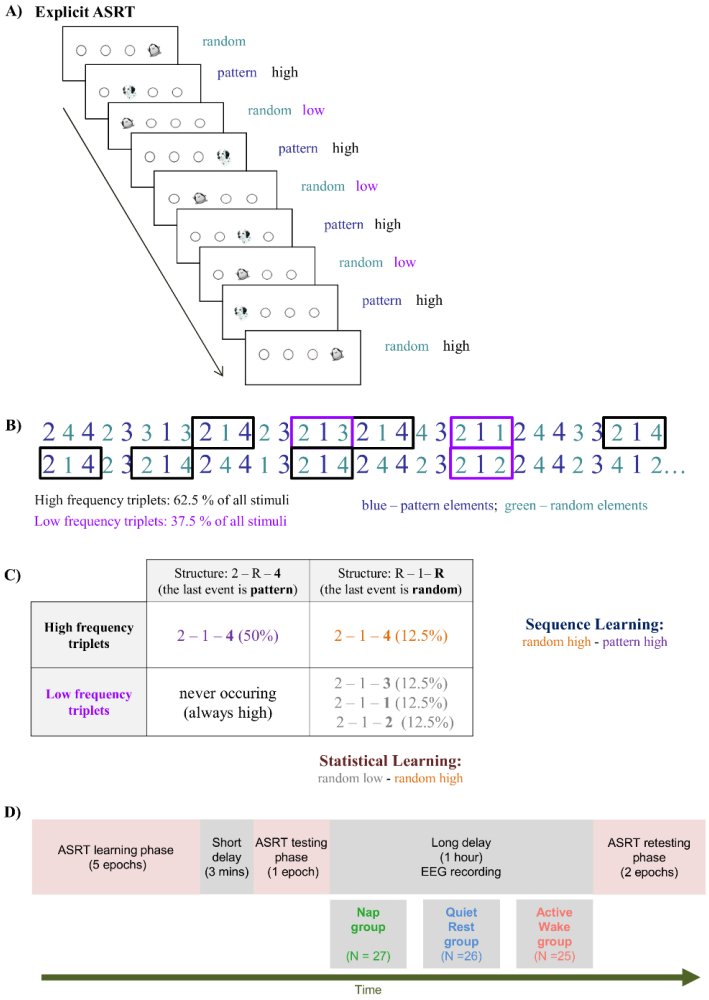

Figure 4.1 The modified Alternating Serial Reaction Time (ASRT) task. A) Pattern and random trials are presented in an alternating fashion. Pattern trials are marked with a picture of a dog, random ones with

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,C 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, Mueller, Gray, Raber, & Piper, 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. On 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 to 10 how stressed are you feeling now?"), as well as an item of the Hungarian version of the Karolinska Sleepiness Scale (KSS, Åkerstedt & Gillberg, 1990) to measure subjective sleepiness were administered. In the Hungarian version of the scale higher scores indicate a more refreshed state, that is, lower sleepiness. 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). In the ASRT task, short breaks were introduced between blocks in the following way: first, at the end of each block, participants were instructed to report the sequence they encountered in that block (which took approximately 6 s on average). Second, they received feedback for their accuracy and RT performance on pattern trials (fixed 3 s). Third, participants were notified (for a fixed 1 s) that the next block can be started by pressing a response button when they are ready; on average, participants continued the next block after approximately 4 s. These breaks were somewhat longer for every fifth blocks (i.e., Block 5, 10, 15, etc.), where participants were instructed to continue the next block after EEG data were saved by the experimenter (which took approximately 20 s on average). Thus, altogether, for the majority of blocks the between-block break was ~ 14 s, and for every fifth block it was ~ 29 s. Additionally, 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 1-h 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 min 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

The 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, Whelan, & Reilly, 2010) implemented in EEGLAB (Delorme & 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 and 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 & 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-s 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-s basis. Wrong channels (N = 6 in the dataset of the Nap group) were replaced by the average of the neighboring channels.

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 and 31 Hz in the Nap group, and between 1.5 and 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: Fp1, Fpz, Fp2, AF3, AF4, F7, F5, F3, F1, Fz, F2, F4, F6, F8, frontocentral and frontotemporal: FT7, FC5, FC3, FC1, FC2, FC4, FC6, FT8), Central (central, centrotemporal and centroparietal: T7, C5, C3, C1, Cz, C2, C4, C6, T8, CP5, CP3, CP1, CPz, CP2, CP4, CP6, TP8), and Posterior (parietal, parietotemporal and occipital: P7, P5, P3, Pz, P2, P4, P6, P8, POz, O1, Oz, O2) electrode derivations.

We quantified sleep spindling activity by the Individual Adjustment Method (IAM, Bódizs, Körmendi, Rigó, & Lázár, 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 spindles 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 fast 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 (R Core 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 and each participant. 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 off-line changes occurring between testing and retesting sessions we used a similar mixed-design ANOVA with EPOCH (6-8) 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 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^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 (7^{th} epoch) session and the testing session (6^{th} 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 8^{th} (2^{nd} half of the retesting session) with those of the 6^{th} 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 & 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 1 h long off-line 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 on the Hungarian version of the KSS scale where higher scores indicate a more refreshed state, that is, lower sleepiness) than the Active Wake group ($p = 0.02$), however, the difference was not significant after FDR correction.

Table 4.1 Descriptive characteristics of groups

Variable	Active Wake group (N = 25) Mean (SD)	Quiet Rest group (N = 26) Mean (SD)	Nap group (N = 27) 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 (before the Learning phase)	2.65 (2.09)	2.55 (1.43)	3.33 (1.98)	$p = 0.35$
Stress scale (after the Learning phase)	2.59 (1.28)	2.00 (1.33)	1.77 (1.41)	$p = 0.17$
KSS (before the Learning phase)	6.44 (1.26)	6.81 (1.13)	6.19 (1.52)	$p = 0.24$
KSS (after the Learning phase)	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. Higher scores in the KSS indicate lower sleepiness.

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

Table 4.2 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)
Fr. fast spindle density	6.37 (0.96)
Cent. fast spindle density	7.45 (0.83)
Post. fast spindle density	7.35 (0.93)
Fr. fast spindle amp.	4.56 (1.32)
Cent. fast spindle amp.	6.01 (1.56)
Post. fast spindle amp.	5.38 (1.38)
Fr. slow spindle density	7.31 (1.12)
Cent. slow spindle density	7.33 (1.19)
Post. slow spindle density	7.4 (1.16)
Fr. slow spindle amp.	3.91 (1.85)
Cent. slow spindle amp.	3.28 (1.49)
Post. slow spindle amp.	2.54 (0.96)

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

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

$F_{2,150} = 52.04$, $p < 0.0001$, $\eta^2_P = 0.41$): 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 ($F_{4,150} = 0.80$, $p = 0.46$, η^2_P =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 ($F_{14,1050} = 11.93$, $p < 0.0001$, $\eta^2_P = 0.14$), 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 ($F_{28,1050} = 0.66$, $p = 0.68$, η^2_P =0.02), suggesting that performance trajectories to the different trial types were similar among the groups.

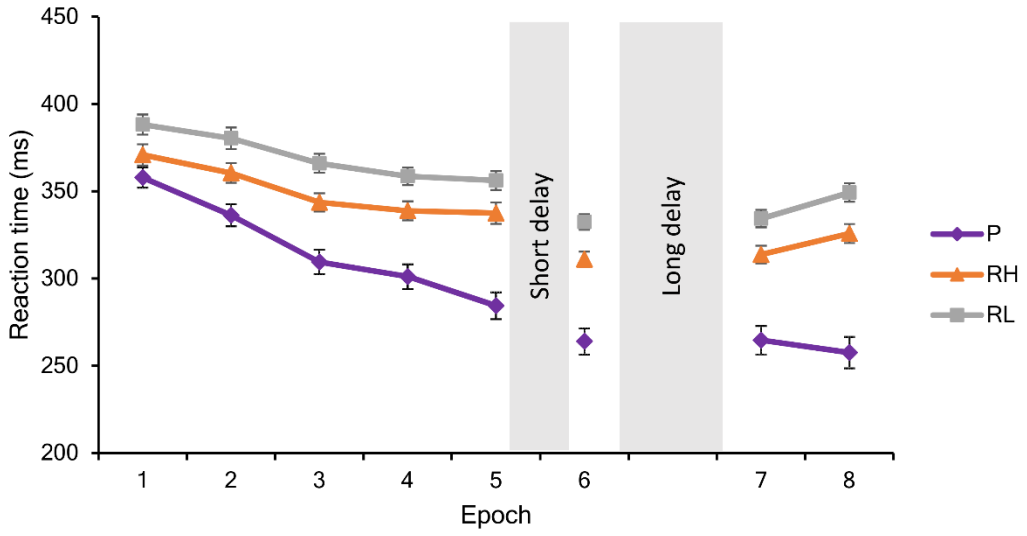

Figure 4.2 Performance during the training (Epochs 1-5), testing (Epoch 6) and retesting (Epochs 7-8) sessions. Mean reaction times and standard errors are visualized in response to pattern (P), random high (RH), and random low (RH) trials during each epoch.

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: $F_{2,75} = 1.10$, $p = 0.34$, $\eta^2_P = 0.03$). Irrespectively of learning type, performance improved across epochs of training (main effect of EPOCH: $F_{4,300} = 10.92$, $p < 0.0001$, $\eta^2_P = 0.13$). The GROUP x EPOCH interaction was not significant ($F_{8,300} = 0.59$, $p = 0.68$, $\eta^2_P = 0.02$), suggesting that improvement during training was similar between the groups. In addition, the main effect of LEARNING TYPE was significant ($F_{1,75} = 3.93$, $p = 0.05$, $\eta^2_P = 0.05$): 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 ($F_{2,75} = 0.81$, $p = 0.45$, $\eta^2_P = 0.02$), suggesting that the difference between Sequence and Statistical Learning were similar among the groups. Furthermore, a significant interaction between EPOCH and LEARNING TYPE emerged ($F_{4,300} = 5.52$, $p = 0.002$, $\eta^2_P = 0.07$): as illustrated in Figure 3, participants, on average, 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 ($F_{8,300} = 0.58$, $p = 0.72$, $\eta^2_P = 0.02$), suggesting that training-dependent patterns of Sequence Learning and Statistical Learning were similar across the groups.

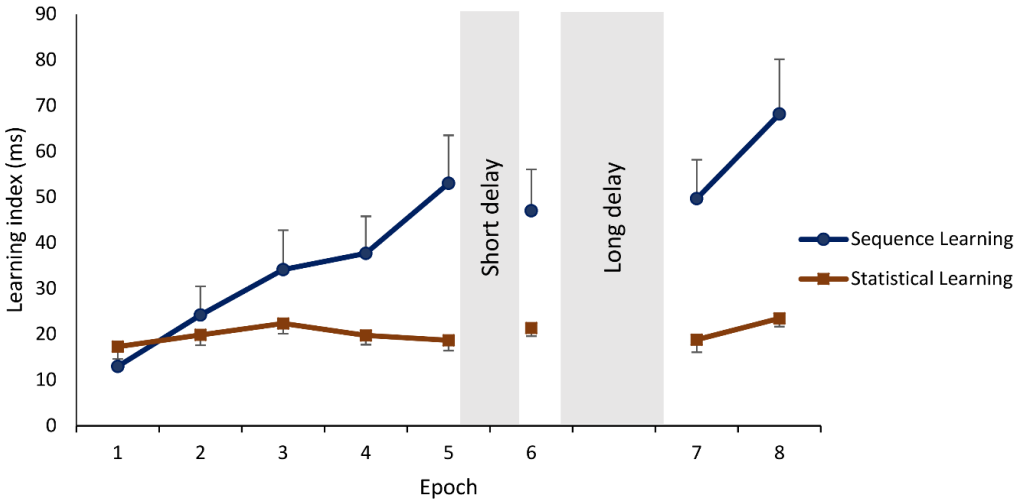

Figure 4.3 Learning and off-line changes in Sequence and Statistical Learning. Sequence Learning is quantified as the difference in reaction times to random high elements vs. pattern elements. Statistical Learning is quantified as the difference in reaction times to random low elements vs. random high elements. 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.

Beyond the group-level results presented in the previous paragraph, we performed an additional analysis to reveal learning trajectories on a subject-by-subject basis. We categorized each subject's learning trajectory during training by a combination of curve fitting and visual inspection. For comparability, we performed the same steps for Sequence and Statistical learning (see Figures 4A,B, respectively) and found that ~33% of participants showed gradually increasing Sequence learning during training, while the trajectory for Statistical learning was gradually increasing only in ~16% of participants [$\chi^2(1) = 3.80$, $p = 0.05$]. Compared to these percentages, a relatively smaller number of participants exhibited a step-like increase in learning performance: ~10% of participants for Sequence learning and ~4% of participants of Statistical learning ($p = 0.15$). Additionally, a small portion of participants exhibited a decreasing pattern, with the best performance at the beginning of the task (~5% of participants for Sequence learning, and ~13% of participants for Statistical learning; $p = 0.42$). The learning trajectory of the majority of participants did not clearly follow any of the patterns described above. These learning trajectories were categorized as 'Other pattern' (~53% of participants for

Sequence learning, and ~66% of participants for Statistical learning; p = 0.81). These participants exhibited relatively large changes in performance from one epoch to another and then returned to the previous performance level. The timing of these larger changes in performance was evenly distributed across epochs. It is plausible that these participants explored different (explicit or implicit) strategies over the course of learning that may have resulted in large changes in some epochs compared to their overall learning performance. Note, however, that the primary focus of our study was not to test these possible strategies but to compare Sequence and Statistical learning trajectories across the three experimental groups (Quiet Rest, Active Wake, and Nap). Importantly, the distribution of subgroups exhibiting different learning trajectories was similar across the three experimental groups both for Sequence learning [$\chi^2(6) = 0.91$, $p = 0.99$] and for Statistical learning [$\chi^2(6) = 1.98$, $p = 0.92$].

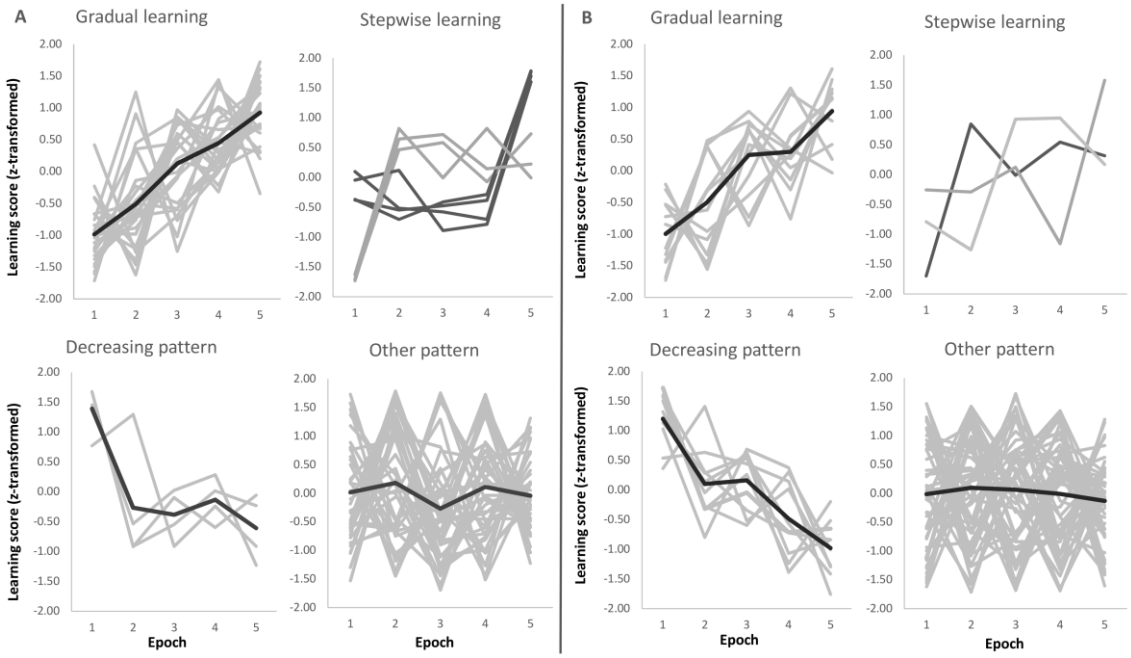

Figure 4.4 Sequence (A) and Statistical (B) learning trajectories for individual subjects. Each participant's learning trajectory is presented in a light grey color, while the average learning trajectory for that subgroup is presented in a darker gray color for the 'Gradual learning', 'Decreasing pattern' and 'Other pattern' panels. For the 'Stepwise learning' panel, the light and dark gray colors represent subgroups of participants depending on the timing of their performance increase (no average learning trajectory is presented).

Early Statistical Learning Effects During Training

To provide further insights into the trajectory of Statistical learning, we performed additional analyses by focusing on block-level and below block-level data. The first set of analyses aimed to determine the time point when participants successfully extracted the statistical regularities from the stimulus stream. First, we computed Statistical learning scores for each block of Epoch 1, and tested if these Statistical learning scores were significantly different from zero. We found significant Statistical learning effect already in Block 1 of the ASRT task [$t(73) = 2.12$, $p = 0.04$, Cohen's $d = 0.25$]. Next, we zoomed into Block 1 to further test this learning effect. In this analysis, we split Block 1 into two halves and computed Statistical learning scores for each participant, for each half. This level of granularity seemed the most appropriate so that all participants had at least a few random-high trials (~ 4 trials on average, ranging from 2 to 9), enabling us to compute learning scores for all participants. These Statistical learning scores were submitted into one sample t-tests, which showed that Statistical learning scores did not reach significance in the first half of Block 1 [$t(73) = 1.11$, $p = 0.269$, Cohen's $d = 0.13$], while they were significant in the second half of Block 1 [$t(73) = 1.99$, $p = 0.05$, Cohen's $d = 0.2$]. This analysis thus demonstrates that statistical regularities are learned (albeit very quickly) and the observed significant Statistical learning scores at the very early phase of the task are not due to other (not learning-related) preexisting tendencies.

This rapid learning effect is in fact not surprising if we consider that 80 trials are presented in the first block, and ~ 50 of those trials can be categorized as high frequency triplets (occurring in pattern or random positions). As there are 16 individual triplets that are high frequency, that means that participants encounter each individual triplet approximately four times in the first block already. In contrast, there are 48 individual triplets that are low frequency, and participants encounter these individual triplets approximately (or less than) once in a block. Thus, the observed significant Statistical learning scores (i.e., the difference between the random-high and random-low frequency trials) suggests that participants are so sensitive to the frequency statistics that as little as, on average, four presentations of the same trials are sufficient to show speeded responses to them.

Nevertheless, it is important to highlight that significant learning does not necessarily mean that participants have a stable knowledge about the statistical regularities. Thus,

even though the Statistical learning scores are already significant at the early phase of learning and these scores numerically do not change as the task progresses, it is reasonable to assume that more practice can help strengthen the acquired knowledge. We ran an additional analysis to test this assumption. In this analysis, we focused on block-level data and computed Cohen's *d* effect sizes for the block-level Statistical learning scores. These effect sizes were substantially smaller in the first five blocks of the ASRT task (0.27 on average for Blocks 1-5, i.e., Epoch 1) compared to the later blocks (blocks of Epoch 2: 0.45, Epoch 3: 0.51, Epoch 4: 0.53, Epoch 5: 0.50). This difference in the effect sizes suggests that, although participants were able to extract the statistical regularities from the stimulus stream very early in the task, additional training helped them strengthen the acquired statistical knowledge.

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, η^2_P = 0.02), nor the interactions GROUP x EPOCH ($F_{4,150}$ = 0.52, p = 0.67, η^2_P = 0.01), GROUP x LEARNING TYPE ($F_{2,75}$ = 0.65, p = 0.53, η^2_P = 0.02), and GROUP x EPOCH x LEARNING TYPE ($F_{4,150}$ = 0.73, p = 0.55, η^2_P = 0.02) emerged as significant predictors. The lack of a group effect is shown in Figure 5 that illustrates off-line changes (7th minus the 6th 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, η^2_P = 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, η^2_P = 0.20). More specifically, overall performances (regardless of learning type) were unchanged from the testing phase (6th epoch) to the first retesting epoch (7th) (p = 0.86), but improved (p < 0.0001) from the testing phase to the end of the retesting session (8th epoch), and from the first retesting epoch to the second (7th epoch vs 8th 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, η^2_P = 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 (6th epoch, M = 47.02 vs. 7th epoch, M = 47.69, p = 0.85). Similarly, Statistical Learning remained unchanged from testing to the first retesting (6th epoch, M = 21.39 vs. 7th 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 (8th epoch, M = 68.19, p = 0.001), whereas Statistical Learning did not show any significant changes by the end of the retesting phase (8th epoch: M = 23.51, p = 0.41).

To further explore potential group differences during the off-line period we ran additional ANOVAs separately for Sequence and Statistical learning scores considering their different learning curves. Based on these ANOVAs, we found no group differences in the consolidation (6th epoch vs. 7th epoch) of the acquired knowledge (Sequence learning: p = 0.35, Statistical learning: p = 0.78). Similarly, no group differences emerged in the additional increase between 7th epoch and 8th epoch (Sequence learning: p = 0.65, Statistical learning: p = 0.36).

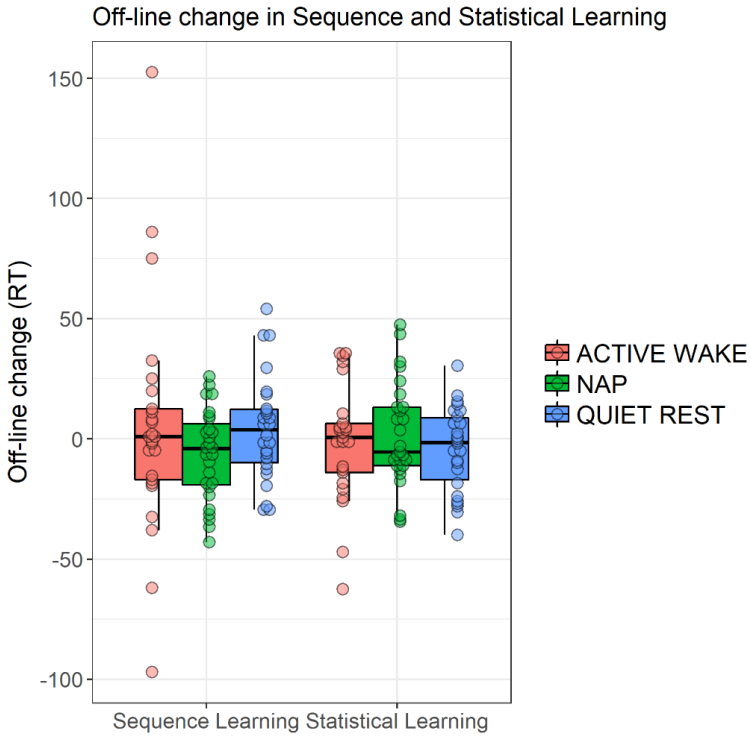

Figure 4.5 Off-line changes in learning indices within the three groups. Off-line changes were calculated by the learning scores of the 7th epoch minus the respective learning scores of the 6th epoch. Dots show individual data points, the vertical line within the boxes show the medians, boxes represent the first and third quartiles, whiskers indicate the interquartile range of 1.5.

Awareness of the Sequence in the Groups

For the analysis of sequence awareness, two participants' data had to be excluded due the technical issues during collection of sequence reports (one data from the active wake and one data from the nap group). Additionally, eleven participants could not report the correct sequence consistently during training (N = 3 in the active wake, N = 3 in the nap, and N = 5 in the quiet rest group), and therefore they were also excluded from the following analyses. Importantly, there were no group differences in the number of participants who could or could not report the correct sequence consistently and were excluded (chi-square = 1.77, p = 0.78).

On average, participants could report the correct sequence consistently from the 6th block (M = 6.58, SD = 7.04), with no differences across the groups ($F_{2,64}$ = 1.53, p = 0.23). Overall, the block number from which participants could consistently report the correct sequence showed a significant negative correlation with the Sequence learning scores (r = -0.28, p = 0.02). Thus, the earlier participants could find the correct sequence and report consistently thereafter, the better their overall Sequence learning was. No association was observed between the block number and the Statistical learning scores (r = -0.06, p = 0.63), suggesting that sequence awareness primarily affected Sequence learning but not Statistical learning.

Finally, we conducted an ANOVA for the Sequence learning scores of the training phase (Epoch 1 to 5), including the block number from which participants could consistently report the correct sequence as a covariate to check how sequence awareness affected the time course of learning across groups. The ANOVA revealed a significant main effect of EPOCH ($F_{4,244}$ = 10.53, p < 0.001, η^2_P = 0.147), indicating better Sequence learning scores as learning progressed. This effect was modulated by the block number on a trend level ($F_{4,244}$ = 2.58, p = 0.08, η^2_P = 0.041), suggesting that the earlier participants could report the correct sequence, the better their Sequence learning became across training. Importantly, no significant group differences emerged either in overall learning or in the trajectory of learning even after taking into account the block number as a covariate (ps > 0.21).

A similar ANOVA was conducted for the consolidation analysis (Epoch 6 to 8). This ANOVA also revealed a significant main effect of EPOCH ($F_{2,122}$ = 8.34, p < 0.001, η^2_P

= 0.120), which is consistent with the previous ANOVA conducted for these epochs, showing increase in Sequence learning scores due to additional training (Epoch 7 vs. Epoch 8, see Figure 3). This effect was not modulated by the block number (p = 0.49). Furthermore, no significant group differences emerged either in overall learning scores or in the trajectory of learning scores across these epochs, even after taking into account the block number as a covariate (ps > 0.32). These results altogether suggest that, although the timing when participants gained explicit knowledge about sequence affects their Sequence learning scores, this effect is similar across the groups both during training and consolidation.

Associations Between EEG Spectra and Off-line changes

Off-line changes in Sequence Learning as indexed by the difference scores between the 7th (first half of retesting phase) and the 6th epochs' (testing phase) scores were positively associated with frontal theta power (r = 0.44 p = 0.028) within the nap group. Off-line changes in Sequence Learning were not associated with spectral EEG power measures in the either of the awake (AW, QR) 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), showed a positive association with frontal theta power (r = 0.52, p = 0.008) within the nap group only. Nevertheless, these correlations did not reach statistical significance after FDR correction of multiple comparisons (all ps > 0.05). 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 changes (7th vs. 6th epoch and 8th vs. 6th epoch) in Sequence Learning within the nap group. As shown in Figure 6**,** associations with theta band power were prominent at frontal electrode sites, peaking at left frontopolar locations in case of immediate off-line changes (Figure 6A), as well as in case of additional off-line changes in performance (Figure 6B). Finally, we examined the associations between off-line (7th vs. 6th epoch and 8th vs. 6th epoch) changes in Sequence Learning and bin-wise EEG spectral power averaged across all electrodes (within the Nap group). Immediate (7th vs. 6th epoch) and delayed (8th vs. 6th epoch) post-sleep improvement in Sequence Learning correlated only with slow frequency activity between 2 and 7.75 Hz (all bins p < 0.01).

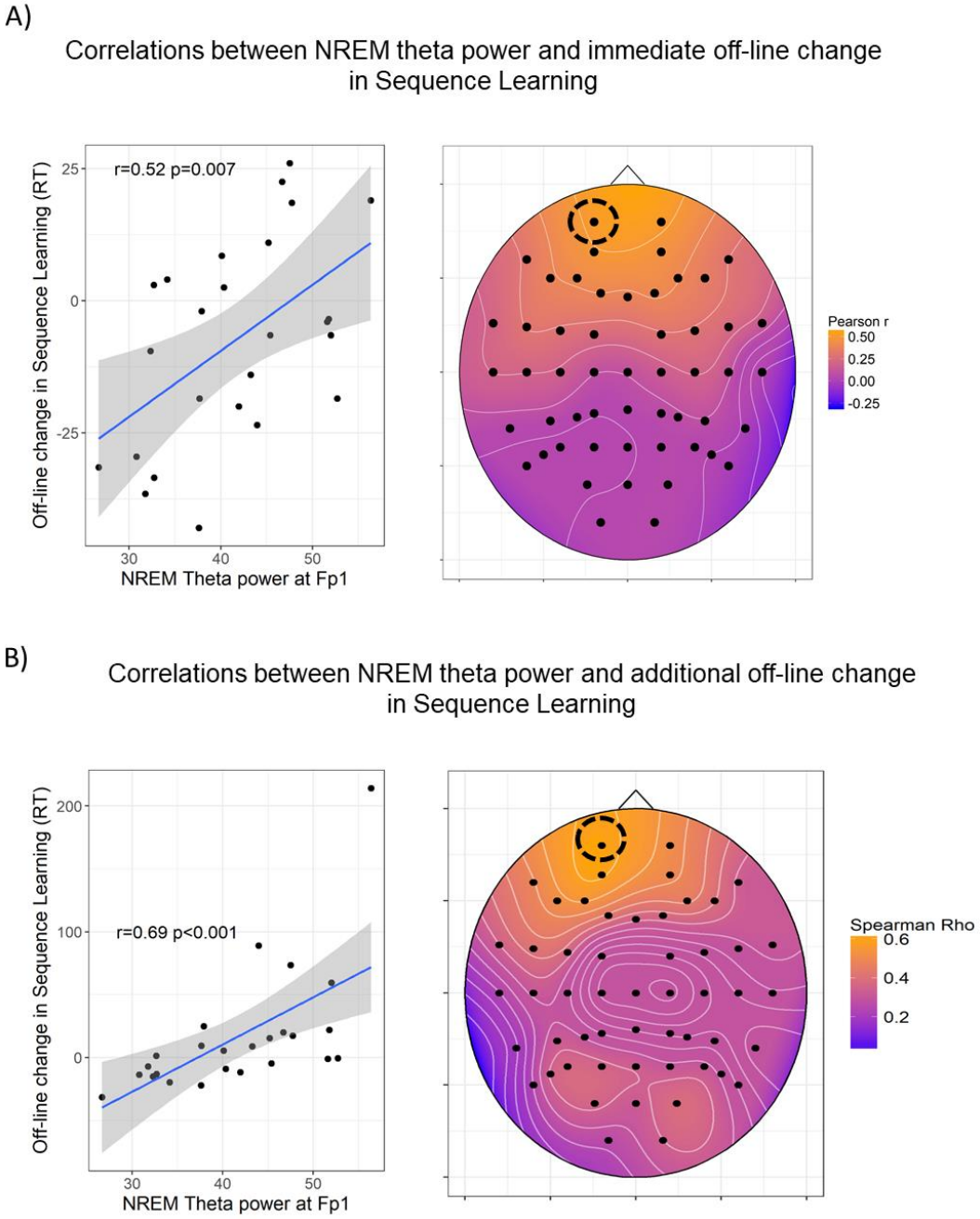

Figure 4.6 Associations between NREM theta power and off-line changes in Sequence Learning. A) Pearson correlations between NREM theta band power and immediate (7th vs. 6th epoch) post-sleep changes in Sequence Learning. **B)** Spearman Rho correlations coefficients between NREM theta band power and delayed (8th vs. 6th epoch) post-sleep changes in Sequence Learning. The heat plots on the right indicate the magnitude of correlation coefficients, the scatterplots on the left show the association in a prominent (left frontal) electrode site. In case of 5B the correlation coefficient remained unchanged (r = 0.64, p < 0.001) after the exclusion of the outlier. The figures show uncorrected p values (before FDR correction). For the immediate off-line changes, only Fp1, Fp2, AF3, AF4 locations remained significant after FDR correction. For the additional off-line changes, frontal channels Fp1, Fpz, Fp2, AF3, AF4, F7, F5, F3, F1, Fz, F2, F4, F6, F8 as well as FC4, FC5, CP5 and P5 locations remained significant after FDR correction.

Immediate and additional off-line changes (7^{th} vs. 6^{th} epoch and 8^{th} vs. 6^{th} epoch) in Statistical Learning were 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 Statistical Learning assessed immediately after the long delay (6^{th} vs. 7^{th} epoch) and after extended practice, (6^{th} vs. 8^{th} epoch) were not associated with spectral EEG power measures in any of the three groups. On the other hand, immediate and delayed post-sleep improvements in Sequence Learning were predicted by high delta and theta activity during sleep within the Nap group. Nevertheless, these correlations did not remain significant after correction for multiple comparisons.

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 at 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 (p = 0.036).

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 (all ps > 0.05).

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 5^{th} and the 1^{th} 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 = 0.006; Central: r = -0.52, p = 0.009; Posterior: r = -0.51, p = 0.005).

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 showed trends after partialling 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 (delayed) 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 different learning curves. Whereas performance in Sequence Learning exhibited an increase during training, Statistical Learning was rapidly acquired and remained unchanged throughout training. During the off-line period, both forms of learning were preserved as no significant off-line changes emerged in either Sequence or Statistical Learning. Nevertheless, Sequence Learning improved after additional practice (i.e., in the retesting phase), 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 consolidation 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 the awake groups. Within the Nap group we found a trend indicating a positive association between frontal theta band power and off-line change in Sequence Learning. In addition, frontal theta power predicted further improvements in Sequence Learning after additional practice. 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.

Our data suggests that sequence and statistical learning are markedly different sub-processes of procedural learning. Frequency-based information is acquired rapidly and appears to undergo less prominent changes during further training compared to the acquisition of order-based information that may exhibit further performance improvements. Our fine-grained analyses revealed that statistical learning occurs already in the first block of the task. This finding suggests that participants are so sensitive to the frequency statistics that as little as, on average, four presentations of the same trials are sufficient to show speeded responses to them. Nevertheless, the further analysis of effect sizes showed that, although participants were able to extract the statistical regularities from the stimulus stream very early in the task, additional training helped them strengthen the acquired statistical knowledge.

Rapid statistical learning has also been reported before: for instance, in the ASRT study of Szegedi-Hallgató and colleagues (2017), statistical learning was apparent already in the first epoch in the Explicit group but seemed to have larger individual differences in the Implicit groups as only one of the two Implicit groups exhibited significant statistical learning in the first epoch (see Supplementary results and figures). Similarly, in Kóbor et al.'s (2018) study, statistical learning was observed in the first epoch of the explicit version of the ASRT task, along with a significant sequence learning as well. Consequently, a possible explanation for the very rapid statistical learning is that, in an explicit condition, the instructions and motivation to learn can have an overarching effect, providing a cognitive state, in which not only the instructed sequential but also the uninstructed statistical regularities can be learned quickly. Although this was not in the primary focus of these previous studies, if we take a closer look at the learning trajectories, it appears that statistical regularities are extracted very early and no (or very little) further gains may be observed during training if explicit instructions are given for the *sequential* information (Szegedi-Hallgató et al., 2017; Kóbor et al., 2018). In contrast, in the implicit conditions, statistical learning may undergo further improvements during training (Szegedi-Hallgató et al., 2017), above and beyond the strengthening of the acquired knowledge as suggested in the previous paragraph. These observations support the interpretation that explicit instructions and the motivation to learn can have an overarching effect in that not only the instructed sequential but also the uninstructed statistical regularities can be learned more quickly. Interestingly, a recent study showed

that, if the task is fix-paced instead of self-paced, no such overarching effect can be observed, suggesting a complex interplay of multiple factors that may influence the effect of explicit instructions on learning (Horvath et al., 2018). Further studies should directly test these factors.

Nevertheless, it is important to note that statistical learning typically occurs implicitly (i.e., without conscious intent to learn and without awareness about the learning situation itself or about the actual regularities) and relatively quickly, already in one learning session (e.g., Song et al., 2007a; Nemeth et al., 2013; Kóbor et al., 2017). In contrast, it has been previously shown that acquiring the alternating sequence structure (frequently referred to as higher-order sequence learning) in the ASRT task typically occurs after 4 days of practice if learning is implicit (Howard and Howard, 1997; Howard et al., 2004), while this can be substantially faster if explicit instruction is provided to the participants (Nemeth et al., 2013). Accordingly, participants quickly formed explicit knowledge about the sequence. Therefore, we think that the current study design was suitable to measure both sequence and statistical learning, bringing them in the same time frame of acquisition (i.e., showing significant learning in one learning session for both measures).

The present study narrows down the concept of statistical learning by regarding it as only one of the processes that is the sensitivity to frequency information. From a theoretical perspective, however, it is important to note that at the level of transitional probabilities, statistical learning (in this narrow sense) and sequence learning could be considered as similar. Namely, both are statistical learning in a broader sense. When acquiring frequency information (statistical learning in the narrow sense), a 2nd order probabilistic sequence should be learned, in which there are always one probable continuation and some less probable continuations for the first two elements of a given three-element stimulus chunk (Szegedi-Hallgató et al., 2017; Kóbor et al., 2018). When acquiring order information (sequence learning), the 2nd order transitional probability is equal to one; namely, consecutive elements in the sequence could be predicted with 100% certainty from the previous sequence element (Kóbor et al., 2018).

Our finding of different learning trajectories within one learning session is in line with the results of Kóbor and colleagues (2018) well as 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. Beyond the group-level results, we performed an additional analysis to characterize learning trajectories on a subject-by-subject basis. This analysis revealed that one-third of participants showed gradually increasing Sequence learning during training, and this proportion was significantly higher than the number of participants who exhibited gradually increasing Statistical learning, confirming differences in learning trajectories for Sequence vs. Statistical learning beyond the group-level findings. Nevertheless, the majority of participants exhibited a learning trajectory other than gradual. It is plausible that these participants explored different strategies over the course of learning that may have resulted in large changes in some epochs compared to their overall learning performance. Further investigations should directly focus on individual level heterogeneity and test which factors/characteristics predict learning trajectories on the individual level.

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 (Hallgato, Győri-Dani, Pekár, Janacsek, & Nemeth, 2013; Nemeth, Janacsek, Londe, et al., 2010; Peigneux et al., 2003; P. Peigneux et al., 2006; Song et al., 2007b), 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 (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 (B. Török, Janacsek, Nagy, Orbán, & Nemeth, 2017) might have a different impact on these tasks. For instance, effects of fatigue are typical to occur in learning tasks (Brawn, Fenn, Nusbaum, & Margoliash, 2010; Pan & Rickard, 2015; Rickard et al., 2008), 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., 2016; West, Vadillo, Shanks, & Hulme, 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, suggesting that post-sleep increases in our case were also largely dependent on further practice. Interestingly, delayed (training-dependent) off-line improvements were associated with slow oscillatory activity within the Nap group. This finding suggests that not sleep *per se*, but low-frequency oscillations are associated with delayed performance gains after sleep and additional practice. Our findings indicate that slower oscillatory activity including the (high) delta and the theta frequency ranges (from 2 to 7.75 Hz) during daytime sleep might be predictive of post-sleep improvements in Sequence Learning. Slow frequency oscillations peaking at anterior locations and spanning between 1 and 8 Hz reflect the homeostatic and restorative capacity of sleep as power in these frequencies is increased after prolonged wakefulness (Borbély, Baumann, Brandeis, Strauch, & Lehmann, 1981; Marzano et al., 2010) in fronto-central derivations. Furthermore, the homeostatic increase in spectral power between 2 and 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, Nakashima, & Nishikawa, 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 measurements 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 may develop more gradually. On the behavioral level, both sequence and statistical knowledge were retained and were independent of whether the off-line period included sleep or not. Although our measures of cortical oscillations assessed during the off-line period showed associations with behavioral performance within the sleep group to some extent, the influence of sleep-

specific oscillations on Sequence and Statistical learning should be examined in future studies. Nevertheless, our findings suggest that sleep does not have 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 related oscillatory activity. 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.

The relationship between subjective sleep quality and cognitive performance in healthy young adults: Evidence from three empirical studies[13]

Abstract

The role of subjective sleep quality in cognitive performance has gained increasing attention in recent decades. In this paper, our aim was to test the relationship between subjective sleep quality and a wide range of cognitive functions in a healthy young adult sample combined across three studies. Sleep quality was assessed by the Pittsburgh Sleep Quality Index, the Athens Insomnia Scale, and a sleep diary to capture general subjective sleep quality, and the Groningen Sleep Quality Scale to capture prior night's sleep quality. Within cognitive functions, we tested working memory, executive functions, and several sub-processes of procedural learning. To provide more reliable results, we included robust frequentist as well as Bayesian statistical analyses. Unequivocally across all analyses, we showed that there is no association between subjective sleep quality and cognitive performance in the domains of working memory, executive functions and procedural learning in healthy young adults. Our paper can contribute to a deeper understanding of subjective sleep quality and its measures, and we discuss various factors that may affect whether associations can be observed between subjective sleep quality and cognitive performance.

[13] Zavecz, Z., Nagy, T., Galkó, A., Nemeth, D., & Janacsek, K. (2020). The relationship between subjective sleep quality and cognitive performance in healthy young adults: Evidence from three empirical studies. *Scientific reports*, *10*(1), 1-12.

Introduction

There is a widely accepted belief that experiencing poor sleep quality, including subjective experiences (e.g., reporting difficulties falling asleep, waking up frequently during the night, or feeling tired during the day), indisputably decreases cognitive performance. We can often hear people complaining about weaker memory and/or attentional performance in relation to their experienced sleep insufficiency. This phenomenon can be particularly prevalent amongst university students since the pressure for academic performance in this population is exceptionally high. The possible overestimation of the importance of one's subjective sleep quality can even lead to placebo or nocebo effects on cognitive performance (Draganich & Erdal, 2014; Gavriloff et al., 2018). However, scientific evidence on the relationship between experienced subjective sleep quality and cognition is still inconclusive (Bastien et al., 2003; Miyata et al., 2013; Nebes, Buysse, Halligan, Houck, & Monk, 2009; Stepanski et al., 1989; van den Noort et al., 2016). Therefore, our aim in the current study was to test whether subjective sleep quality is associated with cognitive performance in healthy young adults.

The role of sleep in cognitive performance has gained increasing attention in neuroscience and sleep research in recent decades (Diekelmann & Born, 2010; Jones & Harrison, 2001). Numerous experimental methods exist that can be employed for examining the association between sleep and cognitive performance. Sleep parameters can be evaluated based on actigraph or electroencephalograph measurements (i.e., objective measures), which are time-consuming and require expensive equipment. Hence, researchers and clinicians often tend to rely on questionnaires (i.e., subjective measures) to assess sleep parameters (e.g., sleep latency, sleep quality, sleep disturbances, or sleep duration). This inclination has also motivated the current study to explore the relationship between sleep questionnaires and cognitive functions.

Previous studies have shown that subjective and objective sleep parameters, such as sleep latency, sleep duration, or sleep efficiency could differ (Armitage et al., 1997; Guedes et al., 2016; Landry et al., 2015b); the strength of correlation between the subjective and objective measures of the same parameters varied between 0.21 and 0.62 for sleep latency and duration, while it was close to 0 for sleep efficiency. Subjective sleep quality can vary from objective sleep quality as it is typically estimated from a combination of parameters, such as sleep initiation, sleep continuity (number of

awakenings), and/or depth of sleep. For instance, extreme deviations can occur between subjective and objective measures in sleep disorders, such as insomnia or sleep-state misperception. According to Zhang and Zhao (2007), the subjective and objective measures together should determine the type of treatment and medication in sleep disorders. Stepanski et al. (1989) showed that, within insomniac patients, the decisive factor of whether a patient seeks medication is their subjective evaluation of their sleep quality and daytime functioning. Furthermore, Gavriloff et al. (2018) found that providing sham feedback about their sleep to patients with insomnia influenced their daytime symptoms and performance in attention and vigilance tasks. Similarly, in a placebo sleep study, young adults were randomly told they had below or above average sleep quality based on their brainwaves and other psychophysiological measures (Draganich & Erdal, 2014). This constructed belief about their sleep quality affected their performance in attentional and executive function tasks. Thus, beyond therapeutic importance, it appears that subjective sleep quality can have further explanatory value for cognitive performance compared to objective measures.

One of the most widely-used sleep questionnaires is the Pittsburgh Sleep Quality Index (PSQI) (Buysse, Reynolds III, Monk, Berman, & Kupfer, 1989), a self-administered questionnaire, in which participants rate their subjective sleep quality based on several questions. These questions deal with various aspects of sleep that range from the average amount of sleep during the night, the difficulty experienced in falling asleep, and other sleep disturbances. Nevertheless, there are other popular measurements, such as the Athens Insomnia Scale (AIS) (Soldatos et al., 2000), which measures difficulties in falling asleep or maintaining sleep, as well as sleep diaries, which capture the sleeping habits of the participants from day to day, spanning a few days or weeks. Sleep questionnaires and sleep diaries are two different types of self-reported measures: while sleep questionnaires are administered at a single point in time, and ask about various aspects of sleep experience in a longer time period retrospectively, sleep diaries are ongoing, daily self-monitoring tools. Libman, Fichten, Bailes, and Amsel (2000) showed that the two measurement types are tapping the same domains but lead to somewhat different results due to methodological differences: questionnaires can be susceptible to memory distortion while sleep diaries may be distorted by atypical sleep experiences during the monitored period.

Previous research on subjective sleep quality and cognitive performance has led to mixed findings. While some studies focusing on healthy participants have shown that poorer sleep quality as measured by the PSQI score was associated with weaker working memory (van den Noort et al., 2016), executive functions (Nebes et al., 2009), and decision-making performance (Telzer, Fuligni, Lieberman, & Galván, 2013), others have failed to find an association between subjective sleep quality and cognitive performance (Miyata et al., 2013; Stepanski et al., 1989). Bastien et al. (2003) showed different associations between subjective sleep quality as measured by a sleep diary and cognitive performance in patients with insomnia who received or did not receive treatment and in elderly participants who reported good sleep quality. Interestingly, in good sleepers, greater subjective depth, quality, and efficiency of sleep were associated with better performance on attention and concentration tasks but poorer memory performance. These findings suggest that further studies are needed to clarify the complex relationship between subjective sleep quality and aspects of cognitive functioning.

Notably, these previous studies focused on diverse populations, including adolescents, elderly and clinical groups, and relied on sample sizes ranging from around 20 to 100, with smaller sample sizes potentially limiting the robustness of the observed results. In these studies, subjective sleep quality was assessed by a combination of self-reported measures, such as difficulty in sleep initiation, sleep continuity, and/or depth of sleep. In contrast to subjective sleep quality captured by a combination of such measures, self-reported sleep duration has been studied more thoroughly. In a large study with more than 100,000 participants, Sternberg et al. (2013) reported a quadratic relationship between self-reported sleep duration and performance in cognitive tasks assessing working memory and arithmetics. Furthermore, a recent powerful meta-analysis focusing on elderly participants also showed that both short and long sleep increased the odds of poor cognitive performance (Lo, Groeger, Cheng, Dijk, & Chee, 2016). A similar association was shown in another study investigating insomnia symptoms and cognitive performance in a large sample of participants (Kyle et al., 2017): self-reported sleep duration extremes were associated with impaired performance. Systematic investigations on the relationship between subjective sleep quality as captured by a combination of parameters (such as sleep latency, subjective sleep quality, sleep disturbances) and cognitive performance using larger sample sizes are, however, still lacking.

Moreover, in previous investigations focusing on the association between subjective sleep quality and various aspects of cognitive performance, the potential relationship with procedural learning/memory has largely been neglected. The procedural memory system underlies the learning, storage, and use of cognitive and perceptual-motor skills and habits (Poldrack et al., 2001). Evidence suggests that the system is multifaceted in that it supports numerous functions that are performed automatically, including sequences, probabilistic categorization, and grammar, and perhaps aspects of social skills (Fiser & Aslin, 2001; J. H. Howard, Jr. & Howard, 1997; Lieberman, 2000; Poldrack & Foerde, 2008; Pothos, 2007). Considering the importance of this memory system, the clarification of its relationship with subjective sleep quality would be indispensable.

Here we aimed to fill the gaps identified in previous research by providing an extensive investigation on the relationship between subjective sleep quality and cognitive performance in healthy young adults. Within cognitive functions, we focused on working memory, executive functions and procedural learning. We chose these domains because 1) the relationship between working memory, executive functions and subjective sleep quality has remained inconclusive, and 2) the relationship between procedural learning/memory and subjective sleep quality has largely been neglected in previous studies. Therefore, in the latter case, we explored several measures of procedural learning in order to obtain a more detailed picture of the potential associations with subjective sleep quality. To increase the robustness of our analyses, we created a database of 235 participants' data by pooling three separate datasets from our lab. We assessed subjective sleep quality by PSQI and AIS (Study 1-3), Groningen Sleep Quality Scale (GSQS, Study 2), and a sleep diary (Study 2). These separate measures capture somewhat different aspects of self-reported sleep quality and thus provide a detailed picture. We tested working memory, executive functions and several sub-processes of procedural learning in all three studies. To control for possible confounding effects, we included age, gender and chronotype as covariates in our analyses. To test the amount of evidence either for associations or no associations between subjective sleep quality and cognitive performance, we calculated Bayes Factors that offer a way of evaluating the evidence against or in favor of the null hypothesis, respectively.

Methods

Participants

Participants were selected from a large pool of undergraduate students from Eötvös Loránd University. The selection procedure was based on the completion of an online questionnaire assessing mental and physical health status. Respondents reporting current or prior chronic somatic, psychiatric or neurological disorders, or the regular consumption of drugs other than contraceptives were 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 or daily rhythms were also excluded from the study.

The data was obtained from three different studies, each with a slightly different focus. Importantly, the analyses presented in the current paper are completely novel, none of the separate studies focused on the relationship between subjective sleep quality and cognitive performance. Forty-seven participants took part in Study 1 (C. Török, Janacsek, & Nemeth, 2016), 103 participants took part in Study 2 (Simor et al., 2019), and 85 participants took part in Study 3 (Á. Takács et al., 2016). Descriptive characteristics of participants in the three studies are listed in Table 1. All participants were white/Caucasian. All participants provided written informed consent and received course credits for taking part. The studies were approved by the Research Ethics Committee of Eötvös Loránd University (201410, 2016/209). The study was conducted in accordance with the Declaration of Helsinki.

Table 5.1 Descriptive characteristics of participants

Study	**N**	**Age** ***Mean (SD)***	**Gender**	**Years in education** ***Mean (SD)***	**MEQ score** ***Mean (SD)***
Study 1	47	21.38 (1.79)	10M/37F	14.36 (1.58)	34.96 (6.69)
Study 2	103	21.62 (2.00)	30M/73F	14.50 (1.74)	33.99 (6.31)
Study 3	85	20.99 (1.59)	23M/62F	14.28 (1.60)	33.61 (5.68)

Note: M = male, F = female, MEQ = Morningness-Eveningness Questionnaire

Procedure

We conducted three separate studies on the association of subjective sleep quality and procedural learning, working memory, and executive functions in healthy young adults. The sleep questionnaires included in the studies and the timing of the procedural learning task slightly differed. While we assessed subjective sleep quality by PSQI and AIS in all three studies, in Study 2, we included further measures of subjective sleep quality as well: (1) a sleep diary to assess day-to-day general sleep quality and (2) Groningen Sleep Quality Scale (GSQS) to assess prior night's sleep quality. To control for the potential confounding effect of chronotype, we also administered the Morningness-Eveningness Questionnaire (MEQ) (Horne & Östberg, 1976; Zavecz, Török, Köteles, Pálosi, & Simor, 2015), henceforth referred to as morningness score because a larger score on this questionnaire indicates greater morningness.

In all three studies, PSQI and AIS sleep quality questionnaires and the MEQ were administered online, while the GSQS in Study 2 and the tasks assessing cognitive performance in all studies were administered in a single session in the lab. Due to technical problems, the data of six participants on executive functions are missing. To ensure that participants do the tests in their preferred time of the day, the timing of the session was chosen by the participants themselves (between 7 am and 7 pm). The timing of the sessions was normally distributed in all three studies, with most participants performing the tasks during the daytime between 11 am and 3 pm. The sleep diary in Study 2 was filled by the participants for at least one week, and to a maximum of two weeks, prior to the cognitive assessment that was scheduled based on the participants' availability.

Questionnaires and tasks

All cognitive performance tasks and subjective sleep questionnaires are well-known and widely used in the field of psychology and neuroscience (for details about each task and questionnaire, see Supplementary methods).

Subjective sleep quality questionnaires – To capture the general sleep quality of the last month, we administered the Pittsburgh Sleep Quality Index (PSQI) (Buysse et al., 1989; J. Takács et al., 2016) and the Athens Insomnia Scale (AIS) (Novak et al., 2004;

Soldatos et al., 2000). Additionally, in Study 2, we administered a Sleep diary (Gilson et al., 2015) to assess the sleep quality of the last one-two weeks, and the Groningen Sleep Quality Scale (GSQS) (Meijman, de Vries-Griever, De Vries, & Kampman, 1988; Simor, Köteles, Bódizs, & Bárdos, 2009) to capture the sleep quality of the night prior testing.

Cognitive performance tasks – Working memory was measured by the Counting Span task (Case, Kurland, & Goldberg, 1982; A. R. Conway et al., 2005; Engle, Tuholski, Laughlin, & Conway, 1999; Virag et al., 2015). Executive functions were assessed by the Wisconsin Card Sorting Test (WCST) (Berg, 1948; Nemeth, Janacsek, Polner, et al., 2013; Piper et al., 2015). The outcome measure of the WCST task was the number of perseverative errors, which shows the inability/difficulty to change the behavior despite feedback. Procedural learning was measured by the explicit version of the Alternating Serial Reaction Time (ASRT) task (Figure S1, see also (Nemeth, Janacsek, & Fiser, 2013)). There are several learning indices that can be acquired from this task. Higher-order sequence learning refers to the acquisition of the sequence order of the stimuli. Statistical learning refers to the acquisition of frequency information embedded in the task. However, previous ASRT studies often assessed Triplet learning, which is a mixed measure of acquiring frequency and sequential information (for details, see Supplementary methods). In addition to these learning indices, we measured the average reaction times (RTs) and accuracy (ACC), which reflect the average general performance of the participants across the task, and the changes in RT and ACC from the beginning to the end of the task, which indicate general skill learning that occurs due to more efficient visuomotor and motor-motor coordination as the task progresses (Hallgato et al., 2013).

Data analysis

Statistical analyses were conducted in R 3.6.1 (R Core Team, 2018) using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015). Bootstrapped confidence intervals and p-values were calculated using the boot package (Canty & Ripley, 2019; Davison & Hinkley, 1997). The data and analysis code can be found on the following link: https://github.com/nthun/performance_sleep_quality/

Analysis of the relationship between subjective sleep quality and cognitive performance – Subjective sleep quality scales (PSQI and AIS) were combined into a single metric, using principal component analysis. Then separate linear mixed-effect

models were created for each outcome measure (i.e., performance metric), where the aggregated sleep quality metric (hereinafter referred to as sleep disturbance) was used as a predictor, and 'Study' (1, 2 or 3) was added as a random intercept. This way we could estimate an aggregated effect while accounting for the potential differences across studies. To control for possible confounding effects, we included age, gender and morningness score as covariates in our analyses. Thus, the estimates reported in the Results section are controlled for these factors.

As the residuals did not show normal distribution, we used bootstrapped estimates and confidence intervals, using 1000 bootstrap samples, from which we calculated the p-values (Canty & Ripley, 2019; Davison & Hinkley, 1997). Bayes Factors (BF_{01}) were calculated by using the exponential of the Bayesian Information Criterion (BIC) of the fitted models minus the BIC of the null models – that contained the confounders only, and a random intercept by study (Wagenmakers, 2007). The BF is a statistical technique that helps conclude whether the collected data favors the null-hypothesis ($H0$) or the alternative hypothesis ($H1$); thus, the BF could be considered as a weight of evidence provided by the data (Wagenmakers, Wetzels, Borsboom, & van der Maas, 2011). It is an effective mathematical approach to show if there is no association between two measures. In Bayesian correlation analyses, $H0$ is the lack of associations between the two measures, and $H1$ states that association exists between the two measures. Here we report BF_{01} values. According to Wagenmakers et al. (2011), BF_{01} values between 1 and 3 indicate anecdotal evidence for $H0$, while values between 3 and 10 indicate substantial evidence for $H0$. Conversely, while values between 1/3 and 1 indicate anecdotal evidence for $H1$, values between 1/10 and 1/3 indicate substantial evidence for $H1$. If the BF is below 1/10, 1/30, or 1/100, it indicates strong, very strong, or extreme evidence for $H1$, respectively. Values around 1 do not support either $H0$ or $H1$. Thus, Bayes Factor is a valuable tool to provide evidence for no associations between constructs as opposed to frequentists analyses, where no such evidence can be obtained based on non-significant results.

To test the association between the additional subjective sleep quality measures and cognitive performance in Study 2, we used robust linear regression, this time without random effects. We included the same potential confounders (age, gender, morningness score), and Bayes factors were calculated in the previously described way.

Analysis of the ASRT data – Performance in the ASRT task was analyzed by repeated-measures analyses of variance (ANOVA) in each study (for details of these analyses, see Supplementary methods). Based on these ANOVAs, Triplet learning, Higher-order sequence learning and Statistical learning occurred in all three studies, both in ACC and RT (all *p*s < .001; for details, see Supplementary results and Figure S2).

Results

Cognitive performance in the three studies

The working memory capacity (measured by the counting span) and executive functions (measured by the number of perseverative errors in the WCST task) of the participants were in the standard range for their age (Heaton, 1981; Racsmány, Lukács, Németh, & Pléh, 2005). The mean counting span for the entire sample was 3.59 (*SD* = 0.85) in the three studies. This average score represents a mid-range cognitive performance, as obtainable scores range from 1 to 6. The mean number of perseverative errors was 14.76 (*SD* = 5.27) in the three studies (no maximum score can be defined in this case). For procedural learning, mean scores were 26.48 (*SD* = 26.37) for RT Triplet learning, 16.63 (*SD* = 40.34) for RT Higher-order sequence learning, 16.74 (*SD* = 9.94) for RT Statistical learning, 359.88 (*SD* = 40.94) for average RT, and 31.13 (*SD* = 30.15) for RT general skill learning. Accuracy scores were as follows: 0.04 (*SD* = 0.03) for ACC Triplet learning, 0.02 (*SD* = 0.03) for ACC Higher-order sequence learning, 0.03 (SD = 0.03) for ACC Statistical learning, 0.90 (*SD* = .10) for average ACC, -0.02 (*SD* = 0.09) for ACC general skill learning, in all three studies. Note that for accuracy, these values represent proportions (e.g., the average ACC was 90%, hence 0.90), and the learning scores are difference scores (e.g., the ACC Triplet learning score shows that participants were on average 4% more accurate on high-frequency triplets compared to the low-frequency ones). All presented RT and ACC scores represent typical values in ASRT studies with healthy young adults.

We also provide descriptive data for Study 2 separately, as additional analyses were run on cognitive performance from this dataset and GSQS and sleep diary scores. In Study 2, the mean counting span was 3.65 (*SD* = 1.01), and the mean number of perseverative errors was 14.46 (*SD* = 6.37). For procedural learning in Study 2, mean scores were 33.04

(*SD* = 27.96) for RT Triplet learning, 28.53 (*SD* = 51.44) for RT Higher-order sequence learning, 18.77 (*SD* = 9.78) for RT Statistical learning, 348.29 (*SD* = 42.26) for average RT, and 39.30 (*SD* = 34.74) for RT general skill learning. Accuracy scores were as follows: 0.03 (*SD* = 0.02) for ACC Triplet learning, 0.01 (*SD* = 0.02) for ACC Higher-order sequence learning, 0.02 (*SD* = 0.02) for ACC Statistical learning, 0.94 (*SD* = 0.03) for average ACC, 0.02 (*SD* = 0.03) for ACC general skill learning.

Overall, these values represent a mid-range cognitive performance with a sufficient level of variability in the sample to conduct the planned analyses.

Subjective sleep questionnaire scores in the three studies

The obtainable scores, means, standard deviations, and proportions of good, moderate and poor sleepers for each questionnaire are presented in Table 2. The mean scores of PSQI in the current sample were higher than the score of 1.91 for the same components in Buysse et al. (1989), and in the range or even higher than the global PSQI score (which aggregates seven components; *M* = 2.67) for the control participants, whose age was between 24 and 83 years. In the same study (Buysse et al., 1989), the participants with sleep disorders had a mean score of 4.78 for the three components of PSQI, suggesting that ~18% of the current sample had a score higher than the average score of sleep-disordered patients. The mean scores of AIS were somewhat higher than the mean score of 3 reported for a representative Hungarian adult sample in Novak et al. (Novak et al., 2004). According to the cut-off score of 10 suggested in that paper, ~5% of our sample would fall in the diagnostic category of insomnia. However, according to a stricter cut-off score of 6 suggested by Soldatos, Dikeos & Paparrigopoulos (Soldatos, Dikeos, & Paparrigopoulos, 2003), up to 23% of the participants would have complaints comparable to those of insomniac patients. The mean of the GSQS score was lower than the mean score reported for a Hungarian sample of young adults (*M* = 4.70, *SD* = 1.78) in Simor et al. (2009). The mean of the Sleep diary score in Study 2 was comparable to the mean PSQI score of 1.3 for the same components for the control participants and lower than the score of 6.36 for the participants with sleep disorders in Buysse et al. (1989).

Although with some differences across questionnaires, these sleep scores suggest a moderate to poor sleep quality of the current sample, with about 15% of participants experiencing very poor sleep quality, comparable to those of patients with sleep disorders.

Overall, all sleep measures used in the current study appear to have a sufficient level of variability to conduct the planned analyses.

Table 5.2 Descriptive statistics of the subjective sleep questionnaire scores

	Obtainable scores	Mean (SD)	Good sleepers	Moderate sleepers	Poor sleepers
			Scores (percentage of participants)		
PSQI	0–9				
All participants		2.99 (1.57)	0–1 (15.3%)	2–4 (66.4%)	5–8 (18.3%)
Study 2		2.54 (1.29)			
AIS	0–24				
All participants		3.98 (2.66)	0–2 (35%)	3–6 (50%)	7–17 (15%)
Study 2		3.41 (2.09)			
GSQS	0–14				
Study2		2.86 (2.87)	0–1 (40%)	2–7 (53%)	8–13 (7%)
Sleep dairy	0–12				
Study 2		1.38 (1.22)	0–1 (60%)	2–5 (40%)	

Note: PSQI = Pittsburgh Sleep Quality Index, AIS = Athens Insomnia Scale, GSQS = Groningen Sleep Quality Scale

Combining sleep quality metrics

Principal component analysis was used to combine PSQI and AIS into a single 'sleep disturbance' metric. The Bartlett's test of sphericity indicated that the correlation between the scales was adequately large for a PCA, $\chi^2(235) = 84.88$, $p < .0001$. One principal factor with an eigenvalue of 1.55 was extracted to represent sleep disturbance. The component explained 83.7% of the variance, and it was named 'sleep disturbance' as higher values of this metric show more disturbed sleep. The aggregated sleep disturbance index across the three studies ranged from -1.9 to 3.86.

Associations between subjective sleep quality and cognitive performance

As described above, to study the associations between subjective sleep quality and cognitive performance, separate linear mixed-effect models were created for each outcome measure (i.e., cognitive performance metric), where sleep disturbance was used as a fixed predictor, and 'Study' was added as a random intercept. Sleep disturbance did not show an association with any of the cognitive performance metrics (see Table 3 and

Fig. 1). Bayes Factors ranged from 5.01 to 14.35, indicating substantial evidence for no association between subjective sleep quality and the measured cognitive processes (Wagenmakers et al., 2011).

Table 5.3 The association of sleep disturbance with cognitive performance metrics

Outcome	***β***	**95% CI**	**df**	***p***	**BF_{01}**
ACC learning indices					
ACC Higher-order sequence learning	-.041	[-0.18, 0.11]	205	.58	12.28
ACC Statistical learning	-.038	[-0.17, 0.09]	205	.56	12.42
ACC Triplet learning	-.067	[-0.19, 0.06]	205	.30	8.50
RT learning indices					
RT Higher-order sequence learning	.014	[-0.15, 0.16]	205	.85	14.29
RT Statistical learning	-.062	[-0.21, 0.07]	205	.39	10.48
RT Triplet learning	-.028	[-0.17, 0.12]	205	.71	13.60
General skill indices					
ACC general skill learning	.037	[-0.06, 0.13]	205	.45	11.06
Average ACC	.065	[-0.04, 0.17]	205	.23	6.79
RT average	-.019	[-0.17, 0.12]	205	.80	14.05
RT general skill learning	-.075	[-0.23, 0.07]	205	.33	8.83
WM and EF indices					
Counting Span	-.013	[-0.17, 0.14]	205	.87	14.35
WCST – perseverative error	.107	[-0.03, 0.24]	199	.13	5.01

Note: The table shows standardized regression coefficients for sleep disturbance, where the 'Study' random intercept was included in separate linear mixed-effect models for each cognitive performance metrics. Age, gender, and morningness score were added as covariates. BF_{01} was derived from BIC (see the 'Data analysis' section for details). ACC = accuracy. RT = reaction time. WM = working memory. EF = executive function. WCST = Wisconsin Card Sorting Test.

To test whether AIS or PSQI scores separately are associated with cognitive performance, we performed similar analyses as for the sleep disturbance metric. Additionally, we also tested whether cognitive performance differed between "good" and "poor" sleepers as defined by the extremes in the overall PSQI score. For this analysis,

we considered those with a score of 0 or 1 as good sleepers (N = 36), while those with a score of 5 to 8 as poor sleepers (N = 43), corresponding to approximately the upper and lower 15% of the data (see Table 2). These additional analyses (reported in the Supplementary results) are consistent with the above findings for the sleep disturbance metric, suggesting no relationship between subjective sleep quality and cognitive performance using these measures.

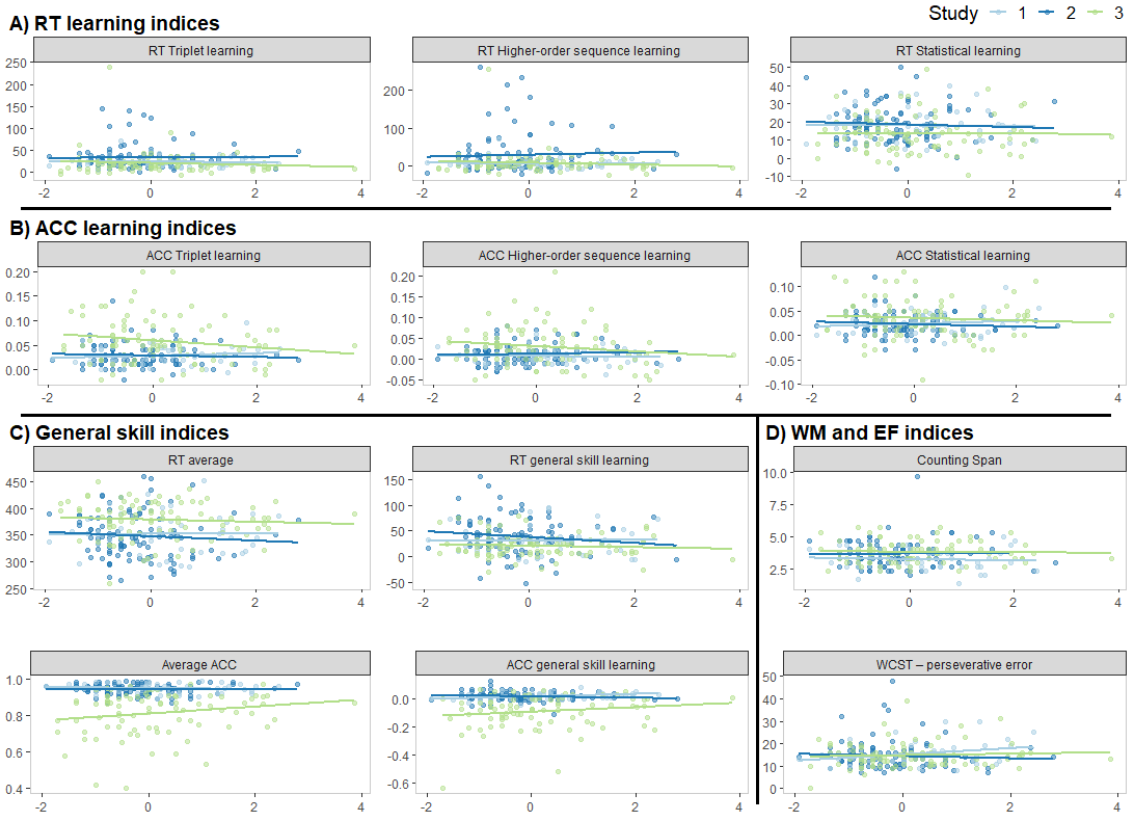

Figure 5.1 Association between sleep disturbance and cognitive performance metrics by study. Horizontal axes represent the sleep disturbance index, while vertical axes represent the outcome variables, with their names shown in the panel titles. The scatterplots and the linear regression trendlines show no association between subjective sleep quality and procedural learning indices in terms of reaction time (RT, **A**), or accuracy (ACC, **B**), general skill indices in terms of RT or ACC (**C**), and working memory and executive function indices (**D**).

In Study 2, to investigate the associations between further subjective sleep quality questionnaires and cognitive performance, we created a separate linear mixed-effect model for each outcome measure (i.e., cognitive performance metric), and each additional sleep questionnaire (i.e., sleep diary and GSQS). Sleep diary scores did not show association with any of the cognitive performance metrics (all *p*s > .05, see Table 4 and

Fig. 2). Bayes Factors ranged from 2.51 to 12.58, indicating, in all but one cases, substantial evidence for no association between subjective sleep quality and measures of cognitive performance (Wagenmakers et al., 2011). The lowest value of 2.51 for ACC general skill learning also pointed to the same direction, indicating slightly weaker evidence for no association with subjective sleep quality.

Table 5.4 The association of sleep diary with cognitive performance metrics in Study 2

Outcome	***β***	**95% CI**	***t***	**df**	***p***	**BF_{01}**
ACC learning indices						
ACC Higher-order sequence learning	-.077	[-0.28, 0.13]	-0.749	97	.46	7.73
ACC Statistical learning	-.031	[-0.24, 0.17]	-0.296	97	.77	8.09
ACC Triplet learning	-.111	[-0.31, 0.09]	-1.092	97	.28	4.46
RT learning indices						
RT Higher-order sequence learning	-.001	[-0.11, 0.11]	-0.025	97	.98	9.76
RT Statistical learning	-.205	[-0.41, 0.00]	-1.955	97	.05	8.96
RT Triplet learning	-.059	[-0.19, 0.07]	-0.917	97	.36	11.28
General skill indices						
ACC general skill learning	-.171	[-0.35, 0.01]	-1.866	97	.07	2.51
Average ACC	.035	[-0.18, 0.25]	0.317	97	.75	8.94
RT average	-.086	[-0.31, 0.13]	-0.764	97	.45	12.79
RT general skill learning	-.064	[-0.26, 0.14]	-0.623	97	.53	7.10
WM and EF indices						
Counting Span	-.065	[-0.26, 0.13]	-0.664	97	.50	5.63
WCST – perseverative error	.005	[-0.13, 0.14]	0.072	96	.94	9.71

Note: The table shows standardized regression coefficients for sleep diary scores in separate linear mixed-effect models for each cognitive performance metrics. Age, gender, and morningness score were added as covariates. BF_{01} was derived from BIC (see 'Data analysis' section for details). ACC = accuracy. RT = reaction time. WM = working memory. EF = executive function. WCST = Wisconsin Card Sorting Test.

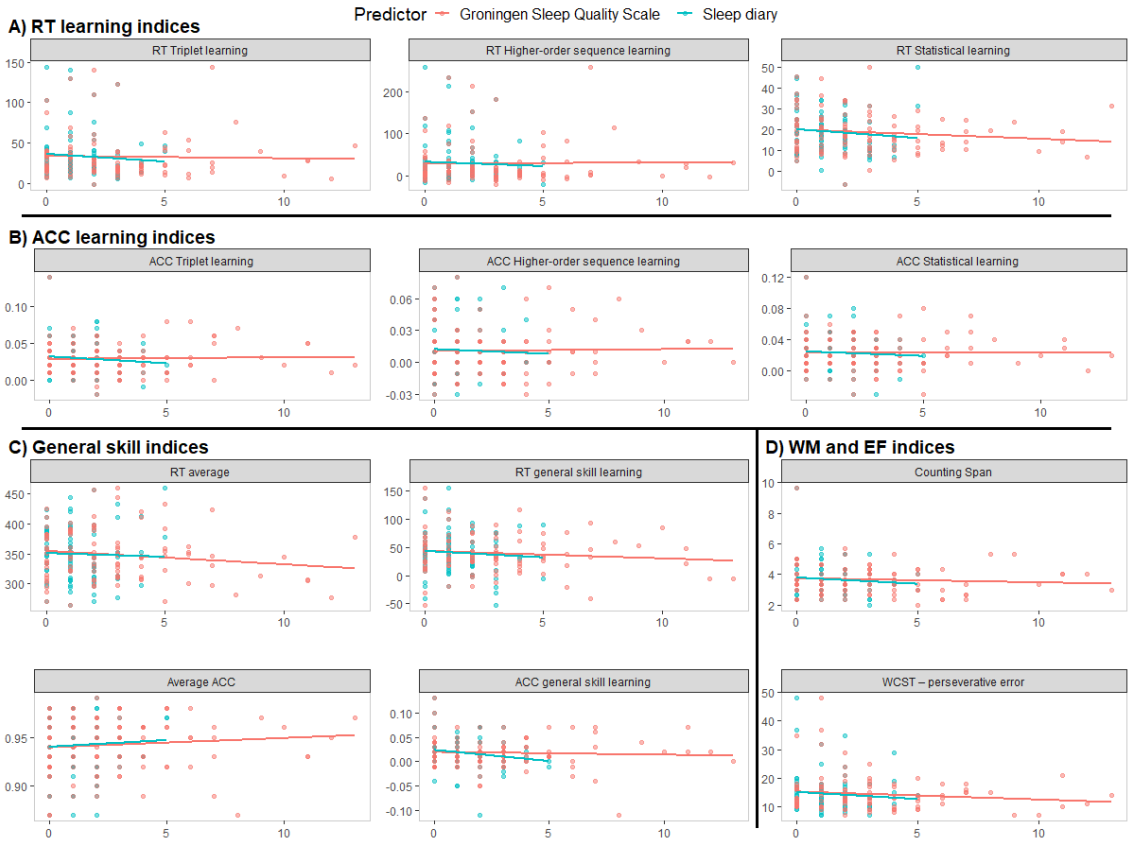

Figure 5.2 Association between sleep diary and GSQS scores and cognitive performance metrics. Horizontal axes represent the sleep disturbance index, while vertical axes represent the outcome variables, with their names shown in the panel titles. The scatterplots and the linear regression trendlines show no association between subjective sleep quality (measured with a sleep diary (blue) or the GSQS (red)) and procedural learning indices in terms of reaction time (RT, **A**), or accuracy (ACC, **B**), general skill indices in terms of RT or ACC (**C**), and working memory and executive function indices (**D**).

Similarly, GSQS scores did not show association with any of the cognitive performance metrics (all *p*s > .11, see Table 5 and Fig. 2). Bayes Factors ranged from 3.46 to 16.46, indicating substantial evidence for no association between subjective sleep quality and the measured cognitive processes (Wagenmakers et al., 2011).

Table 5.5 The association of GSQS with cognitive performance metrics in Study 2

Outcome	*β*	95% CI	*t*	df	*p*	BF_{01}
ACC learning indices						
ACC Higher-order sequence learning	.029	[-0.17, 0.23]	0.278	102	.78	10.87
ACC Statistical learning	-.001	[-0.20, 0.20]	-0.013	102	.99	10.08
ACC Triplet learning	.000	[-0.20, 0.20]	0.000	102	1.00	10.15
RT learning indices						
RT Higher-order sequence learning	-.004	[-0.11, 0.10]	-0.070	102	.94	10.14
RT Statistical learning	-.105	[-0.32, 0.11]	-0.973	102	.33	5.39
RT Triplet learning	-.054	[-0.17, 0.07]	-0.866	102	.39	16.46
General skill indices						
ACC general skill learning	.040	[-0.13, 0.21]	0.452	102	.65	12.28
Average ACC	.156	[-0.05, 0.36]	1.466	102	.15	5.16
RT average	-.176	[-0.39, 0.04]	-1.617	102	.11	3.46
RT general skill learning	-.104	[-0.30, 0.09]	-1.039	102	.30	5.85
WM and EF indices						
Counting Span	-.062	[-0.26, 0.13]	-0.632	102	.53	6.07
WCST – perseverative error	-.009	[-0.13, 0.14]	-0.133	101	.89	9.22

Note: The table shows standardized regression coefficients for GSQS scores in separate linear mixed-effect models for each cognitive performance metrics. Age, gender, and morningness score were added as covariates. BF_{01} was derived from BIC (see the 'Data analysis' section for details). ACC = accuracy. RT = reaction time. WM = working memory. EF = executive function. WCST = Wisconsin Card Sorting Test.

Discussion

Our aim was to investigate the relationship between subjective sleep quality and cognitive performance in healthy young adults. Cognitive performance was tested in the domains of working memory, executive functions, and procedural learning. To provide more reliable results, we pooled data from three different studies, controlled for possible confounders, such as age, gender and chronotype, and performed robust frequentists as well as Bayesian statistical analyses. We did not find associations between subjective sleep quality and cognitive performance measures using the robust frequentist statistical analyses. Moreover, the Bayes factors provided substantial evidence for no association between subjective sleep quality and measures of working memory, executive functions,

and procedural learning. This pattern held when subjective sleep quality was reported retrospectively for a longer period (i.e., a month; with PSQI and AIS), as well as when monitored daily (for one to two weeks; with the sleep diary) or reported for the night prior to testing (with GSQS). These results suggest that neither moderately persistent nor transient subjective sleep quality is associated with cognitive performance in healthy young adults.

There are several factors to consider why subjective sleep quality showed no associations with cognitive performance in our sample of healthy young adults. First, it is possible that methodological issues contributed to the null effects. For example, having a lower range of obtainable scores on the selected subjective sleep quality and cognitive performance measures can limit the possibility of finding a relationship between these measures. Importantly, all measures that we used in the current study have been well-established in previous research and have a reasonable range of obtainable values. Although the sample choice of healthy young adults has naturally limited the range of scores on the used measures, our analyses showed a sufficient level of variability in all measures. Therefore, the obtained null results seem unlikely to be explained by such methodological issues.

Second, as we studied healthy university students, there may be a ceiling effect in subjective sleep quality. Sleep disturbance can be more prevalent in elderly populations and in clinical disorders (Buysse et al., 1989; Novak et al., 2004). Consequently, variance and extremities in subjective sleep quality could be greater in these populations, while it can remain relatively low in healthy young adults. Nevertheless, previous research has found that university students are also prone to sleep disturbances, and in particular to chronic sleep deprivation (Gaultney, 2010). Although with some variation across sleep questionnaires, most participants' subjective sleep quality ranged from moderate to poor in our sample, with about 15% of participants experiencing very poor sleep quality similar to those of patients with sleep disorders. Thus, it seems unlikely that the obtained results are due to a ceiling effect in subjective sleep quality.

Third, it is possible that because young adults typically show a peak cognitive performance, poor subjective sleep quality may not have a substantial impact on it. In line with this explanation, the studies that reported associations between subjective sleep quality and cognitive performance (Nebes et al., 2009; van den Noort et al., 2016)(Telzer

et al., 2013) focused primarily on adolescents, older adults, or clinical populations, where cognitive performance has not yet peaked or have declined. Further supporting this explanation, Saksvik, Bjorvatn, Hetland, Sandal, and Pallesen (2011) found in their meta-analysis that young adults are not as prone to the negative consequences of shift work as the elderly. Moreover, Gao, Terlizzese, and Scullin (2019) in a recent study showed that above-average cognitive abilities buffer against insufficient sleep durations. However, not all cognitive functions peak in adulthood: while previous studies have reported the best performance in working memory and executive functions in young adulthood (Craik & Bialystok, 2006; Tanczos, Janacsek, & Nemeth, 2013a, 2013b; Zelazo, Craik, & Booth, 2004), some aspects of procedural learning (as measured by the ASRT task) has been shown to peak in childhood and to decline already around adolescents (Janacsek et al., 2012; Juhasz, Nemeth, & Janacsek, 2019; Nemeth, Janacsek, & Fiser, 2013). Consequently, a cognitive peak may explain finding no relationship between subjective sleep quality and aspects of working memory and executive functions, while this explanation for the measures of procedural learning seems unlikely.

Fourth, the conditions under which the data collection took place could have also contributed to the null results. We conducted our experiments during the term-time when the workload in the university is typically moderate. Moreover, students could choose the time of day for cognitive testing, and they may have chosen a time when they typically felt well-rested. There is evidence that performing in a preferred circadian time period can attenuate the effect of sleep disturbances (Goel, Basner, Rao, & Dinges, 2013). Consistently, previous studies showed that participants exhibit better performance on working memory and executive functions tasks in their preferred time of day (Matchock & Mordkoff, 2009; Rowe, Hasher, & Turcotte, 2009). However, a recent study found that participants, in fact, exhibit weaker performance in procedural learning in their preferred time of day, and better performance in their non-preferred time of day, suggesting variability in the relationship between circadian effects and cognitive functions (Delpouve, Schmitz, & Peigneux, 2014). Additionally, independent of the time of day, participants may have perceived the session with the cognitive tasks as a testing situation and may have been motivated to show their best performance, compensating for any possible effect of poor subjective sleep quality. Indeed, there is evidence that highly motivated participants are less prone to the effect of sleep deprivation (Hull, Wright Jr,

& Czeisler, 2003). Thus, the time of testing and participants' motivation may have contributed to our findings by potentially compensating for any negative effects of poor subjective sleep quality on cognitive performance.

Fifth, the relationship between sleep and cognitive performance can vary depending on what parameters of sleep are assessed. Associations between objective sleep quality (measured by actigraphy or electroencephalography) and various aspects of working memory, executive functions and procedural learning have been frequently reported in previous studies (for a review, see (Diekelmann & Born, 2010; Jones & Harrison, 2001)). Here we showed that subjective sleep quality is not associated with these cognitive functions, at least under the circumstances described above. As outlined in the Introduction, this dissociation suggests that objective and subjective sleep quality, although measure the same domains, do not necessarily capture the same aspects of sleep quality and sleep disturbances (Armitage et al., 1997). Subjective sleep quality may be estimated based on a combination of objective sleep parameters. Moreover, some objective parameters of sleep that contribute to cognitive performance may not be captured with self-reported instruments. For example, it is often reported that spindle activity or time spent in slow-wave sleep (SWS) or in REM sleep is essential for memory consolidation (Clemens, Fabo, & Halasz, 2005; Siegel, 2001; Walker, 2009). Also, in laboratory sleep examinations, sleep quality is usually carefully controlled for several days prior to the examination. Potentially, the objective sleep parameters showing associations with cognitive performance may only be measured in these carefully controlled conditions (i.e., when sleep quality on the night of testing as well as in the preceding days are good). Hence, it is possible that while results with objective sleep quality may show how healthy sleep is related to cognitive functioning, results with subjective sleep quality may reflect how aspects of sleep disturbances are related to cognitive functioning.

Sixth, and relatedly, there could be differences in the association with cognitive performance within self-reported measures of sleep as well. In our study, we captured the perceived disturbances in initiating and maintaining sleep rather than the self-reported duration of sleep. While we found no associations between these measures of subjective sleep quality and cognitive performance, there is solid evidence that self-reported extreme sleep durations (both long and short sleep times) are associated with worse cognitive

performance (Kyle et al., 2017; Lo et al., 2016; Sternberg et al., 2013). These findings suggest a dissociation between sleep quality as measured by extreme self-reported sleep durations and other types of sleep quality disturbances.

Seventh, it is possible that while interindividual differences in subjective sleep quality do not contribute to at least some aspects of cognitive performance, intraindividual fluctuations do. The possible importance of intraindividual rather than interindividual differences was also suggested by Ackermann, Hartmann, Papassotiropoulos, de Quervain, and Rasch (2015) in a large study, in which contrary to previous studies they showed no associations between declarative memory consolidation and objective sleep parameters. Further studies are warranted to test whether day-to-day variations in subjective sleep quality predict day-to-day changes in cognitive performance.

Finally, our paper has some limitations. As mentioned above, it is possible that investigating populations more susceptible to sleep disturbances or cognitive performance problems could yield different results, and the lack of associations could be specific to healthy young adults. Furthermore, it would be interesting to test whether individual differences in other factors (for example, interoceptive ability, i.e., how accurately one perceives their own body sensations) influence the relationship between subjective sleep quality and cognitive performance.

Conclusions

In conclusion, we showed that self-reported, subjective sleep quality is not associated with working memory, executive functions, and various aspects of procedural learning in a relatively large sample of healthy young adults. These findings were supported not only by frequentist statistical analyses but also by Bayes factors that provided substantial evidence for no associations between these functions. Importantly, however, our findings do not imply that sleep per se has no relationship with these cognitive functions; instead, it emphasizes the dissociation between subjective and objective sleep quality. We believe that our approach of systematically testing the relationship between self-reported sleep questionnaires and a relatively wide range of cognitive functions can inspire future systematic studies on the relationship between subjective/objective sleep parameters and cognition. Within healthy young adults, future studies are warranted to probe the

relationship between subjective sleep quality and cognitive performance assessed in the non-preferred time of day, include other aspects of cognitive functions, and test intraindividual, day-to-day variations in the relationship between sleep and cognitive performance.

General discussion

The aim of the dissertation was twofold: to better characterize procedural memory on the behavioral and on the neural level. In order to do this, we aimed to answer different questions in the four studies of the dissertation. In Study 1, we examined both declarative and non-declarative memory in a sleep-disordered population to determine their relation to sleep. Study 2 focused on the neural background of procedural memory by examining the causal role of theta oscillations. In Study 3, besides aiming to characterize two subprocesses of procedural memory, we also explored the neural background of its consolidation, in particular, whether sleep has a prominent role in it. Lastly, in Study 4, we focused on subjective sleep parameters and their possible associations with different subprocesses of procedural memory.

How declarative and non-declarative memory is related to sleep in a sleep-disordered population?

In Study 1, we investigated the differential association of declarative and non-declarative memory with sleep in a sleep-disordered population. More precisely, our aim was to investigate whether Slow-wave sleep (SWS) spectral power is associated with learning capacity and overnight memory consolidation within a group of children with Sleep-disordered breathing (SDB). Moreover, we applied both a declarative and a non-declarative memory task in order to explore the specificity of sleep-related memory impairments in SDB. Our results confirmed that slow frequency activity was associated with declarative learning capacity: delta power (1-4 Hz) during post-learning SWS was positively, whereas theta power (4-8 Hz) was negatively associated with declarative learning capacity. However, we did not find any associations between the spectral composition of SWS and non-declarative learning capacity. Apart from learning capacity, we also did not find associations between the spectral composition of SWS and the overnight memory consolidation neither in the declarative nor in the non-declarative task.

This finding might suggest that the memory impairments associated with sleep disorders are the result of chronic insufficient sleep quality affecting learning capacity rather than overnight insufficient sleep quality affecting consolidation mechanisms. However, specific studies are warranted to test this hypothesis, as lower learning capacity

could lead to floor effects in consolidation, i.e., individuals suffering from sleep disorders learn less, and therefore have the capacity to consolidate all learned content. In contrast, if they were to learn as much as their good-sleeper peers, we might saw impairments in consolidation performance as well. However, it is also possible that the relationship between sleep parameters and learning impairments appears due to a third factor that leads to both sleep disturbances and cognitive impairments. As we studied a sleep disorder in which the sleep fragmentation is a result of breathing difficulties, a likely candidate for a third factor is abnormal respiratory patterns. Studies have shown that the reduced oxygen delivery during sleeping in SDB can result in neuronal damage, especially in the prefrontal cortex (Beebe & Gozal, 2002; Blunden & Beebe, 2006). Thus breathing difficulties can result both in sleep fragmentation and through affecting the maturation of the prefrontal cortex, cognitive impairment. However, the severity of the respiratory problems does not predict the severity of the cognitive impairments, i.e. children with milder or severe SDB symptoms show similar cognitive impairments (Archbold et al., 2004; Bourke et al., 2011a, 2011b; Csabi et al., 2015; Gottlieb et al., 2004).

Our results of associations between sleep parameters and declarative but not with non-declarative memory performance are in line with previous studies comparing the performance of children with and without sleep disorders, showing declarative memory impairment together with intact non-declarative memory performance (Csábi et al., 2013; Gottlieb et al., 2004; Nemeth et al., 2012; Nemeth, Janacsek, Londe, et al., 2010). Albeit some studies reported sleep-dependent memory effects in case of similar procedural learning tasks (Durrant et al., 2013; Durrant et al., 2011; Urbain et al., 2013), these tasks differ in their methodology and presumably, also in their neural correlates (Durrant et al., 2013; Durrant et al., 2011; Janacsek et al., 2015; Nemeth, Janacsek, Király, et al., 2013; Urbain et al., 2013). Several studies indicate that sleep-related benefits of memory consolidation are restricted to skill-learning paradigms that require attention, intentional learning or contain an explicit representation of the sequence (Robertson et al., 2004; Song & Cohen, 2014; Wilhelm et al., 2011). In contrast, learning within the ASRT task is implicit, and occurs without explicit awareness (J. H. Howard, Jr. & Howard, 1997; Nemeth, Janacsek, Londe, et al., 2010; Song et al., 2007b). To sum up, our results indicate that declarative and non-declarative memory is differently associated with sleep: we found oscillatory mechanisms during sleep that showed associations with declarative

memory performance, but we did not reveal associations between sleep parameters with non-declarative memory performance. We cannot rule out however, that the lack of associations between sleep and non-declarative memory is specific to the measured type of procedural memory.

Even though this was not the primary focus of the study, it is important to note, that these associations and effects of sleep disruption could be especially impactful during development. Research showed a stronger association between sleep quality and neurobehavioral functioning in younger children than in older children (Sadeh, Gruber, & Raviv, 2002). This higher vulnerability to poor sleep in early childhood could be explained by the important prefrontal cortex development occurring during early adolescence (Casey, Tottenham, Liston, & Durston, 2005). This early childhood period is especially characterized by dramatic prefrontal cortex changes in structural architecture and functional organization that decline throughout adolescence. Based on these findings, we can assume that the influence of low sleep quality on prefrontal cortex functions impacts cognitive and school performance during development. Considering this vulnerability in brain maturation in childhood, the question of reversibility of neurocognitive symptoms of children with sleep disorders raises. Yet, studies investigating the effect of treatment in pediatric sleep disorders are scarce. In a follow-up study that is currently ongoing, we are aiming to answer these important questions as well.

Thesis statement 1: Declarative memory is associated with parameters of sleep in a sleep-disordered population, whereas non-declarative is not.

Thesis statement 2: Slow oscillations (delta and theta) of sleep are relevant for declarative memory.

Is theta oscillation crucial for procedural memory?

In Study 2, we investigated the neural background of procedural memory by directly manipulating oscillatory activity during learning: we tested the causal relationship between fronto-parietal midline theta synchronization and procedural learning with non-invasive transcranial alternating current stimulation (tACS). We could not find stimulation effects, the overall learning performance and the time course of learning did

not differ between the stimulation and the control conditions. This does not mean, however, that theta synchronization is irrelevant for procedural memory measured by the ASRT task. In the lack of an EEG recording simultaneously with or following the stimulation, we could not confirm that the brain stimulation induced the expected changes in brain activity. We chose the stimulation parameters based on studies that could influence brain activity/cognitive performance with a similar setting (Chander et al., 2016; Polanía et al., 2015; Violante et al., 2017; Vosskuhl et al., 2015). However, it is likely that the chosen parameters for the tACS stimulation, such as the fronto-parietal midline montage, the relatively weak (1 mA) current intensity, and/or the chosen theta frequency were not appropriate to influence learning in our study. Regarding the current intensity, there have been several recent animal studies suggesting that the usual intensities applied in transcranial electric stimulation studies (including our study) are not sufficient to elicit significant changes in brain activity (Khatoun, Asamoah, & Mc Laughlin, 2019; Krause, Vieira, Csorba, Pilly, & Pack, 2019; Vöröslakos et al., 2018). To be able to confirm if certain brain dynamics are crucial for procedural memory, the changes in brain activity induced by the brain stimulation must be measured and proved. Unfortunately, such devices that can simultaneously measure the electrical activity of the brain and apply electrical stimulation are hardly accessible and were not available for our research group for this study. It would worth replicating this study (possibly with higher current intensity) with methods that enable us to confirm the induced changes in brain activity due to the stimulation.

Thesis statement 3: Non-invasive brain stimulation is a powerful tool to test the causal relationship between brain dynamics and memory performance.

Is sleep essential for the consolidation of different subprocesses of procedural memory?

In Study 3, we investigated the consolidation of sequence and statistical knowledge in case of the off-line period spent sleeping or awake. We could not find differences in memory consolidation in the groups of participants who, after learning, either slept, rested quietly or watched a movie. Both sequence and statistical learning were preserved during

the off-line period independent of the activity of the period. These results are in line with previous studies that did not find beneficial effects of sleep on procedural memory tasks involving regularities (Peigneux et al., 2003, 2006; Song et al., 2007a; Nemeth et al., 2010; Hallgató et al., 2013). However, EEG spectral power assessed during the off-line period showed associations with memory consolidation during sleep, but not during awake states. This association was selective for sequence learning, namely frontal theta band power during sleep showed a positive association with the consolidation of sequence knowledge. In addition, frontal theta power also (and more strongly) predicted further improvements in sequence learning after additional practice following sleep. This finding suggests that not sleep per se, but low-frequency oscillations (2-8 Hz) are associated with memory consolidation and delayed performance gains.

Albeit the primary focus of the study was to explore the effects of sleep and the neural background of procedural memory, this was also one of the first studies characterizing the subprocesses of procedural memory, sequence and statistical learning, separately within the same paradigm. Our findings indicate that sequence and statistical learning have different learning trajectories. Sequence learning exhibited a steady increase with practice, even after returning to the task following the off-line period. In contrast, statistical learning was acquired rapidly and remained stable throughout practice. Our fine-grained analyses showed that statistical learning occurs already after very little exposure to regularities, although additional training is required to strengthen the acquired statistical knowledge. Regarding consolidation, both forms of learning were preserved as no significant off-line changes emerged in either sequence or statistical learning. This latter could indicate, that the consolidation of these two subprocesses might rely on more similar mechanisms than their acquisition. However, the dissociation of the relevant oscillations during the off-line period indicates otherwise. In a follow-up study, we are aiming to reveal in greater detail whether the studied subprocesses rely on distinct neural oscillations.

Thesis statement 4: Sequence and statistical learning show different learning trajectories.

Thesis statement 5: Sleep does not seem to benefit the consolidation of procedural memory more than wakefulness.

Thesis statement 6: Sleep-specific slow oscillations however are associated with the consolidation of sequence knowledge.

Are subjective measures of sleep associated with different subprocesses of procedural memory?

In Study 4, we investigated whether subjective sleep quality is associated with procedural memory. According to our results, none of the various aspects of procedural learning that we investigated (alongside with working memory and executive functions) showed associations with self-reported sleep quality. Our findings do not imply that sleep per se has no relationship with these cognitive functions. There is a great dissociation between subjective and objective sleep quality (Armitage et al., 1997; Guedes et al., 2016; Landry et al., 2015b) therefore generalization from one aspect of sleep to the other should be avoided. Instead, this is another aspect of sleep that does not seem to affect procedural memory.

Reviewing the literature on the relationship between subjective sleep quality and cognition revealed a pattern in studies that found and studies that did not find associations. Studies that reported associations between subjective sleep quality and cognitive performance (Nebes et al., 2009; Telzer et al., 2013; van den Noort et al., 2016), focused primarily on adolescents, older adults, or clinical populations, where cognitive performance has not yet peaked or have declined. In contrast, we studied healthy university students, who are at their peak cognitive performance. Saksvik et al. (2011) found in their meta-analysis that young adults are not as prone to the negative consequences of shift work as the elderly. Moreover, Gao et al. (2019) in a recent study showed that above-average cognitive abilities buffer against insufficient sleep durations. Differential associations between populations more or less susceptible to negative consequences of sleep (or other environmental factors) can affect the findings of all studies investigating sleep and memory, including the studies in this dissertation. Therefore, the lack of associations showed in this dissertation (and other studies) investigating healthy young adults should be treated with caution and should not be automatically generalized to sleep and memory overall. Another good approach would be to probe the relationship between (subjective) sleep quality and cognitive performance

within healthy young adults in a non-ideal setting for cognitive performance, for instance in a non-preferred time of day.

Thesis statement 7: Subjective sleep quality is not associated with procedural memory, as in sequence, statistical or general skill learning and motor abilities.

Conclusions of the studies on the behavioral level

On the behavioral level, we provided evidence for the differentiation of the investigated two subprocesses of procedural memory: sequence and statistical learning have different learning trajectories. Importantly, in our studies, we defined sequence and statistical learning as processes sensitive to serial-order and frequency/probability information, respectively. However, from a theoretical perspective sequence and statistical learning could be both considered as statistical learning at the level of transitional probabilities. The difference between these two forms of learning at the level of transitional probabilities is that in case of sequence learning, the second order transitional probability is one, i.e. a sequence element can be predicted with a 100% certainty from the previous sequence element. In contrast, in case of statistical learning (as defined in our studies) the second order transitional probability is less than one, as there is always one probable continuation and some less probable continuations for the elements. However, our results in Study 3 and 4, as well as previous studies (Kóbor et al., 2018; Nemeth, Janacsek, & Fiser, 2013; Szegedi-Hallgató et al., 2017) suggest that this difference is significant enough to consider these two processes as separate.

However, even in these subprocesses of procedural memory, modification within the same task can still cause great changes in learning trajectories. For instance, sequence learning has been previously shown to occur slowly in the implicit ASRT task, over several sessions and days typically (Howard and Howard, 1997; Howard et al., 2004), while this can be substantially faster (one session) if explicit instruction is provided to the participants (Kóbor et al., 2018; Nemeth, Janacsek, & Fiser, 2013). In contrast, statistical learning occurs very quickly in both task versions (Horvath, Torok, Pesthy, Nemeth, & Janacsek, 2019; Kóbor et al., 2018; Szegedi-Hallgató et al., 2017). The further trajectory of statistical learning, however, also differs in the implicit and explicit version of the task. The explicit instruction about the sequence seems to suppress the gradual increase of

statistical knowledge: while the trajectory of statistical learning remains stable in the explicit ASRT, it gradually increases throughout practice in the implicit version. Similarly to the implicit version, in the explicit version, when the task is fix-paced instead of self-paced, a gradual increase of statistical knowledge can be observed (Horvath et al., 2019; Szegedi-Hallgató et al., 2017). Considering that in the fixed-paced ASRT task, the participants have to give very quick answers, it is possible that altogether it is not the explicit knowledge, but the use of that knowledge is the decisive factor regarding the trajectory of statistical learning. In conditions where participants cannot rely on strategies containing the explicit sequence (in a lack of the explicit knowledge, or the lack of the time to use that knowledge), they show an increase in learning throughout practice, whereas their performance remains stable when they rely on such strategies. These dynamics all suggest that there might be a competition between the explicit focus on the sequential information and the implicit acquisition of the statistical regularities (Hardwick, Forrence, Krakauer, & Haith, 2019). This competition has also been shown in a developmental setting (Nemeth, Janacsek, & Fiser, 2013). In conclusion, the different parameters within the same task cause great variability in the behavioral characteristics of even well dissected subprocesses of memory. This variability could indicate that our taxonomy for memory might not consider (all) the relevant factors.

Is it time to replace the classical taxonomies of memory?

There have been several suggestions for alternative memory taxonomies over the past decades (Henke, 2010; Konkel & Cohen, 2009; P. J. Reber, 2013; Reder, Park, & Kieffaber, 2009; Shanks & St. John, 1994). P. J. Reber (2013) drew attention to the problem of the mere definition of non-declarative or implicit memory relying on an absence: the absence of MTL dependence in case of non-declarative and absence of consciousness in case of implicit memory. The MTL dependence is problematic as it seems to play a role in memory processes that were assumed to be independent of it (Chun & Phelps, 1999; Hannula & Greene, 2012). The definition by consciousness is problematic for two reasons: one reason is that it is not a binary construct, but instead a continuous scale. Relatedly, several memory processes can occur with or without conscious awareness as well (see sequence learning in our studies). The other reason why the definition by consciousness is problematic is that the variety of how implicit memory

can be observed is so great, that it is impossible to find a common ground to them. For this reason, Willingham and Preuss (1995) suggested abandoning the term entirely.

A common feature of the alternative taxonomies is to account more for the brain's information processing activities, as memory is a fundamental property of that (Konkel & Cohen, 2009; P. J. Reber, 2013; Reder et al., 2009). However, most of the alternative taxonomies only tried to characterize the non-declarative/implicit memory system better in order to have a processing characteristic that fully distinguishes this system from declarative/explicit memory. In contrast, Henke (2010) proposed a taxonomy where the boundaries of declarative and non-declarative memory were not considered any more: subtypes of memory that were traditionally characterized as declarative or non-declarative memory became part of the same category in this taxonomy. The taxonomy distinguishes three basic processing modes: rapid encoding of flexible associations, slow encoding of rigid associations and rapid encoding of single or unitized items. In this way, for instance, semantic and procedural memory both fall in the same (second) category. However, this classification still encompasses the same memory types (episodic, semantic, procedural, etc.), ultimately.

These definitional problems present a setback in memory research. It is futile to try to characterize memory types without a good working definition of those types. Due to this problem, memory research is currently characterizing behavioral and neural correlates of specific memory tasks rather than memory types themselves. The misleading classification also makes it hard to oversee which memory processes have more similar behavioral characteristics and neural correlates as these could be memory processes traditionally categorized into distinct memory types (e.g., declarative and non-declarative). This rigidity prohibits a new, better taxonomy to emerge. While it was not the aim of this dissertation to provide a novel memory taxonomy, the findings of the studies emphasize the problems with the current classification.

Are phases of memory clear?

A second important classical division that can deteriorate memory research and should be considered in future research is the phases of memory. Traditionally, we differentiate learning, consolidation and retrieval. However, memory consolidation is a poorly defined construct. At this time, it is unclear how memories are altered after initial

encoding, as well as there is no consensus as to which of the processes contributing to this alteration should be included under the umbrella term of memory consolidation. Instead of a single process, we can differentiate phases or types of consolidation as well, such as stabilization, enhancement or integration (Stickgold & Walker, 2007). More recently, the concept of memory reconsolidation was also introduced, referring to the phenomenon of when previously stabilized memories are reactivated, they return to a labile state in which they are again susceptible to destructive interference, and need to be stabilized in a subsequent off-line period again (Nader, 2003). The focus of the current study was learning and consolidation, therefore, the phenomenon of reconsolidation did not affect the results of our studies. However, recently another consolidation mechanism has been identified, that could be relevant in our studies. Consolidation traditionally thought to take an extensive amount of time, however, recent papers show that it can occur in an extremely short period as well, on a scale of seconds (Bönstrup et al., 2019; Robertson, 2019). This phenomenon was referred to as ultra-fast offline improvement (Robertson, 2019). As the learning period in our studies took a relatively long time, the ASRT task was administered in blocks. These ultra-fast consolidation processes likely occurred between the blocks, leading to a mixed measure of acquisition and consolidation in each learning period. According to a new study, the extraction of regularities in a procedural memory task (i.e., the memory performance we were primarily interested in) is not affected by these consolidation processes (Fanuel et al., 2020). Nevertheless, future studies investigating memory processes that are measured by tasks that contain short resting periods should address this question specifically.

The relationship between sleep and procedural memory on the behavioral level

Other than characterizing the processes of procedural memory, we also wanted to explore its relationship with sleep. On the behavioral level, we did not find a beneficial effect of sleep on learning capacity or consolidation of procedural memory. In Study 3, we did not find differences in memory consolidation following post-learning sleep or wakefulness. In Study 4, we showed that there is no association between subjective sleep quality and procedural memory capacity. In both studies, the lack of association between memory performance and sleep was present both for sequence and statistical learning.

Importantly, we did not find associations between sleep and procedural memory with neither objective nor subjective sleep parameters. This is in line with previous studies that also did not show associations between sleep and procedural memory (Csabi et al., 2015; Nemeth et al., 2012; Wilhelm et al., 2008).

Conclusions of the studies on the neural level

The relationship between sleep and procedural memory on the neural level

In contrast to the behavioral level, it is not clear if sleep, or more specifically sleep-related oscillations on the neural level could affect procedural memory. In Study 1, the spectral composition of SWS did not correlate with procedural memory (in contrast with declarative memory that showed associations with the slow oscillations). In Study 3, however, where we dissected sequence and statistical learning, the findings were mixed. Statistical memory performance did not show associations with neural activity during sleep, whereas, we found associations between the spectral composition of sleep and the consolidation of sequence knowledge. Interestingly, this association emerged between slow oscillations of sleep and memory performance, similar to the association between declarative memory and sleep in Study 1.

This similarity between declarative memory and explicit sequence learning could occur for two reasons: 1) Declarative memory and explicit sequence learning share some characteristics, such as consciousness or MTL dependence and/or 2) slow oscillations during sleep support different types of memory as well. There is some evidence for both of these possibilities.

Declarative memory by definition relies on the MTL (Squire, 1992b) and some studies showed that MTL is relevant for sequence learning as well (Albouy, King, Maquet, & Doyon, 2013; Schapiro et al., 2012; Schendan et al., 2003). Indeed, theta oscillations have been observed consistently in the hippocampus and have been suggested to reflect cortico-hippocampal interactions underlying memory processes (Bastiaansen & Hagoort, 2003; Buzsáki & Moser, 2013; Sauseng, Griesmayr, Freunberger, & Klimesch, 2010). However, it is not clear whether the theta oscillations measured on the scalp via EEG originate from the hippocampus.

Concerning consciousness, the declarative memory task was explicit, as well as sequence learning in the cued ASRT might have included explicit processes due to the instruction. Theta oscillations have also been associated with consciousness (Klimesch et al., 2001; Matsuoka, 1990). However, explicit processes can manifest on different levels, such as the instruction, awareness of learning, learning strategies, and representation of the acquired knowledge. However, these levels are somewhat independent of each other, i.e., explicit instruction does not necessarily lead to explicit learning strategies. Additionally, in the cued ASRT task, sequence learning is measured as a difference in reaction times (and accuracy) between the triplets constituted from the sequence as first and last elements and triplets with elements appearing in the same order, but starting and ending with random elements (pattern high-frequency vs. random high-frequency triplets). Even if the knowledge of the sequence helps the former triplet type, it does not affect the latter. This means that the sequence learning measured in the task could not be fully explicit. To conclude, these two memory processes (story recall and sequence learning) are still fundamentally different.

Therefore, besides common characteristics between the two memory processes, slow oscillations having a more general benefit to memory could also explain the similar associations. Slow frequency activity of NREM sleep was consistently linked to better cognitive outcomes in healthy adults (Arico et al., 2010; Drago et al., 2011; Ferri et al., 2010) and children (Bruni et al., 2012). Studies showed a power increase in these slow frequency oscillations (1-8 Hz) after prolonged wakefulness, suggesting they reflect the homeostatic and restorative capacity of sleep (Borbély et al., 1981; Marzano et al., 2010). Serving the function of homeostasis and restoration, the role of slow oscillations does not seem to be specific to support only a narrow band of memory processes. Instead, it is likely that different types of memory can equally benefit from this neural mechanism.

Finally, it is also important to note, that while the associated frequency ranges (slow oscillations, 1-8 Hz) were identical in the two studies with the different memory processes, the exact associations differed. In Study 1, oscillations around 1 Hz positively, whereas oscillations between 4 and 7 Hz negatively associated with the declarative learning capacity. In Study 3, oscillations between 2 and 8 Hz positively correlated with the consolidation of sequence knowledge. Therefore, the precise frequency bins, direction of the association and the associated memory phase all varied between the two findings.

Thus, further studies are warranted to explore whether it is the same mechanism (and if yes, what mechanism it is exactly) underlying the associations between the slow oscillations of sleep and declarative and explicit sequence learning.

Neural background of procedural memory independently of sleep

While our findings regarding the neural correlates of procedural memory independent of sleep are inconclusive, there is still some enlightenment to be found in our studies. We introduced transcranial electric stimulation as a potential way to test causal relationship between brain activity and behavior. This is an important technique, as in humans, brain stimulation is the main method to test causality between brain activity and cognition. Regarding transcranial electric stimulation, we also enumerated several stimulation parameters that could help to influence (procedural) memory.

Furthermore, it also became clear that more sophisticated measures of EEG are necessary to reveal the neural correlates of procedural memory consolidation in awake states. A more sophisticated analysis could be functional connectivity analysis, which we used in a previous EEG study investigating the brain activity during the learning phase of a procedural memory task, and indeed we did find associations between functional networks and memory performance (Tóth et al., 2017). Motivated by this, we have several ongoing studies where we aim to explore associations between functional networks measured by EEG and procedural memory performance. With this approach, we do find correlations between brain activity during post-learning wakefulness and consolidation of procedural memory.

Summary

The aim of the dissertation was to provide a deeper understanding of procedural memory processes by investigating their behavioral characteristics, neural background, and their relationship to sleep. We provided evidence that procedural memory is not unitary, and at least two subprocesses, sequence and statistical learning should be differentiated. These subprocesses have different learning trajectories and neural correlates. Moreover, procedural memory seems to be independent of sleep, however, sleep-specific oscillations might have a role in the consolidation of sequential

information. We used a wide variety of methods to explore associations between memory, sleep and their underlying brain activity: brain stimulation, electrophysiology, objective and subjective sleep measures. We also provided a list of factors that should be tested systematically in further studies, such as ultra-fast consolidation and reconsolidation, the differential associations between sleep and memory in different populations and the adequacy of alternative memory taxonomies.

Implications

Besides the theoretical implications that could interest sleep and memory researchers, these findings also have practical importance. They could be relevant for those in the fields of medicine and educational sciences, providing essential community benefits via potential public applications. A precise understanding of the neural background of a cognitive process could enable us to directly modify those via for instance the presented transcranial electric stimulation. If we unravel the neural background of memory, it could be the foundation for improving everyday memory performance, and for clinical therapies of several memory-related disorders (amnesia, post-traumatic stress disorder or dementia). Furthermore, showing differential associations of sleep with different types of memory could be of importance for training: for instance, if one form of memory is impaired in a sleep-disordered population, but the other one is intact, the latter can be used as a compensation technique. For a similar approach with a different disorder, see Ullman and Pullman (2015) and Ullman and Pierpont (2005).

References

Achermann, P., Dijk, D.-J., Brunner, D. P., & Borbély, A. A. (1993). A model of human sleep homeostasis based on EEG slow-wave activity: quantitative comparison of data and simulations. *Brain Research Bulletin, 31*(1-2), 97-113.

Ackermann, S., Hartmann, F., Papassotiropoulos, A., de Quervain, D. J., & Rasch, B. (2015). No associations between interindividual differences in sleep parameters and episodic memory consolidation. *Sleep, 38*(6), 951-959.

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience, 52*(1-2), 29-37.

Albouy, G., Fogel, S., King, B. R., Laventure, S., Benali, H., Karni, A., . . . Doyon, J. (2015). Maintaining vs. enhancing motor sequence memories: respective roles of striatal and hippocampal systems. *Neuroimage, 108*, 423-434.

Albouy, G., King, B. R., Maquet, P., & Doyon, J. (2013). Hippocampus and striatum: Dynamics and interaction during acquisition and sleep-related motor sequence memory consolidation. *Hippocampus, 23*(11), 985-1004.

Albouy, G., Sterpenich, V., Balteau, E., Vandewalle, G., Desseilles, M., Dang-Vu, T., . . . Maquet, P. (2008). Both the hippocampus and striatum are involved in consolidation of motor sequence memory. *Neuron, 58*(2), 261-272.

Alekseichuk, I., Pabel, S. C., Antal, A., & Paulus, W. (2017). Intrahemispheric theta rhythm desynchronization impairs working memory. *Restorative Neurology and Neuroscience, 35*(2), 147-158.

Ali, N. J., Pitson, D. J., & Stradling, J. R. (1993). Snoring, sleep disturbance, and behaviour in 4-5 year olds. *Archives of Disease in Childhood, 68*(3), 360-366.

Ambrus, G. G., Pisoni, A., Primaßin, A., Turi, Z., Paulus, W., & Antal, A. (2015). Bi-frontal transcranial alternating current stimulation in the ripple range reduced overnight forgetting. *Frontiers in Cellular Neuroscience, 9*, 374.

Ambrus, G. G., Vékony, T., Janacsek, K., Trimborn, A. B., Kovacs, G., & Nemeth, D. (2019). When less is more: enhanced statistical learning of non-adjacent dependencies after disruption of bilateral DLPFC. *bioRxiv*, 198515. doi: https://doi.org/10.1101/198515

Andreano, J. M., & Cahill, L. (2006). Glucocorticoid release and memory consolidation in men and women. *Psychological Science, 17*(6), 466-470. doi: 10.1111/j.1467-9280.2006.01729.x

Antal, A., Boros, K., Poreisz, C., Chaieb, L., Terney, D., & Paulus, W. (2008). Comparatively weak after-effects of transcranial alternating current stimulation (tACS) on cortical excitability in humans. *Brain stimulation, 1*(2), 97-105.

Antal, A., & Herrmann, C. S. (2016). Transcranial alternating current and random noise stimulation: possible mechanisms. *Neural Plasticity, 2016*.

Antal, A., & Paulus, W. (2013). Transcranial alternating current stimulation (tACS). *Frontiers in Human Neuroscience, 7*, 317.

Archbold, K. H., Giordani, B., Ruzicka, D. L., & Chervin, R. D. (2004). Cognitive executive dysfunction in children with mild sleep-disordered breathing. *Biolological Research For Nursing, 5*(3), 168-176. doi: 10.1177/1099800403260261

Arico, D., Drago, V., Foster, P. S., Heilman, K. M., Williamson, J., & Ferri, R. (2010). Effects of NREM sleep instability on cognitive processing. *Sleep Medicine, 11*(8), 791-798. doi: 10.1016/j.sleep.2010.02.009

Armitage, R., Trivedi, M., Hoffmann, R., & Rush, A. J. (1997). Relationship between objective and subjective sleep measures in depressed patients and healthy controls. *Depression and anxiety, 5*(2), 97-102.

Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Science, 9*(2).

Bartlett, F. C. (1932). *Remembering: An experimental and social study*: Cambridge: Cambridge University.

Bastiaansen, M. C. M., & Hagoort, P. (2003). Event-induced theta responses as a window on the dynamics of memory. *Cortex, 39*(4-5), 967-992.

Bastien, C. H., Fortier-Brochu, É., Rioux, I., LeBlanc, M., Daley, M., & Morin, C. M. (2003). Cognitive performance and sleep quality in the elderly suffering from chronic insomnia: relationship between objective and subjective measures. *Journal of Psychosomatic Research, 54*(1), 39-49.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Beebe, D. W., & Gozal, D. (2002). Obstructive sleep apnea and the prefrontal cortex: towards a comprehensive model linking nocturnal upper airway obstruction to daytime cognitive and behavioral deficits. *Journal of Sleep Research, 11*(1), 1-16. doi: 10.1046/j.1365-2869.2002.00289.x

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation, 7*(6), 1129-1159.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B, 57*, 289-300.

Berg, E. A. (1948). A simple objective treatment for measuring flexibility in thinking. *Journal of General Psychology, 39*, 15-22.

Bergman, E. T., & Roediger, H. L. (1999). Can Bartlett's repeated reproduction experiments be replicated? *Memory & Cognition, 27*(6), 937-947. doi: 10.3758/BF03201224

Berry, R. B., Brooks, R., Gamaldo, C. E., Harding, S. M., Marcus, C. L., & Vaughn, B. V. (2012). The AASM manual for the scoring of sleep and associated events. *Rules, Terminology and Technical Specifications, Darien, Illinois, American Academy of Sleep Medicine, 176*, 2012.

Biel, A. L., & Friedrich, E. V. (2018). Why you should report bayes factors in your transcranial brain stimulation studies. *Frontiers in Psychology, 9*, 1125.

Bixler, E. O., Vgontzas, A. N., Lin, H. M., Liao, D., Calhoun, S., Vela-Bueno, A., . . . Graff, G. (2009). Sleep disordered breathing in children in a general population sample: prevalence and risk factors. *Sleep, 32*(6), 731-736.

Blunden, S., & Beebe, D. W. (2006). The contribution of intermittent hypoxia, sleep debt and sleep disruption to daytime performance deficits in children: consideration of respiratory and non-respiratory sleep disorders. *Sleep Medicine Reviews, 10*(2), 109-118. doi: 10.1016/j.smrv.2005.11.003

Blunden, S., Lushington, K., Lorenzen, B., Martin, J., & Kennedy, D. (2005). Neuropsychological and psychosocial function in children with a history of snoring or behavioral sleep problems. *The Journal of pediatrics, 146*(6), 780-786.

Bódizs, R., Békésy, M., Szűcs, A., Barsi, P., & Halász, P. (2002). Sleep-Dependent Hippocampal Slow Activity Correlates with Waking Memory Performance in

Humans. *Neurobiology of Learning and Memory, 78*(2), 441-457. doi: 10.1006/nlme.2002.4078

Bódizs, R., Kis, T., Lázár, A. S., Havran, L., Rigó, P., Clemens, Z., & Halász, P. (2005). Prediction of general mental ability based on neural oscillation measures of sleep. *Journal of Sleep Research, 14*(3), 285-292. doi: 10.1111/j.1365-2869.2005.00472.x

Bódizs, R., Körmendi, J., Rigó, P., & Lázár, A. S. (2009). The individual adjustment method of sleep spindle analysis: methodological improvements and roots in the fingerprint paradigm. *Journal of Neuroscience Methods, 178*(1), 205-213.

Borbély, A. A., Baumann, F., Brandeis, D., Strauch, I., & Lehmann, D. (1981). Sleep deprivation: effect on sleep stages and EEG power density in man. *Electroencephalography and Clinical Neurophysiology, 51*(5), 483-493.

Bourke, R. S., Anderson, V., Yang, J. S., Jackman, A. R., Killedar, A., Nixon, G. M., . . . Horne, R. S. (2011a). Cognitive and academic functions are impaired in children with all severities of sleep-disordered breathing. *Sleep Medicine, 12*(5), 489-496. doi: 10.1016/j.sleep.2010.11.010

Bourke, R. S., Anderson, V., Yang, J. S., Jackman, A. R., Killedar, A., Nixon, G. M., . . . Horne, R. S. (2011b). Neurobehavioral function is impaired in children with all severities of sleep disordered breathing. *Sleep Medicine, 12*(3), 222-229. doi: 10.1016/j.sleep.2010.08.011

Bönstrup, M., Iturrate, I., Thompson, R., Cruciani, G., Censor, N., & Cohen, L. G. (2019). A rapid form of offline consolidation in skill learning. *Current Biology, 29*(8), 1346-1351. e1344.

Brawn, T. P., Fenn, K. M., Nusbaum, H. C., & Margoliash, D. (2010). Consolidating the effects of waking and sleep on motor-sequence learning. *The Journal of neuroscience, 30*(42), 13977-13982.

Brokaw, K., Tishler, W., Manceor, S., Hamilton, K., Gaulden, A., Parr, E., & Wamsley, E. J. (2016). Resting state EEG correlates of memory consolidation. *Neurobiology of Learning and Memory, 130*, 17-25.

Brunetti, L., Rana, S., Lospalluti, M. L., Pietrafesa, A., Francavilla, R., Fanelli, M., & Armenio, L. (2001). Prevalence of obstructive sleep apnea syndrome in a cohort of 1,207 children of southern Italy. *CHEST Journal, 120*(6), 1930-1935.

Bruni, O., Kohler, M., Novelli, L., Kennedy, D., Lushington, K., Martin, J., & Ferri, R. (2012). The role of NREM sleep instability in child cognitive performance. *Sleep, 35*(5), 649-656. doi: 10.5665/sleep.1824

Buchmann, A., Ringli, M., Kurth, S., Schaerer, M., Geiger, A., Jenni, O. G., & Huber, R. (2011). EEG sleep slow-wave activity as a mirror of cortical maturation. *Cerebral Cortex, 21*(3), 607-615. doi: 10.1093/cercor/bhq129

Buysse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Research, 28*(2), 193-213.

Buzsáki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science, 304*(5679), 1926-1929.

Buzsáki, G., & Moser, E. I. (2013). Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nature Neuroscience, 16*(2), 130-138.

Cajochen, C., Foy, R., & Dijk, D. J. (1999). Frontal predominance of a relative increase in sleep delta and theta EEG activity after sleep loss in humans. *Sleep Res Online, 2*(3), 65-69.

Canty, A., & Ripley, B. (2019). boot: Bootstrap R (S-Plus) Functions.

Caplan, J. B., & Glaholt, M. G. (2007). The roles of EEG oscillations in learning relational information. *Neuroimage, 38*(3), 604-616.

Carrier, J., Land, S., Buysse, D. J., Kupfer, D. J., & Monk, T. H. (2001). The effects of age and gender on sleep EEG power spectral density in the middle years of life (ages 20-60 years old). *Psychophysiology, 38*(2), 232-242.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*(3), 386-404.

Casey, B., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: what have we learned about cognitive development? *Trends in cognitive sciences, 9*(3), 104-110.

Castronovo, V., Zucconi, M., Nosetti, L., Marazzini, C., Hensley, M., Veglia, F., . . . Ferini-Strambi, L. (2003). Prevalence of habitual snoring and sleep-disordered breathing in preschool-aged children in an Italian community. *Journal of Pediatrics, 142*(4), 377-382. doi: 10.1067/mpd.2003.118

Chander, B. S., Witkowski, M., Braun, C., Robinson, S. E., Born, J., Cohen, L. G., . . . Soekadar, S. R. (2016). tACS phase locking of frontal midline theta oscillations disrupts working memory performance. *Frontiers in Cellular Neuroscience, 10*, 120.

Chatburn, A., Coussens, S., Lushington, K., Kennedy, D., Baumert, M., & Kohler, M. (2013). Sleep spindle activity and cognitive performance in healthy children. *Sleep, 36*(2), 237-243.

Chervin, R. D., Archbold, K. H., Dillon, J. E., Panahi, P., Pituch, K. J., Dahl, R. E., & Guilleminault, C. (2002). Inattention, hyperactivity, and symptoms of sleep-disordered breathing. *Pediatrics, 109*(3), 449-456.

Chun, M. M., & Phelps, E. A. (1999). Memory deficits for implicit contextual information in amnesic subjects with hippocampal damage. *Nature Neuroscience, 2*(9), 844-847.

Clemens, Z., Fabo, D., & Halasz, P. (2005). Overnight verbal memory retention correlates with the number of sleep spindles. *Neuroscience, 132*(2), 529-535.

Conway, A. R., Kane, M. J., Bunting, M., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review, 12*(5), 769-786.

Conway, C., Bauernschmidt, A., Huang, S., & Pisoni, D. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition, 113*(3), 356-371.

Craig, M., & Dewar, M. (2018). Rest-related consolidation protects the fine detail of new memories. *Scientific Reports, 8*(1), 6857.

Craik, F. I., & Bialystok, E. (2006). Cognition through the lifespan: mechanisms of change. *Trends in cognitive sciences, 10*(3), 131-138.

Curcio, G., Ferrara, M., & De Gennaro, L. (2006). Sleep loss, learning capacity and academic performance. *Sleep Medicine Reviews, 10*(5), 323-337.

Csábi, E., Benedek, P., Janacsek, K., Katona, G., & Nemeth, D. (2013). Sleep disorder in childhood impairs declarative but not nondeclarative forms of learning. *Journal of Clinical and Experimental Neuropsychology, 35*(7), 677-685.

Csabi, E., Benedek, P., Janacsek, K., Zavecz, Z., Katona, G., & Nemeth, D. (2015). Declarative and Non-declarative Memory Consolidation in Children with Sleep

Disorder. *Frontiers in Human Neuroscience, 9*, 709. doi: 10.3389/fnhum.2015.00709

Csabi, E., Varszegi-Schulz, M., Janacsek, K., Malecek, N., & Nemeth, D. (2014). The consolidation of implicit sequence memory in obstructive sleep apnea. *PloS One, 9*(10), e109010.

Darien, I. (2014). International classification of sleep disorders. *American Academy of Sleep Medicine*.

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application* (Vol. 1): Cambridge university press.

Daw, Niv, & Dayan. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience, 8*(12), 1704-1711.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9-21.

Delpouve, J., Schmitz, R., & Peigneux, P. (2014). Implicit learning is better at subjectively defined non-optimal time of day. *Cortex, 58*, 18-22.

Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage, 53*(1), 1-15.

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

Diekelmann, S., Wilhelm, I., & Born, J. (2009). The whats and whens of sleep-dependent memory consolidation. *Sleep Medicine Reviews, 13*(5), 309-321.

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. *Experimental Brain Research, 195*(1), 15-26.

Draganich, C., & Erdal, K. (2014). Placebo sleep affects cognitive functioning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(3), 857.

Drago, V., Foster, P. S., Heilman, K. M., Arico, D., Williamson, J., Montagna, P., & Ferri, R. (2011). Cyclic alternating pattern in sleep and its relationship to creativity. *Sleep Medicine, 12*(4), 361-366. doi: 10.1016/j.sleep.2010.11.009

Durrant, S. J., Cairney, S. A., & Lewis, P. A. (2013). Overnight consolidation aids the transfer of statistical knowledge from the medial temporal lobe to the striatum. *Cerebral Cortex, 23*(10), 2467-2478. doi: 10.1093/cercor/bhs244

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

Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. *Nature Reviews Neuroscience, 1*(1), 41-50.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology, 128*(3), 309-331.

Ersu, R., Arman, A. R., Save, D., Karadag, B., Karakoc, F., Berkem, M., & Dagli, E. (2004). Prevalence of snoring and symptoms of sleep-disordered breathing in primary school children in istanbul. *Chest, 126*(1), 19-24. doi: 10.1378/chest.126.1.19

Fanuel, L., Pleche, C., Vekony, T., Quentin, R., Janacsek, K., & Nemeth, D. (2020). The longer the better? General skill but not probabilistic learning improves with the duration of short rest periods. *bioRxiv*, 2020.2005.2012.090886. doi: 10.1101/2020.05.12.090886

Feld, G. B., & Diekelmann, S. (2015). Sleep smart—optimizing sleep for declarative learning and memory. *Frontiers in Psychology, 6*, 622.

Fenn, K. M., Nusbaum, H. C., & Margoliash, D. (2003). Consolidation during sleep of perceptual learning of spoken language. *Nature, 425*(6958), 614-616.

Ferrara, M., & De Gennaro, L. (2011). Going local: insights from EEG and stereo-EEG studies of the human sleep-wake cycle. *Current Topics in Medicinal Chemistry, 11*(19), 2423-2437. doi: 10.2174/156802611797470268

Ferreira, A. M., Clemente, V., Gozal, D., Gomes, A., Pissarra, C., Cesar, H., . . . Azevedo, M. H. (2000). Snoring in Portuguese primary school children. *Pediatrics, 106*(5), E64. doi: 10.1542/peds.106.5.e64

Ferri, R., Bruni, O., Miano, S., & Terzano, M. G. (2005). Topographic mapping of the spectral components of the cyclic alternating pattern (CAP). *Sleep Medicine, 6*(1), 29-36. doi: 10.1016/j.sleep.2004.06.010

Ferri, R., Drago, V., Arico, D., Bruni, O., Remington, R. W., Stamatakis, K., & Punjabi, N. M. (2010). The effects of experimental sleep fragmentation on cognitive processing. *Sleep Medicine, 11*(4), 378-385. doi: 10.1016/j.sleep.2010.01.006

Ferri, R., Huber, R., Arico, D., Drago, V., Rundo, F., Ghilardi, M. F., . . . Tononi, G. (2008). The slow-wave components of the cyclic alternating pattern (CAP) have a role in sleep-related learning processes. *Neuroscience Letters, 432*(3), 228-231. doi: 10.1016/j.neulet.2007.12.025

Filoteo, J. V., Lauritzen, S., & Maddox, W. T. (2010). Removing the Frontal Lobes The Effects of Engaging Executive Functions on Perceptual Category Learning. *Psychological Science, 21*(3), 415-423.

Fischer, S., Hallschmid, M., Elsner, A. L., & Born, J. (2002). Sleep forms memory for finger skills. *PNAS, 99*(18), 11987-11991.

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12*(6), 499-504.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, 99*(24), 15822-15826. doi: 10.1073/pnas.232472899

Fletcher, P. C., Zafiris, O., Frith, C. D., Honey, R. A. E., Corlett, P. R., Zilles, K., & Fink, G. R. (2005). On the benefits of not trying: brain activity and connectivity reflecting the interactions of explicit and implicit sequence learning. *Cerebral Cortex, 15*(7), 1002-1015.

Fogel, S. M., & Smith, C. T. (2011). The function of the sleep spindle: a physiological index of intelligence and a mechanism for sleep-dependent memory consolidation. *Neuroscience and Biobehavioral Reviews, 35*(5), 1154-1165.

Fox, C. J., Mueller, S. T., Gray, H. M., Raber, J., & Piper, B. J. (2013). Evaluation of a short-form of the Berg Card Sorting Test. *PloS One, 8*(5), e63885.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences, 9*(10), 474-480.

Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin, 145*(12), 1128.

Gao, C., Terlizzese, T., & Scullin, M. K. (2019). Short sleep and late bedtimes are detrimental to educational learning and knowledge transfer: An investigation of individual differences in susceptibility. *Chronobiology International, 36*(3), 307-318.

Gauld, A., & Stephenson, G. M. (1967). Some experiments relating to Bartlett's theory of remembering. *British Journal of Psychology, 58*(1-2), 39-49.

Gaultney, J. F. (2010). The prevalence of sleep disorders in college students: impact on academic performance. *Journal of American College Health, 59*(2), 91-97.

Gavriloff, D., Sheaves, B., Juss, A., Espie, C. A., Miller, C. B., & Kyle, S. D. (2018). Sham sleep feedback delivered via actigraphy biases daytime symptom reports in people with insomnia: Implications for insomnia disorder and wearable devices. *Journal of Sleep Research*, e12726.

Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral cortex (New York, NY: 1991), 7*(4), 374-385.

Gilson, M., Deliens, G., Leproult, R., Bodart, A., Nonclercq, A., Ercek, R., & Peigneux, P. (2015). REM-enriched naps are associated with memory consolidation for sad stories and enhance mood-related reactivity. *Brain sciences, 6*(1), 1.

Goel, N., Basner, M., Rao, H., & Dinges, D. F. (2013). Circadian rhythms, sleep deprivation, and human performance. *Progress in Molecular Biology and Translational Science, 119*, 155-190.

Gottlieb, D. J., Chase, C., Vezina, R. M., Heeren, T. C., Corwin, M. J., Auerbach, S. H., . . . Lesko, S. M. (2004). Sleep-disordered breathing symptoms are associated with poorer cognitive function in 5-year-old children. *Journal of Pediatrics, 145*(4), 458-464. doi: 10.1016/j.jpeds.2004.05.039

Gottlieb, D. J., Vezina, R. M., Chase, C., Lesko, S. M., Heeren, T. C., Weese-Mayer, D. E., . . . Corwin, M. J. (2003). Symptoms of sleep-disordered breathing in 5-year-old children are associated with sleepiness and problem behaviors. *Pediatrics, 112*(4), 870-877. doi: 10.1542/peds.112.4.870

Gozal, D., & O'Brien, L. M. (2004). Snoring and obstructive sleep apnoea in children: why should we treat? *Paediatric Respiratory Reviews, 5 Suppl A*, S371-376. doi: 10.1016/S1526-0542(04)90066-8

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(3), 501-518.

Guedes, L. G., Abreu, G. d. A., Rodrigues, D. F., Teixeira, L. R., Luiz, R. R., & Bloch, K. V. (2016). Comparison between self-reported sleep duration and actigraphy among adolescents: gender differences. *Revista Brasileira de Epidemiologia, 19*(2), 339-347.

Guo, M., Igue, M., Malhotra, A., Stickgold, R., & Djonlagic, I. (2013). The effect of obstructive sleep apnea on declarative memory consolidation. *Sleep Medicine, 14*, e33-e34. doi: 10.1016/j.sleep.2013.11.041

Halbower, A. C., Degaonkar, M., Barker, P. B., Earley, C. J., Marcus, C. L., Smith, P. L., . . . Mahone, E. M. (2006). Childhood obstructive sleep apnea associates with neuropsychological deficits and neuronal brain injury. *PLoS Medicine, 3*(8), e301. doi: 10.1371/journal.pmed.0030301

Hallgato, E., Győri-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.

Hannula, D. E., & Greene, A. J. (2012). The hippocampus reevaluated in unconscious learning and memory: at a tipping point? *Frontiers in Human Neuroscience, 6*, 171.

Hardwick, R. M., Forrence, A. D., Krakauer, J. W., & Haith, A. M. (2019). Time-dependent competition between goal-directed and habitual response preparation. *Nature Human Behaviour, 3*(12), 1252-1262. doi: 10.1038/s41562-019-0725-0

Heaton, R. K. (1981). *A manual for the Wisconsin card sorting test*: Western Psycological Services.

Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews Neuroscience, 11*, 523-532.

Hikosaka, O., Nakamura, K., Sakai, K., & Nakahara, H. (2002). Central mechanisms of motor skill learning. *Current Opinion in Neurobiology, 12*(2), 217-222.

Horne, J. A., & Östberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology*.

Horvath, K., Torok, C., Pesthy, O., Nemeth, D., & Janacsek, K. (2019). Explicit instruction differentially affects subcomponents of procedural learning and consolidation. *bioRxiv*, 433243. doi: https://doi.org/10.1101/433243

Howard, D. V., Howard, J. H., Jr., Japikse, K., DiYanni, C., Thompson, A., & Somberg, R. (2004). Implicit sequence learning: effects of level of structure, adult age, and extended practice. *Psychology and Aging, 19*(1), 79-92.

Howard, J. H., Jr., & Howard, D. V. (1997). Age differences in implicit learning of higher-order dependencies in serial patterns. *Psychology and Aging, 12*(4), 634-656.

Hsieh, L.-T., & Ranganath, C. (2014). Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval. *Neuroimage, 85*, 721-729.

Hull, J. T., Wright Jr, K. P., & Czeisler, C. A. (2003). The influence of subjective alertness and motivation on human performance independent of circadian and homeostatic regulation. *Journal of Biological Rhythms, 18*(4), 329-338.

Hunter, S. J., & Gozal, D. (2016). Effect of Sleep-disordered Breathing Severity on Cognitive Performance Measures in a Large Community Cohort of Young School-aged Children. *American Journal of Respiratory and Critical Care Medicine, 194*(6), 739-747. doi: 10.1164/rccm.201510-2099OC

Janacsek, K., Ambrus, G. G., Paulus, W., Antal, A., & Nemeth, D. (2015). Right hemisphere advantage in statistical learning: evidence from a probabilistic sequence learning task. *Brain stimulation, 8*(2), 277-282. doi: 10.1016/j.brs.2014.11.008

Janacsek, K., Borbély-Ipkovich, E., Nemeth, D., & Gonda, X. (2018). How can the depressed mind extract and remember predictive relationships of the environment? Evidence from implicit probabilistic sequence learning. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 81*, 17-24.

Janacsek, K., Fiser, J., & Nemeth, D. (2012). The best time to acquire new skills: age-related differences in implicit sequence learning across the human lifespan. *Developmental science, 15*(4), 496-505.

Janacsek, K., Shattuck, K., Tagarelli, K., Lum, J., Turkeltaub, P., & Ullman, M. T. (2020). Sequence learning in the human brain: A functional neuroanatomical meta-analysis of serial reaction time studies. *Neuroimage, 207*, 116387.

Jasper, H. (1958). Report of the committee on methods of clinical examination in electroencephalography. *Electroencephalography and Clinical Neurophysiology, 10*, 370-375.

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience, 15*(8), 1395-1399.

Jones, K., & Harrison, Y. (2001). Frontal lobe function, sleep loss and fragmented sleep. *Sleep Medicine Reviews, 5*(6), 463-475.

Juhasz, D., Nemeth, D., & Janacsek, K. (2019). Is there more room to improve? The lifespan trajectory of procedural learning and its relationship to the between-and within-group differences in average response times. *PloS One*. doi: doi: 10.1371/journal.pone.0215116

Jussila, M., Saarenpaa-Heikkila, O., Huupponen, E., Lapinlampi, A. M., Penttala, J., & Himanen, S. L. (2016). Sleep apnea reduces the amount of computational deep sleep in the right frontopolar area in school-aged children. *Clinical Neurophysiology, 127*(5), 2167-2174. doi: 10.1016/j.clinph.2016.02.016

Karni, A., Meyer, G., Rey-Hipolito, C., Jezzard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1998). The acquisition of skilled motor performance: Fast and slow experience-driven changes in primary motor cortex. *Proceedings of the National Academy of Sciences of the United States of America, 95*(3), 861-868.

Kaufman, S. B., Deyoung, C. G., Gray, J. R., Jimenez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition, 116*(3), 321-340. doi: 10.1016/j.cognition.2010.05.011

Khatoun, A., Asamoah, B., & Mc Laughlin, M. (2019). How does transcranial alternating current stimulation entrain single-neuron activity in the primate brain? *Proceedings of the National Academy of Sciences, 116*(45), 22438-22439.

Kheirandish-Gozal, L., Miano, S., Bruni, O., Ferri, R., Pagani, J., Villa, M. P., & Gozal, D. (2007). Reduced NREM sleep instability in children with sleep disordered breathing. *Sleep, 30*(4), 450-457. doi: 10.1093/sleep/30.4.450

Kim, R., Seitz, A., Feenstra, H., & Shams, L. (2009). Testing assumptions of statistical learning: is it long-term and implicit? *Neuroscience Letters, 461*(2), 145-149.

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

Klimesch, W., Doppelmayr, M., Yonelinas, A., Kroll, N. E., Lazzara, M., Röhm, D., & Gruber, W. (2001). Theta synchronization during episodic retrieval: neural correlates of conscious awareness. *Cognitive Brain Research, 12*(1), 33-38.

Klimesch, W., Freunberger, R., Sauseng, P., & Gruber, W. (2008). A short review of slow phase synchronization and memory: evidence for control processes in different memory systems? *Brain Research, 1235*, 31-44.

Knowlton, B. J., Ramus, S. J., & Squire, L. R. (1992). Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for specific instances. *Psychological Science, 3*(3), 172-179.

Kobor, A., Janacsek, K., Takacs, A., & Nemeth, D. (2017). Statistical learning leads to persistent memory: Evidence for one-year consolidation. *Scientific Reports, 7*(1), 760.

Kóbor, A., Janacsek, K., Takács, Á., & Nemeth, D. (2017). Statistical learning leads to persistent memory: Evidence for one-year consolidation. *Scientific Reports, 7*(1), 760. doi: 10.1038/s41598-017-00807-3

Kóbor, A., Takács, Á., Kardos, Z., Janacsek, K., Horváth, K., Csépe, V., & Nemeth, D. (2018). ERPs differentiate the sensitivity to statistical probabilities and the learning of sequential structures during procedural learning. *Biological Psychology, 135*, 180-193.

Kohler, M. J., Lushington, K., van den Heuvel, C. J., Martin, J., Pamula, Y., & Kennedy, D. (2009). Adenotonsillectomy and neurocognitive deficits in children with Sleep Disordered Breathing. *PloS One, 4*(10), e7343. doi: 10.1371/journal.pone.0007343

Konkel, A., & Cohen, N. J. (2009). Relational memory and the hippocampus: representations and methods. *Frontiers in Neuroscience, 3*, 23.

Krakauer, J. W., & Shadmehr, R. (2006). Consolidation of motor memory. *Trends in Neuroscience, 29*(1), 58-64.

Krause, M. R., Vieira, P. G., Csorba, B. A., Pilly, P. K., & Pack, C. C. (2019). Transcranial alternating current stimulation entrains single-neuron activity in the primate brain. *Proceedings of the National Academy of Sciences, 116*(12), 5747-5755.

Kyle, S. D., Sexton, C. E., Feige, B., Luik, A. I., Lane, J., Saxena, R., . . . Little, M. A. (2017). Sleep and cognitive performance: cross-sectional associations in the UK Biobank. *Sleep Medicine, 38*, 85-91.

Landry, G. J., Best, J. R., & Liu-Ambrose, T. (2015a). Measuring sleep quality in older adults: a comparison using subjective and objective methods. *Frontiers in aging neuroscience, 7*.

Landry, G. J., Best, J. R., & Liu-Ambrose, T. (2015b). Measuring sleep quality in older adults: a comparison using subjective and objective methods. *Frontiers in Aging Neuroscience, 7*, 166.

Libman, E., Fichten, C. S., Bailes, S., & Amsel, R. (2000). Sleep questionnaire versus sleep diary: which measure is better? *International Journal of Rehabilitation and Health, 5*(3), 205-209.

Lieberman, M. D. (2000). Intuition: a social cognitive neuroscience approach. *Psychological Bulletin, 126*(1), 109-137.

Lo, J. C., Groeger, J. A., Cheng, G. H., Dijk, D.-J., & Chee, M. W. (2016). Self-reported sleep duration and cognitive performance in older adults: a systematic review and meta-analysis. *Sleep Medicine, 17*, 87-98.

Lustenberger, C., Maric, A., Durr, R., Achermann, P., & Huber, R. (2012). Triangular relationship between sleep spindle activity, general cognitive ability and the efficiency of declarative learning. *PloS One, 7*(11), e49561. doi: 10.1371/journal.pone.0049561

Maestri, M., Carnicelli, L., Tognoni, G., Di Coscio, E., Giorgi, F. S., Volpi, L., . . . Bonanni, E. (2015). Non-rapid eye movement sleep instability in mild cognitive

impairment: a pilot study. *Sleep Medicine, 16*(9), 1139-1145. doi: 10.1016/j.sleep.2015.04.027

Maheu, M., Meyniel, F., & Dehaene, S. (2020). Rational arbitration between statistics and rules in human sequence learning. *bioRxiv*.

Mander, B. A., Rao, V., Lu, B., Saletin, J. M., Lindquist, J. R., Ancoli-Israel, S., . . . Walker, M. P. (2013). Prefrontal atrophy, disrupted NREM slow waves and impaired hippocampal-dependent memory in aging. *Nature Neuroscience, 16*(3), 357-364. doi: 10.1038/nn.3324

Mander, B. A., Reid, K. J., Baron, K. G., Tjoa, T., Parrish, T. B., Paller, K. A., . . . Zee, P. C. (2010). EEG measures index neural and cognitive recovery from sleep deprivation. *Journal of Neuroscience, 30*(7), 2686-2693. doi: 10.1523/jneurosci.4010-09.2010

Mantua, J. (2018). Sleep physiology correlations and human memory consolidation: where do we go from here? *Sleep, 41*(2), zsx204.

Maquet, P., Laureys, S., Peigneux, P., Fuchs, S., Petiau, C., Phillips, C., . . . Cleeremans, A. (2000). Experience-dependent changes in cerebral activation during human REM sleep. *Nature Neuroscience, 3*(8), 831-836.

Marcus, C. L. (2001). Sleep-disordered breathing in children. *American Journal of Respiratory and Critical Care Medicine, 164*(1), 16-30. doi: 10.1164/ajrccm.164.1.2008171

Marshall, L., Helgadottir, H., Molle, M., & Born, J. (2006). Boosting slow oscillations during sleep potentiates memory. *Nature, 444*(7119), 610-613. doi: 10.1038/nature05278

Marzano, C., Ferrara, M., Curcio, G., & Gennaro, L. D. (2010). The effects of sleep deprivation in humans: topographical electroencephalogram changes in non-rapid eye movement (NREM) sleep versus REM sleep. *Journal of Sleep Research, 19*(2), 260-268.

Matchock, R. L., & Mordkoff, J. T. (2009). Chronotype and time-of-day influences on the alerting, orienting, and executive components of attention. *Experimental Brain Research, 192*(2), 189-198.

Matsuoka, S. (1990). Theta rhythms: state of consciousness. *Brain Topography, 3*(1), 203-208.

McGaugh, J. L. (2000). Memory--A century of consolidation. *Science, 287*(5451), 248-251.

McGaugh, J. L., & Izquierdo, I. (2000). The contribution of pharmacology to research on the mechanisms of memory formation. *Trends in Pharmacological Sciences, 21*, 208-210.

Mednick, S. C., Cai, D. J., Shuman, T., Anagnostaras, S., & Wixted, J. T. (2011). An opportunistic theory of cellular and systems consolidation. *Trends in Neurosciences, 34*(10), 504-514.

Meijman, T., de Vries-Griever, A., De Vries, G., & Kampman, R. (1988). The evaluation of the Groningen sleep quality scale. *Groningen: Heymans Bulletin (HB 88-13-EX), 2006*.

Meyer, L., Grigutsch, M., Schmuck, N., Gaston, P., & Friederici, A. D. (2015). Frontal–posterior theta oscillations reflect memory retrieval during sentence comprehension. *Cortex, 71*, 205-218.

Miyata, S., Noda, A., Iwamoto, K., Kawano, N., Okuda, M., & Ozaki, N. (2013). Poor sleep quality impairs cognitive performance in older adults. *Journal of Sleep Research, 22*(5), 535-541.

Molle, M., Marshall, L., Gais, S., & Born, J. (2004). Learning increases human electroencephalographic coherence during subsequent slow sleep oscillations. *Proceedings of the National Academy of Sciences of the United States of America, 101*(38), 13963-13968. doi: 10.1073/pnas.0402820101

Munch, M., Knoblauch, V., Blatter, K., Schroder, C., Schnitzler, C., Krauchi, K., . . . Cajochen, C. (2004). The frontal predominance in human EEG delta activity after sleep loss decreases with age. *European Journal of Neuroscience, 20*(5), 1402-1410. doi: 10.1111/j.1460-9568.2004.03580.x

Nader, K. (2003). Memory traces unbound. *Trends in Neurosciences, 26*(2), 65-72.

Nebes, R. D., Buysse, D. J., Halligan, E. M., Houck, P. R., & Monk, T. H. (2009). Self-reported sleep quality predicts poor cognitive performance in healthy older adults. *The Journals of Gerontology: Series B, 64*(2), 180-187.

Nemeth, D., Csábi, E., Janacsek, K., Varszegi, M., & Mari, Z. (2012). Intact implicit probabilistic sequence learning in obstructive sleep apnea. *Journal of Sleep Research, 21*(4), 396-401. doi: doi: 10.1111/j.1365-2869.2011.00983.x

Nemeth, D., & Janacsek, K. (2011). The dynamics of implicit skill consolidation in young and elderly adults. *Journal of Gerontology Psychological Science, 66*(1), 15-22.

Nemeth, D., Janacsek, K., Balogh, V., Londe, Z., Mingesz, R., Fazekas, M., . . . Vetro, A. (2010). Learning in Autism: Implicitly Superb. *PloS One, 5*(7), e11731.

Nemeth, D., Janacsek, K., & Fiser, J. (2013). Age-dependent and coordinated shift in performance between implicit and explicit skill learning. *Frontiers in Computational Neuroscience, 7*. doi: 10.3389/fncom.2013.00147

Nemeth, D., Janacsek, K., Király, K., Londe, Z., Németh, K., Fazekas, K., . . . Csányi, A. (2013). Probabilistic sequence learning in mild cognitive impairment. *Frontiers in Human Neuroscience, 7*, 318. doi: 10.3389/fnhum.2013.00318

Nemeth, D., Janacsek, K., Londe, Z., Ullman, M. T., Howard, D. V., & Howard, J. H., Jr. (2010). Sleep has no critical role in implicit motor sequence learning in young and old adults. *Experimental Brain Research, 201*(2), 351-358. doi: 10.1007/s00221-009-2024-x

Nemeth, D., Janacsek, K., Polner, B., & Kovacs, Z. A. (2013). Boosting Human Learning by Hypnosis. *Cerebral Cortex, 23*(4), 801-805. doi: doi:10.1093/cercor/bhs068

Nishida, M., Nakashima, Y., & Nishikawa, T. (2016). Slow sleep spindle and procedural memory consolidation in patients with major depressive disorder. *Nature and science of sleep, 8*, 63.

Nissen, M. J., Nissen, M. J., Willingham, D., & Hartman, M. (1989). Explicit and implicit remembering: When is learning preserved in amnesia? *Neuropsychologia, 27*(3), 341.

Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: fully automated statistical thresholding for EEG artifact rejection. *Journal of Neuroscience Methods, 192*(1), 152-162.

Novak, M., Mucsi, I., Shapiro, C. M., Rethelyi, J., & Kopp, M. S. (2004). Increased utilization of health services by insomniacs—an epidemiological perspective. *Journal of Psychosomatic Research, 56*(5), 527-536.

Nyberg, L., Cabeza, R., & Tulving, E. (1996). PET studies of encoding and retrieval: The HERA model. *Psychonomic bulletin & review, 3*(2), 135-148.

O'Brien, L. M., Mervis, C. B., Holbrook, C. R., Bruner, J. L., Smith, N. H., McNally, N., . . . Gozal, D. (2004). Neurobehavioral correlates of sleep-disordered breathing in

children. *Journal of Sleep Research, 13*(2), 165-172. doi: 10.1111/j.1365-2869.2004.00395.x

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Onton, J., Delorme, A., & Makeig, S. (2005). Frontal midline EEG dynamics during working memory. *Neuroimage, 27*(2), 341-356.

Pan, S. C., & Rickard, T. C. (2015). Sleep and motor learning: is there room for consolidation? *Psychological Bulletin, 141*(4), 812.

Paraskevopoulos, E., Chalas, N., & Bamidis, P. (2017). Functional connectivity of the cortical network supporting statistical learning in musicians and non-musicians: an MEG study. *Scientific Reports, 7*(1), 16268.

Peigneux, P., Laureys, S., Fuchs, S., Destrebecqz, A., Collette, F., Delbeuck, X., . . . Degueldre, C. (2003). Learned material content and acquisition level modulate cerebral reactivation during posttraining rapid-eye-movements sleep. *Neuroimage, 20*(1), 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 Biology, 4*(4), e100.

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 Biology, 4*(4), e100.

Penhune, V. B., & Steele, C. J. (2012). Parallel contributions of cerebellar, striatal and M1 mechanisms to motor sequence learning. *Behavioural Brain Research, 226*(2), 579-591.

Pikovsky, A., Kurths, J., Rosenblum, M., & Kurths, J. (2003). *Synchronization: a universal concept in nonlinear sciences* (Vol. 12): Cambridge university press.

Piper, B. J., Mueller, S. T., Geerken, A. R., Dixon, K. L., Kroliczak, G., Olsen, R. H., & Miller, J. K. (2015). Reliability and validity of neurobehavioral function on the Psychology Experimental Building Language test battery in young adults. *PeerJ, 3*, e1460.

Polanía, R., Moisa, M., Opitz, A., Grueschow, M., & Ruff, C. C. (2015). The precision of value-based choices depends causally on fronto-parietal phase coupling. *Nature communications, 6*, 8090.

Polanía, R., Nitsche, M. A., Korman, C., Batsikadze, G., & Paulus, W. (2012). The importance of timing in segregated theta phase-coupling for cognitive performance. *Current Biology, 22*(14), 1314-1318.

Poldrack, R. A. (2012). The future of fMRI in cognitive neuroscience. *Neuroimage, 62*(2), 1216-1220.

Poldrack, R. A., Clark, J., Pare-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature, 414*(6863), 546-550.

Poldrack, R. A., & Foerde, K. (2008). Category learning and the memory systems debate. *Neuroscience and Biobehavioral Reviews, 32*(2), 197-205.

Pollok, B., Boysen, A.-C., & Krause, V. (2015). The effect of transcranial alternating current stimulation (tACS) at alpha and beta frequency on motor learning. *Behavioural Brain Research, 293*, 234-240.

Pothos, E. M. (2007). Theories of artificial grammar learning. *Psychological Bulletin, 133*(2), 227-244.

Prinzmetal, W., McCool, C., & Park, S. (2005). Attention: Reaction time and accuracy reveal different mechanisms. *Journal of Experimental Psychology: General, 134*(1), 73-92. doi: 10.1037/0096-3445.134.1.73

R Core Team. (2018). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Austria, 2015: ISBN 3-900051-07-0: URL http://www. R-project. org.

Racsmány, M., Lukács, Á., Németh, D., & Pléh, C. (2005). A verbális munkamemória magyar nyelvű vizsgálóeljárásai (Hungarian Diagnostic Tools of Verbal Working Memory Functions). *Magyar Pszichológiai Szemle (Hungarian Review of Psychology), 60*(4), 479-506.

Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological Reviews, 93*(2), 681-766. doi: 10.1152/physrev.00032.2012

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*, 855-863.

Reber, P. J. (2013). The neural basis of implicit learning and memory: a review of neuropsychological and neuroimaging research. *Neuropsychologia, 51*(10), 2026-2042.

Reder, L. M., Park, H., & Kieffaber, P. D. (2009). Memory systems do not divide on consciousness: Reinterpreting memory in terms of activation and binding. *Psychological Bulletin, 135*(1), 23.

Remillard, G. (2008). Implicit learning of second-, third-, and fourth-order adjacent and nonadjacent sequential dependencies. *The Quarterly journal of experimental psychology, 61*(3), 400-424. doi: 10.1080/17470210701210999

Rickard, T. C., Cai, D. J., Rieth, C. A., Jones, J., & Ard, M. C. (2008). Sleep does not enhance motor sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition; Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 834.

Rieth, C. A., Cai, D. J., McDevitt, E. A., & Mednick, S. C. (2010). The role of sleep and practice in implicit and explicit motor learning. *Behavioural Brain Research, 214*(2), 470-474.

Robertson, E. M. (2019). Skill memory: mind the ever-decreasing gap for offline processing. *Current Biology, 29*(8), R287-R289.

Robertson, E. M., Pascual-Leone, A., & Press, D. Z. (2004). Awareness modifies the skill-learning benefits of sleep. *Current Biology, 14*(3), 208-212. doi: 10.1016/j.cub.2004.01.027

Rosen, C. L., Storfer-Isser, A., Taylor, H. G., Kirchner, H. L., Emancipator, J. L., & Redline, S. (2004). Increased behavioral morbidity in school-aged children with sleep-disordered breathing. *Pediatrics, 114*(6), 1640-1648. doi: 10.1542/peds.2004-0103

Rouder, J. N., Engelhardt, C. R., McCabe, S., & Morey, R. D. (2016). Model comparison in ANOVA. *Psychonomic bulletin & review, 23*(6), 1779-1786.

Rowe, G., Hasher, L., & Turcotte, J. (2009). Short article: Age and synchrony effects in visuospatial working memory. *Quarterly Journal of Experimental Psychology, 62*(10), 1873-1880.

Rózsa, S., Szádóczky, E., & Furedi, J. (2001). Psychometric properties of the Hungarian version of the shortened Beck Depression Inventory. *Psychiatria Hungarica, 16*(4), 384-402.

Sadeh, A., Gruber, R., & Raviv, A. (2002). Sleep, neurobehavioral functioning, and behavior problems in school-age children. *Child Development, 73*(2), 405-417.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928.

Saksvik, I. B., Bjorvatn, B., Hetland, H., Sandal, G. M., & Pallesen, S. (2011). Individual differences in tolerance to shift work–a systematic review. *Sleep Medicine Reviews, 15*(4), 221-235.

Salinas, E., & Sejnowski, T. J. (2001). Correlated neuronal activity and the flow of neural information. *Nature Reviews Neuroscience, 2*(8), 539.

Sanchez, D. J., & Reber, P. J. (2013). Explicit pre-training instruction does not improve implicit perceptual-motor sequence learning. *Cognition, 126*(3), 341-351.

Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: a possible function of EEG theta oscillations. *Neuroscience and Biobehavioral Reviews, 34*(7), 1015-1022.

Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Current Biology*, *22*(17), 1622-1627.

Scheeringa, R., Petersson, K. M., Oostenveld, R., Norris, D. G., Hagoort, P., & Bastiaansen, M. C. (2009). Trial-by-trial coupling between EEG and BOLD identifies networks related to alpha and theta EEG power increases during working memory maintenance. *Neuroimage, 44*(3), 1224-1238.

Schendan, H. E., Searl, M., Melrose, R., & Stern, C. (2003). An FMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron, 37*(6), 1013-1025.

Schwabe, L., Römer, S., Richter, S., Dockendorf, S., Bilak, B., & Schächinger, H. (2009). Stress effects on declarative memory retrieval are blocked by a b-adrenoceptor antagonist in humans. *Psychoneuroendocrinology, 34*, 446—454. doi: 10.1016/j.psyneuen.2008.10.009

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences, 17*(3), 367-395.

Siegel, D. J. (2001). Memory: an overview, with emphasis on developmental, interpersonal, and neurobiological aspects. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(9), 997-1011.

Silber, M. H., Ancoli-Israel, S., Bonnet, M. H., Chokroverty, S., Grigg-Damberger, M. M., Hirshkowitz, M., . . . Iber, C. (2007). The visual scoring of sleep in adults. *Journal of Clinical Sleep Medicine, 3*(2), 121-131.

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. *Mentálhigiéné és Pszichoszomatika, 10*(3), 249-261.

Simor, P., Zavecz, Z., Horvath, K., Elteto, N., Török, C., Pesthy, O., . . . Nemeth, D. (2019). Deconstructing procedural memory: Different learning trajectories and consolidation of sequence and statistical learning. *Frontiers in Psychology, 9*, 2708.

Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annual Review of Physiology, 55*(1), 349-374.

Soldatos, C. R., Dikeos, D. G., & Paparrigopoulos, T. J. (2000). Athens Insomnia Scale: validation of an instrument based on ICD-10 criteria. *Journal of Psychosomatic Research, 48*(6), 555-560.

Soldatos, C. R., Dikeos, D. G., & Paparrigopoulos, T. J. (2003). The diagnostic validity of the Athens Insomnia Scale. *Journal of Psychosomatic Research, 55*(3), 263-267.

Song, S., & Cohen, L. G. (2014). Practice and sleep form different aspects of skill. *Nat Commun, 5*, 3407. doi: 10.1038/ncomms4407

Song, S., Howard, J. H., Jr., & Howard, D. V. (2007a). Implicit probabilistic sequence learning is independent of explicit awareness. *Learning and Memory, 14*, 167–176.

Song, S., Howard, J. H., Jr., & Howard, D. V. (2007b). Sleep does not benefit probabilistic motor sequence learning. *Journal of Neuroscience, 27*(46), 12475-12483. doi: 10.1523/jneurosci.2062-07.2007

Spencer, R. M., Sunm, M., & Ivry, R. B. (2006). Sleep-Dependent Consolidation of Contextual Learning. *Current Biology, 16,* 1001-1005.

Squire, L. R. (1992a). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience, 4,* 232-243.

Squire, L. R. (1992b). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review, 99*(2), 195-231.

Squire, L. R. (2009). The legacy of patient H.M. for neuroscience. *Neuron, 61*(1), 6-9.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences USA, 93,* 13515-13522.

Stark-Inbar, A., Raza, M., Taylor, J. A., & Ivry, R. B. (2016). Individual differences in implicit motor learning: task specificity in sensorimotor adaptation and sequence learning. *Journal of Neurophysiology, 117*(1), 412-428.

Stepanski, E., Koshorek, G., Zorick, F., Glinn, M., Roehrs, T., & Roth, T. (1989). Characteristics of individuals who do or do not seek treatment for chronic insomnia. *Psychosomatics, 30*(4), 421-427.

Sternberg, D. A., Ballard, K., Hardy, J. L., Katz, B., Doraiswamy, P. M., & Scanlon, M. (2013). The largest human cognitive performance dataset reveals insights into the effects of lifestyle factors and aging. *Frontiers in Human Neuroscience, 7,* 292.

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature, 437*(7063), 1272-1278.

Stickgold, R., James, L., & Hobson, J. A. (2000). Visual discrimination learning requires sleep after training. *Nature Neuroscience, 3*(12), 1237-1238. doi: 10.1038/81756

Stickgold, R., & Walker, M. P. (2007). Sleep-dependent memory consolidation and reconsolidation. *Sleep Medicine, 8*(4), 331-343.

Stillman, C. M., Gordon, E. M., Simon, J. R., Vaidya, C. J., Howard, D. V., & Howard Jr, J. H. (2013). Caudate resting connectivity predicts implicit probabilistic sequence learning. *Brain connectivity, 3*(6), 601-610.

Summerfield, C., & Mangels, J. A. (2005). Coherent theta-band EEG activity predicts item-context binding during encoding. *Neuroimage, 24*(3), 692-703.

Szegedi-Hallgató, E., Janacsek, K., Vékony, T., Tasi, L. A., Kerepes, L., Hompoth, E. A., . . . Németh, D. (2017). Explicit instructions and consolidation promote rewiring of automatic behaviors in the human mind. *Scientific Reports, 7*(1), 4365.

Takács, Á., Kóbor, A., Chezan, J., Éltető, N., Tárnok, Z., Nemeth, D., . . . Janacsek, K. (2018). Is procedural memory enhanced in Tourette syndrome? Evidence from a sequence learning task. *Cortex, 100*, 84-94.

Takács, Á., Kóbor, A., Janacsek, K., Honbolygó, F., Kardos, Z., & Nemeth, D. (2016). *Controlled and automatic functions in risk-taking behavior*. Paper presented at the 3rd Conference of the European Society for Cognitive and Affective Neuroscience, Porto, Portugal.

Takács, J., Bódizs, R., Ujma, P. P., 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 and Breathing, 20*(3), 1045-1051.

Tanczos, T., Janacsek, K., & Nemeth, D. (2013a). [Verbal fluency tasks I. Investigation of the Hungarian version of the letter fluency task between 5 and 89 years of age]. *Psychiatria Hungarica: A Magyar Pszichiátriai Társaság Tudományos Folyóirata, 29*(2), 158-180.

Tanczos, T., Janacsek, K., & Nemeth, D. (2013b). [Verbal fluency tasks II. Investigation of the Hungarian version of the semantic fluency task between 5 and 89 years of age]. *Psychiatria Hungarica: A Magyar Pszichiátriai Társaság Tudományos Folyóirata, 29*(2), 181-207.

Team, J. (2019). JASP (Version 0.11.1)[Computer software].

Telzer, E. H., Fuligni, A. J., Lieberman, M. D., & Galván, A. (2013). The effects of poor quality sleep on brain function and risk taking in adolescence. *Neuroimage, 71*, 275-283.

Thut, G., Bergmann, T. O., Fröhlich, F., Soekadar, S. R., Brittain, J.-S., Valero-Cabré, A., . . . Siebner, H. R. (2017). Guiding transcranial brain stimulation by EEG/MEG to interact with ongoing brain activity and associated functions: a position paper. *Clinical Neurophysiology, 128*(5), 843-857.

Thut, G., Schyns, P., & Gross, J. (2011). Entrainment of perceptually relevant brain oscillations by non-invasive rhythmic stimulation of the human brain. *Frontiers in Psychology, 2,* 170.

Tononi, G., & Cirelli, C. (2006). Sleep function and synaptic homeostasis. *Sleep Medicine Reviews, 10*(1), 49-62. doi: 10.1016/j.smrv.2005.05.002

Tóth, B., Janacsek, K., Takács, Á., Kóbor, A., Zavecz, Z., & Nemeth, D. (2017). Dynamics of EEG functional connectivity during statistical learning. *Neurobiology of Learning and Memory*.

Tóth, B., Kardos, Z., File, B., Boha, R., Stam, C. J., & Molnár, M. (2014). Frontal midline theta connectivity is related to efficiency of WM maintenance and is affected by aging. *Neurobiology of Learning and Memory, 114*, 58-69.

Török, B., Janacsek, K., Nagy, D. G., Orbán, G., & Nemeth, D. (2017). Measuring and filtering reactive inhibition is essential for assessing serial decision making and learning. *Journal of Experimental Psychology: General, 146*(4), 529.

Török, C., Janacsek, K., & Nemeth, D. (2016). *How to boost consolidation of implicit probabilistic learning by manipulation of reconsolidation*. Paper presented at the Internetional Conference On Memory, Budapest, Hungary.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381-403). New York: Academic Press.

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: role of the hippocampus. *Hippocampus, 8*(3), 198-204.

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience, 21*(10), 1934-1945.

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *The Journal of Neuroscience, 30*(33), 11177-11187.

Ujma, P. P., Gombos, F., Genzel, L., Konrad, B. N., Simor, P., Steiger, A., . . . Bódizs, R. (2015). A comparison of two sleep spindle detection methods based on all night averages: individually adjusted vs. fixed frequencies. *Frontiers in Human Neuroscience, 9,* 52.

Ujma, P. P., Konrad, B. N., Genzel, L., Bleifuss, A., Simor, P., Potari, A., . . . Dresler, M. (2014). Sleep spindles and intelligence: evidence for a sexual dimorphism. *Journal of Neuroscience, 34*(49), 16358-16368. doi: 10.1523/jneurosci.1857-14.2014

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*(1-2), 231-270.

Ullman, M. T. (2015). The declarative/procedural model: a neurobiological model of language learning, knowledge, and use *Neurobiology of language* (pp. 953-968): Elsevier.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: the procedural deficit hypothesis. *Cortex, 41*(3), 399-433.

Ullman, M. T., & Pullman, M. Y. (2015). A compensatory role for declarative memory in neurodevelopmental disorders. *Neuroscience and Biobehavioral Reviews, 51*, 205-222.

Ulrich, D. (2016). Sleep spindles as facilitators of memory formation and learning. *Neural Plasticity, 2016*.

Urbain, C., Schmitz, R., Schmidt, C., Cleeremans, A., Van Bogaert, P., Maquet, P., & Peigneux, P. (2013). Sleep-dependent Neurophysiological Processes in Implicit Sequence Learning. *Journal of Cognitive Neuroscience*. doi: doi:10.1162/jocn_a_00439

van den Noort, M., Struys, E., Perriard, B., Staudte, H., Yeo, S., Lim, S., & Bosch, P. (2016). Schizophrenia and depression: The relation between sleep quality and working memory. *Asian Journal of Psychiatry, 24*, 73-78.

Varela, F., Lachaux, J.-P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nature Reviews Neuroscience, 2*(4), 229-239.

Violante, I. R., Li, L. M., Carmichael, D. W., Lorenz, R., Leech, R., Hampshire, A., . . . Sharp, D. J. (2017). Externally induced frontoparietal synchronization modulates network dynamics and enhances working memory performance. *Elife, 6*, e22001.

Virag, M., Janacsek, K., Horvath, A., Bujdoso, Z., Fabo, D., & Nemeth, D. (2015). Competition between frontal lobe functions and implicit sequence learning:

evidence from the long-term effects of alcohol. *Experimental Brain Research, 233*(7), 2081-2089.

Vosskuhl, J., Huster, R. J., & Herrmann, C. S. (2015). Increase in short-term memory capacity induced by down-regulating individual theta frequency via transcranial alternating current stimulation. *Frontiers in Human Neuroscience, 9*, 257.

Vöröslakos, M., Takeuchi, Y., Brinyiczki, K., Zombori, T., Oliva, A., Fernández-Ruiz, A., . . . Buzsáki, G. (2018). Direct effects of transcranial electric stimulation on brain circuits in rats and humans. *Nature communications, 9*(1), 483.

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p values. *Psychon Bull Rev, 14*(5), 779-804. doi: 10.3758/BF03194105

Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. (2011). Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). *Journal of Personality and Social Psychology, 100*(3), 426-432. doi: 10.1037/a0022790

Walker, M. P. (2009). The Role of Slow Wave Sleep in Memory Processing. *Journal of Clinical Sleep Medicine, 5*(2 Suppl), S20-S26.

Walker, M. P., Brakefield, T., Morgan, A., Hobson, J. A., & Stickgold, R. (2002). Practice with sleep makes perfect: sleep-dependent motor skill learning. *Neuron, 35*(1), 205-211.

Walker, M. P., & Stickgold, R. (2004). Sleep-dependent learning and memory consolidation. *Neuron, 44*(1), 121-133.

West, G., Vadillo, M. A., Shanks, D. R., & Hulme, C. (2017). The procedural learning deficit hypothesis of language learning disorders: we see some problems. *Developmental science*.

Wilhelm, I., Diekelmann, S., & Born, J. (2008). Sleep in children improves memory performance on declarative but not procedural tasks. *Learning and Memory, 15*(5), 373-377.

Wilhelm, I., Diekelmann, S., Molzow, I., Ayoub, A., Molle, M., & Born, J. (2011). Sleep selectively enhances memory expected to be of future relevance. *Journal of Neuroscience, 31*(5), 1563-1569. doi: 10.1523/jneurosci.3575-10.2011

Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological Review, 105*(3), 558-584.

Willingham, D. B., & Preuss, L. (1995). The death of implicit memory. *Psyche, 2*(15), 1-10.

Willingham, D. B., Salidis, J., & Gabrieli, J. D. (2002). Direct comparison of neural systems mediating conscious and unconscious skill learning. *Journal of Neurophysiology, 88*, 1451-1460.

Wischnewski, M., Zerr, P., & Schutter, D. J. (2016). Effects of theta transcranial alternating current stimulation over the frontal cortex on reversal learning. *Brain stimulation, 9*(5), 705-711.

Yang, J. S., Nicholas, C. L., Nixon, G. M., Davey, M. J., Anderson, V., Walker, A. M., . . . Horne, R. S. (2010). Determining sleep quality in children with sleep disordered breathing: EEG spectral analysis compared with conventional polysomnography. *Sleep, 33*(9), 1165-1172. doi: 10.1093/sleep/33.9.1165

Zavecz, Z., Török, C., Köteles, F., Pálosi, V., & Simor, P. (2015). The psychometric properties of the Hungarian version of the Morningness-Eveningness Questionnaire (MEQ-H): The separate factors of morning freshness and circadian rhythmicity. *Psychiatria Hungarica: A Magyar Pszichiátriai Társaság Tudományos Folyóirata, 30*(3), 318-331.

Zelazo, P. D., Craik, F. I., & Booth, L. (2004). Executive function across the life span. *Acta Psychologica, 115*(2-3), 167-183.

Zhang, L., & Zhao, Z.-X. (2007). Objective and subjective measures for sleep disorders. *Neuroscience Bulletin, 23*(4), 236-240.

Zola-Morgan, S., & Squire, L. R. (1984). Preserved learning in monkeys with medial temporal lesions: Sparing of motor and cognitive skills. *Journal of Neuroscience, 4*(4), 1072-1085.

Zola-Morgan, S., Squire, L. R., & Amaral, D. G. (1986). Human Amnesia and the medial temporal region: Enduring memory impairment following a bilateral lesion limited to field CA1 of the hippocampus. *Journal of Neuroscience, 6*(10), 2950-2967.

Zwart, F. S., Vissers, C. T. W., Kessels, R. P., & Maes, J. H. (2017). Procedural learning across the lifespan: A systematic review with implications for atypical development. *Journal of Neuropsychology*.

Supplementary materials for Study 1 titled

"Delta and theta activity during slow wave sleep are associated with declarative, but not with non-declarative learning in children with sleep-disordered breathing"

Analyses excluding children with OSA

In order to verify whether the above correlations were not produced due to impaired learning specifically within the OSA (n = 4) subgroup, we performed the same analyses based on the data of the primary snoring subgroup only (n = 23).

Declarative Memory (Story Recall)

SWS spectral power within the delta range showed a positive correlation with evening story recall (r = .62, p = .002, *Fig.S1A*), whereas a negative correlation was found within the theta band (r = -.67 p = .001, *Fig.S1B*). All other frequency bands showed non-significant (*p*s > .45) correlations with the evening score. No significant correlations were found between spectral power measures (all *p*s > .52) and overnight memory consolidation (i.e., the change in performance from evening to morning).

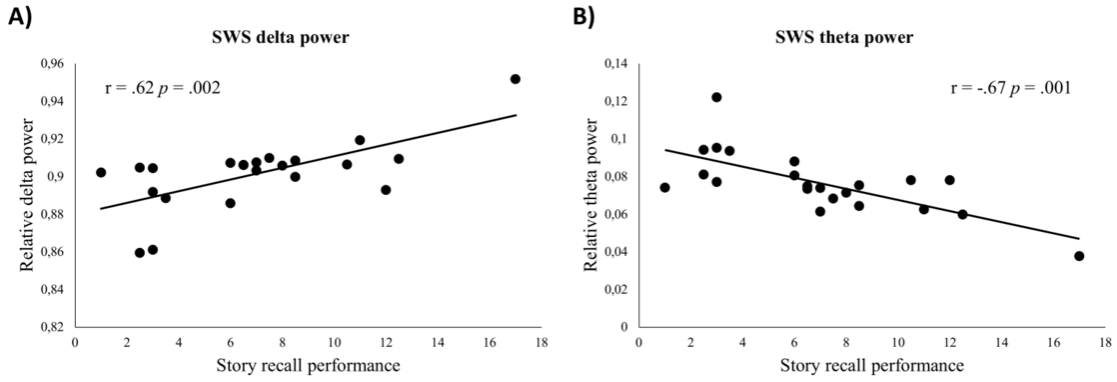

Fig.S1. Correlation between slow wave sleep delta (A) and theta (B) power spectrum and immediate (evening) story recall performance in the case of only primary snoring patients.

To control for the confounding factor of age that might influence both memory performance and SWS we conducted a regression analysis with evening (immediate) story recall performance as the dependent factor, and age and SWS delta spectral power as separately entered independent variables. In the first model, performance in story recall was significantly associated with age (Std. beta = .58, p = .005). In the second model, age (Std. beta = .39, p = .024), and delta power (Std.beta = .53, p = .004) were both significant predictors of immediate story recall. We conducted the same regression analysis with evening story recall performance as the dependent variable, and age and SWS theta spectral power as separately entered independent variables. In the final model, age was not significantly associated with story recall performance (Std. beta = .22, p = .13), but theta power remained a significant predictor (Std.beta = -.56, p = .005). Both delta and theta power increased the explained variance of evening recall beyond the explained variance of age.

Non-declarative memory (ASRT)

SWS spectral power measures were not associated with the statistical learning score in the evening (based on the last, 5th epoch) (all *p*s > .57), or in the morning session (all *p*s > .19) in terms of ACCs. Moreover, spectral power measures were not associated to memory consolidation (all *p*s >.19) of statistical learning indexed by overnight change in ACCs. Similarly, no significant correlations emerged between statistical learning performance in the evening (all *p*s > .54), or in the morning session (all *p*s > .08) in terms of RT. Spectral power measures were not associated with memory consolidation (all *p*s >.31) indexed by overnight change in RT.

Unlike statistical learning, SWS spectral power measures were associated with general skill learning in case of ACCs. Similarly to story recall, SWS spectral power in the delta range showed a positive correlation with the average ACCs (averaged across high and low frequency triplets) assessed in the evening (based on the last, 5th epoch, r = .59, p = .004), whereas a negative correlation was found with theta band power (r = - .47, p = .03). All other frequency bands showed non-significant (*p*s > .12) correlations with the average ACCs measured in the evening session. Similar, although stronger correlations were found between the morning ACCs and band-wise spectral power measures (delta: r = .82, p < .001, theta: r = -.70, p = .001, all other *p*s > .33). No

significant correlations were found between spectral power measures (all *p*s > .51) and overnight change in average ACCs (i.e. consolidation of general skill learning).

In the case of general skill learning indexed by averaged RTs for high and low frequency triplets, no significant correlations emerged between skill learning and spectral power (all *p*s > .25). Neither we found significant correlations between the overnight RTs change and spectral power measures, although delta (r = .04, p = .08) and theta (r = -.44, p = .05) band power correlated with overnight change on a trend level (all other *p*s > .58).

Similarly to story recall, we controlled for the confounding factor of age that might influence both memory performance and SWS. First, we conducted a regression analysis with average evening ACCs as the dependent factor, and age and slow wave delta spectral power as separately entered independent variables. In the first model, ACCs was significantly associated with age (Std. beta = .48, p = .023; Adj. R^2 = .20, F (1,20) = 6.10; p = .023). In the second model, the influence of age remained significant (Std. beta = .38, p = .05), but delta power was not a significant predictor (Std.beta = .24, p = .24) of ACCs (Model: Adj. R^2 = .40, F (2,21) = 7.92 p = .003). We conducted the same regression analysis with average evening ACCs as dependent variable, and age and slow wave theta spectral power as separately entered independent variables. In the third model in which both age and theta spectral power were entered, neither age (Std. beta = .26, p = .29), nor theta power (Std.beta = -.39, p = .11) were significant predictors of ACCs (Model: Adj. R^2 = .26, F (2,21) = 4.71 p = .022).

Supplementary materials for Study 2 titled

"Frontal-midline theta frequency and probabilistic learning: A transcranial Alternating Current Stimulation study"

Does baseline performance influence the effects of the stimulation?

Recent studies showed that baseline performance could interact with the effect of stimulation, namely that low-performing participants usually benefit more from transcranial electric stimulation compared to high-performing participants [1-4]. As a post-hoc analysis, we wanted to explore whether the participants' initial performance interact with the effect of the stimulation. We conducted four mixed-design analyses of variance (ANOVAs) with STIMULATION (Sham vs. Active) and EPOCH (1-4) as within-subject factors and four different grouping variables based on initial performance as a between-subject factor on the reaction time (RT) and accuracy (ACC) learning scores. The four grouping variables for high- vs. low-performing participants were the following: 1) BASELINE RT GROUP, 2) BASELINE RT LEARNING GROUP, 3) BASELINE ACC GROUP, 4) BASELINE ACC LEARNING GROUP. For each of these grouping variables, we divided participants into two groups with equal size (N = 13) based on their initial performance on the respective variable measured in the first epoch of the first session (independent of the stimulation condition): average RT was used for the BASELINE RT GROUP, RT learning score for the BASELINE RT LEARNING GROUP, average ACC for the BASELINE ACC GROUP, and ACC learning scores for the BASELINE ACC LEARNING GROUP. ANOVAs with grouping variables based on RT data were computed on standardized RT learning scores, and ANOVAs with grouping variables based on ACC data were computed on the ACC learning scores. Thus, we conducted the two mixed-design ANOVAs with the RT grouping variables on the standardized RT learning scores (for the calculation of the learning scores, see the main text) and the two ANOVAs with the ACC grouping variables on the ACC learning scores. Furthermore, we conducted the same mixed-design ANOVAs with a Bayesian approach, as well. To summarize the importance of the effects (in particular, the interaction between the groups and the stimulation) across all models, we performed model averaging and

report inclusion Bayes factors (for further details on the Bayesian ANOVA, see main text).

Baseline RT performance. First, we investigated whether the effect of stimulation differed depending on participants' initial RTs (Table S1-S2, Baseline RT column). Similarly to the results reported in the main text, the main effect of EPOCH was significant, indicating that the learning scores increased throughout training, irrespective of the stimulation. We did not find any significant differences between groups with fast vs. slow baseline RTs either in overall learning scores (main effect of BASELINE RT GROUP) or the trajectory of these learning scores (BASELINE RT GROUP x EPOCH interaction). We also did not find any significant differences between the learning scores in the active stimulation vs. sham conditions (main effect of STIMULATION, and STIMULATION x EPOCH interaction). Furthermore, initial baseline RTs did not seem to influence the stimulation effects (STIMULATION x BASELINE RT GROUP interaction and STIMULATION x EPOCH x BASELINE RT GROUP interaction).

The analysis of effects (model-averaged results) of the Bayesian ANOVA on the RT learning scores showed that the effect of EPOCH should be included in the model ($BF_{inclusion}$ = 74.523), while other main effects or interactions should not (all $BF_{inclusion}$ < 1, Table S2). Importantly, there was substantial evidence that the STIMULATION x BASELINE RT GROUP and STIMULATION x EPOCH x BASELINE RT GROUP interactions should not be included in the model ($BF_{inclusion}$ = 0.072, $BF_{inclusion}$ = 0.0002 respectively). To conclude, we did not find different stimulation effects in participants with fast or slow initial RTs.

Baseline RT learning performance. Second, we tested whether initial probabilistic learning performance measured by RTs could interact with the effect of stimulation (Table S1-S2, Baseline RT learning column). The main effect of EPOCH was again significant, indicating increasing learning scores as the task progressed. The BASELINE RT LEARNING GROUP x EPOCH interaction was also significant, indicating that the trajectory of the learning scores differed between the two groups, irrespective of the stimulation. The post-hoc analysis showed that the groups significantly differed in the first epoch ($p < .001$, possibly influenced by the grouping criterion itself), but not in the subsequent epochs ($ps > .12$) indicating that the initial performance difference diminished throughout the task. The overall learning scores did not differ

significantly between the two groups (main effect of BASELINE RT LEARNING GROUP). Again, we did not find any significant differences between the learning scores in the active stimulation vs. sham conditions (main effect of STIMULATION and STIMULATION x EPOCH interaction). Furthermore, initial learning scores did not seem to influence the stimulation effects either (STIMULATION x BASELINE RT LEARNING GROUP interaction and STIMULATION x EPOCH x BASELINE RT LEARNING GROUP interaction).

The analysis of effects (model-averaged results) of the Bayesian ANOVA on RT learning scores showed that the effect of EPOCH, BASELINE RT LEARNING GROUP and BASELINE RT LEARNING GROUP x EPOCH should be included in the model ($BF_{inclusion}$ = 337.966, $BF_{inclusion}$ = 2.367, $BF_{inclusion}$ = 7.672 respectively), while other main effects or interactions should not (all $BF_{inclusion}$ < 1, Table S2). Importantly, there was substantial evidence that the STIMULATION x BASELINE RT GROUP and STIMULATION x EPOCH x BASELINE RT GROUP interactions should not be included in the model ($BF_{inclusion}$ = 0.127, $BF_{inclusion}$ = 0.025 respectively). To conclude, we did not find different stimulation effects in participants with high or low initial learning scores measured by RT.

Baseline ACC performance. Next, we tested whether the effect of stimulation differed depending on participants' initial ACC (Table S1-S2, Baseline ACC column). Contrary to RT learning scores, ACC learning scores appeared to be stable throughout the task (non-significant main effect of EPOCH). We did not find any significant differences between groups with high vs. low baseline ACC either in the overall learning scores (main effect of BASELINE ACC GROUP) or the trajectory of these learning scores (BASELINE ACC GROUP x EPOCH interaction). We also did not find any significant differences between the learning scores of the active stimulation vs. sham conditions (main effect of STIMULATION and STIMULATION x EPOCH interaction). Furthermore, initial baseline ACC did not seem to influence the stimulation effects (STIMULATION x BASELINE ACC GROUP interaction and STIMULATION x EPOCH x BASELINE ACC GROUP interaction).

The analysis of effects (model-averaged results) of the Bayesian ANOVA on ACC learning scores showed that none of the studied effects should be included in the model (all $BF_{inclusion}$ < 1, Table S2). Importantly, there was substantial evidence that the

STIMULATION x BASELINE RT GROUP and STIMULATION x EPOCH x BASELINE RT GROUP interactions should not be included in the model ($BF_{inclusion}$ = 0.019, $BF_{inclusion}$ = 0.00006 respectively). To conclude, we did not find different stimulation effects in participants with high vs. low initial ACC.

Baseline ACC learning performance. Lastly, we tested whether initial probabilistic learning performance measured by ACC could interact with the effect of stimulation (Table S1-S2, Baseline ACC learning column). Again, the ACC learning scores appeared to be stable throughout the task (non-significant EPOCH main effect). However, there was a significant difference in the overall learning scores between those who had low vs. high baseline ACC learning scores (main effect of BASELINE ACC LEARNING GROUP): the difference in the initial learning performance that was the basis of the grouping, appeared to have remained throughout the task (M = 1.8% and M = 3.4% for the groups with low vs. high initial ACC learning scores, respectively). We did not find significant differences between the trajectory of learning scores in the two groups (BASELINE ACC LEARNING GROUP x EPOCH interaction). Again, we did not find any significant differences between the learning scores of the active stimulation and sham conditions (main effect of STIMULATION and STIMULATION x EPOCH interaction). Furthermore, initial learning scores did not seem to influence the stimulation effects either (main effect of BASELINE ACC LEARNING GROUP, STIMULATION x BASELINE ACC LEARNING GROUP interaction and STIMULATION x EPOCH x BASELINE ACC LEARNING GROUP interaction).

The analysis of effects (model-averaged results) of the Bayesian ANOVA on the ACC learning scores showed that none of the studied effects should be included in the model (all $BF_{inclusion} < 1$, Table S2). Importantly, there was substantial evidence that the STIMULATION x BASELINE RT GROUP and STIMULATION x EPOCH x BASELINE RT GROUP interactions should not be included in the model ($BF_{inclusion}$ = 0.047, $BF_{inclusion}$ = 0.004 respectively). To conclude, we did not find different stimulation effects in participants with high or low initial learning scores measures by ACC.

Table S1. Results from frequentist ANOVAs performed with different initial performance groups

Effect	Statistics	Baseline RT	Baseline RT learning	Baseline ACC	Baseline ACC learning
Stimulation	*F*	0.051	0.053	0.015	0.015
	p	.824	.819	.902	.903
	η^2_p	.002	.002	.001	.001
Epoch	*F*	6.945	8.030	2.208	2.317
	p	< .001*	< .001*	.095	.083
	η^2_p	.224	.251	.084	.088
Stimulation x Epoch	*F*	0.574	0.590	1.170	1.284
	p	.634	.623	.322	.287
	η^2_p	.023	.024	.046	.051
Group	*F*	0.238	2.165	1.562	7.699
	p	.630	.154	.223	.011*
	η^2_p	.010	.083	.061	.243
Group x Epoch	*F*	0.814	4.688	0.302	1.506
	p	.490	.005*	.824	.220
	η^2_p	.033	.163	.012	.059
Stimulation x Group	*F*	0.110	1.343	0.272	0.172
	p	.743	.258	.607	.682
	η^2_p	.005	.053	.011	.007
Stimulation x Group x Epoch	*F*	0.525	1.203	0.113	2.457
	p	.666	.315	.912	.089
	η^2_p	.021	.048	.005	.093

Note: Results of four mixed-design analyses of variance (ANOVAs) with STIMULATION (Sham vs. Active) and EPOCH (1-4) as within-subject factors and 1) BASELINE RT GROUP, 2) BASELINE RT LEARNING GROUP, 3) BASELINE ACC GROUP, or 4) BASELINE ACC LEARNING GROUP as a between-subject factor on the respective RT/ACC learning scores. Relevant interaction effects that show whether the baseline performance groups interacted with the stimulation effects are boldfaced. * $p < .05$.

Table S2. Model-averaged results from Bayesian ANOVAs performed with different initial performance groups

Effect	Baseline RT	Baseline RT learning	Baseline ACC	Baseline ACC learning
Stimulation	0.066	0.086	0.060	0.067
Epoch	74.523	337.966	0.185	0.225
Stimulation x Epoch	0.131	0.040	0.016	0.021
Group	0.029	2.367	0.132	0.560
Group x Epoch	0.018	7.672	0.014	0.145
Stimulation x Group	0.072	0.127	0.019	0.047
Stimulation x Group x Epoch	2.392e -4	0.025	6.461e -5	0.004

Note: We report inclusion Bayes Factors. Results of four mixed-design analyses of variance (ANOVAs) with STIMULATION (Sham vs. Active) and EPOCH (1-4) as within-subject factors and 1) BASELINE RT GROUP, 2) BASELINE RT LEARNING GROUP, 3) BASELINE ACC GROUP, or 4) BASELINE ACC LEARNING GROUP as a between-subject factor on the respective RT/ACC learning scores. Relevant interaction effects that show whether the baseline performance groups interacted with the stimulation effects are boldfaced.

Altogether, our results show that the individual differences in the initial baseline performance (as defined by the average speed and accuracy, and RT and ACC probabilistic learning scores at the beginning of the task) did not interact with the effects of the stimulation. However, as these tests were post-hoc, and the number of participants was low in the groups for each comparison, these results should be treated with caution and further studies are warranted to probe whether initial performance in a probabilistic learning task could be a relevant factor when testing the effect of brain stimulation on learning.

Does the partial overlap between the sequences practiced during the two stimulation sessions influence the effects of the stimulation?

To ensure that the partially overlapping sequence in the task between the two sessions did not distort the effects of the stimulation, we computed learning scores excluding the responses (RT and ACC) to those triplets that were high-probability in both sessions. As the retention of probabilistic information is intact after one week [5], or even after one year [6], responses during the second session can be faster for those high-probability triplets that were practiced in the previous session. By eliminating these high-probability triplets, we ensure that the difference between the performance for high- and low-probability triplets will be due to the probabilistic information acquired during that session. Thus, for both sessions, we included data only from high-probability triplets that were unique for the given session.

After the elimination of overlapping high-probability triplets, overall RTs and ACC significantly differed between the two sessions (as revealed by the main effect of SESSION in the repeated-measures ANOVA with SESSION (First vs. Second), EPOCH (1-4) and TRIPLET TYPE (High vs. Low) as within-subject factors for RT: $F(1, 25) = 142.879$, $p < .001$, $\eta^2_p = .851$ and for ACC: $F(1, 25) = 5.821$, $p = .024$, $\eta^2_p = .022$). Therefore, we calculated standardized scores within each subject in each session. For RTs, we calculated z-score within each subject in each session, while for ACC, we corrected the performance of each session with the initial (i.e., during the 1st epoch) performance of each subject. Then, for each epoch, we calculated learning scores both for RT and ACC data. For RT, the learning score was calculated as the difference between the z-transformed RTs for low-probability triplets minus the z-transformed RTs for high-probability triplets. For ACC, the learning score was calculated as the standardized ACC for high-probability triplets minus the standardized ACC for low-probability triplets. In both cases, a higher learning score indicates better learning.

To evaluate changes in probabilistic learning as a function of stimulation, we conducted repeated-measures ANOVAs separately for the RT and ACC learning scores with STIMULATION (Sham vs. Active) and EPOCH (1-4) as within-subject factors. 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^2_p) as the measure of effect size. Furthermore, we conducted the same repeated-measures

ANOVAs with STIMULATION and EPOCH as within-subject factors for RT and ACC learning scores with a Bayesian approach. The Bayesian ANOVA is a model comparison approach (for details on the interpretation of the Bayesian ANOVA, see the main text).

RT performance. The frequentist repeated-measures design ANOVA on the z-transformed RT learning scores revealed a significant Intercept ($F(1, 25) = 32.552$, $p < .001$, $\eta^2_p = .566$), suggesting that learning occurred in the ASRT task. The main effect of EPOCH was also significant ($F(3, 75) = 6.102$, $p = .001$, $\eta^2_p = .196$), indicating that the learning scores increased throughout the task, independent of the stimulation condition (Fig. S1A). However, we did not find any significant differences between the active stimulation and sham conditions either in overall learning (main effect of STIMULATION: $F(1, 25) = 0.048$, $p = .829$, $\eta^2_p = .002$) or in the trajectory of learning (STIMULATION * EPOCH interaction: $F(3, 75) = 0.471$, $p = .704$, $\eta^2_p = .018$).

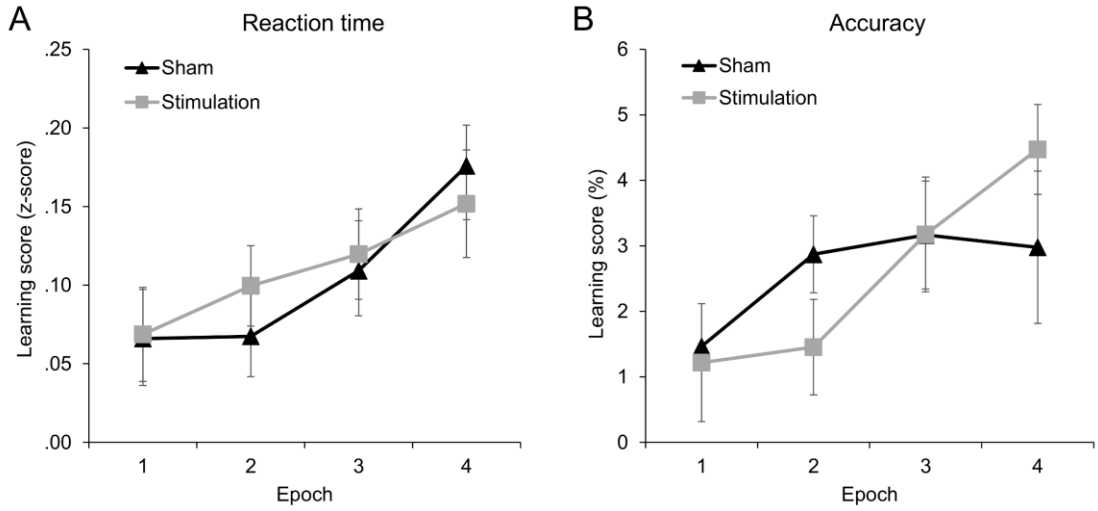

Figure S1. Probabilistic learning in terms of reaction times (A) and accuracy (B) in the active stimulation vs. sham conditions across the four epochs of the ASRT task, excluding the performance for overlapping high-probability triplets over the sessions. There was no significant difference between the active stimulation in theta frequency (grey squares) and sham (black triangles) conditions either in overall learning or in the time course of learning. Error bars indicate Standard Error of Mean (SEM).

The analysis of effects (model-averaged results) of the Bayesian repeated-measures ANOVA on the RT learning scores showed that the effect of EPOCH should be included in the model ($BF_{inclusion} = 14.05$), while the main effect of STIMULATION and the STIMULATION x EPOCH interaction should not (all $BF_{inclusion} < 1$, Table S3).

Thus, based on the Bayesian analysis of effects, the learning scores changed throughout the task, but they were independent of the stimulation condition or the order of the stimulation.

Table S3. Model-averaged results of Bayesian ANOVA for RT learning scores

Effects	P(incl)	P(incl|data)	$BF_{inclusion}$
Stimulation	0.600	0.146	0.114
Epoch	0.600	0.955	14.050
Stimulation x Epoch	0.200	0.011	0.043

Note: The Effects column denotes predictors of interest, the column P(incl) shows the prior inclusion probability, P(incl | D) shows the posterior inclusion probability, and $BF_{Inclusion}$ shows the inclusion Bayes factor.

Based on the Bayesian model comparison, the best model for our data contained only the main effect of EPOCH (Table S4). More specifically, the model with the main effect of EPOCH is ~6 times more likely than any model including the effect of the STIMULATION. This suggests that while the learning scores changed during the task, this was independent of the stimulation condition.

Table S4. Bayesian model comparisons for RT learning scores

Models	P(M)	P(M|data)	BF_M	BF_{10}	error %
Epoch	0.200	0.814	17.551	1.000	
Stimulation + Epoch	0.200	0.130	0.596	0.159	2.428
Null model	0.200	0.039	0.163	0.048	0.753
Stimulation + Epoch + Stim. x Epoch	0.200	0.011	0.043	0.013	3.773
Stimulation	0.200	0.006	0.024	0.007	1.481

Note: All models include Subject. The Model column shows the predictors included in each model, the P(M) column the prior model probability, the P(M | D) column the posterior model probability, the BF_M column the posterior model odds, and the BF_{10} column the Bayes factors of all models compared to the best model. The final column, 'error' is an estimate of the numerical error in the computation of the Bayes factor. All models are compared to the best model and are sorted from highest Bayes factor to lowest.

ACC performance. The frequentist repeated-measures design ANOVA on the ACC learning scores revealed a significant Intercept ($F(1, 25) = 47.205$, $p < .001$, $\eta^2_p = .654$), suggesting that learning occurred in the ASRT task. The main effect of EPOCH was also significant ($F(3, 75) = 3.490$, $p = .020$, $\eta^2_p = .122$), indicating that the learning scores increased throughout the task, independent of the stimulation condition (Fig. S2B). However, we did not find any significant differences between the active stimulation and sham conditions either in overall learning (main effect of STIMULATION: $F(1, 25) = 0.004$, $p = .952$, $\eta^2_p < .001$) or in the time course of learning (STIMULATION * EPOCH interaction: $F(3, 75) = 1.072$, $p = .361$, $\eta^2_p = .041$).

The analysis of effects (model-averaged results) of the Bayesian repeated-measures ANOVA on the ACC learning scores showed that the main effect of STIMULATION and the STIMULATION x EPOCH interaction should not be included in the model (all $BF_{inclusion} < 1$, Table S5). Inclusion of the main effect of EPOCH remained inconclusive ($BF_{inclusion} = 1.098$). Thus, based on the Bayesian analysis of effects, the learning scores were stable throughout the task and they were independent of the stimulation condition or the order of the stimulation.

Table S5. Model-averaged results of Bayesian ANOVA for ACC learning scores

Effects	P(incl)	P(incl|data)	$BF_{inclusion}$
Stimulation	0.600	0.141	0.109
Epoch	0.600	0.622	1.098
Stimulation x Epoch	0.200	0.014	0.055

Note: The Effects column denotes predictors of interest, the column P(incl) shows the prior inclusion probability, P(incl | D) shows the posterior inclusion probability, and $BF_{Inclusion}$ shows the inclusion Bayes factor.

Based on the Bayesian model comparison, the best model for our data is with only the main effect of EPOCH (Table S6). This model is ~7 times more likely than any model including the STIMULATION factor. This suggests that while the learning scores changed during the task, this was independent of the stimulation condition.

Table S6. Bayesian model comparisons for ACC learning scores

Models	P(M)	P(M|data)	BF_M	BF_{10}	error %
Epoch	0.200	0.530	4.507	1.000	
Null model	0.200	0.329	1.965	0.622	0.753
Stimulation + Epoch	0.200	0.079	0.342	0.149	1.393
Stimulation	0.200	0.048	0.203	0.091	1.313
Stimulation + Epoch + Stim. x Epoch	0.200	0.014	0.055	0.026	1.913

Note: All models include Subject. The Model column shows the predictors included in each model, the P(M) column the prior model probability, the P(M | D) column the posterior model probability, the BF_M column the posterior model odds, and the BF_{10} column the Bayes factors of all models compared to the best model. The final column, 'error' is an estimate of the numerical error in the computation of the Bayes factor. All models are compared to the best model and are sorted from highest Bayes factor to lowest.

Altogether, our results after the elimination of the overlapping high-probability triplets are identical to the results without the elimination of these triplets: the stimulation did not influence the learning scores and their trajectories neither in case of RT nor ACC learning scores. Therefore, we can conclude that the overlapping sequences in the two sessions did not influence the effect of the stimulation considerably.

Are there any obvious patterns in the stimulation effects for different individuals?

To explore whether the stimulation affected individuals differently, we plotted the average RT and ACC learning scores for each epoch in the two stimulation conditions for each participant (Fig. S2-S3). We could not identify any obvious subgroups of participants as a function of their learning performance in the stimulation vs. sham conditions. Furthermore, to explore visually whether the order of the conditions influenced the effect of stimulation, we grouped the participants based on whether they completed the sham condition (Fig. S2A and S3A), or the active stimulation condition first (Fig. S2B and S3B). Again, no obvious patterns emerged based on the stimulation order.

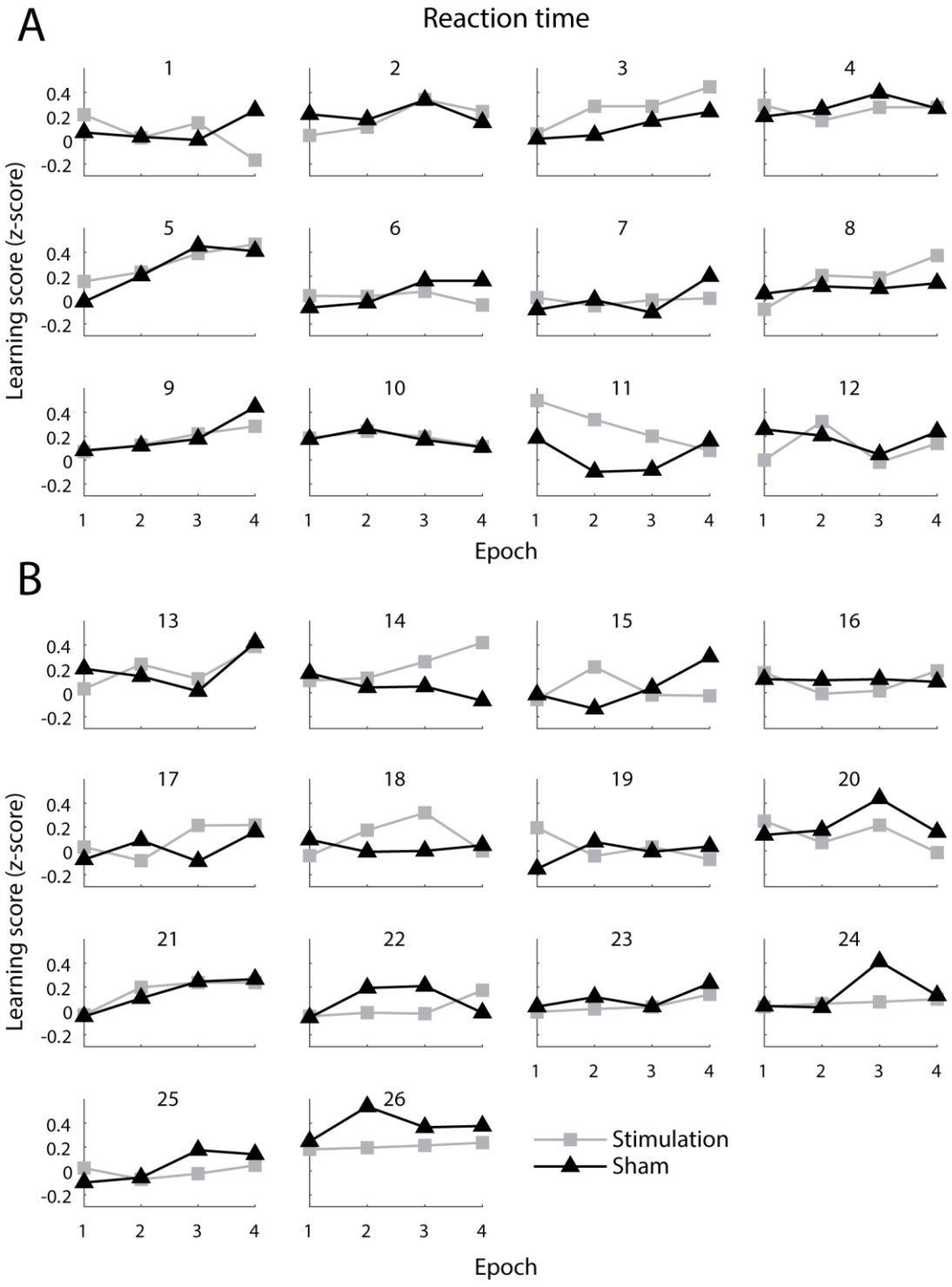

Figure S2. Probabilistic learning in terms of reaction times in the active stimulation vs. sham conditions across the four epochs of the ASRT task for each participant separately for those who completed the sham (A) or the active stimulation (B) condition first. There were no obvious subgroups based on the difference between the active stimulation in theta frequency (grey squares) and sham (black triangles) conditions either in overall learning or in the time course of learning. Furthermore, the order of the stimulation did not seem to interact with the effects of the stimulation.

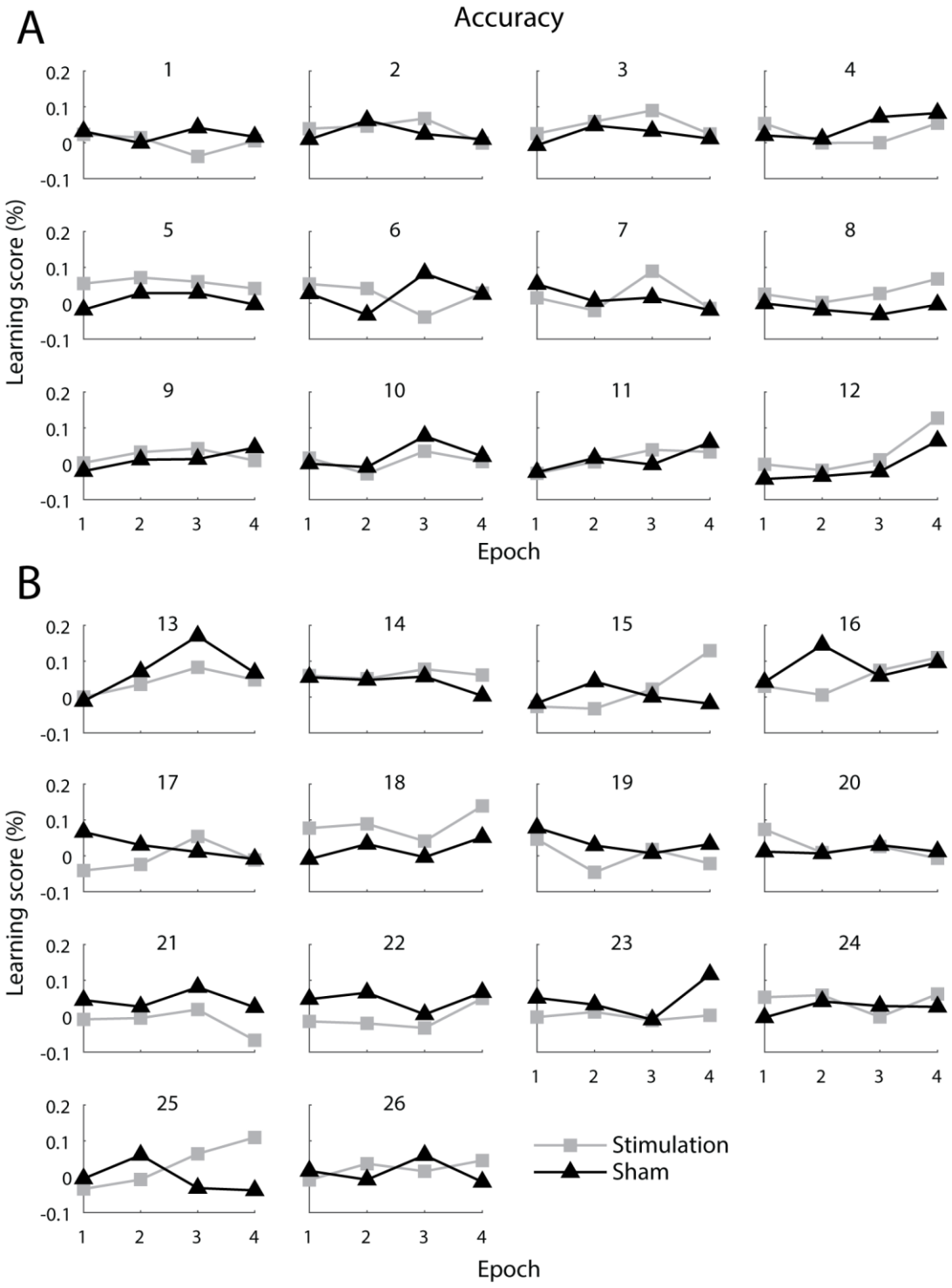

Figure S3. Probabilistic learning in terms of accuracy in the active stimulation vs. sham conditions across the four epochs of the ASRT task for each participant separately for those who completed the sham (A) or the active stimulation (B) condition first. There were no obvious subgroups based on the difference between the active stimulation in theta frequency (grey squares) and sham (black triangles) conditions either in overall learning or in the time course of learning. Furthermore, the order of the stimulation did not seem to interact with the effects of the stimulation.

Altogether, the plots did not unravel any obvious subgroups based on differences between the active stimulation and sham conditions either in overall learning or in the time course of learning. Furthermore, the order of the stimulation did not seem to interact with the effects of the stimulation.

References

[1] E. Santarnecchi, T. Muller, S. Rossi, A. Sarkar, N. Polizzotto, A. Rossi, R.C. Kadosh, Individual differences and specificity of prefrontal gamma frequency-tACS on fluid intelligence capabilities, Cortex 75 (2016) 33-43.

[2] P. Tseng, Y.-T. Chang, C.-F. Chang, W.-K. Liang, C.-H. Juan, The critical role of phase difference in gamma oscillation within the temporoparietal network for binding visual working memory, Sci. Rep. 6 (2016) 32138.

[3] P. Tseng, T.-Y. Hsu, C.-F. Chang, O.J. Tzeng, D.L. Hung, N.G. Muggleton, V. Walsh, W.-K. Liang, S.-k. Cheng, C.-H. Juan, Unleashing potential: transcranial direct current stimulation over the right posterior parietal cortex improves change detection in low-performing individuals, J. Neurosci. 32(31) (2012) 10554-10561.

[4] G. Learmonth, G. Thut, C.S. Benwell, M. Harvey, The implications of state-dependent tDCS effects in aging: behavioural response is determined by baseline performance, Neuropsychologia 74 (2015) 108-119.

[5] D. Nemeth, K. Janacsek, The dynamics of implicit skill consolidation in young and elderly adults, Journal of Gerontology Psychological Science 66(1) (2011) 15-22.

[6] A. Kobor, K. Janacsek, A. Takacs, D. Nemeth, Statistical learning leads to persistent memory: Evidence for one-year consolidation, Sci. Rep. 7(1) (2017) 760.

Supplementary materials for Study 4 titled

"The relationship between subjective sleep quality and cognitive performance in healthy young adults: Evidence from three empirical studies"

Supplementary methods

Cognitive performance assessments

Procedural learning. Procedural learning performance was measured by the explicit version of the Alternating Serial Reaction Time (ASRT) task (Figure S1, see also [1]). In the explicit version of the Alternating Serial Reaction Time (ASRT) 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 Chronos response box (Psychology Software Tools, INC) in Study 1 and Study 2, and a special keyboard with four heightened keys (Z, C, B, and M on a QWERTY keyboard) in Study 3 when the stimulus occurred. The appearance of stimuli followed a predetermined alternating sequence order, such that every second element was part of the sequence and every second element was randomly selected: the dog stimulus always corresponded to sequence elements, and the penguin stimulus indicated random elements (Figure S1A). Participants were informed about this underlying structure of the sequence, and their attention was drawn to the alternation of sequence and random elements by the different visual cues (i.e., dogs vs. penguins). Participants were instructed to respond as quickly and accurately as they could, and to find the hidden pattern defined by the dog in order to improve their performance.

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 timing of the stimulus differed in the three studies. In Study 1, the stimulus remained on the screen until the participant pressed the correct response button,

and the next stimulus was presented 250 ms following the previous response. In Study 2, the stimulus remained on the screen until the participant pressed the correct response button, and the next stimulus was presented 120 ms following the previous response. In Study 3, the stimulus remained on the screen for 580 ms, the participant was asked to respond within this time window, and the next stimulus was presented 120 ms following the previous stimulus. Thus, the task was self-paced with different response-to-stimulus intervals (RSI) in Study 1 and 2, while it was fix-paced (inter-stimulus interval, ISI, of 700 ms) in Study 3. The timing parameters of Study 3 was determined based on previous ASRT studies showing that healthy young adults' average RT performance is around 370-430 ms during the task [1,2]. We used different settings to explore which timing parameters promote better learning performance. It has been suggested that longer RSI/ISI (e.g., 250 as opposed to 120 ms) can lead to better learning performance as participants have more time to process and elaborate the stimuli [3]. Nevertheless, it is also plausible that the shorter the time between subsequent stimuli, the easier to find the association among them, which is essential in the ASRT task to achieve a good learning performance.

In all three studies, the ASRT task consisted of 20 blocks. As one block took approximately 1-1.5 min, the session took approximately 20-25 min. For each participant, one of the six unique permutations of the four possible stimulus positions was selected in a pseudo-random manner, so that the six different sequences were used equally often across participants [4,5].

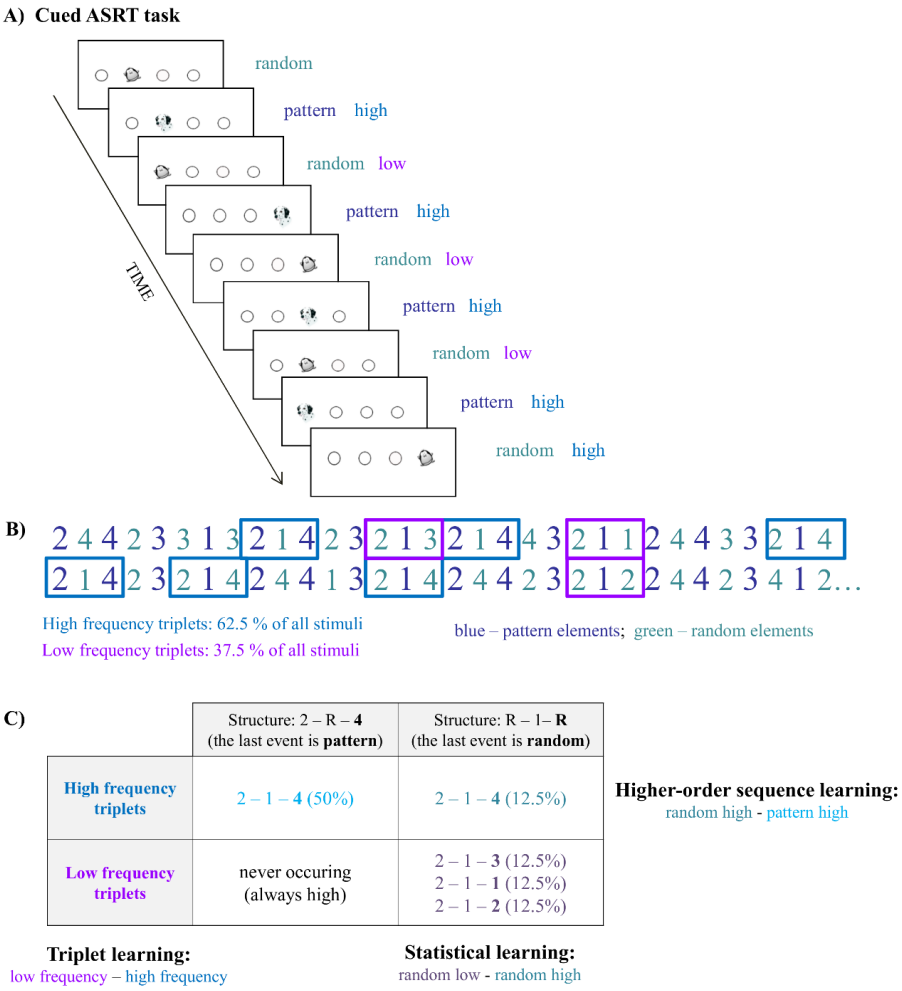

Figure S1. **Schematic diagram of the procedural learning (ASRT) task design**. A) In this task, the appearance of stimuli is based on a predetermined sequence order, in which pattern and random elements alternate (e.g., 2r4r3r1r, where numbers correspond to the four locations on the screen and the 'r' represents randomly chosen locations). The pattern and random elements are cued differently: the dog stimulus always corresponded to pattern elements, and the penguin stimulus indicated random elements. B) Numbers (corresponding to the locations on the screen) in blue represent elements of the pattern trials (e.g., appearing in the sequential order 2, 4, 3, 1 throughout the task), which were alternating with random elements (green). Because of this alternating structure, some runs of three consecutive trials (triplets) occur more frequently than others (high- vs. low-frequency triplets). For each element, we determined whether it was the last element of a high-frequency triplet (one example in blue frame) or low-frequency triplet (examples in magenta frames). C) *Triplet learning* (see text) was calculated as the difference in responses for the last elements of high-frequency triplets (irrespective of random or pattern position) compared to the last elements of low-frequency triplets. *Statistical learning* was assessed by comparing the responses for those *random* elements that were the last elements of high-frequency triplets vs. those that were the last elements of low-frequency triplets (right column). *Higher-order sequence learning* was assessed as a difference between responses for pattern elements (which are always high-frequency triplets) vs. random-high frequency triplet elements (top row).

Trial types and procedural learning indices in the ASRT task. The alternating sequence of the ASRT task forms a sequence structure in which some of the runs of three successive trials (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 a pattern or a random stimulus. However, 3X2 and 4X2 occur less frequently since the first and the third elements can only be a random stimulus. Figure S1B and S1C 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 termed as *high-frequency* triplets, whereas the latter types are termed as *low-frequency* triplets (Figure S1C, see also [1]). The third element of a high-frequency triplet is highly predictable (with 62.5% probability) based on the first element of the triplet. In contrast, in low-frequency triplets, the predictability of the third element is less predictable (with 12.5 % probability) based on the first element of the triplet. According to this principle, each trial was categorized as either the third element of a high- or a low-frequency triplet.

Additionally, trials are differentiated by the visual cues (dog vs. penguin) indicating whether a pattern or a random stimulus was presented in that given trial. In case of pattern trials, participants can use their explicit knowledge of the sequence to predict that trial. Consequently, we further differentiate the previously defined high-frequency triplets into two categories based on whether the last element of the triplet was a pattern or a random stimulus. This way, the task consists of three trial types: 1) trials that belong to the explicitly cued sequential pattern and, at the same time, appear as the last element of a high-frequency triplet are termed *pattern* trials; 2) trials of random stimuli that appear as the last element of a high-frequency triplet are termed *random high* trials; and 3) trials of random stimuli that appear as the last element of a low-frequency triplet are termed *random low* trials (see the example in Figure S1C).

Previous studies have shown that as people practice the ASRT task, they come to respond more quickly and more accurately to the high-frequency triplets (irrespective of whether it was for a pattern or a random stimulus) compared to low-frequency triplets (always random), revealing *Triplet learning* [5,6]. *Triplet learning* is measured as the difference in reaction time (RT) and accuracy (ACC) between high- and low-frequency triplets (RTs of low-frequency triplets minus RTs of high-frequency triplets; ACC of

high-frequency triplets minus ACC of low-frequency triplets). Thus, greater Triplet learning is defined as faster/more accurate responses to high-frequency triplets compared to low-frequency triplets. Importantly, however, the comparison of RT and ACC of high- vs. low-frequency triplets does not take into account whether the last elements of the high-frequency triplets are pattern or random stimuli and consequently, provides a mixed measure of at least two separate learning processes.

The two key learning processes that can be disentangled in the explicit ASRT task are the so-called *Higher-order sequence learning* and the so-called *Statistical learning* (Figure S1C). *Higher-order sequence learning* is measured as the difference in RTs between random high and pattern trials (RTs for random high trials minus RTs for pattern trials; ACC for pattern trials minus ACC for random high trials). These trials 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 trials). Thus, greater Higher-order sequence learning is defined as faster/more accurate responses to pattern trials compared to random high trials. This learning measure thus can reflect the knowledge about the alternating sequential structure that the participants explicitly acquired during the task.

Statistical learning is assessed by comparing the responses for those random trials that were the last elements of a high-frequency triplet vs. those that were the last elements of a low-frequency triplet (RTs for random low trials minus RTs for random high trials; ACC for random high trials minus ACC for random low trials). These trials 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. While Higher-order sequence learning quantifies the acquisition of the sequential pattern, Statistical learning captures purely frequency-based learning [1,7]. Based on previous findings, the cueing of pattern and random stimuli is necessary to promote Higher-order sequence learning, otherwise, it occurs more slowly, and cannot be acquired during a single session [1,4].

Additionally, more general changes in RT and ACC performance can be measured in the ASRT task. These changes occur similarly for all trial types, thus are not related to acquiring the sequential or statistical structure embedded in the stimulus stream. Instead,

these general changes indicate general skill improvements, such as more efficient visuo-motor and motor-motor coordination as the task progresses [8], combined with potential fatigue effects that can accumulate during practice [9,10]. General skill improvements in terms of RT are assessed as the difference of speed in the beginning and at the end of the task (RTs of the first five blocks minus RTs of the last five blocks, see also Statistical analysis). Similarly, general changes can be quantified in ACC between the beginning and the end of the task (ACC of the first five blocks minus ACC of the last five blocks).

In our study, we first report the Triplet learning results because this has been the most common analysis method in the ASRT studies and thus it enables to directly compare our results with those of previous studies. Next, we report Higher-order sequence learning and Statistical learning measures to obtain a more detailed picture of the underlying processes within procedural learning. Finally, we report average RTs and ACCs and their change from the beginning to the end of the end of the task to test whether these more general aspects of performance have a differential association pattern with sleep compared to the learning scores.

Working memory. The Counting Span task [11-14] was used to assess working memory (WM) performance. The task consisted of three series. In each series, each trial included three to nine blue circles as targets, one to nine blue squares, and one to five yellow circles as distractors on a grey background. Participants counted aloud the number of blue circles in each trial, and when finished with counting, they repeated the total number. When presented with a recall cue, participants recalled each total from the preceding set of trials, in the order in which they appeared. The number of presented trials (i.e., set) ranged from two to six. A participant's counting span capacity was calculated as the average of the highest set sizes of the three series at which the participant was able to recall the totals in the correct serial order.

Executive functions. The Wisconsin Card Sorting Test (WCST) [15,16] was used to assess executive functions. In this task, participants are asked to find out a sorting rule for cards based on the feedback they receive for their card-sorting choices. During the task, there are four decks on the screen with symbols on them, which differ in three features: number, shape, and color. On the bottom of the screen, a stimulus card appears, and participants are asked to match this card to one of the decks (based on a sorting rule of their choice). After the choice, participants receive a feedback whether the choice was

correct or not. Based on the feedback, participants have to find out the correct sorting rule. In each trial, only one sorting rule is correct (e.g., number, shape or color), and the rule changes several times during the task, allowing to measure adaptation to changing rules. The outcome measure of the task is the number of *perseverative errors*, which shows the inability to change the behavior despite feedback, so the higher values of this measure indicate weaker executive functions.

Subjective sleep quality assessments

Pittsburgh Sleep Quality Index. The Pittsburgh Sleep Quality Index (PSQI) [17,18] is one of the most commonly used questionnaires measuring self-reported sleep habits and sleep disturbances over the last month. Here we focused on three components of the questionnaire, which were obtained in all three studies: subjective sleep quality, sleep latency and sleep disturbances. Item 6 (referring to the original coding of PSQI) measured the participant's perceived sleep quality, item 5a indicated sleep latency and items 5b-5j (9 items) showed sleep disturbances. We chose the aforementioned items because the factors they form are those that contribute most to the overall PSQI score (besides daytime dysfunction which we didn't include, because we wanted to measure daytime functioning with the tests included). These three components range from 0 to 3 and form a global score that ranges between 0 and 9, a higher score indicating poorer sleep quality. Henceforth we refer to this 11-item long PSQI as PSQI.

Athens Insomnia Scale. The Athens Insomnia Scale (AIS) [19,20] was administered in all three studies. AIS is a self-reported questionnaire assessing general sleep quality (over a one month time period), and consists of eight items; the first five items assess difficulty with falling asleep, awakening during the night, early morning awakening, total sleep time, and overall quality of sleep, while the last 3 items pertain to the sense of well-being, overall functioning and sleepiness during the day. Each item of AIS can be rated from 0 to 3 and the total score ranges from 0 to 24, where higher scores indicate poorer sleep quality.

Groningen Sleep Quality Scale. In Study 2, subjective sleep quality of the night before cognitive testing was assessed by the Groningen Sleep Quality Scale (GSQS) [21,22], which is a 15-item self-administered questionnaire. Every item is a yes or no question,

scoring 0 or 1, thus GSQS scores range from 0 to 14 (the first item is typically not scored), a higher score indicating poorer quality of sleep.

Sleep diary. In Study 2, we also asked participants to keep a sleep diary for 1-2 weeks prior the testing session [23]. In this diary, participants had to mark the time they went to bed, the time they got up, and the hours they spent with sleep during this period. After each night, participants also had to rate how good their sleep was (on a scale from 1 to 5), report how long it took them to fall asleep (in minutes), and how many times they woke up during the night. We evaluated data from sleep diaries similarly to PSQI component scores. The average subjective sleep quality was scored as the first component of PSQI; the average sleep latency as the second component of PSQI, the average time spent with sleep as the third component of PSQI, and the average sleep time divided with the time spent in bed (i.e., sleep efficiency) as the fourth component of PSQI. Altogether, based on the sleep diary, we had 4 component scores, ranging from 0 to 3 and form a global score that ranges between 0 and 12, a higher score indicating poorer quality of sleep.

Chronotype assessment

Morningness-Eveningness Questionnaire. The Morningness–Eveningness Questionnaire [24] is a widely used and reliable scale assessing individual differences in morningness–eveningness. In the current study, we used the shortened 13 item Hungarian version [25] of this questionnaire. The items focus on subjective preferences of sleep–wake schedules, such as preferred rising and sleep times, peak times, morning freshness, as well as optimal time for intellectually or physically demanding activities. In the short Hungarian version, scores range between 10 and 59, with higher scores indicating greater morningness.

Statistical analysis

Analysis of the ASRT data. To facilitate data processing and to reduce intra-individual variability, the blocks of ASRT were collapsed into epochs of five blocks, following previous ASRT studies [1,14]. The first epoch contained blocks 1–5, the second epoch contained blocks 6–10, etc. We calculated mean accuracy (ACC) for all responses, and median reaction times (RTs) for correct responses only, separately for pattern,

random high and random low trials for each epoch. As in previous ASRT studies [5], two kinds of low-frequency triplets were eliminated: repetitions (e.g. 222, 333) and trills (e.g. 212, 343). Repetitions and trills are low-frequency for all participants, and people often show pre-existing response tendencies to them [26]. By eliminating these triplets, we attempted to ensure that differences between high- vs. low-frequency triplets emerged due to learning and not to pre-existing response tendencies.

Performance in the ASRT task was analyzed by repeated measures analyses of variance (ANOVA) with median RTs or mean ACCs as the outcome measure, and EPOCH (1^{st}-4^{th} epochs) and TRIAL TYPE (pattern, random-high, and random-low) as within-subject factors. To evaluate the effect of TRIAL TYPE, and thus to confirm that Higher-order sequence learning (pattern vs. random-high difference) and Statistical learning (random-high vs. random-low difference) occurred, Fisher's LSD post-hoc comparisons were performed. 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. Note that for the sake of brevity, we do not report ANOVAs for the mixed measure of Triplet learning (RT/ACC difference in responses to high- vs. low-frequency triplets), which does not clearly differentiate between acquiring the sequential and statistical structure embedded in the task. We conducted these ANOVAs, and Triplet learning occurred in all three studies, both in ACC and RT (significant main effect of TRIPLET: all *p*s < .001).

Supplementary results

Procedural learning across the three studies

Study 1. The repeated-measures ANOVA on RT data (Figure S2A) revealed a significant main effect of EPOCH ($F_{3,138} = 33.84$, $p < .0001$, $\eta_p^2 = .42$), such that RTs decreased as the learning progressed indicating general skill improvements. The main effect of TRIAL TYPE was significant as well ($F_{2,92} = 60.89$, $p < .001$, $\eta_p^2 = .57$). The post-hoc analysis revealed that responses (averaged across epochs) to pattern trials were faster (M = 340.27 ms) compared to random high (M = 348.70 ms, $p = .002$) and random low trials (M = 366.68 ms, $p < .0001$), and responses to random high trials were faster compared to random low trials ($p < .0001$). The different RTs for the different trial types indicate Higher-order sequence learning (difference in pattern vs. random high trials) and

Statistical learning (difference in random low vs. random high trials) occurred during the task. The interaction between EPOCH and TRIAL TYPE was not significant ($F_{6,276} = .89$, $p = .41$, $\eta_p^2 = .02$).

The repeated-measures ANOVA on accuracy data (Figure S2B) again revealed a significant main effect of EPOCH ($F_{3,138} = 13.08$, $p < .0001$, $\eta_p^2 = .22$): ACC decreased during the task. Similarly, to RTs, there was a significant main effect of TRIAL TYPE ($F_{2,92} = 51.06$, $p < .0001$, $\eta_p^2 = .53$). The post-hoc analysis revealed that responses (averaged across epochs) to pattern trials were more accurate (M = 96%) compared to random high (M = 95%, $p = .08$) and random low trials (M = 93%, $p < .0001$), and responses to random high trials were more accurate compared to random low trials ($p < .0001$). This again indicates that Higher-order sequence learning and Statistical learning occurred during the task. The EPOCH x TRIAL TYPE interaction was also significant ($F_{6,276} = 4.15$, $p = .001$, $\eta_p^2 = .08$), the ACC for random low trial type decreased more during the task (3.4% decrease on average) than ACC for pattern (1.2% decrease) or random high trial types (1% decrease).

Study 2. Similarly to Study 1, the ANOVA for RTs (Figure S2C) revealed a significant main effect of EPOCH ($F_{3,306} = 93.13$, $p < .0001$, $\eta_p^2 = .48$): RTs decreased as the learning progressed. We also found a significant main effect of TRIAL TYPE ($F_{2,204} = 63.39$, $p < .0001$, $\eta_p^2 = .38$). The post-hoc analysis revealed that responses (averaged across epochs) to pattern trials were faster (M = 323.01 ms) compared to random high (M = 351.55 ms, $p = .002$) and random low trials (M = 370.32, $p < .001$), and responses to random high trials were faster compared to random low trials ($p < .001$), indicating Higher-order sequence learning and Statistical learning occurred during the task. The EPOCH x TRIAL TYPE interaction was also significant ($F_{6,612} = 8.74$, $p < .001$, $\eta_p^2 = .08$), RTs for pattern trials decreased more (58 ms on average) than for random high (31 ms) and random low (28 ms) trials.

Again, for accuracy data (Figure S2D), there was a significant main effect of EPOCH ($F_{3,306} = 18.12$, $p < .0001$, $\eta_p^2 = .15$): ACC decreased during the task. There was also a significant main effect of TRIAL TYPE ($F_{2,204} = 97.55$, $p < .0001$, $\eta_p^2 = .49$). The post-hoc analysis revealed that responses (averaged across epochs) to pattern trials were more accurate (M = 96%) compared to random high (M = 95%, $p < .001$) and random low trials (M = 92%, $p < .001$), and responses to random high trials were more accurate compared

to random low trials ($p < .001$), indicating Higher-order sequence learning and Statistical learning occurred during the task. The EPOCH x TRIAL TYPE interaction was also significant ($F_{6,612} = 5.47$, $p = .0001$, $\eta_p^2 = .05$), the ACC for random low trial type decreased more during the task (3.3%) than ACC for pattern (0.7% decrease) or random high trial (1.3% decrease) types.

Study 3. As in the previous studies, the ANOVA for RTs (Figure S2E) revealed a significant main effect of EPOCH ($F_{3,252} = 45.60$, $p < .0001$, $\eta_p^2 = .35$): RTs decreased during the task. The main effect of TRIAL TYPE was significant as well ($F_{2,168} = 31.58$, $p < .0001$, $\eta_p^2 = .27$). The post-hoc analysis revealed that responses (averaged across epochs) to pattern trials were faster (M = 367.03 ms) compared to random high (M = 375.92 ms, $p = .008$) and random low trials (M = 388.93, $p < .0001$), and responses to random high trials were faster compared to random low trials ($p < .0001$). The different RTs for the different trial types again indicate that Higher-order sequence learning and Statistical learning occurred during the task. The EPOCH x TRIAL TYPE interaction was also significant ($F_{6,504} = .4.50$, $p < .001$, $\eta_p^2 = .05$), the RTs for pattern and random high trial types decreased more (27 ms and 22 ms respectively) than RTs for random low trials (13 ms).

Again, the ANOVA on accuracy data (Figure S2F) revealed a significant main effect of EPOCH ($F_{3,252} = 33.03$, $p < .0001$, $\eta_p^2 = .28$), such as ACC decreased during the task, and a significant main effect of TRIAL TYPE ($F_{2,168} = 78.97$, $p < .001$, $\eta_p^2 = .49$). The post-hoc analysis revealed that responses (averaged across epochs) to pattern trials were more accurate (M = 84%) compared to random high (M = 82%, $p < .0001$) and random low trials (M = 78%, $p < .0001$), and responses to random high trials were more accurate compared to random low trials ($p < .0001$) indicating that Higher-order sequence learning and Statistical learning occurred during the task. The EPOCH x TRIAL TYPE interaction was not significant ($F_{6,504} = .91$, $p = .47$, $\eta_p^2 = .01$).

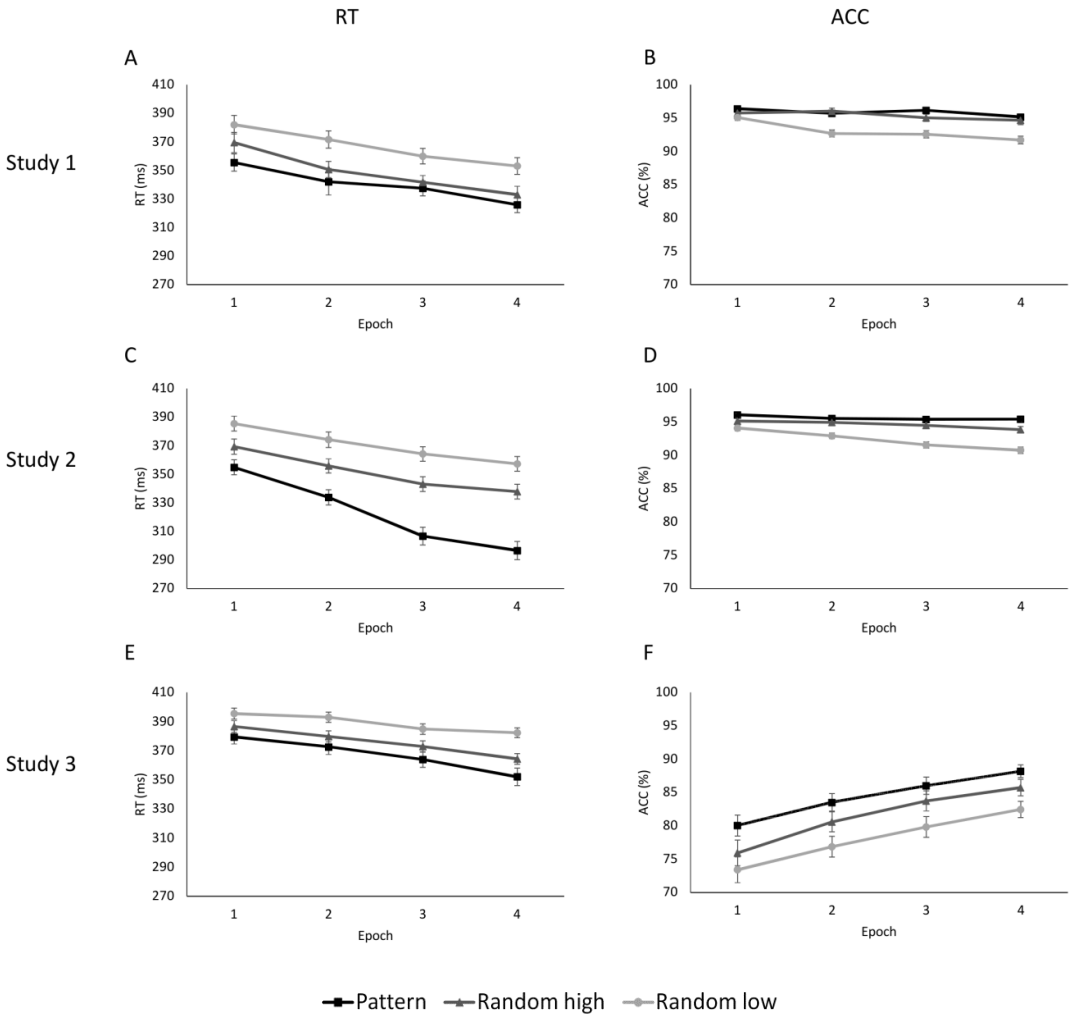

Figure S2. RT for correct responses (A, C, E) and accuracy for all responses (B, D, F) as a function of epoch (1-4) and trial type (pattern, random high- and low-frequency trials) in the ASRT task assessing procedural learning. The gap between the curves of pattern and random high-frequency trials indicates Higher-order sequence learning, the gap between the curves of random high and low-frequency indicates Statistical learning. Error bars denote standard error of mean.

Associations between subjective sleep quality and cognitive performance

To explore the associations between subjective sleep quality and cognitive performance, separate linear mixed-effect models were created for each outcome measure (i.e., cognitive performance metric), and PSQI or AIS was used as a fixed predictor, and 'Study' was added as a random intercept. To control for possible confounding effects, we included age, gender and morningness score as covariates. As the residuals were not normally distributed we used bootstrapped estimates and confidence intervals, using 1000

bootstrap samples, from which we calculated the p-values [27,28]. Bayes Factors (BF_{01}) were calculated by using the exponential of the Bayesian Information Criterion (BIC) of the fitted models minus the BIC of the null models – that contained the confounders only, and a random intercept by study [29]. Neither PSQI nor AIS showed an association with any of the cognitive performance metrics (all *p*s > .15, see Table S1 and S2, respectively). For PSQI, Bayes Factors ranged from 5.83 to 14.52, indicating substantial evidence for no association between subjective sleep quality and the measured cognitive processes [30].

Table S1. The association of Pittsburgh Sleep Quality Index with cognitive performance metrics

Outcome	*β*	95% CI	df	*p*	BF_{01}
ACC learning indices					
ACC Triplet learning	-.04	[-0.16, 0.08]	205	.48	11.48
ACC Higher-order sequence learning	-.03	[-0.17, 0.11]	205	.66	13.22
ACC Statistical learning	-.03	[-0.16, 0.10]	205	.61	12.90
RT learning indices					
RT Triplet learning	-.05	[-0.19, 0.10]	205	.50	11.72
RT Higher-order sequence learning	-.02	[-0.17, 0.12]	205	.76	13.90
RT Statistical learning	-.06	[-0.20, 0.09]	205	.42	10.87
General skill indices					
Average ACC	.03	[-0.08, 0.14]	205	.57	12.26
ACC general skill learning	.00	[-0.09, 0.10]	205	.94	14.52
RT average	-.05	[-0.19, 0.09]	205	.49	11.46
RT general skill learning	-.10	[-0.24, 0.05]	205	.16	5.83
WM and EF indices					
Counting Span	.01	[-0.12, 0.17]	205	.86	14.34
WCST – perseverative error	.08	[-0.06, 0.23]	199	.29	7.98

Note: The table shows standardized regression coefficients for PSQI, where the 'Study' random intercept was included in separate linear mixed-effect models for each cognitive performance metrics. Age, gender, and morningness score was added as covariates. BF_{01} was derived from BIC (for details, see the 'Data analysis' section in the main text). ACC = accuracy. RT = reaction time. WM = working memory. EF = executive function. WCST = Wisconsin Card Sorting Test.

For AIS, Bayes Factors ranged from 4.47 to 14.56, suggesting substantial evidence for no association between subjective sleep quality and the measured cognitive processes [30].

Table S2. The association of Athens Insomnia Scale with cognitive performance metrics

Outcome	β	95% CI	df	p	BF_{01}
ACC learning indices					
ACC Triplet learning	-.07	[-0.19, 0.04]	205	.25	7.95
ACC Higher-order sequence learning	-.04	[-0.18, 0.10]	205	.57	12.35
ACC Statistical learning	-.03	[-0.17, 0.10]	205	.65	12.96
RT learning indices					
RT Triplet learning	.00	[-0.14, 0.14]	205	1.00	14.56
RT Higher-order sequence learning	.04	[-0.09, 0.18]	205	.52	12.01
RT Statistical learning	-.05	[-0.18, 0.10]	205	.52	11.96
General skill indices					
Average ACC	.08	[-0.02, 0.19]	205	.15	4.47
ACC general skill learning	.06	[-0.04, 0.15]	205	.23	7.21
RT average	.01	[-0.12, 0.15]	205	.84	14.24
RT general skill learning	-.03	[-0.17, 0.11]	205	.68	13.47
WM and EF indices					
Counting Span	-.03	[-0.18, 0.10]	205	.65	13.13
WCST – perseverative error	.10	[-0.04, 0.26]	199	.17	5.07

Note: The table shows standardized regression coefficients for AIS, where the 'Study' random intercept was included in separate linear mixed-effect models for each cognitive performance metrics. Age, gender, and morningness score was added as covariates. BF_{01} was derived from BIC (for details, see the 'Data analysis' section in the main text). ACC = accuracy. RT = reaction time. WM = working memory. EF = executive function. WCST = Wisconsin Card Sorting Test.

Cognitive performance in subjective sleep quality extremes

We also tested whether cognitive performance differed between "good" and "poor" sleepers as defined by the extremes in the overall PSQI score. As reported in the main text, we considered those with a score of 0 or 1 as good sleepers (N = 36), while

those with a score of 5 to 8 as poor sleepers (N = 43), corresponding to approximately the upper and lower 15% of the data. As the cognitive performance metrics were not normally distributed, we compared the two groups' performance using robust frequentist as well as Bayesian Mann-Whitney U tests (see Table S3, U and p values were obtained from the frequentist, while BF_{01} values were obtained from the Bayesian Mann-Whitney tests).

We did not find a significant difference between good and poor sleepers in any of the cognitive performance metrics (all ps > .10, see Table S3). Bayes Factors indicated substantial evidence for no difference in Statistical learning (both RT and ACC measures), average ACC, ACC general skill learning, Counting Span and Perseverative error (WCST). Furthermore, Bayes Factors indicated anecdotal evidence for no difference in Higher-order sequence learning (both ACC and RT measures), and Triplet learning (both ACC and RT measures). The Bayes Factor for general skill learning in RT remained inconclusive.

Table S3. Cognitive performance in PSQI extremes

Outcome	***U***	***p***	**BF_{01}**
ACC learning indices			
ACC Higher-order sequence learning	639	.18	1.84
ACC Statistical learning	730	.67	3.48
ACC Triplet learning	652	.23	2.26
RT learning indices			
RT Higher-order sequence learning	679	.35	2.72
RT Statistical learning	696	.45	3.11
RT Triplet learning	611	.11	1.53
General skill indices			
ACC general skill learning	747	.79	4.21
Average ACC	715	.56	3.92
RT average	656	.80	2.38
RT general skill learning	604	.10	1.03
WM and EF indices			
Counting Span	758	.87	4.19

WCST – perseverative error	760	.89	4.23

Note: ACC = accuracy. RT = reaction time. WM = working memory. EF = executive function. WCST = Wisconsin Card Sorting Test.

References

1 Nemeth, D., Janacsek, K. & Fiser, J. Age-dependent and coordinated shift in performance between implicit and explicit skill learning. *Front. Comput. Neurosci.* **7**, doi:10.3389/fncom.2013.00147 (2013).

2 Kóbor, A., Janacsek, K., Takács, Á. & Nemeth, D. Statistical learning leads to persistent memory: Evidence for one-year consolidation. *Sci. Rep.* **7**, 760, doi:10.1038/s41598-017-00807-3 (2017).

3 Destrebecqz, A. & Cleeremans, A. in *Attention and Implicit Learning* (ed L Jiménez) 181-213 (John Benjamins, 2003).

4 Howard, J. H., Jr. & Howard, D. V. Age differences in implicit learning of higher-order dependencies in serial patterns. *Psychol. Aging* **12**, 634-656 (1997).

5 Nemeth, D. *et al.* Sleep has no critical role in implicit motor sequence learning in young and old adults. *Exp. Brain Res.* **201**, 351-358, doi:10.1007/s00221-009-2024-x (2010).

6 Song, S., Howard, J. H., Jr. & Howard, D. V. Sleep does not benefit probabilistic motor sequence learning. *J. Neurosci.* **27**, 12475-12483, doi:10.1523/jneurosci.2062-07.2007 (2007).

7 Simor, P. *et al.* Deconstructing procedural memory: Different learning trajectories and consolidation of sequence and statistical learning. *Front. Psychol.* **9**, 2708 (2019).

8 Hallgato, E., Győri-Dani, D., Pekár, J., Janacsek, K. & Nemeth, D. The differential consolidation of perceptual and motor learning in skill acquisition. *Cortex* **49**, 1073-1081 (2013).

9 Siegelman, N., Bogaerts, L., Christiansen, M. H. & Frost, R. Towards a theory of individual differences in statistical learning. *Phil. Trans. R. Soc. B* **372**, 20160059 (2017).

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

11 Case, R., Kurland, D. M. & Goldberg, J. Operational efficiency and the growth of short-term memory span. *J. Exp. Child Psychol.* **33**, 386-404 (1982).

12 Engle, R. W., Tuholski, S. W., Laughlin, J. E. & Conway, A. R. A. Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *J. Exp. Psychol.* **128**, 309-331 (1999).

13 Conway, A. R. *et al.* Working memory span tasks: A methodological review and user's guide. *Psychon. Bull. Rev.* **12**, 769-786 (2005).

14 Virag, M. *et al.* Competition between frontal lobe functions and implicit sequence learning: evidence from the long-term effects of alcohol. *Exp. Brain Res.* **233**, 2081-2089 (2015).

15 Piper, B. J. *et al.* Reliability and validity of neurobehavioral function on the Psychology Experimental Building Language test battery in young adults. *PeerJ* **3**, e1460 (2015).

16 Berg, E. A. A simple objective treatment for measuring flexibility in thinking. *J. Gen. Psychol.* **39**, 15-22 (1948).

17 Buysse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R. & Kupfer, D. J. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res.* **28**, 193-213 (1989).

18 Takács, J. *et al.* Reliability and validity of the Hungarian version of the Pittsburgh Sleep Quality Index (PSQI-HUN): comparing psychiatric patients with control subjects. *Sleep. Breath.* **20**, 1045-1051 (2016).

19 Diekelmann, S., Biggel, S., Rasch, B. & Born, J. Offline consolidation of memory varies with time in slow wave sleep and can be accelerated by cuing memory reactivations. *Neurobiol. Learn. Mem.* **98**, 103-111 (2012).

20 Novak, M., Mucsi, I., Shapiro, C. M., Rethelyi, J. & Kopp, M. S. Increased utilization of health services by insomniacs—an epidemiological perspective. *J. Psychosom. Res.* **56**, 527-536 (2004).

21 Meijman, T., de Vries-Griever, A., De Vries, G. & Kampman, R. The evaluation of the Groningen sleep quality scale. *Groningen: Heymans Bulletin (HB 88-13-EX)* **2006** (1988).

22 Simor, P., Köteles, F., Bódizs, R. & Bárdos, G. A questionnaire based study of subjective sleep quality: the psychometric evaluation of the Hungarian version of the Groningen Sleep Quality Scale. *Mentálhigiéné és Pszichoszomatika* **10**, 249-261 (2009).

23 Gilson, M. *et al.* REM-enriched naps are associated with memory consolidation for sad stories and enhance mood-related reactivity. *Brain Sci.* **6**, 1 (2015).

24 Horne, J. A. & Östberg, O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *Int. J. Chronobiol.* (1976).

25 Zavecz, Z., Török, C., Köteles, F., Pálosi, V. & Simor, P. The psychometric properties of the Hungarian version of the Morningness-Eveningness Questionnaire (MEQ-H): The separate factors of morning freshness and circadian rhythmicity. *Psychiatria Hungarica: A Magyar Pszichiátriai Társaság Tudományos Folyóirata* **30**, 318-331 (2015).

26 Howard, D. V. *et al.* Implicit sequence learning: effects of level of structure, adult age, and extended practice. *Psychol. Aging* **19**, 79-92 (2004).

27 Canty, A. & Ripley, B. boot: Bootstrap R (S-Plus) Functions. (2019).

28 Davison, A. C. & Hinkley, D. V. *Bootstrap methods and their application*. Vol. 1 (Cambridge university press, 1997).

29 Wagenmakers, E. J. A practical solution to the pervasive problems of p values. *Psychon Bull Rev* **14**, 779-804, doi:10.3758/BF03194105 (2007).

30 Wagenmakers, E. J., Wetzels, R., Borsboom, D. & van der Maas, H. L. Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). *J. Pers. Soc. Psychol.* **100**, 426-432, doi:10.1037/a0022790 (2011).