

Pitfalls in sleep and memory research and how to avoid them: A consensus paper

Dezso Nemeth^{1,2,3*}, Emilie Gerbier^{4*}, Jan Born⁵, Timothy Rickard⁶, Genevieve Albouy⁷,
Susanne Diekelmann⁵, Stuart Fogel⁸, Lisa Genzel⁹, Alexander Prehn-Kristensen¹¹, Jessica Payne¹⁰,
Martin Dresler¹², Peter Simor^{2,13}, Stephanie Mazza¹⁴, Kerstin Hoedlmoser¹⁵, Perrine Ruby¹,
Rebecca M. C. Spencer¹⁶, Manuel Schabus¹⁵, Karolina Janacsek^{17,2*}

¹ Lyon Neuroscience Research Center (CRNL), INSERM U1028, CNRS UMR5292, Université Claude Bernard Lyon 1, Université de Lyon, Lyon, France

² Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary

³ Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Budapest, Hungary

⁴ Université Côte d'Azur, CNRS, BCL, Nice, France

⁵ Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Tübingen, Germany

⁶ University of California, San Diego, USA

⁷ KU Leuven, Department of Movement Sciences, Movement Control and Neuroplasticity Research Group, Leuven, Belgium

⁸ School of Psychology, University of Ottawa, Ottawa, Canada

⁹ Donders Institute for Brain Cognition and Behaviour, Radboud University, Nijmegen/Netherlands

¹⁰ University of Notre Dame, Department of Psychology, Notre Dame, USA

¹¹ Department of Child and Adolescent Psychiatry and Psychotherapy, Centre for Integrative Psychiatry, University Hospital Schleswig-Holstein, Kiel, Germany

¹² Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands

¹³ Université Libre de Bruxelles, Brussels, Belgium

¹⁴ Research on Healthcare Performance lab, INSERM U1290 RESHAPE, Université Lyon 1, Lyon, France

¹⁵ Department of Psychology, Centre for Cognitive Neuroscience, University of Salzburg, Salzburg, Austria

¹⁶ Department of Psychological and Brain Sciences, University of Massachusetts, Amherst, USA

¹⁷ Centre for Thinking and Learning, Institute for Lifecourse Development, School of Human Sciences, Faculty of Education, Health and Human Sciences, University of Greenwich, London, United Kingdom

* These authors contributed equally to this work.

Author Note: Correspondence and requests for material should be addressed to Prof. Dezso Nemeth, CRNL – Lyon Neuroscience Research Center, Université Claude Bernard Lyon 1, Lyon, France. Address: CRNL – Centre Hospitalier Le Vinatier – Bâtiment 462 – 95 Bd Pinel – 69675 Bron Cedex – France. Phone: +33 4 81 10 65 46.

Abstract

Understanding the complex relationship between sleep and memory is a major challenge in neuroscience. Thousands of studies on memory consolidation in humans suggest that sleep triggers offline memory processes, resulting in less forgetting of declarative memory and performance stabilization in non-declarative memory. However, an increasing number of contradictory findings reveal potential issues with how research is conducted in this field and call into question the reliability and interpretation of the results. In this consensus paper, we describe four sets of prevalent methodological pitfalls in human sleep and memory research: (i) non-optimal experimental designs, (ii) task complexity, (iii) fatigue effects in repetitive tasks, and (iv) inappropriate data analysis practices. We then offer solutions to each of these pitfalls. We believe that implementing these solutions in future research of sleep and memory will lead to more reliable results and significantly advance our understanding in this field.

Keywords: long-term memory, declarative memory, procedural memory, sleep, consolidation, fatigue effect, circadian effect, napping, data analysis recommendations, open science

Introduction

There is a great interest in sleep in both the general public and the scientific community. The critical influence of sleep on health and some aspects of cognition is well established; moreover, our modern lifestyle and technologies affect our sleep habits and quality in new ways every day, increasing the prevalence of sleep deprivation and bad sleep habits. Memory is also a focus of societal interest, with respect to education and learning on one side of the developmental spectrum, and to aging and age-related memory decline on the other. As such, the effect of sleep on memory has gained much attention in psychology and neuroscience research over the last two decades, with thousands of dedicated publications. Additionally, a number of theories and models explaining the effect of sleep on memory have been developed (e.g., Ackermann & Rasch, 2014; Antony et al., 2019; Boyce et al., 2017; Diekelmann & Born, 2010; Diekelmann et al., 2009; Feld & Born, 2017; Lewis & Durrant, 2011; Mednick et al., 2011; Saletin & Walker, 2012; Siegel, 2001; Stickgold & Walker, 2005, 2013; Tononi & Cirelli, 2006; Tononi & Cirelli, 2014; Walker, 2005). The focus of this paper is on the effect of sleep on memory consolidation, that is, on how sleeping after having learned something (e.g., new vocabulary or playing the piano) benefits subsequent memory, compared to an equivalent time spent without sleep.

According to the oft-cited empirical studies (e.g., Gais et al., 2006; Walker et al., 2003) and reviews on this topic (Diekelmann et al., 2009; King et al., 2017; Rasch & Born, 2013), in healthy adults declarative memory appears more resistant to forgetting when encoding is followed by a period of sleep compared to a period of wakefulness, whereas non-declarative memory performance can even be improved when sleep follows training. The evidence for sleep-related memory consolidation appears to be so convincing that it has been claimed that “*While memory formation is not the only function of sleep, it seems to be the most important (...)*” (Born & Wilhelm, 2012, p. 192) or that “*(...) active system consolidation might be an evolutionary conserved function of sleep.*” (Vorster & Born, 2015, p. 103). While such judgements may be justified if based on a large body of

rather heterogeneous experimental approaches, we claim here that the support such statements receive from individual experiments is not yet compelling, due to a multitude of experimental and methodological issues that pertain. Indeed, there has been critical discussion as to the actual impact of sleep on memory consolidation (e.g., Mantua, 2018; Pan & Rickard, 2015; Vertes & Siegel, 2005). A non-negligible number of studies have not found sleep-related consolidation effects, especially for non-declarative memory (e.g., Csabi et al., 2014; Nemeth et al., 2010; Robertson et al., 2004; Song et al., 2007; Viczko et al., 2018; Wilson et al., 2012). There is also increasing evidence that the effect of sleep on memory consolidation is perhaps more multifaceted than initially thought (e.g., King et al., 2017). In addition, the correlations observed between sleep physiology and memory consolidation are often not replicated across studies and are sometimes too numerous to reliably interpret the significant ones (Pan & Rickard, 2015). They sometimes even go in the opposite direction than what is expected (Mantua, 2018; Payne et al., 2009). Contradictory findings in the field are not an issue *per se*, as they can highlight the complexity of the effect of sleep on memory consolidation. Those findings become an issue, however, if they point to systematic problems with the literature. Thus, providing guidelines for future studies is crucial to maintaining progress in the field.

In this consensus paper, we propose a guideline for future research on sleep and memory. Such a comprehensive guideline is lacking so far (but see King et al., 2017), hindering the progress toward a better understanding of the effect of sleep on memory in fields ranging from psychology to biology and neuroscience. We highlight four sets of critical methodological pitfalls that could be responsible for some of the contradictory findings in the literature, and then propose solutions to prevent them and guide future research.

Pitfall 1: Non-optimal experimental designs

In this section, we identify five areas that could benefit from improvements in the experimental designs and suggest solutions for each of them. Figure 1 illustrates the main study

conditions that can be included in the experimental designs testing the effect of sleep on memory consolidation. It is important to note from the outset that it is difficult to address all of these methodological caveats in one parsimonious experimental design. Thus, several types of studies may be necessary to draw strong conclusions about the function of a particular type of sleep for a particular type of memory (Peigneux & Smith, 2010).

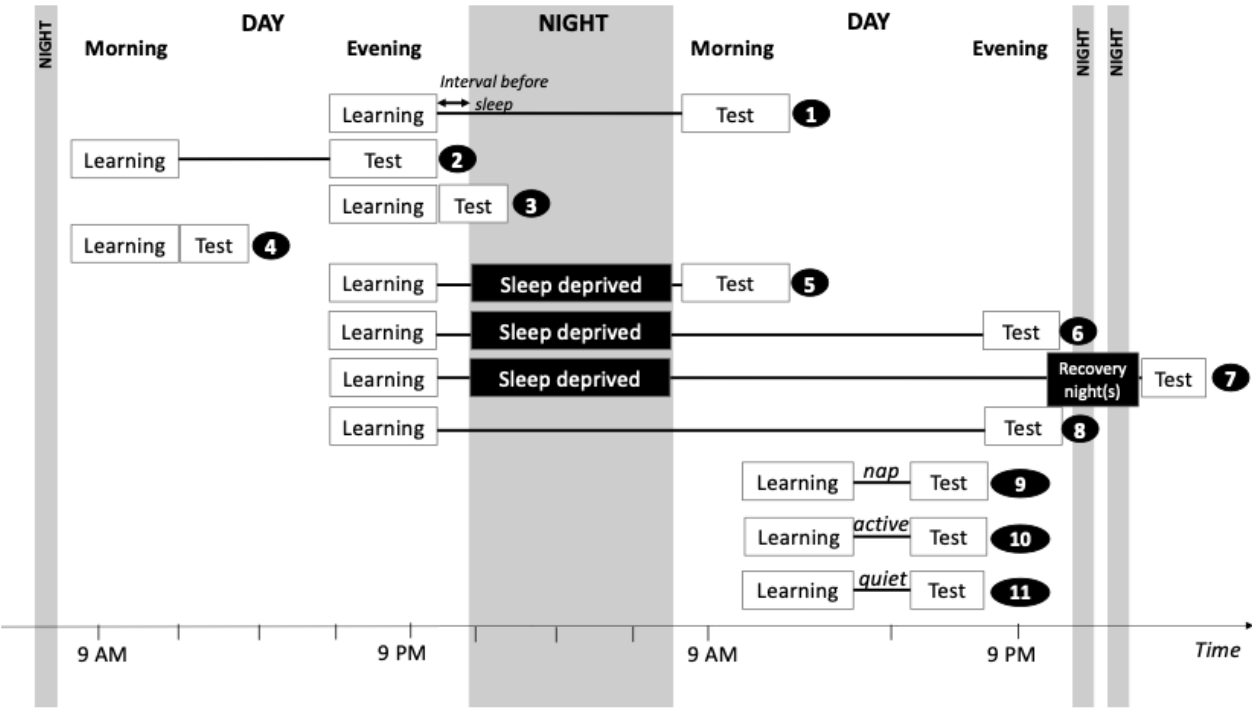


Figure 1. Illustration of 11 different study conditions for testing the effect of sleep on memory consolidation (for further details, see Pitfall 1).

The issues with non-optimal experimental designs can be organized into three sets. The first one is related to the influence of the time of day when the tasks are performed (see section 1a). The second is about whether the observed benefit of sleep over wake intervals is just due to the fact that, during sleep, interference is much diminished (see sections 1b, 1c, and 1d). This issue is related to the fundamental question whether sleep passively (through reduced interference) or actively

(through sleep-specific neural processes) contributes to memory consolidation (for in depth discussion see Ellenbogen et al., 2006). If sleep has only a passive role in this regard, then managing to create wake intervals with reduced external interferences during the post-learning interval may be sufficient to trigger a level of consolidation comparable to that of a sleep interval. Finally, we also briefly discuss the effect of baseline measurements and feedback on performance changes in the retention interval as differences in these aspects of the experimental design can further confound the observed relationship between sleep and memory (see section 1e).

Pitfall 1a. Time-of-day (circadian) effects

It has been shown that learning and memory performance is affected by the time of day when the task is performed (Schmidt et al., 2007). Time-of-day (circadian) effects can lead to confounds in sleep-related consolidation studies. A typical approach in these studies is to compare performance change in a Sleep condition (i.e., learning in the evening and testing memory the next morning; Figure 1, condition 1) with that in a Wake condition (i.e., learning in the morning and testing memory the next evening; condition 2). Importantly, however, a greater off-line improvement in a Sleep condition compared to a Wake condition in such a design may be, at least partially, explained by two confounds: 1) worse performance in the evening (i.e., when learning takes place in the Sleep condition) due to circadian effects, including day-long buildup of fatigue (Keisler et al., 2007), and/or 2) a better performance in the morning (i.e., when testing takes place in the Sleep condition) when participants are likely well rested. Probing circadian effects, Pan and Rickard (2015) have shown in their meta-analysis of sleep-related motor memory consolidation that performance is best if the test session occurs in the early afternoon. Notably, the issue of circadian effects can be even more pronounced over the course of the human lifespan, especially when comparing young vs. older groups of individuals with differences in chronotype and homeostatic sleep pressure (Taillard et al., 2021). In either case, it is unclear how much sleep *per se* contributes to improved performance compared to these circadian effects.

Moreover, hormones like cortisol strongly influence memory processes, and their release is subject to distinct circadian rhythms. For instance, growth hormone (GH) has its peak in the first half of the night, whereas GH concentrations in the second half are very low. By contrast, cortisol has its daily nadir in the first half of the night and its peak in the second half (Dresler et al., 2014). This confounds not only evening-morning vs. morning-evening (Figure 1, conditions 1 vs. 2, respectively) comparisons but also within-night comparisons that are permitted by the classical split-night paradigm, which compares the SWS-rich first half of the night with the REM sleep-rich second half of the night (Genzel & Robertson, 2015). Thus, the confounds of endocrine fluctuations across the circadian rhythm are difficult to avoid, rendering the split-night paradigm problematic.

Solutions to control for time-of-day (circadian) effects

These potential circadian effects are not easy to control because, in humans, sleep typically occurs during the night (although see Tucker et al., 2017 for an inverted 12-hour schedule with sleep occurring during daytime). Nevertheless, several solutions have been proposed to address this potential caveat. None are ideal nor exhaustive, but a combination of converging results across studies employing these different strategies can accrue confidence in the conclusions.

When focusing on the effects of nighttime sleep on memory consolidation, e.g., in an evening-morning (Sleep) vs. morning-evening (Wake) design (Figure 1, conditions 1 vs. 2, respectively), additional control groups should be tested to disentangle the circadian effects from the effect of sleep *per se*. Thus, a full design would include evening-morning (i.e., 12h Sleep; condition 1) and morning-evening (i.e., 12h Wake; condition 2) conditions together with evening-alone (condition 3) and morning-alone (condition 4) conditions, the latter ones with immediate testing. A difference in *learning* performance (i.e., how long it takes to learn or the overall performance during training) and/or in the immediate *retrieval* performance when learning/immediate testing takes place during the evening vs. in the morning would indicate potential circadian effects and preclude

further interpretation about the effect of sleep on consolidation (e.g., Fenn et al., 2003; Hallgato et al., 2013; Talamini et al., 2008; Tucker et al., 2011). In such a design, only the following pattern of results would demonstrate a beneficial effect of sleep without any time-of-day effect: 1) an absence of difference in learning performance in the four conditions; 2) an absence of difference in testing performance between the morning-alone and the evening-alone conditions; and 3) a better retrieval in the evening-morning over the morning-evening condition.

Another possible solution is to include sleep-deprived wake control groups in evening-morning (12h Sleep-deprived) or evening-evening (24h Sleep-deprived) conditions (see Figure 1, conditions 5 vs. 6, respectively) and compare their performance with that of an evening-morning (12h Sleep; condition 1) group. In these sleep deprivation controls, learning takes place in the evening and testing takes place either in the morning or in the next evening, with subjects staying awake during the night (in the evening-morning condition) or during both the night and the following day (in the evening-evening condition). In both cases, testing can take place either after sleep deprivation, with subjects being acutely sleep deprived at testing, or testing can be delayed by another 24-48 hours to allow for one or two nights of recovery sleep (condition 7), with the sleep conditions likewise being tested after comparable delays. The advantage of sleep deprivation control designs is that learning and testing take place at the same time of day in the 12h-Sleep and 12h Sleep-deprived groups, thus controlling for potential circadian differences that are an issue with the typical morning-evening (12h Wake; condition 2) controls. The same holds for the comparison of evening-evening groups, where one group could sleep (24h Sleep), while the other stayed awake (24h Sleep-deprived) in the delay period (see Figure 1 conditions 8 vs. 6, respectively). Including recovery sleep (condition 7) ensures that participants are not acutely sleep-deprived at testing, which reduces the negative impact of sleep deprivation on test performance. However, including recovery sleep comes at the price of extending the retention interval, which may likewise affect test performance through processes of decay, forgetting or memory restructuring. Both immediate or

delayed testing solutions are limited because acute sleep deprivation introduces confounding influences on test performance, and recovery sleep may exert additional confounds with regard to longer retention intervals and potential compensatory effects of recovery sleep on memory consolidation, or confounds related to sleep rebound effects. That is, sleep during the recovery night(s) may compensate for the missed opportunity for sleep consolidation during the first night of sleep deprivation, thus masking the original effect on memory consolidation. Only few studies applied such sleep deprivation designs. For instance, Gais, Rasch, Dahmen, Sara, and Born (2011) examined the role of noradrenaline for sleep-dependent memory consolidation by pharmacologically blocking noradrenaline via clonidine administration in an evening-evening sleep vs. sleep deprivation design without recovery sleep (Figure 1 conditions 8 vs. 6, respectively). They observed impaired memory retention after clonidine administration compared to placebo in the sleep condition but no difference between clonidine and placebo in the sleep-deprived wake condition, suggesting that noradrenaline supports memory consolidation specifically during sleep but not during wakefulness.

A third possible solution is to focus on the effect of daytime sleep (i.e., napping) on learning and memory performance, as in this case training and testing occur at the same time in the nap and awake groups (Figure 1 conditions 9 vs. 10, respectively) (Mednick et al., 2003) (see also section 1c). Note, however, that daytime and nighttime sleep can affect memory differently (Payne et al., 2015), and daytime naps might largely vary across participants with regards to the duration, depth and composition of sleep (e.g., appearance of REM stage in some, but not all participants), which should be taken into account during analysis and interpretation of these studies.

Finally, beyond the inclusion of additional groups/conditions to control for circadian effects, further questionnaires/tasks should also be used to assess subjective sleepiness (e.g., Stanford Sleepiness Scale) and objective vigilance (e.g., Psychomotor Vigilance Test) before and after the encoding and testing sessions. These assessments could provide further useful information about the

subjective states and vigilance of participants, and could be included in the analyses as covariates, or extreme values could be used for participant exclusion. It is important to note, however, that questionnaires/tasks with good psychometric properties should be selected for this purpose, and a null effect with these assessments *alone* (i.e., without the inclusion of control groups/conditions discussed above) should not be used to dismiss an alternative explanation entirely since null effects could be due to low statistical power (see also Pitfall 4).

Pitfall 1b. No control conditions in overnight studies

Studies investigating the effect of sleep on memory consolidation sometimes compare experimental conditions across sleep intervals only. For example, studies in clinical populations with patients suffering from sleep disorders (e.g., primary insomnia, obstructive sleep apnea, sleep-disordered breathing) often compare a pathological group with a control healthy group and often use *only* an evening-morning sleep condition (Figure 1, condition 1) (Backhaus et al., 2006; Csabi et al., 2014) (for reviews, see Ahuja et al., 2018; Cellini, 2017). Importantly, however, if patients show a smaller overnight performance benefit compared to control participants, one cannot disentangle whether it is caused by the specific effect of that overnight sleep (i.e., state-dependent consolidation) or by a trait-dependent effect of sleep disturbances on memory processes involving not only consolidation, but perhaps also encoding and retrieval or even other cognitive limitations (Ahuja et al., 2018; Csábi et al., 2013; Rosenzweig et al., 2015; Wallace & Bucks, 2013). An additional issue is that pathologies can also influence circadian cycles (Wulff et al., 2010), thus time of peak performance may be shifted in such populations. Many of these issues need to be likewise considered in aging populations, as aging impacts the prevalence of sleep disorders, cognitive performance, daytime functioning, and mnemonic functioning.

Similarly, studies investigating sleep interventions (e.g., using pharmacological agents or electrical stimulation) typically only compare sleep conditions with vs. without intervention in an evening-morning design (Figure 1, condition 1). However, with this design, it cannot be ascertained

that any observed effects are sleep-specific or whether the intervention exerts general effects that are independent of sleep.

Solutions: using appropriate control conditions in overnight studies

To assess the specificity of sleep-related consolidation and sleep interventions, it is essential to include appropriate control groups/conditions in which participants stay awake for a comparable period and, in the case of intervention studies, also receive the same experimental manipulations as in the sleep groups/conditions. There are two main classes of wake controls in overnight studies: morning-evening (wake) controls (Figure 1, condition 2), and evening-morning (condition 1) or evening-evening (sleep deprived) controls (condition 6) (see section 1a).

For example, to demonstrate that a pathology (or an intervention) *specifically* affects sleep-related consolidation, we need to observe not only that test performance in the sleep condition is different in the pathological (or intervention) group than in the control group, but also that test performance or improvement in the wake control condition is similar in both groups. In other words, there should be a group difference in the sleep condition, but no group difference in the wake condition (i.e., a group-by-condition interaction; see Figure 2, panel a). However, if test performance in the sleep condition is similar in the pathological/intervention group and in the control group, then one cannot conclude that the pathology/intervention affects sleep-related consolidation processes (i.e., no interaction; Panel c), unless there is also significant group difference in the wake condition (i.e., an interaction; Panel b). Finally, if test performance in the pathological/intervention group is different from that in the control group *both* in the Sleep and in the Wake conditions without a group-by-condition interaction (i.e., no interaction; Panel d), then one would conclude that the pathology or the intervention has a more general effect, possibly influencing other processes (such as encoding or retrieval) but not sleep-related consolidation *per se*.

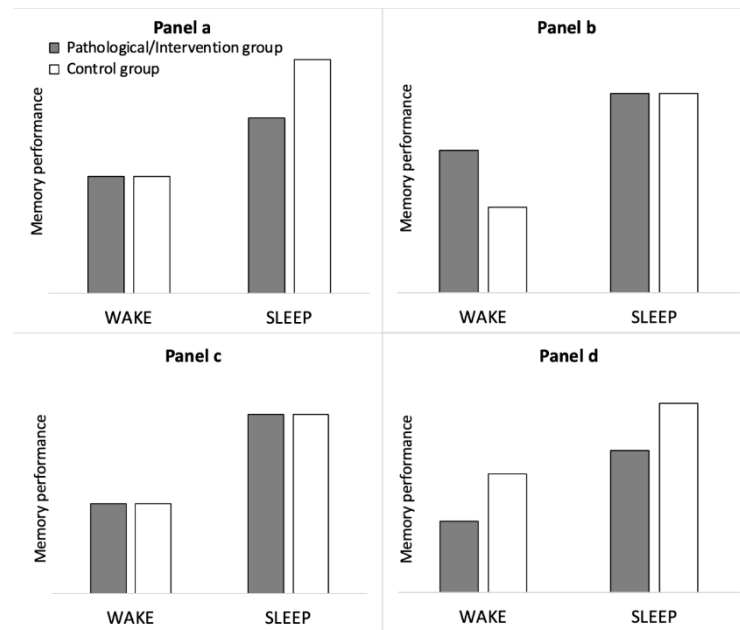


Figure 2. Schematic illustration of the different patterns of results and their conclusion in the case of studies comparing a pathological or intervention group vs. a control group in Wake vs. Sleep conditions. Vertical axis represents memory performance. Patterns depicted in Panels a and b display cases where it can be concluded that the pathology or intervention specifically affects sleep-related memory consolidation, compared to a control group (i.e., group-by-condition interaction). Patterns depicted in Panels c and d display cases where it cannot be concluded that the pathology or intervention specifically affects sleep-related memory consolidation, compared to a control group (i.e., no group-by-condition interaction). Note that in panels b and d, pathology and intervention may have opposite effects (i.e., hindering or improving performance, respectively) compared to the control group, without changing the overall logic of the figure.

Such a design has been used in a handful of studies only. For instance, Nissen et al. (2011) compared the effect of evening-morning (sleep) vs. morning-evening (wake) conditions (Figure 1, conditions 1 vs. 2) on consolidation of non-declarative/procedural and declarative memory in insomniac and healthy control participants. For procedural memory, they observed similar retention over the morning-evening (wake) interval in both groups (condition 2). However, the healthy control group showed better retention over the evening-morning (sleep) interval (condition 1) compared to the insomniac group. This corresponds to Figure 2, panel A. The authors concluded that insomnia specifically impaired sleep-related consolidation in procedural memory. Since declarative memory retention did not differ significantly between the two groups either in the wake or in the sleep condition (although it showed the same overall pattern as in Figure 2, panel A), no

conclusion about the effect of insomnia on sleep-related consolidation of declarative memory could have been drawn in this case.

Pitfall 1c. No or inappropriate control conditions in napping studies

Napping studies typically compare performance changes following an interval that includes daytime sleep with an interval that includes awake activities at the same time of day (Figure 1, conditions 9 vs. 10). Therefore, they allow to control for circadian effects (see section 1a). The duration of naps varies greatly across studies, ranging from 6 minutes (Lahl et al., 2008) and 40-90 minutes (Diekelmann et al., 2012; Nishida & Walker, 2007; Payne et al., 2015) to 3 hours (Gais et al., 2000; Plihal & Born, 1997; Schönauer et al., 2014). Most nap studies take place around noon (Mednick et al., 2008; Nishida & Walker, 2007; Payne et al., 2015), while others are conducted in the early morning (Sopp et al., 2018) or at night (Diekelmann et al., 2012). Split-night designs have also been applied, with 3-hour sleep periods either during the first (SWS-rich) or second (REM-rich) half of the night (Gais et al., 2000; Plihal et al., 1996; Yaroush et al., 1971).

Importantly, as described in section 1a, naps differ significantly with respect to the composition of sleep stages and hormonal concentrations depending on the time of day and the duration of the nap. Moreover, including wake controls at the same time of day and of the same duration does not rule out the possibility that other factors than sleep *per se* affect memory consolidation. For example, the general concern that reducing external interferences during the post-learning interval may be sufficient to aid off-line consolidation also pertains to nap studies (e.g., Mednick et al., 2011; Wamsley, 2019).

Solutions: using appropriate control conditions in napping studies

First, depending on the aim of the study, it should be carefully considered at which time of day the nap is scheduled and for how long participants are allowed to nap. Attending to these

variables allows, for example, comparing naps with and without REM sleep (e.g., McDevitt et al., 2015; Mednick et al., 2003).

Second, adding a carefully controlled quiet rest condition (e.g., keeping eyes open to avoid falling asleep; condition 11) helps to distinguish whether sleep is a specific state that actively triggers off-line memory improvement (i.e., Sleep > Quiet Rest) or a non-specific state that only passively protects memories from interference (i.e., Sleep = Quiet Rest) (Ellenbogen et al., 2006). The benefits of such a design have been just recently recognized. Studies using such a design have led to mixed findings, with some studies showing better consolidation in the nap condition than in the quiet rest condition (Piosczyk et al., 2013; Schichl et al., 2011; Schönauer et al., 2014) and others showing that quiet rest produced effects on memory consolidation similar to those observed in nap conditions (Mednick et al., 2009; Simor, Zavecz, et al., 2019), suggesting that sleep *per se* may not be necessary for consolidation but rather only provides a favorable environment. Observations that memory reactivation occurs not only during sleep but also during quiet rest (e.g., Schapiro et al., 2018) further highlight the need for such control conditions. Therefore, monitoring the wake (quiet rest) condition with polysomnography is essential to rule out any sleep-like activity during the quiet rest interval, as well as to examine whether polysomnographic indicators during wake (quiet rest) are specifically associated with memory consolidation (see also section 4b).

Note that a quiet rest condition would also be informative in a classic overnight design (Figure 1, condition 1) that employs control conditions, such as morning-evening (wake) or evening-morning with sleep deprivation (Figure 1 conditions 2 vs. 5, respectively; see also section 1a). However, this solution is not feasible since it would be extremely difficult to stay in quiet wakefulness for 12 hours in the morning-evening condition and it would be even more difficult (and stressful) to avoid falling asleep in a quiet environment in the evening-morning condition. Therefore, overnight studies typically use an *active* wake evening-morning (sleep-deprived; condition 5) control condition if they want to control for circadian effects in their design.

Finally, few studies have included both an overnight sleep (condition 1), a daytime nap (condition 9), and a quiet rest (wake; condition 11) condition in a single experimental design to gain a better understanding of the specific effect of sleep on consolidation (den Berg van et al., 2021; Van den Berg et al., 2019). While this approach introduces its own set of challenges (e.g., the architecture of a nap and overnight sleep are not comparable), it can be advantageous as it allows: 1) the direct comparison of the relative benefit of an overnight sleep vs. a nap, 2) a better control for time-of-day effects, and, 3) the examination of specific benefits of napping and the minimum amount of sleep necessary to afford a benefit to memory consolidation.

Pitfall 1d. Time interval between memory encoding and sleep onset

The time interval between the learning task and sleep onset (see Figure 1 condition 1) may vary across experiments, conditions and individuals, potentially hindering the assessment of the true effect of sleep on memory consolidation. Although consolidation of procedural memory appears rather insensitive to such effects (King et al., 2017), it has been shown in declarative memory that the more time elapses between the end of the learning task and sleep onset, the smaller the sleep-related memory benefit (Gais et al., 2006). A longer wake interval before sleep onset may hinder the manifestation of the beneficial effect of sleep due to the participant's involvement in activities that may interfere with recently learned information by re-engaging the same cognitive processes and/or recruiting the same neural networks.

Solutions: experimental designs/procedures controlling the time interval between memory encoding and sleep onset

Experiments should control the duration of the interval between memory encoding and bedtime/sleep onset, as well as the participants' activities during this interval. To minimize interference during this interval, participants should go to bed as soon as possible after memory encoding. Such designs are more feasible when participants sleep in the lab during the experiment.

If, however, participants sleep at home after the learning session, then mobile actigraphy or, as a less compelling substitute, sleep diaries and post-experiment questionnaires, should be employed to assess the duration of this interval and the activities performed. This information then should be appropriately considered in data analysis.

Pitfall 1e. Baseline measurements and feedback effects in declarative memory paradigms

When designing a declarative memory paradigm, a critical question is what procedure to use to ensure that participants encode a sufficient number of items for a later reliable and valid test of retrieval performance. In most sleep-related declarative memory studies that use cued or free recall, a certain learning criterion is defined, for instance 60% of recall success. (Other methods involve restrictions of study time (e.g. Lahl et al., 2008) or of number of trials during encoding (e.g., Mikutta et al., 2019)). If the learning criterion is not met after the first run of trials, a common strategy is to repeat the whole run, until the learning criterion is met (e.g. Cordi et al., 2014; Prehn-Kristensen et al., 2014). An advantage of this procedure is that all participants encode a sufficient number of items for later retrieval testing. However, the quality of encoding can be significantly different between participants: A participant who met the learning criterion within the very first run studies all items only once and, therefore, encodes them rather weakly. Another participant, who needed several repetitions to meet the criterion, studies all items multiple times. In this latter example, a difference can arise in the strength/quality of encoding of different items: some items may have been successfully encoded already in the first run and further practiced during successive repetitions, while other items may have been encoded only in the final repetition. These differences in repetitions and encoding level could have an enormous impact on later retrieval performance (Ebbinghaus, 2013; Xue et al., 2010; Young & Bellezza, 1982). This effect deserves even more consideration in studies comparing different populations (e.g., healthy participants vs. patients, or children vs. adults) that presumably learn at a different pace.

Another issue that arises in declarative memory paradigms is that the performance level observed during the last run of encoding (i.e., just when the learning criterion is met) is frequently used as a baseline measurement to evaluate recall performance after the retention interval. However, encoding runs are often designed to give direct feedback, often in the form of providing the correct answer after each item, potentially resulting in further, unmeasured encoding. Therefore, the baseline measurement does not reflect the exact memory state at the end of the learning phase, but rather probably underestimates it in such cases (Chan & McDermott, 2007; Karpicke & Roediger, 2008; Soderstrom et al., 2016; Wiklund-Hornqvist et al., 2014).

Solutions: experimental procedures to minimize differences across participants/studies in baseline measurements and feedback effects

Most researchers investigating sleep-related consolidation of declarative memory choose a learning criterion between 40% and 80% (e.g., Fowler et al., 1973; Klinzing et al., 2016). While systematic studies are lacking, an unsystematic Pubmed search yielded a mode of 60%-criterion for word-pair learning or visuo-spatial learning tasks (e.g., Cordi et al., 2014; Fenn & Hambrick, 2012; Marshall et al., 2004; Payne et al., 2012; Prehn-Kristensen et al., 2014; Rasch et al., 2007; Wilhelm et al., 2008). Overall, based on these studies, the 60% learning criterion seems to be a reasonable choice to account for possible floor and ceiling effects.

An option to circumvent the use of a predefined learning criterion is the so-called selective reminding procedure (Buschke, 1973). All items are presented to the participant during a first study run. Subsequently, a first test run is conducted where all items are tested. In a second study run, only those items that were not recalled correctly during the first test run are presented. Then, in a third study run, only those items are presented that were not recalled correctly during the second test run and so on. New study runs proceed until all items are remembered correctly once. This procedure enables all participants to encode the same number of items while no item is ‘over-

learned’ (for examples, see Mazza et al., 2016; Quan et al., 2018; Uguccioni et al., 2013). This procedure could be used to reach 100% for baseline encoding level in all participants. One limitation of such approach (i.e., 100% learning criterion) is, however, that it is only suitable for those memory studies where a loss in declarative memory is expected over the retention interval. In other cases, using the selective reminding procedure with a lower predefined learning criterion (e.g., 60%) could ensure similar encoding strength/quality across participants (no over-learned items), while avoiding potential floor and ceiling effects.

Additionally, a test run of all items could be introduced immediately after the learning criterion is met, without any corrective feedback. This could provide an even more precise measure of the baseline. However, one should remember that test runs can also boost learning and subsequent consolidation even if no feedback is given—a phenomenon called ‘test-enhanced learning’ (McDermott, 2021; Roediger III & Karpicke, 2006).

Pitfall 2: Task complexity

A significant difficulty in sleep and memory research, and in cognitive neuroscience and psychology in general, is that practically every task involves several cognitive processes (e.g., Jacoby, 1991; Sigman & Dehaene, 2005). The learning/memory scores that are used to assess behavioral performances typically reflect a mixture of these cognitive processes (Cohen et al., 2005). For example, even a simple perceptual-motor learning task requires at least the processing of perceptual stimuli, acquisition of their serial order and/or transitional probabilities, perceptual-motor coordination, and selective attention. As learning progresses (i.e., as the involved neurocognitive system is being fine-tuned to the task), these processes could improve at a different pace and, therefore, could contribute to the behavioral performance assessed at the end of the learning session to a varying degree. Importantly, consolidation can differentially affect these processes involved in the task (e.g., Conte & Ficca, 2013; King et al., 2017; Stickgold, 2013; Stickgold & Walker, 2013),

and this could be even further exacerbated by individual differences in their contribution to behavioral performance.

There are not only different types of learning and cognitive processes but also different types of retrieval processes that may determine whether beneficial effects of sleep on memory consolidation are detected or not. Specifically, declarative memory paradigms most frequently probe recall and recognition, with recall referring to the ability to retrieve a stimulus from memory with or without a cue, and recognition referring to the ability to decide whether a given stimulus has been previously encountered. Some evidence suggests that recall is more sensitive to sleep effects than recognition, possibly because sleep facilitates the integration of new memories into pre-existing knowledge networks, thereby increasing potential access routes for recall (Diekelmann et al., 2009).

Differences in retrieval may be evident for procedural memory as well. For example, tasks may tap into explicit (i.e., conscious) vs. implicit (i.e., unconscious) aspects of the acquired knowledge to a varying degree (e.g., Fischer et al., 2006; Schendan et al., 2003). This could potentially lead to contradictory findings across studies and mask the differential effect of sleep on the consolidation of different aspects of knowledge.

Another factor to consider is that different memory systems are thought to interact with each other during learning and possibly also during consolidation (Freedberg et al., 2020). However, to optimize use of resources, researchers sometimes include memory tasks tapping into different memory systems (e.g., a non-declarative/procedural and a declarative memory task) in a given experiment. This is problematic because acquiring memories that tap into different memory systems shortly after one another may cause interference in their consolidation. This could potentially alter the observed effects of post-learning sleep and their interpretation. For instance, it has been shown that acquiring procedural memories just after a declarative memory task was affected by participants' memory performance in the latter and led to differences in consolidation over the wake

vs. sleep conditions. A complementary pattern has been observed when the declarative memory task was performed after the acquisition of procedural memories (Brown & Robertson, 2007).

Solutions: disentangling and contrasting different cognitive processes and aspects of memory involved in a given task

Since there are no process-pure learning/memory tasks, attention should be paid to the specific cognitive processes that are involved in a particular task. We recommend that future research aiming at understanding the specific effect of sleep on consolidation should use tasks and designs that could help disentangle these different cognitive processes, and examine whether they are differentially affected by sleep. For example, in declarative memory tasks, different aspects of retrieval (free recall, cued recall, recognition) should be systematically compared within the same experimental design to examine how they can reveal (potentially different) sleep effects. These effects may also vary depending on the type of information to be encoded, for example, paired-associates learning (Feld et al., 2016; Plihal & Born, 1997), word-list learning (Abel & Bauml, 2012; Diekelmann et al., 2010), emotional picture learning (Cairney et al., 2015; Payne et al., 2015) and object-location memory (Rasch et al., 2007; Rudoy et al., 2009), suggesting that they involve at least partially distinct cognitive processes. To better understand the differential effect of sleep on aspects of memory, contrasting the encoding/consolidation of different types of information within the same experimental design is warranted. In procedural learning/memory, for instance, research has disentangled and contrasted allocentric vs. egocentric representations (Viczkó et al., 2018), perceptual vs. motor components of learning (Hallgato et al., 2013), transition vs. ordinal representations (Song & Cohen, 2014), and acquisition of statistical vs. sequential regularities (Simor, Zavecz, et al., 2019), and showed differential effects of sleep in some of these aspects (Albouy et al., 2013; Cohen et al., 2005; Song & Cohen, 2014).

Furthermore, to minimize the potential interactions between different memory systems, which could confound the identification of sleep effects, it may be beneficial to administer tasks tapping into different memory systems using a between-subject design. If, however, a within-subject design is chosen, the order of task administration should be counterbalanced across participants and included in data analysis as a separate factor.

Pitfall 3: Fatigue effect in repetitive tasks

Some studies, particularly those investigating non-declarative/procedural learning, use tasks that involve continuous practice with a series of repetitions of the same action, for example, pressing keys (Nissen & Bullemer, 1987). Learning is measured as the improvement in accuracy or in reaction times as the task progresses. Usually, the performance at the end of the training session serves as a baseline to measure improvement at the test session that takes place after an interval involving sleep or wakefulness. Yet, after a certain amount of time spent performing the task, the subject's observed improvement is less marked, which can be interpreted as a *reactive inhibition effect* that reflects the build-up of fatigue over the trials (e.g., Brawn et al., 2010; Pan & Rickard, 2015). This effect often results in smaller improvement or even a decrease in performance as the task progresses. Thus, the measured performance after longer/extended practice is not representative of the level of expertise gained in the task and, therefore, comparing the performance at the test session with that of the end of the training session may lead to illusory sleep-related improvement and may also bias the quantification of the sleep benefit. Figure 3 upper panel illustrates this issue, which can be even further exacerbated by using performance measures averaged across multiple trials, instead of a trial-by-trial analysis. In several cases, after eliminating the reactive inhibition effect by releasing the presumed fatigue, the sleep-related off-line improvement was no longer observed (e.g., Cai & Rickard, 2009; Rickard et al., 2008). Rather than an actual performance improvement, after elimination of the reactive inhibition effect, the benefit of sleep was expressed

as a stabilization of performance (Nettersheim et al., 2015; Rickard et al., 2008). Although this issue is primarily relevant in procedural learning studies, it is possible that reactive inhibition also affects performance in declarative memory studies, particularly if they include repetitive presentations of the same items or a long period of memorization (Abel & Bauml, 2012).

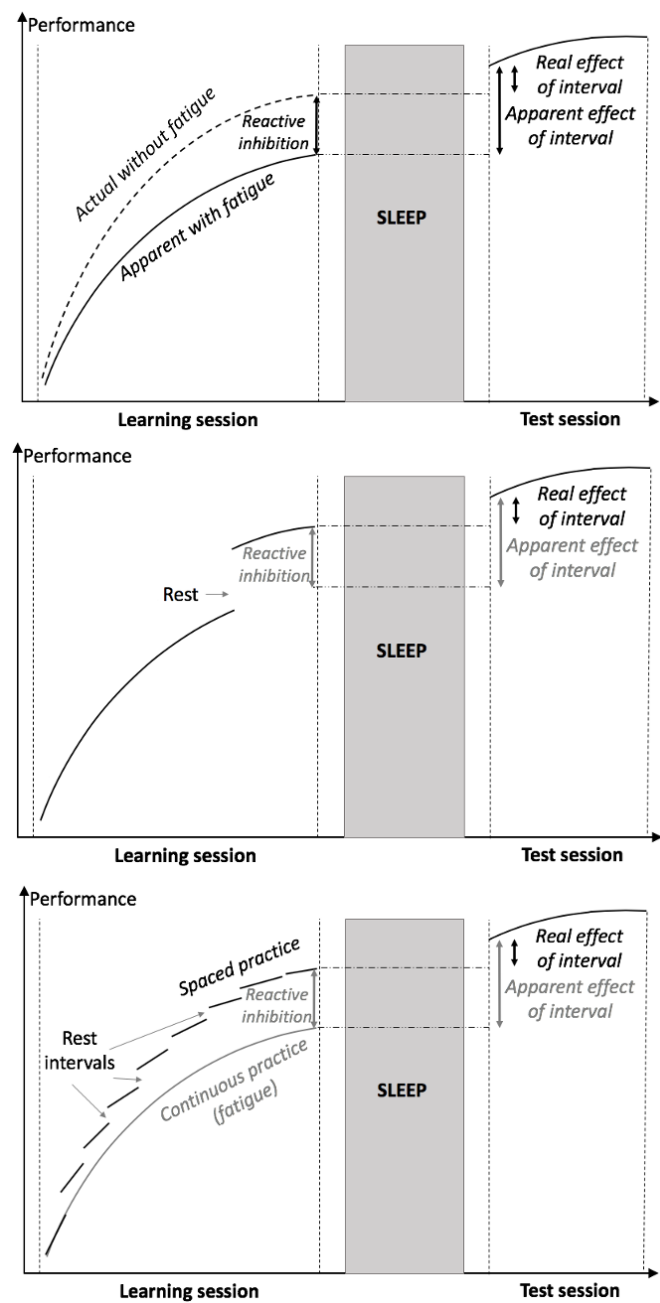


Figure 3. The fatigue effect (upper panel), the short rest solution (middle) and the spaced practice solution (lower).

Solutions: modified experimental designs and data analysis approaches can control for fatigue effect in repetitive tasks

Experimental design solutions

Using post-rest performance at the end of the training session as a baseline. Resting for a few minutes after the training session appears to be sufficient to wash out the effect of reactive inhibition on performance. Measuring performance after a break is therefore a more appropriate baseline to assess subsequent off-line consolidation (e.g., Brawn et al., 2010; Simor, Zavecz, et al., 2019). Figure 3 middle panel illustrates this solution.

Learning through spaced rather than massed practice. Use of short (e.g., 10 s) performance intervals between longer (e.g., 30 s) rest intervals during the training session (often termed spaced practice) impedes the accumulation of reactive inhibition compared to experimental designs that use massed practice in which there are longer task intervals (e.g., Brawn et al., 2010; Rickard et al., 2008; Rieth et al., 2010). Figure 3 lower panel illustrates this solution.

Data analysis solutions

Using curve fitting methods. Here we highlight two such methods. First, a function-based model (e.g., a power function for reaction time improvement) can be fitted to the training session data and used to predict future performance (under the null hypothesis that the delay between training and test sessions has no effect on performance). This method enables a comparison between the predicted (under H_0) and the actual outcomes measured during the test session. This way, one avoids averaging over data points to compute a pre-post gain—a procedure which may yield illusory off-line performance gains if performance is improving between the end of the training session and the beginning of the test session, wherein the data averaging is done (Pan & Rickard, 2015). As a second and more formal approach, a function can be fitted to the training and test session data and then a continuity test can be used to infer whether the performance is a simple

continuation of that function from the training session to the test session, or whether there is an abrupt change between the sessions (see details on these approaches in Pan & Rickard, 2015).

Performing computational modeling. Using computational models on trial-by-trial data can help overcome the issue of fatigue by directly including reactive inhibition as a separate parameter in the model. For instance, in a probabilistic sequence learning task, Török, Janacsek, Nagy, Orbán, & Nemeth (2017) used such model allowing the estimation of the actual magnitude of learning, independent of the effect of reactive inhibition. Such models can be used in a wide range of learning and memory tasks, including finger tapping and other sequence learning tasks.

Pitfall 4: Inappropriate data analysis practices

The studies of sleep and memory suffer from similar problems as the whole field of psychology and neuroscience discussed in recent years as the ‘replication crisis’ (Ioannidis, 2005; Maxwell et al., 2015; OpenScienceCollaboration, 2015), and thus could similarly benefit from an update in practices that are currently evolving in the scientific community in general (see e.g., Rickard et al., in press about publication bias in sleep and motor sequence learning literature).

Pitfall 4a. Questionable practices related to sample size and interpretation of non-significant results

In the field of sleep-related consolidation, studies have typically used samples with 12-20 participants per group (e.g., Gais et al., 2006; Rickard et al., in press; Wagner et al., 2006), or in some cases even smaller samples (e.g., Csabi et al., 2015), which may be due to complicated or demanding study designs, difficulties recruiting clinical populations, and/or drop-outs of participants (i.e., experimental attrition). Moreover, sample sizes have usually not been determined by *a priori* power analysis based on expected effect sizes. Importantly, small sample sizes could result in low statistical power, potentially increasing Type 2 errors (i.e., not detecting an existing

effect), as well as could lead to non-replicable, spurious findings. On the other hand, a common but questionable practice of collecting additional data until a significant effect is reached could increase Type 1 errors (i.e., detecting an effect that does not exist), again, leading to non-replicable findings.

Another issue arises from the interpretation of non-significant findings. For example, non-significant effects could be observed in pre-sleep vs. post-sleep comparisons when consolidation results in stabilization of the acquired knowledge without forgetting or off-line performance improvement (i.e., no performance change). If one wants to conclude that sleep promotes stabilization of the acquired knowledge or has no effect on some aspects of memory consolidation compared to wakefulness, such conclusion cannot be drawn by showing non-significant results in classical statistical approaches (e.g., frequentist t-test, ANOVA, correlation, etc.).

Solutions: a priori power analysis and Bayesian statistical approaches

Before data collection. It has long been recommended in guidelines (e.g., published by the American Psychological Association) that experimenters should determine the sample size before starting the experiment by computing power analyses based on the expected effect size estimated or found in previous studies that observed similar effects.

During data collection. For particularly costly experimental protocols, Bayesian statistical analyses (Dienes, 2016; Dienes et al., 2018; Dienes & Mclatchie, 2018; Wagenmakers et al., 2018) computed in the course of data collection can be used to determine whether there is enough evidence in favor of a given a priori defined effect so that one can stop data collection (Rouder, 2014; Schönbrodt et al., 2017).

After data collection. Bayesian analyses, in particular the Bayes Factors, are rarely reported in the field of sleep-related memory consolidation (see Brown & Maylor, 2017 for an exception) whereas they are increasingly reported in other areas of psychology and neuroscience. The Bayes Factor indicates an odds ratio of relative probabilities in favor of the null hypothesis (i.e., the

absence of a difference between the conditions or groups) vs. in favor of the alternative hypothesis (i.e., the difference between the conditions or groups) (Jarosz & Wiley, 2014; Rouder et al., 2009). Bayesian statistics thus allow a more fine-grained quantitative evaluation of the effect of sleep on memory consolidation. Additionally, effect size measures should always be reported to provide an estimate of the relevance of the observed effect. Effect sizes can also be indicative of the true effect in cases of non-significant results with small samples and potential Type 2 errors.

Pitfall 4b. Spurious correlations between sleep parameters and memory consolidation

Beyond the comparison of groups or conditions, conclusions for the effect of sleep are often based on correlations between behavioral performance and sleep polysomnographic parameters (e.g., Scullin, 2013; Simor, Zavecz, et al., 2019) (see also section 4c). However, the concern has been raised that some of these correlation analyses are a consequence of suboptimal statistical practices, leading to spurious correlations (e.g., Mantua, 2018; Pan & Rickard, 2015). Specifically, polysomnography provides a wealth of parameter combinations, including different spindle parameters (absolute/relative number, amplitude, length, activity) at different frequencies (slow/fast) at different scalp regions/electrodes during different sleep stages and different fractions of the night. The resulting massive ‘researcher’s degrees of freedom’ easily allow for (intentional or unintentional) p-hacking.

Small sample sizes are a further source of spurious correlations. When sufficiently large sample sizes are used, these correlations may be greatly reduced (Ujma, 2021) or even disappear. For example, between sleep parameters and episodic memory consolidation—an area that has received much attention in the past decades of sleep and memory research—Ackermann et al. (2015) did not find any significant correlation in a large sample of 929 participants.

Solutions: a priori planning of appropriate statistical analyses and reporting all (significant and non-significant) findings

The correlations to be computed should be planned *a priori* (see also the section on further recommendations below) and corrected for multiple comparisons in order to avoid increases in Type 1 errors (Abdi, 2007). Non-significant planned correlations should also be systematically reported (Forstmeier et al., 2017). If no relationship is expected between certain sleep parameters and behavioral performance, Bayesian approaches should be used to draw conclusions in favor of the null hypothesis instead of (or in addition to) reporting non-significant p-values (see section 4a).

Pitfall 4c. Individual differences in general cognitive abilities are not controlled for in correlational studies of sleep and memory

Certain features of sleep (e.g., sleep spindles) appear to be highly correlated with trait-like individual differences in cognitive abilities. Particularly strong relationships have been identified for cognitive abilities related to reasoning, problem solving, the ability to identify complex patterns and relationships, and the use of logic (i.e., ‘fluid intelligence’) (Bódizs et al., 2005; Fang et al., 2019; Fang et al., 2017; Fogel & Smith, 2006; Ujma et al., 2015). Since these cognitive abilities are not only associated with certain features of sleep but they have been shown to support memory functions as well, they may confound the associations revealed between sleep and memory consolidation. Therefore, when the specific effect of sleep on memory consolidation is tested, associations between sleep (e.g., spindles) and these cognitive abilities (e.g., intelligence) should be controlled for.

Solutions: employing appropriate control tasks and conditions

The problem of disentangling individual differences in the associations between sleep and general cognitive abilities from the associations between sleep and memory can be addressed by at

least two ways. First, one can employ neurocognitive assessments (e.g., intelligence testing) and include these scores as covariates to statistically control for possible confounding effects when testing the specific associations between sleep and memory consolidation. Second, a comparable baseline night of sleep together with an appropriate control task can be included in the study design. This control task should be comparable to the experimental task without engaging the targeted specific processes that are the focus of sleep-related memory consolidation. Comparing the two experimental conditions can reveal the specific effect of sleep on the memory process of interest.

Further recommendations: Engaging in open science practices

Other fields of neuroscience using techniques such as neuroimaging (e.g., fMRI) have engaged in open science initiatives by, for example, depositing raw data in open-access databases (e.g., OpenfMRI; Poldrack & Gorgolewski, 2017), which have recently been extended to other neuroimaging and electrophysiological methods such as EEG as well (OpenNEURO; Gorgolewski et al., 2017). However, sleep and memory research is lagging behind with respect to this practice. Research transparency is further hindered by the lack of pre-registration of the studies on sleep and memory.

Depositing research data on open-access platforms

Making research data publicly available enables the re-analysis of old data when new analysis techniques and/or new theories are developed. For example, there are at least two different types of REM microstates (tonic vs. phasic), each with different characteristics (Simor et al., 2017; Simor, van Der Wijk, et al., 2019). Access to previous sleep EEG data would make it possible to specifically test the role of REM microstructure in memory consolidation on previous datasets. Publicly available sleep EEG and behavioral data could also provide solutions for at least some of the pitfalls discussed above by, for example, enabling data re-analysis to evaluate evidence for

potentially non-significant results (see section 4a) and resolve at least some of the previous issues of spurious results (see section 4b). These open databases can also support reliable synthesis of the data through meta-analyses with larger sample sizes.

To maximize the benefits of previous research in the scientific community, we recommend that sleep researchers engage in open science (Nosek et al., 2015) and make data publicly available. There are several online open repositories available for such purposes, such as Open Science Framework, OpenNeuro.org, or sleepdata.org. Developing a specific open database for sleep research with EEG, polysomnographic and behavioral data would further benefit this field.

Pre-registration of studies

Another way of increasing transparency of research is to pre-register studies before data collection (Lindsay et al., 2016; Nosek et al., 2018). Pre-registration includes the specification of the research question, experimental design, subject population, sample size as well as planned analysis methods. Pre-registration is already the gold standard in many fields of research, including for clinical trials in medical research. The neurosciences and psychology fields increasingly recognize the importance of pre-registration as well. Yet, this option has been largely neglected in sleep and memory research so far.

Studies can be pre-registered in different ways. One option is pre-registration in independent online registries like the Open Science Framework or ClinicalTrials.gov. In these registries, researchers provide a detailed description of their planned study that can be accessed by other researchers as well as journal editors and reviewers to determine whether the pre-specified plan was followed adequately. Another option is to write a registered report, which is a novel publication type offered by an increasing number of journals (e.g., Plos Biology, eLife, eNeuro, Cortex). A registered report undergoes two stages of peer review, first before data collection to determine the appropriateness of the research plan and methodology, and then after data collection covering the full research report including the results. If the first round of peer review is successful, the authors

are typically offered ‘in principle acceptance’ by the journal, allowing the results to be published irrespective of the actual findings.

Both procedures, pre-registration in online registries and registered reports, increase the quality of research by reducing inappropriate data analysis practices, including p-hacking, HARKing (hypothesizing after the results are known), and the application of unplanned statistical tests. Thus, pre-registration can promote the implementation of solutions to Pitfall 4 discussed above (see e.g., sections 4a and 4b). Additionally, registered reports could also reduce the file-drawer problem because the study, irrespective of finding significant or non-significant results, could be published in the target journal.

Table 1. Summary of pitfalls and their solutions

Pitfalls	Solutions
1. Non-optimal experimental designs May lead to inaccurate conclusions due to confounding variables	<p>a. Time-of-day (circadian) effects: Using multiple control conditions: morning-evening; evening-morning; morning alone; evening alone; evening-morning and evening-evening with sleep deprivation (with and without recovery sleep). Napping conditions. Assessing sleepiness and vigilance.</p> <p>b. Overnight studies (pathological population or intervention): Including a morning-evening condition and/or an evening-morning deprivation condition. Including a control group.</p> <p>c. Napping studies: Considering time of day and duration of nap. Including a quiet-wake control condition. Monitoring the nap with polysomnography.</p> <p>d. Time interval between memory encoding and sleep onset: Controlling for duration of and subject's activity during the interval between end of task and sleep. Monitoring the activities during the interval with actigraphy and/or questionnaires.</p> <p>e. Baseline measurements and feedback effects (declarative memory): Using a selective reminding procedure, possibly combined with a predefined learning criterion.</p>
2. Task complexity Every task involves multiple cognitive processes that need to be disentangled to better understand the specific effect of sleep	Disentangling and contrasting different cognitive processes and aspects of memory involved in a given task.
3. Fatigue effect in repetitive tasks (non-declarative memory) May lead to a spurious beneficial effect of sleep by negatively affecting performance after a longer practice	Using appropriate experimental designs, e.g., including post-rest performance at the end of the training session as a baseline, and promoting learning through spaced rather than massed practice. Using appropriate data analysis methods, such as curve fitting and computational modeling.
4. Inappropriate data analysis practices Small sample sizes and inappropriate analyses/reporting may lead to spurious correlations and incorrect conclusions.	<p>a. Small sample size and reporting only significant results: Determining the required sample size a priori. Using Bayesian analyses to decide when to stop data collection. Reporting Bayes Factors and effect sizes.</p> <p>b. Spurious correlations between sleep parameters and memory consolidation: Planning correlation analyses of interest in advance,</p>

	<div>correcting for multiple comparisons, and reporting non-significant planned comparisons.</div> <div><i>c. Not controlling for individual differences in general cognitive abilities in correlational studies of sleep and memory:</i> Including neurocognitive assessments of general cognitive abilities as covariates. Including a baseline night of sleep with an appropriate control task.</div>
--	--

Conclusion

In this consensus paper, we highlighted four sets of critical methodological pitfalls that impede research in the field of sleep and memory, and we offered solutions to avoid them (for a summary, see Table 1). We believe that implementing the solutions presented in this paper will lead to more reliable results and significantly advance our understanding of the complex relationship between sleep and memory. Since some of the pitfalls described in this paper (such as those related to fatigue effects, task complexity, and data analysis practices) are relevant not only in sleep and memory research but also in other fields of psychology and neuroscience, applying these solutions where appropriate could benefit the broader scientific community as well.

Acknowledgements

This research was supported by the National Brain Research Program (project 2017-1.2.1-NKP-2017-00002); Hungarian Scientific Research Fund (NKFIH-OTKA K 128016, PI: D. N.; NKFIH-OTKA PD 124148, PI: K.J.; NKFI FK 128100, PI: P.S.); Janos Bolyai Research Fellowship of the Hungarian Academy of Sciences (to K. J.); IDEXLYON Fellowship (to D.N); French National Agency for Research (ANR, grant n°ANR-15-CE33-0003, PI: S.M.). We are grateful to Kate Schipper and Teodora Vekony for their help and comments on the manuscript.

References

- Abdi, H. (2007). Bonferroni and Šidák corrections for multiple comparisons. *Encyclopedia of measurement and statistics*, 3, 103-107.
- Abel, M., & Bauml, K. H. (2012, Jul). Retrieval-induced forgetting, delay, and sleep. *Memory*, 20(5), 420-428. <https://doi.org/10.1080/09658211.2012.671832>
- 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.
- Ackermann, S., & Rasch, B. (2014). Differential effects of non-REM and REM sleep on memory consolidation? *Current Neurology and Neuroscience Reports*, 14(2), 430.
- Ahuja, S., Chen, R. K., Kam, K., Pettibone, W. D., Osorio, R. S., & Varga, A. W. (2018). Role of normal sleep and sleep apnea in human memory processing. *Nature and science of sleep*, 10, 255.
- Albouy, G., Fogel, S., Pottiez, H., Nguyen, V. A., Ray, L., Lungu, O., Carrier, J., Robertson, E., & Doyon, J. (2013). Daytime sleep enhances consolidation of the spatial but not motoric representation of motor sequence memory. *PloS One*, 8(1), e52805.
- Antony, J. W., Schönauer, M., Staresina, B. P., & Cairney, S. A. (2019). Sleep spindles and memory reprocessing. *Trends in Neurosciences*, 42(1), 1-3.
- Backhaus, J., Junghanns, K., Born, J., Hohaus, K., Faasch, F., & Hohagen, F. (2006). Impaired declarative memory consolidation during sleep in patients with primary insomnia: influence of sleep architecture and nocturnal cortisol release. *Biological Psychiatry*, 60(12), 1324-1330.
- Bódizs, R., Kis, T., Lázár, A. S., Havran, L., Rigó, P., Clemens, Z., & Halász, P. (2005, Sep). Prediction of general mental ability based on neural oscillation measures of sleep. *Journal of Sleep Research*, 14(3), 285-292. <https://doi.org/10.1111/j.1365-2869.2005.00472.x>
- Born, J., & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological Research*, 76(2), 192-203.
- Boyce, R., Williams, S., & Adamantidis, A. (2017). REM sleep and memory. *Current Opinion in Neurobiology*, 44, 167-177.

- 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.
- Brown, H., & Maylor, E. A. (2017). Memory consolidation effects on memory stabilization and item integration in older adults. *Psychonomic bulletin & review*, 24(4), 1032-1039.
- Brown, R. M., & Robertson, E. M. (2007). Off-line processing: Reciprocal interactions between declarative and procedural memories. *Journal of Neuroscience*, 27(39), 10468-10475.
- Buschke, H. (1973). Selective reminding for analysis of memory and learning. *Journal of Verbal Learning and Verbal Behavior*, 12, 543-550.
- Cai, D. J., & Rickard, T. C. (2009). Reconsidering the role of sleep for motor memory. *Behavioral Neuroscience*, 123(6), 1153.
- Cairney, S. A., Durrant, S. J., Power, R., & Lewis, P. A. (2015, Jun). Complementary roles of slow-wave sleep and rapid eye movement sleep in emotional memory consolidation. *Cereb Cortex*, 25(6), 1565-1575. <https://doi.org/10.1093/cercor/bht349>
- Cellini, N. (2017). Memory consolidation in sleep disorders. *Sleep Medicine Reviews*, 35, 101-112.
- Chan, J. C., & McDermott, K. B. (2007, Mar). The testing effect in recognition memory: a dual process account. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 33(2), 431-437. <https://doi.org/10.1037/0278-7393.33.2.431>
- Cohen, D. A., Pascual-Leone, A., Press, D. Z., & Robertson, E. M. (2005). Off-line learning of motor skill memory: A double dissociation of goal and movement. *PNAS*, 102(50), 18237-18241.
- Conte, F., & Ficca, G. (2013). Caveats on psychological models of sleep and memory: a compass in an overgrown scenario. *Sleep Medicine Reviews*, 17(2), 105-121.
- Cordi, M. J., Diekelmann, S., Born, J., & Rasch, B. (2014). No effect of odor-induced memory reactivation during REM sleep on declarative memory stability. *Front Syst Neurosci*, 8, 157. <https://doi.org/10.3389/fnsys.2014.00157>
- Csábi, E., Benedek, P., Janacek, 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. <https://doi.org/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.
- den Berg van, N. H., Pozzobon, A., Fang, Z., Al-Kuwatli, J., Toor, B., Ray, L. B., & Fogel, S. M. (2021). Sleep Enhances Consolidation of Memory Traces for Complex Problem-Solving Skills. *Cerebral Cortex*. <https://doi.org/10.1093/cercor/bhab216>
- Diekelmann, S., Biggel, S., Rasch, B., & Born, J. (2012). Offline consolidation of memory varies with time in slow wave sleep and can be accelerated by cuing memory reactivations. *Neurobiology of Learning and Memory*, 98(2), 103-111.
- Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, 11(2), 114-126.
- Diekelmann, S., Born, J., & Wagner, U. (2010, Apr 2). Sleep enhances false memories depending on general memory performance. *Behav Brain Res*, 208(2), 425-429. <https://doi.org/10.1016/j.bbr.2009.12.021>
- Diekelmann, S., Wilhelm, I., & Born, J. (2009). The whats and whens of sleep-dependent memory consolidation. *Sleep Medicine Reviews*, 13(5), 309-321.
- Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical Psychology*, 72, 78-89.
- Dienes, Z., Coulton, S., & Heather, N. (2018). Using Bayes factors to evaluate evidence for no effect: examples from the SIPS project. *Addiction*, 113(2), 240-246.
- Dienes, Z., & Mclatchie, N. (2018). Four reasons to prefer Bayesian analyses over significance testing. *Psychonomic bulletin & review*, 25(1), 207-218.
- Dresler, M., Spoormaker, V., Beiting, P., Czisch, M., Kimura, M., Steiger, A., & Holsboer, F. (2014). Neuroscience-driven discovery and development of sleep therapeutics. *Pharmacology & therapeutics*, 141(3), 300-334.
- Ebbinghaus, H. (2013, Oct). Memory: a contribution to experimental psychology. *Ann Neurosci*, 20(4), 155-156. <https://doi.org/10.5214/ans.0972.7531.200408>

- Ellenbogen, J. M., Payne, J. D., & Stickgold, R. (2006). The role of sleep in declarative memory consolidation: passive, permissive, active or none? *Current Opinion in Neurobiology*, 16(6), 716-722.
- Fang, Z., Ray, L. B., Owen, A. M., & Fogel, S. M. (2019). Brain Activation Time-Locked to Sleep Spindles Associated With Human Cognitive Abilities. *Front Neurosci*, 13, 46. <https://doi.org/10.3389/fnins.2019.00046>
- Fang, Z., Sergeeva, V., Ray, L. B., Viczko, J., Owen, A. M., & Fogel, S. M. (2017, Jan). Sleep Spindles and Intellectual Ability: Epiphenomenon or Directly Related? *J Cogn Neurosci*, 29(1), 167-182. https://doi.org/10.1162/jocn_a_01034
- Feld, G. B., & Born, J. (2017). Sculpting memory during sleep: concurrent consolidation and forgetting. *Current Opinion in Neurobiology*, 44, 20-27.
- Feld, G. B., Weis, P. P., & Born, J. (2016). The Limited Capacity of Sleep-Dependent Memory Consolidation. *Front Psychol*, 7, 1368. <https://doi.org/10.3389/fpsyg.2016.01368>
- Fenn, K. M., & Hambrick, D. Z. (2012, Aug). Individual differences in working memory capacity predict sleep-dependent memory consolidation. *Journal of Experimental Psychology: General*, 141(3), 404-410. <https://doi.org/10.1037/a0025268>
- Fenn, K. M., Nusbaum, H. C., & Margoliash, D. (2003). Consolidation during sleep of perceptual learning of spoken language. *Nature*, 425(6958), 614-616.
- Fischer, S., Drosopoulos, S., Tsen, J., & Born, J. (2006). Implicit Learning-Explicit Knowing: A Role for Sleep in Memory System Interaction. *Journal of Cognitive Neuroscience*, 18(3), 311.
- Fogel, S. M., & Smith, C. T. (2006, Sep). Learning-dependent changes in sleep spindles and Stage 2 sleep. *Journal of Sleep Research*, 15(3), 250-255. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16911026
- Forstmeier, W., Wagenmakers, E. J., & Parker, T. H. (2017). Detecting and avoiding likely false-positive findings—a practical guide. *Biological Reviews*, 92(4), 1941-1968.
- Fowler, M. J., Sullivan, M. J., & Ekstrand, B. R. (1973, Jan 19). Sleep and memory. *Science*, 179(4070), 302-304. <https://doi.org/10.1126/science.179.4070.302>
- Freedberg, M., Toader, A. C., Wassermann, E. M., & Voss, J. L. (2020, Jan). Competitive and cooperative interactions between medial temporal and striatal learning systems. *Neuropsychologia*, 136, 107257. <https://doi.org/10.1016/j.neuropsychologia.2019.107257>

- Gais, S., Lucas, B., & Born, J. (2006). Sleep after learning aids memory recall. *Learning and Memory*, 13(3), 259-262.
- Gais, S., Plihal, W., Wagner, U., & Born, J. (2000, Dec). Early sleep triggers memory for early visual discrimination skills. *Nat Neurosci*, 3(12), 1335-1339. <https://doi.org/10.1038/81881>
- Gais, S., Rasch, B., Dahmen, J. C., Sara, S., & Born, J. (2011, Sep). The memory function of noradrenergic activity in non-REM sleep. *J Cogn Neurosci*, 23(9), 2582-2592. <https://doi.org/10.1162/jocn.2011.21622>
- Genzel, L., & Robertson, E. M. (2015). To Replay, Perchance to Consolidate. *PLoS Biology*, 13(10), e1002285. <https://doi.org/10.1371/journal.pbio.1002285>
- Gorgolewski, K., Esteban, O., Schaefer, G., Wandell, B., & Poldrack, R. (2017). OpenNeuro—a free online platform for sharing and analysis of neuroimaging data. *Organization for Human Brain Mapping. Vancouver, Canada, 1677*.
- Hallgato, E., Györi-Dani, D., Pekár, J., Janacsek, K., & Nemeth, D. (2013). The differential consolidation of perceptual and motor learning in skill acquisition. *Cortex*, 49(4), 1073-1081.
- Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), e124.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30, 513-541.
- Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving*, 7(1), 2.
- Karipicke, J. D., & Roediger, H. L., 3rd. (2008, Feb 15). The critical importance of retrieval for learning. *Science*, 319(5865), 966-968. <https://doi.org/10.1126/science.1152408>
- Keisler, A., Ashe, J., & Willingham, D. T. (2007). Time of day accounts for overnight improvement in sequence learning. *Learning and Memory*, 14(10), 669-672.
- 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.
- Klinzing, J. G., Rasch, B., Born, J., & Diekelmann, S. (2016, Dec). Sleep's role in the reconsolidation of declarative memories. *Neurobiol Learn Mem*, 136, 166-173. <https://doi.org/10.1016/j.nlm.2016.10.004>

- Lahl, O., Wispel, C., Willigens, B., & Pietrowsky, R. (2008, Mar). An ultra short episode of sleep is sufficient to promote declarative memory performance. *J Sleep Res*, 17(1), 3-10. <https://doi.org/10.1111/j.1365-2869.2008.00622.x>
- Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in cognitive sciences*, 15(8), 343-351.
- Lindsay, D. S., Simons, D. J., & Lilienfeld, S. O. (2016). Research preregistration 101. *APS observer*, 29(10).
- Mantua, J. (2018). Sleep Physiology Correlations and Human Memory Consolidation: Where Do We Go From Here? *Sleep*, 41(2), zsx204.
- Marshall, L., Mölle, M., Hallschmid, M., & Born, J. (2004). Transcranial Direct Current Stimulation during Sleep Improves Declarative Memory. *Journal of Neuroscience*, 24(44), 9985-9992. <https://doi.org/10.1523/JNEUROSCI.2725-04.2004>
- Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does “failure to replicate” really mean? *American Psychologist*, 70(6), 487.
- Mazza, S., Gerbier, E., Gustin, M. P., Kasikci, Z., Koenig, O., Toppino, T. C., & Magnin, M. (2016, Oct). Relearn Faster and Retain Longer. *Psychol Sci*, 27(10), 1321-1330. <https://doi.org/10.1177/0956797616659930>
- McDermott, K. B. (2021, Jan 4). Practicing Retrieval Facilitates Learning. *Annu Rev Psychol*, 72, 609-633. <https://doi.org/10.1146/annurev-psych-010419-051019>
- McDevitt, E. A., Duggan, K. A., & Mednick, S. C. (2015, Jul). REM sleep rescues learning from interference. *Neurobiol Learn Mem*, 122, 51-62. <https://doi.org/10.1016/j.nlm.2014.11.015>
- 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.
- Mednick, S. C., Makovski, T., Cai, D., & Jiang, Y. V. (2009). Sleep and rest facilitate implicit memory in a visual search task. *Vision Research*, 49(21), 2557-2565.
- Mednick, S. C., Mednick, S. C., Cai, D. J., Kanady, J., & Drummond, S. P. A. (2008). Comparing the benefits of caffeine, naps and placebo on verbal, motor and perceptual memory. *Behavioural Brain Research*, 193(1), 79.
- Mednick, S. C., Nakayama, K., & Stickgold, R. (2003, Jul). Sleep-dependent learning: a nap is as good as a night. *Nat Neurosci*, 6(7), 697-698. <https://doi.org/10.1038/nn1078>

- Mikutta, C., Feige, B., Maier, J. G., Hertenstein, E., Holz, J., Riemann, D., & Nissen, C. (2019, Mar 7). Phase-amplitude coupling of sleep slow oscillatory and spindle activity correlates with overnight memory consolidation. *Journal of Sleep Research*.
<https://doi.org/10.1111/jsr.12835>
- 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. <https://doi.org/10.1007/s00221-009-2024-x>
- Nettersheim, A., Hallschmid, M., Born, J., & Diekelmann, S. (2015). The role of sleep in motor sequence consolidation: stabilization rather than enhancement. *Journal of Neuroscience*, 35(17), 6696-6702.
- Nishida, M., & Walker, M. P. (2007). Daytime naps, motor memory consolidation and regionally specific sleep spindles. *PloS One*, 2(4), e341.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17406665
- Nissen, C., Kloepper, C., Feige, B., Piosczyk, H., Spiegelhalter, K., Voderholzer, U., & Riemann, D. (2011). Sleep-related memory consolidation in primary insomnia. *Journal of Sleep Research*, 20(1pt2), 129-136.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19, 1-32.
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., & Christensen, G. (2015). Promoting an open research culture. *Science*, 348(6242), 1422-1425.
- Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, 115(11), 2600-2606.
- OpenScienceCollaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251).
- Pan, S. C., & Rickard, T. C. (2015). Sleep and motor learning: is there room for consolidation? *Psychological bulletin*, 141(4), 812.
- Payne, J. D., Kensinger, E. A., Wamsley, E. J., Spreng, R. N., Alger, S. E., Gibler, K., Schacter, D. L., & Stickgold, R. (2015). Napping and the selective consolidation of negative aspects of scenes. *Emotion*, 15(2), 176.

- Payne, J. D., Schacter, D. L., Propper, R. E., Huang, L. W., Wamsley, E. J., Tucker, M. A., Walker, M. P., & Stickgold, R. (2009, Oct). The role of sleep in false memory formation. *Neurobiol Learn Mem*, 92(3), 327-334. <https://doi.org/10.1016/j.nlm.2009.03.007>
- Payne, J. D., Tucker, M. A., Ellenbogen, J. M., Wamsley, E. J., Walker, M. P., Schacter, D. L., & Stickgold, R. (2012). Memory for semantically related and unrelated declarative information: the benefit of sleep, the cost of wake. *PLoS ONE*, 7(3), e33079. <https://doi.org/10.1371/journal.pone.0033079>
- Peigneux, P., & Smith, C. (2010). Memory processing in relation to sleep. *Principles and practice of sleep medicine*, 5.
- Piosczyk, H., Holz, J., Feige, B., Spiegelhalter, K., Weber, F., Landmann, N., Kuhn, M., Frase, L., Riemann, D., & Voderholzer, U. (2013). The effect of sleep-specific brain activity versus reduced stimulus interference on declarative memory consolidation. *Journal of Sleep Research*, 22(4), 406-413.
- Plihal, W., & Born, J. (1997, Jul). Effects of early and late nocturnal sleep on declarative and procedural memory. *J Cogn Neurosci*, 9(4), 534-547. <https://doi.org/10.1162/jocn.1997.9.4.534>
- Plihal, W., Weaver, S., Molle, M., Fehm, H. L., & Born, J. (1996, Sep). Sensory processing during early and late nocturnal sleep. *Electroencephalogr Clin Neurophysiol*, 99(3), 247-256. [https://doi.org/10.1016/0013-4694\(96\)95539-x](https://doi.org/10.1016/0013-4694(96)95539-x)
- Poldrack, R. A., & Gorgolewski, K. J. (2017). OpenfMRI: Open sharing of task fMRI data. *Neuroimage*, 144, 259-261.
- Prehn-Kristensen, A., Munz, M., Göder, R., Wilhelm, I., Korr, K., Vahl, W., Wiesner, C. D., & Baving, L. (2014, November - December). Transcranial Oscillatory Direct Current Stimulation During Sleep Improves Declarative Memory Consolidation in Children With Attention-deficit/hyperactivity Disorder to a Level Comparable to Healthy Controls. *Brain Stimul*, 7(6), 793-799. <https://doi.org/10.1016/j.brs.2014.07.036>
- Quan, S. F., Budhiraja, R., & Kushida, C. A. (2018, Jul-Aug). Associations Between Sleep Quality, Sleep Architecture and Sleep Disordered Breathing and Memory After Continuous Positive Airway Pressure in Patients with Obstructive Sleep Apnea in the Apnea Positive Pressure Long-term Efficacy Study (APPLES). *Sleep Sci*, 11(4), 231-238. <https://doi.org/10.5935/1984-0063.20180037>
- Rasch, B., & Born, J. (2013, Apr). About sleep's role in memory. *Physiological Reviews*, 93(2), 681-766. <https://doi.org/10.1152/physrev.00032.2012>

- Rasch, B., Buchel, C., Gais, S., & Born, J. (2007, Mar 9). Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science*, 315(5817), 1426-1429.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17347444
- 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.
- Rickard, T. C., Pan, S. C., & Gupta, M. W. (in press). Severe Publication Bias Contributes to Illusory Sleep Consolidation in the Motor Sequence Learning Literature. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- 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., Pascual-Leone, A., & Press, D. Z. (2004). Awareness modifies the skill-learning benefits of sleep. *Current Biology*, 14(3), 208-212.
<https://doi.org/10.1016/j.cub.2004.01.027>
- Roediger III, H. L., & Karpicke, J. D. (2006). The power of testing: basic research and implications for educational practice. *Perspectives on Psychological Science*, 1(3), 181-210.
- Rosenzweig, I., Glasser, M., Polsek, D., Leschziner, G. D., Williams, S. C., & Morrell, M. J. (2015). Sleep apnoea and the brain: a complex relationship. *The Lancet Respiratory Medicine*, 3(5), 404-414.
- Rouder, J. N. (2014). Optional stopping: No problem for Bayesians. *Psychonomic bulletin & review*, 21(2), 301-308.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic bulletin & review*, 16(2), 225-237.
<https://doi.org/doi:10.3758/pbr.16.2.225>
- Rudoy, J. D., Voss, J. L., Westerberg, C. E., & Paller, K. A. (2009). Strengthening individual memories by reactivating them during sleep. *Science*, 326(5956), 1079-1079.
- Saletin, J. M., & Walker, M. P. (2012). Nocturnal mnemonics: sleep and hippocampal memory processing. *Frontiers in Neurology*, 3, 59.
- Schapiro, A. C., McDevitt, E. A., Rogers, T. T., Mednick, S. C., & Norman, K. A. (2018). Human hippocampal replay during rest prioritizes weakly learned information and predicts memory performance. *Nature communications*, 9(1), 3920.

- 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.
- Schichl, M., Ziberi, M., Lahl, O., & Pietrowsky, R. (2011). The influence of midday naps and relaxation-hypnosis on declarative and procedural memory performance. *Sleep and Hypnosis*, 13(1-2), 7-14.
- Schmidt, C., Collette, F., Cajochen, C., & Peigneux, P. (2007). A time to think: circadian rhythms in human cognition. *Cognitive neuropsychology*, 24(7), 755-789.
- Schönauer, M., Geisler, T., & Gais, S. (2014). Strengthening procedural memories by reactivation in sleep. *Journal of Cognitive Neuroscience*, 26(1), 143-153.
- Schönbrodt, F. D., Wagenmakers, E.-J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, 22(2), 322.
- Scullin, M. K. (2013). Sleep, memory, and aging: the link between slow-wave sleep and episodic memory changes from younger to older adults. *Psychology and Aging*, 28(1), 105.
- Siegel, D. J. (2001, Sep). 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.
- Sigman, M., & Dehaene, S. (2005). Parsing a cognitive task: a characterization of the mind's bottleneck. *PLoS Biology*, 3(2), e37.
- Simor, P., Gombos, F., Blaskovich, B., & Bódizs, R. (2017). Long-range alpha and beta and short-range gamma EEG synchronization distinguishes phasic and tonic REM periods. *Sleep*, 41(3), zsx210.
- Simor, P., van Der Wijk, G., Gombos, F., & Kovács, I. (2019). The paradox of rapid eye movement sleep in the lights of oscillatory activity and cortical synchronization during phasic and tonic microstates. *Neuroimage*, 116066.
- Simor, P., Zavecz, Z., Horvath, K., Elteto, 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.
- Soderstrom, N. C., Kerr, T. K., & Bjork, R. A. (2016, Feb). The Critical Importance of Retrieval--and Spacing--for Learning. *Psychol Sci*, 27(2), 223-230.
<https://doi.org/10.1177/0956797615617778>

- Song, S., & Cohen, L. G. (2014). Practice and sleep form different aspects of skill. *Nature communications*, 5, 3407.
- Song, S., Howard, J. H., Jr., & Howard, D. V. (2007). Sleep does not benefit probabilistic motor sequence learning. *Journal of Neuroscience*, 27(46), 12475-12483.
<https://doi.org/10.1523/jneurosci.2062-07.2007>
- Sopp, M. R., Michael, T., & Mecklinger, A. (2018, Nov 1). Effects of early morning nap sleep on associative memory for neutral and emotional stimuli. *Brain Res*, 1698, 29-42.
<https://doi.org/10.1016/j.brainres.2018.06.020>
- Stickgold, R. (2013). Parsing the role of sleep in memory processing. *Current Opinion in Neurobiology*, 23(5), 847-853.
- Stickgold, R., & Walker, M. P. (2005, Aug). Memory consolidation and reconsolidation: what is the role of sleep? *Trends in Neurosciences*, 28(8), 408-415.
- Stickgold, R., & Walker, M. P. (2013). Sleep-dependent memory triage: Evolving generalization through selective processing. *Nature Neuroscience*, 16(2), 139-145.
- [Record #57295 is using a reference type undefined in this output style.]
- Talamini, L. M., Nieuwenhuis, I. L., Takashima, A., & Jensen, O. (2008). Sleep directly following learning benefits consolidation of spatial associative memory. *Learning and Memory*, 15(4), 233-237.
- Tononi, G., & Cirelli, C. (2006, Feb). Sleep function and synaptic homeostasis. *Sleep Medicine Reviews*, 10(1), 49-62. <https://doi.org/10.1016/j.smrv.2005.05.002>
- Tononi, G., & Cirelli, C. (2014). Sleep and the price of plasticity: from synaptic and cellular homeostasis to memory consolidation and integration. *Neuron*, 81(1), 12-34.
- 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.
- Tucker, M., McKinley, S., & Stickgold, R. (2011). Sleep optimizes motor skill in older adults. *Journal of the American Geriatrics Society*, 59(4), 603-609.
- Tucker, M., Morris, C. J., Morgan, A., Yang, J., Myers, S., Pierce, J. G., Stickgold, R., & Scheer, F. A. (2017). The relative impact of sleep and circadian drive on motor skill acquisition and memory consolidation. *Sleep*, 40(4), zsx036.

- Uguccioni, G., Pallanca, O., Golmard, J. L., Dodet, P., Herlin, B., Leu-Semenescu, S., & Arnulf, I. (2013). Sleep-related declarative memory consolidation and verbal replay during sleep talking in patients with REM sleep behavior disorder. *PLoS ONE*, 8(12), e83352. <https://doi.org/10.1371/journal.pone.0083352>
- Ujma, P. P. (2021). Sleep spindles and general cognitive ability—A meta-analysis. *Sleep Spindles & Cortical Up States*, 2(1), 1-17.
- Ujma, P. P., Gombos, F., Genzel, L., Konrad, B. N., Simor, P., Steiger, A., Dresler, M., & Bódizs, R. (2015). A comparison of two sleep spindle detection methods based on all night averages: individually adjusted vs. fixed frequencies. *Frontiers in Human Neuroscience*, 9, 52.
- Van den Berg, N., Al-Kuwatli, J., Paulin, J., Ray, L., Owen, A., & Fogel, S. (2019). Sleep preferentially enhances memory for a cognitive strategy but not the implicit motor skills used to acquire it. *Neurobiology of Learning and Memory*, 161, 135-142.
- Vertes, R. P., & Siegel, J. M. (2005). Time for the sleep community to take a critical look at the purported role of sleep in memory processing. *Sleep*, 28(10), 1228-1229.
- Viczko, J., Sergeeva, V., Ray, L. B., Owen, A. M., & Fogel, S. M. (2018). Does sleep facilitate the consolidation of allocentric or egocentric representations of implicitly learned visual-motor sequence learning? *Learning and Memory*, 25(2), 67-77.
- Vorster, A. P., & Born, J. (2015). Sleep and memory in mammals, birds and invertebrates. *Neuroscience and Biobehavioral Reviews*, 50, 103-119.
- Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., & Epskamp, S. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic bulletin & review*, 25(1), 35-57.
- Wagner, U., Hallschmid, M., Rasch, B., & Born, J. (2006). Brief sleep after learning keeps emotional memories alive for years. *Biological Psychiatry*, 60(7), 788-790.
- Walker, M. P. (2005). A refined model of sleep and the time course of memory formation. *Behavioral and Brain Sciences*, 28(1), 51-104.
- Walker, M. P., Brakefield, T., Hobson, J. A., & Stickgold, R. (2003, Oct 9). Dissociable stages of human memory consolidation and reconsolidation. *Nature*, 425(6958), 616-620.
- Wallace, A., & Bucks, R. S. (2013). Memory and obstructive sleep apnea: a meta-analysis. *Sleep*, 36(2), 203-220.

- Wamsley, E. J. (2019). Memory consolidation during waking rest. *Trends in cognitive sciences*, 23(3), 171-173.
- Wiklund-Hornqvist, C., Jonsson, B., & Nyberg, L. (2014, Feb). Strengthening concept learning by repeated testing. *Scandinavian Journal of Psychology*, 55(1), 10-16.
<https://doi.org/10.1111/sjop.12093>
- Wilhelm, I., Diekelmann, S., & Born, J. (2008). Sleep in children improves memory performance on declarative but not procedural tasks. *Learn Mem*, 15(5), 373-377.
- Wilson, J. K., Baran, B., Pace-Schott, E. F., Ivry, R. B., & Spencer, R. (2012). Sleep modulates word-pair learning but not motor sequence learning in healthy older adults. *Neurobiology of Aging*, 33(5), 991-100.
- Wulff, K., Gatti, S., Wettstein, J. G., & Foster, R. G. (2010). Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease. *Nature Reviews Neuroscience*, 11(8), 589-599.
- Xue, G., Dong, Q., Chen, C., Lu, Z., Mumford, J. A., & Poldrack, R. A. (2010, Oct 1). Greater neural pattern similarity across repetitions is associated with better memory. *Science*, 330(6000), 97-101. <https://doi.org/10.1126/science.1193125>
- Yaroush, R., Sullivan, M. J., & Ekstrand, B. R. (1971, Jun). Effect of sleep on memory. II. Differential effect of the first and second half of the night. *J Exp Psychol*, 88(3), 361-366.
<https://doi.org/10.1037/h0030914>
- Young, D. R., & Bellezza, F. S. (1982). Encoding variability, memory organization, and the repetition effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8(6), 545-559.