

Targeted Memory Reactivation During Sleep Improves Next-Day Problem Solving

Kristin E. G. Sanders , Samuel Osburn, Ken A. Paller ,
and Mark Beeman

Department of Psychology, Northwestern University

Psychological Science
1–9

© The Author(s) 2019

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0956797619873344

www.psychologicalscience.org/PS



Abstract

Many people have claimed that sleep has helped them solve a difficult problem, but empirical support for this assertion remains tentative. The current experiment tested whether manipulating information processing during sleep impacts problem incubation and solving. In memory studies, delivering learning-associated sound cues during sleep can reactivate memories. We therefore predicted that reactivating previously unsolved problems could help people solve them. In the evening, we presented 57 participants with puzzles, each arbitrarily associated with a different sound. While participants slept overnight, half of the sounds associated with the puzzles they had not solved were surreptitiously presented. The next morning, participants solved 31.7% of cued puzzles, compared with 20.5% of uncued puzzles (a 55% improvement). Moreover, cued-puzzle solving correlated with cued-puzzle memory. Overall, these results demonstrate that cuing puzzle information during sleep can facilitate solving, thus supporting sleep's role in problem incubation and establishing a new technique to advance understanding of problem solving and sleep cognition.

Keywords

problem solving, incubation, sleep, targeted memory reactivation, restructuring, creative cognition

Received 3/1/19; Revision accepted 8/9/19

Some problems are so difficult that one reaches an impasse rather than a solution. Remarkably, *incubation*—a pause in actively working on a problem—can increase the chance of finding a solution (Sio & Ormerod, 2009). After incubation, people often report having an insight—a solution comes to them suddenly, seemingly from nowhere. In these cases, some unconscious reorganization or restructuring of the problem may have transpired during the intervening period (Schooler & Melcher, 1995).

Although research on incubation has focused on awake periods, some empirical evidence tentatively suggests that problem incubation may be especially beneficial during sleep. In one experiment, participants who slept were more likely to discover a hidden shortcut to a tedious numerical task compared with participants who spent an equivalent time awake (Wagner, Gais, Haider, Verleger, & Born, 2004). Similarly, participants in another experiment were more likely to use solution hints from an allegedly unrelated task to solve

tricky word problems if they took an intervening nap with REM sleep (Cai, Mednick, Harrison, Kanady, & Mednick, 2009). However, other recent studies failed to find sleep benefits for problem solving (Brod, Pöhlchen, Täumer, Gais, & Schönauer, 2018; Schönauer et al., 2018), suggesting that unspecified study parameters are critical, as in the broader literature on incubation (Sio & Ormerod, 2009).

Carefully and robustly characterizing sleep incubation may help elucidate problem incubation more generally by targeting processing that is primarily outside conscious control. Sleep has a demonstrated impact on subsequent waking performance, even though what occurs during sleep seems less volitional than what occurs during waking hours. For example, sleep appears to strengthen and

Corresponding Author:

Kristin E. G. Sanders, Northwestern University, Department of Psychology, Swift Hall 102, 2029 Sheridan Rd., Evanston, IL 60208
E-mail: kgrunewald@u.northwestern.edu

potentially transform memory (Paller, Mayes, Antony, & Norman, in press; Payne, 2011). Indeed, these memory-consolidation processes could be conducive to problem incubation (Stickgold & Walker, 2013), and similar paradigms could be adapted to problem solving.

Numerous studies in humans and other animals demonstrate better memory after sleep compared with a similar time awake (Rasch & Born, 2013). Replay of recently learned memories during sleep putatively strengthens memory. Indeed, neurons that fire when rats initially learn a maze fire again in similar patterns while the rats sleep (Wilson & McNaughton, 1994). In humans, hippocampal areas active during daytime route learning reactivate during sleep, and the degree of sleep reactivation correlates with route retrieval after sleep (Peigneux et al., 2004).

Sleep-related memory processes not only strengthen but also can optimize, organize, and transform information (Ellenbogen, Hu, Payne, Titone, & Walker, 2007; Stickgold & Walker, 2013). These processes could be particularly beneficial when solving problems that require some form of restructuring. Broadly construed, restructuring implies forming a new problem representation or approach by ignoring compelling but incorrect ideas, combining information previously viewed as unrelated, or otherwise reorganizing problem elements. Therefore, we examined problems thought to require restructuring in our experiment.

Memory can be selectively modified during sleep using cues that trigger or bias memory replay. For example, odor cues can be associated with learning some spatial information, and subsequently delivering the cues again during slow-wave sleep (SWS) produces better learning compared with control conditions (Rasch, Büchel, Gais, & Born, 2007). Moreover, in a targeted-memory-reactivation (TMR) paradigm, specific knowledge can be selectively cued. For example, if participants learn object locations, each associated with a unique sound, and then half of those sounds are quietly presented during SWS without disrupting sleep (Rudoy, Voss, Westerberg, & Paller, 2009), participants on awakening can recall locations more accurately for cued compared with uncued objects. TMR can apparently bias memory reactivation and consolidation, yielding a disproportionate benefit for cued items. Indeed, playing task-related sound cues to sleeping rats influences the firing of task-related neurons (Bendor & Wilson, 2012). TMR can also reorganize memories, such as improving explicit recognition of a learned tone sequence (Cousins, El-Dereby, Parkes, Hennies, & Lewis, 2014).

If memories are reorganized through reactivation during sleep, then sleep replay may also contribute to problem incubation by facilitating reorganization of the

problem representation. With this possibility in mind, we adapted the TMR paradigm to test whether playing puzzle-associated sound cues during sleep would improve people's ability to solve cued compared with uncued puzzles in the morning after they failed to solve them the night before.

Prior TMR studies demonstrate memory strengthening and reorganization using cues presented during SWS (Hu, Cheng, Chiu, & Paller, 2019); therefore, we used the same experimental strategy here. Our procedure does not address whether other sleep stages, such as REM, contribute to problem solving. Indeed, some anecdotes link problem solving with dreaming (Mazzarello, 2000), and REM has been linked with a broadening of semantic associations (Stickgold, Scott, Rittenhouse, & Hobson, 1999) and the incorporation of purportedly unrelated hints into problem solutions (Cai et al., 2009). Of course, SWS and REM may both be relevant, if for example, memories for problems are reactivated during SWS and then reorganized during REM (Giuditta et al., 1995; Lewis, Knoblich, & Poe, 2018). Although much remains to be learned about these sleep stages, our intention was not to identify which sleep stage is relevant but rather to determine generally whether sleep could be more strongly linked with problem solving.

In two consecutive evening sessions, participants attempted to solve puzzles one at a time. Each puzzle was associated with a unique sound cue (Fig. 1). Evening sessions continued until participants failed to solve 6 puzzles, yielding 12 incubation-ready puzzles across two evenings. Each night, participants slept in their own homes, and sounds from 3 of the incubation-ready puzzles from that evening were played during SWS. Each subsequent morning, participants returned to the lab to try to solve the 6 puzzles they did not solve the night before. To overcome typical methodological challenges of incubation and sleep studies, we employed a within-subjects contrast, included a large number of participants, presented a large set of puzzles that were distinctive from each other (to allow better incubation), and spread data collection over two nights to avoid potential confusion between puzzles (see the Supplemental Material available online for further details). We predicted that participants would solve more cued than uncued puzzles.

Method

Participants

A total of 61 participants (38 female) from the Northwestern University community enrolled in the experiment (age: range: 18–29 years, $M = 20.01$, $SD = 1.92$).

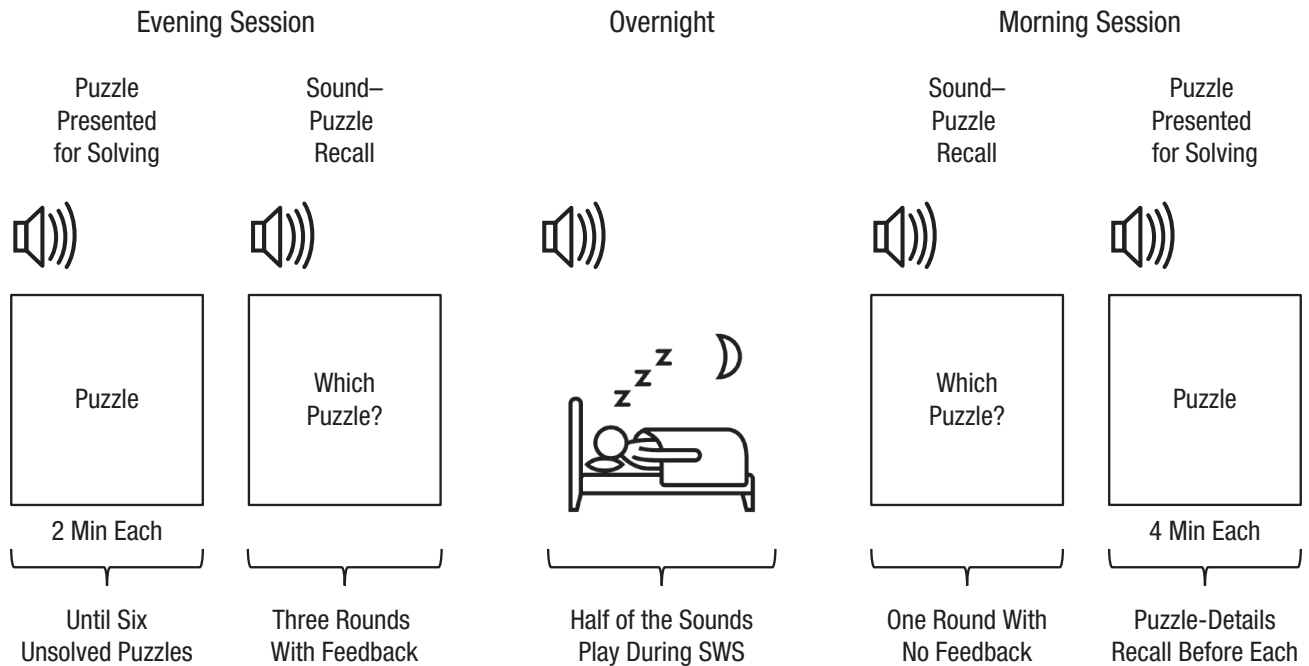


Fig. 1. Experimental procedure (repeated twice). Each evening, participants attempted to solve puzzles while a distinct 15-s sound clip looped for each one. New puzzles were presented until six puzzles remained unsolved. Then participants were tested, with feedback, on their recall for which puzzle went with each sound. Participants took home a portable electroencephalogram sleep monitoring and cuing system. When it detected slow-wave sleep (SWS), the system presented sound cues associated with half of the unsolved puzzles. The following morning, participants returned to the laboratory to complete recall tests on puzzle sounds and puzzle details, after which they attempted to solve the puzzles that they had failed to solve the previous night. Analyses contrasted solving rates for cued versus uncued puzzles.

Our research design and predicted effects were novel, so we decided in advance to use a sample size of 60 (for scheduling reasons, we actually tested 61). Our goal was to provide robust power and reproducibility in a within-subjects design, so we more than doubled the standard number of participants used in prior research. Participants were fluent English speakers who reported no history of neurological or sleep disorders. We excluded data from 3 participants who did not complete all four sessions (two evening, two morning) and a fourth participant who did not receive sound cues because of equipment failure. Thus, data from 57 participants were analyzed. Participants provided informed consent and were monetarily compensated. Procedures were approved by the Northwestern University Institutional Review Board.

Materials and procedure

Participants completed four sessions over 3 days—two evening sessions (mean start time = 6:01 p.m.) and two morning sessions (mean start time = 8:47 a.m.)—with overnight sleep monitoring and cuing in participants' own homes after each evening session.

Evening session. During the first evening session, we instructed participants on using a sleep-monitoring device and introduced the procedure and the types of puzzles. Then, in both evening sessions, participants attempted puzzles one at a time. Before each puzzle, a randomly paired sound clip played once (15 s), to familiarize participants with the sound. Participants then read the puzzle. If it was familiar, the experimenter skipped the puzzle; otherwise, participants attempted to solve the puzzle for 2 min while the sound continuously looped. During the solving period, if participants had an idea, they pressed the space bar and spoke their solution aloud. If the offered solution was incorrect, participants used the remaining time to continue working on the puzzle. If participants correctly solved the puzzle, it was excluded from the remainder of the experiment and replaced with a new puzzle. If participants did not solve the puzzle within 2 min, the sound played one more time, and participants were instructed to try to memorize the sound-puzzle pairing. Puzzle order was randomized for each participant, and participants continued to attempt puzzles until 6 puzzles remained unsolved. On average, participants attempted 7.24 puzzles per evening session ($SD = 1.69$). This solving rate was by design

(because we wanted to cue unsolved puzzles) and is typical for puzzles of this genre. For example, when not given hints, participants' solving rates for the nine-dot problem range from 0% (e.g., MacGregor, Ormerod, & Chronicle, 2001) to 10% when given 10 min for the attempt (Chein, Weisberg, Streeter, & Kwok, 2010).

After the puzzle-solving phase, memory for the sound–puzzle pairings was tested to further reinforce the association between each of the six remaining puzzles and its corresponding sound. In three rounds, participants heard each sound and reported the paired puzzle. Then, regardless of accuracy, they saw the correct puzzle presented with the sound. By the third round, participants correctly recalled 88.0% of puzzles ($SD = 12.8\%$, 95% confidence interval, or CI = [84.6%, 91.4%]). Accuracy was comparable for puzzles that were later cued ($M = 87.9\%$, $SD = 20.0\%$, 95% CI = [82.6%, 93.2%]) and those that were not cued ($M = 87.5\%$, $SD = 15.1\%$, 95% CI = [83.5%, 91.5%]), $t(56) = 0.15$, $p = .882$, $d_z = 0.02$, 95% CI = [-0.40, 0.44]. At the end of each evening session, participants were told to avoid thinking about or trying to solve the puzzles until the morning session, when they would attempt each unsolved puzzle again.

Half of the unsolved puzzles were automatically selected for the cued condition using a programmed algorithm, and the corresponding sounds were presented during sleep. Neither the experimenter nor the participant knew which puzzles were cued. The algorithm was designed to select puzzles such that, across participants, each puzzle was presented equally often in the cued and uncued conditions (puzzles were cued anywhere from 40% to 63% of the times they were used; given the imperatives of maintaining double-blind conditions and variations in puzzle solution rates, a fully balanced design was unobtainable).

Overnight. After the evening session, participants took a sleep monitoring and cuing system (SMCS) with them so sound cues could be administered in their own home. The SMCS includes a transmitter and algorithm (Zeo, Boston, MA) as well as a laptop computer adapted to wirelessly receive signals from the transmitter. Before going to sleep each night, participants placed the laptop on a level surface near their bed, snapped three single-use silver/silver-chloride electrodes into the Zeo wireless transmitter, filled each sensor with a conductive electrolyte gel (signa gel), and affixed the adhesive transmitter to their forehead.

Every 30 s, the SMCS employed an algorithm to determine a participant's sleep stage (see Sleep Monitoring and Cuing System Validation in the Supplemental Material). When the SMCS first detected sleep, it began

quietly playing pink noise, which reduced the likelihood that the sound cues would provoke arousal. When the SMCS detected SWS for at least 1.5 min within a 2-min period, it presented a designated sound for 3 min before switching to the next sound. However, if the SMCS detected a stage other than SWS, the sound immediately stopped until SWS was again detected. If participants awoke and heard sounds playing, they pressed a button to stop the sound. Sounds resumed when SWS was detected again. The pink noise continued throughout the night. We excluded sounds that were reportedly heard during the night, but this occurred very rarely (3.5% of the trials), suggesting that the sound level was low enough to not provoke awakenings.

The SMCS thus allowed us to present sounds in participants' typical sleeping environments without disrupting their sleep. Allowing participants to sleep in their own beds rather than in a laboratory setting increased comfort levels and sleep quality, reduced the need for an adaptation night, and more closely mimicked real-world problem-solving situations. Previous validation efforts (e.g., Shambroom, Fábregas, & Johnstone, 2012) showed that the SMCS sleep-staging algorithm can reliably classify sleep, in a manner resembling standard polysomnographic staging. We did, however, introduce the modification of wet electrodes, which generally would improve signal quality compared with the dry electrodes used for validation (see the Supplemental Material). There was a strong, positive correlation between the amount of participants' self-reported and SMCS-detected sleep, $r(55) = .82$, $p < .001$, 95% CI = [.71, .89]. On average, participants self-reported spending 411 min in bed (95% CI = [395, 427]) and getting 395 min of sleep per night (95% CI = [380, 410]). Calculating from the first detection of sleep to the last detection of sleep, the SMCS detected an average of 393 min of sleep per night (95% CI = [377, 409]).

Morning session. Each morning, participants returned to the lab with the SMCS. They first answered a questionnaire about which sounds (if any) they heard during the night and whether they solved any of the puzzles between the evening and morning sessions. They then heard each of the six sound cues while looking at a blank screen and reported which puzzle was paired with each sound. Participants correctly remembered 71.9% ($SD = 21.3\%$, 95% CI = [66.2%, 77.5%]) of the sound–puzzle pairings. Accuracy was similar for cued puzzles ($M = 70.3\%$, $SD = 27.7\%$, 95% CI = [63.0%, 77.7%]) and uncued puzzles ($M = 73.5\%$, $SD = 25.4\%$, 95% CI = [66.7%, 80.2%]), $t(56) = -0.80$, $p = .428$, $d_z = -0.12$, 95% CI = [-0.71, 0.48]. Next, participants were tested on the puzzles one at a time. For each puzzle, the title was shown, and participants reported aloud

all the details they could remember. Then the full text of the puzzle appeared for 4 min while participants attempted to solve the puzzle. If participants did not solve the puzzle within the single 4-min attempt, they reported any final thoughts. Then the experimenter explained the solution to the puzzle, to reduce the likelihood that participants continued to think about it.

Stimuli. A total of 42 puzzles were gathered from various websites, books, and previous literature on solving problems with insight, and these puzzles were adapted to achieve optimal difficulty levels. Puzzles were specifically chosen to elicit an initial incorrect problem representation or to require combining knowledge in an uncommon way to solve the puzzle. Each puzzle had a unique solution, and the information needed to solve each puzzle was either common knowledge or included in the puzzle. Puzzles were selected from four categories (rebus, matchstick, spatial, and verbal; see Fig. S1 in the Supplemental Material), and we had participants sleep on at least one puzzle of each type per night to maximize distinctiveness. These 42 puzzles formed a pool from which we selected puzzles until a participant had 6 unsolved puzzles per night (12 total over the course of the experiment—three rebus, three spatial, three matchstick, and three verbal).

For each participant, each puzzle was randomly paired with a unique sound clip. We used 42 nonverbal sound clips representing a variety of music genres and instrumentation selected to be highly distinctive from each other. Sound clips were (a) unrelated to the content of the puzzles, so they did not provide hints to the solution; (b) unlikely to be encountered in participants' everyday activities, so cuing would occur only during sleep; and (c) generally unfamiliar to the participants, so the sound would uniquely reactivate its paired puzzle. Each sound was edited into a 15-s clip consisting of 8 s to 12 s of sound plus 3 s to 7 s of silence (to reduce habituation).

Data preparation. To ensure that any effects we found were driven by cue-related processing during sleep, we excluded data on the basis of preplanned exclusion criteria. Including all usable data in the analyses without these exclusions did not alter the results (see Table S1 in the Supplemental Material). We excluded data when (a) a puzzle was solved after the evening session but before the participant slept (1.3% of trials), (b) the participant pressed a button to turn off the sound during the night or reported in the morning questionnaire hearing the specific cued sound (3.5% of trials), or (c) the experimenter incorrectly marked the puzzle as correctly solved during the morning session (0.3% of trials). Occasionally, sounds were not presented as planned because the SMCS detected very little or no SWS (perhaps because the

electrodes did not adhere properly). If some, but not all, sounds were played during sleep, all puzzles whose sounds were not played were marked as uncued and remained in the analysis (3.4% of trials). If no sounds were played (7 nights of 114), all puzzles associated with these nights were excluded from analysis (6.2% of trials), given that sleep on such nights may have been qualitatively different from sleep on other nights.

To quantify memory for puzzle details, we created a scoring guide that listed all the elements in each puzzle. Responses were audio-recorded and later transcribed (except for responses that were scored from the experimenter's notes from 1 participant who did not consent to being recorded). Raters blind to cuing status scored each response for the presence or absence of each element. We restricted our analysis to puzzle elements relevant to the solution that were not part of the puzzle's title. Four rebus puzzles were excluded from this analysis because their critical elements were completely contained in the puzzle's title.

Results

Our primary contrast compared the percentage of puzzles solved in morning sessions in the cued and uncued conditions (Figs. 2a and 2b). Across the two mornings, participants reliably solved more cued puzzles ($M = 31.7\%$, $SD = 28.3\%$, 95% CI = [24.2%, 39.2%]) than uncued puzzles ($M = 20.5\%$, $SD = 18.9\%$, 95% CI = [15.5%, 25.5%]), $t(56) = 3.22$, $p = .002$, $d_z = 0.47$, 95% CI = [0.14, 0.79]. Using the uncued-puzzle solution rate as a baseline, we calculated that participants solved 55% more puzzles when the corresponding sound was surreptitiously played while they slept. Solutions were spread relatively evenly across the 4-min solving period, with 22% to 29% of solutions occurring in each 1-min segment, showing that solutions did not disproportionately appear quickly. Data analyzed separately for the two nights showed minimal differences (see Fig. S2 in the Supplemental Material). The design included two nights not specifically to enable this comparison but because combining data across multiple nights increased statistical power for the primary contrast while also maintaining the distinctiveness of puzzles and sounds for the participants.

We also analyzed data separately for different types of puzzles. The purpose of including multiple puzzle types was to allow for distinct sound-puzzle associations and to reduce any possible confusion between the puzzles. However, to test the generalizability of the cuing effect across puzzles and puzzle types, we used puzzles as the random factor in a 2 (cuing) \times 4 (puzzle type) mixed analysis of variance (ANOVA). We found a main effect of cuing, $F(1, 37) = 9.85$, $p = .003$, $\eta_p^2 =$

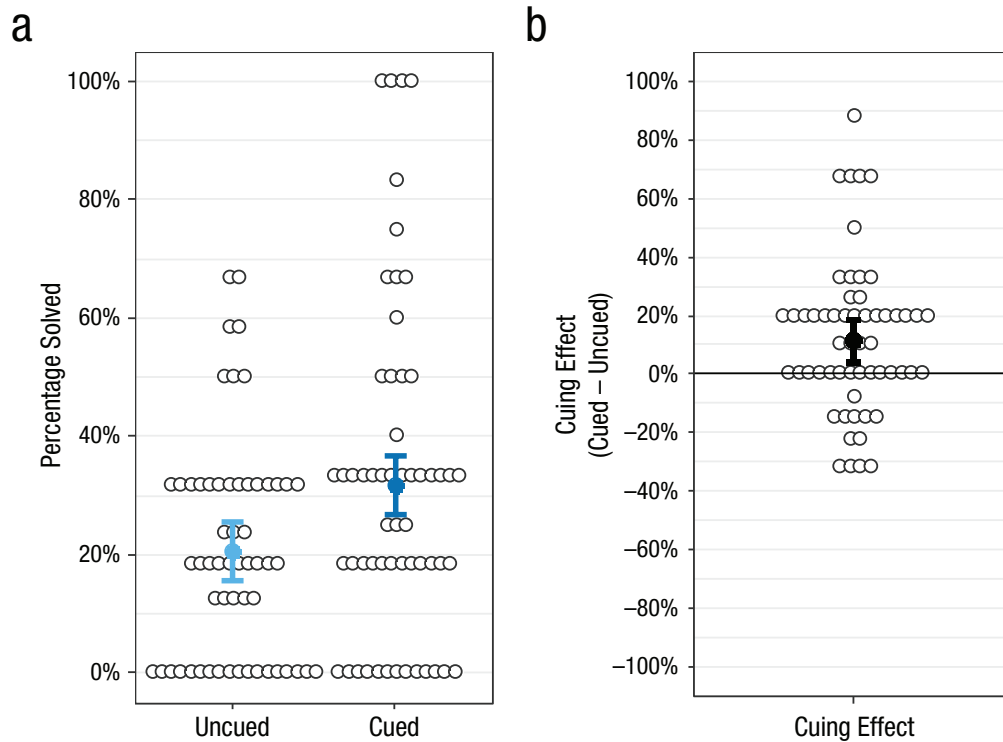


Fig. 2. Solving rate in the morning sessions for uncued and cued puzzles (a) and cuing effect (b; the percentage of cued puzzles solved – the percentage of uncued puzzles solved). Open circles are individual data points, and filled circles are group averages. Error bars represent (a) 95% within-subjects confidence intervals and (b) standard 95% confidence intervals.

.21; no main effect of puzzle type, $F(3, 37) = 0.96$, $p = .422$, $\eta_p^2 = .07$; and no interaction between puzzle type and cuing, $F(3, 37) = 0.52$, $p = .673$, $\eta_p^2 = .04$. Participants exhibited a similarly higher solving rate for cued compared with uncued puzzles for all four types: matchstick (11.3%), rebus (11.5%), spatial (16.7%), and verbal (4.8%) puzzles (see Table S3 in the Supplemental Material).

We also assessed memory for puzzles. However, puzzles were selected to achieve an appropriate level of solvability and so were not ideal for testing memory. Because puzzles ranged from being very easy to recall to being quite difficult to recall, and because responses were not equivalently graded by degree of recall for the different puzzles, the memory test may have low sensitivity. Indeed, in the morning, participants recalled a similar number of details for cued and uncued puzzles ($M = 76.8\%$, $SD = 17.4\%$, 95% CI = [72.1%, 81.4%] vs. $M = 74.3\%$, $SD = 15.7\%$, 95% CI = [70.2%, 78.5%], respectively), $t(56) = 0.84$, $p = .405$, $d_z = 0.15$, 95% CI = [-0.07, 0.37].

Although the retrieval of puzzle details was not required for solving, given that each puzzle was presented when participants attempted to solve it, sleep-related memory processes may nonetheless have

benefited problem solving. Thus, we examined the relation among cuing, puzzle solving, and puzzle memory with exploratory analyses (uncorrected for multiple comparisons). Across participants, overall puzzle-recall reliably correlated with overall solving, $r(55) = .27$, $p = .039$, 95% CI = [.01, .50]. Additionally, recall of cued puzzles positively correlated with solving of cued puzzles, $r(55) = .37$, $p = .004$, 95% CI = [.12, .58], whereas recall of uncued puzzles did not correlate with solving of uncued puzzles, $r(55) = .07$, $p = .61$, 95% CI = [-.20, .32]. The difference between these two correlations was not reliable, $z = 1.67$, $p = .094$, and the overall cuing effect on memory did not correlate with the overall cuing effect on solving, $r(55) = .11$, $p = .424$, 95% CI = [-.16, .36]. Although these correlational findings are tentative, they are consistent with the hypothesis that processing related to memory reactivation and consolidation during sleep may also contribute to problem-solving incubation.

Discussion

As predicted, participants solved more cued than uncued puzzles after one night of sleep during which puzzle-associated sounds were played. We infer that these sounds

reactivated puzzle memories during sleep, given that associated puzzle information must have been accessed in order to differentially alter solving rates—the sounds per se provided no information relevant to finding solutions. This cuing effect occurred within subjects, was robust to data-exclusion criteria, and emerged across numerous puzzle types. The most parsimonious explanation is that playing sounds influenced the utility of incubation via a relative increase in reactivation of associated puzzle information, which helped participants eventually reach the solution. More precise sleep-physiology data would be needed to implicate memory reactivation more directly, as in several recent studies (Belal et al., 2018; Schönauer et al., 2017; Wang et al., 2019). Nonetheless, these results are the first to successfully extend the TMR methodology to problem solving.

We propose that memory reactivation during sleep can allow people to restructure their conception of a problem. That is, sleep can be beneficial both for producing enduring memories and for altering stored information (Lewis et al., 2018), in keeping with how sleep (Ellenbogen et al., 2007) and TMR (Cousins et al., 2014) have been found to help people recognize new patterns in recently learned information. Reactivation may preferentially enhance memory of the problem in a way that allows solvers to creatively forge new links between problem elements as well as between problem elements and relevant prior knowledge. If so, incubation during sleep would most likely benefit the solving of problems that require restructuring, as is typical for problems that are often solved with sudden insight. Future research should address the extent to which sleep might facilitate solving as a function of how much and what type of restructuring is engaged.

A slightly different interpretation is that reactivating puzzle memories during sleep allows puzzle elements to be processed with low attentional selectivity. Withdrawing selective attention from closely related (but ineffective) associations could enable broader associations to percolate until new connections and structures are forged. Indeed, participants in one study generated more creative ideas in response to a prompt when that prompt was cued by an odor during sleep, suggesting that broader associations were facilitated (Ritter, Strick, Bos, Van Baaren, & Dijksterhuis, 2012).

Although presenting cues during SWS typically enhances memory (Hu et al., 2019), we did not find general improvement for recalling puzzle details. Nonetheless, the tentative relation found between recall and solving for cued puzzles is consistent with the notion that reactivation of puzzle memories during sleep contributes to problem solving. The memory assessment may lack the sensitivity to demonstrate memory improvement; alternatively, memory reactivation during

sleep may weaken specific details in the process of reorganization, as has been found in studies of gist memory (Diekelmann, Born, & Wagner, 2010; Payne et al., 2009).

Although past research on problem solving after sleep versus during wakefulness has sometimes yielded minimal effects (Schönauer et al., 2018; Sio, Monaghan, & Ormerod, 2013), the current study clearly implicates sleep in problem solving through targeting a sleep-related process in a powerful within-subjects design. Furthermore, design characteristics such as a large sample of participants, a dozen puzzles of several types, and ample overnight targeted reactivation may have also contributed to the robust findings produced in the present experiment.

Our results suggest that increasing the amount of puzzle-specific memory reactivation during sleep through TMR improves problem solving. However, we cannot determine whether TMR boosted puzzle reactivation for cued puzzles above naturally occurring levels, biased reactivation away from uncued puzzles, or perhaps both. Conceivably, a relative benefit for cued puzzles may parallel a decreased incubation effect for uncued puzzles. The present results also do not address how wakefulness and sleep incubation periods may compare or whether TMR during wakefulness could also be beneficial. During quiet awake periods, for example, TMR can produce better memory for subliminally cued compared with uncued items (Tambini, Berners-Lee, & Davachi, 2017).

Although we endeavored to present cues primarily during SWS, studies with standard polysomnographic recordings would provide further neurophysiological information. We cannot address whether REM might also be beneficial. Indeed, Lewis and colleagues (2018) proposed that a combination of SWS and REM sleep is necessary to improve creative problem solving. Furthermore, REM sleep has been associated with increased priming of weakly related concepts (Stickgold et al., 1999), improved solving of newly presented anagrams without incubation (Walker, Liston, Hobson, & Stickgold, 2002), and utilization of hints to solve problems (Cai et al., 2009). Future experiments could thus use the TMR method to search for various physiological signals that may reflect processing relevant for creative problem solving.

The present findings establish a novel convergence of the sleep, memory, and problem-solving fields. By demonstrating for the first time that targeted reactivation of puzzle memories during sleep improves problem solving the following morning, we add to a growing literature recognizing sleep as useful for both strengthening and reorganizing memory. Future extensions of this research can further shape our understanding of the hidden benefits of sleep for cognition and of the

mechanisms through which incubation leads to creative solutions when one is asleep or awake.

Action Editor

Caren Rotello served as action editor for this article.

Author Contributions

K. E. G. Sanders, M. Beeman, K. A. Paller, and S. Osburn designed the experiment, and K. E. G. Sanders and S. Osburn conducted the experiment. K. E. G. Sanders analyzed the data. K. E. G. Sanders and M. Beeman wrote the manuscript with input from K. A. Paller.

ORCID iDs

Kristin E. G. Sanders  <https://orcid.org/0000-0001-8310-059X>
Ken A. Paller  <https://orcid.org/0000-0003-4415-4143>

Acknowledgments

We thank Susan Florczak, Todd Anderson, Daniel Wetmore, and Sheepdog Sciences for help with the sleep monitoring and cuing system devices. We additionally thank Katherine George, Kira Riley, Megan Conlon, and Lena Goren for data-collection assistance and Sophia McCullough and Carlie Cope for memory-coding assistance.

Declaration of Conflicting Interests

The author(s) declared no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This work was supported by a grant from the National Institutes of Health's National Institute of Child Health and Human Development (R03HD087111) awarded to M. Beeman.

Supplemental Material

Additional supporting information can be found at <http://journals.sagepub.com/doi/suppl/10.1177/0956797619873344>

Open Practices

The experiment reported in this article was not preregistered; neither the data nor the materials have been made available on a permanent third-party archive. Requests for the data or materials can be e-mailed to the first author at kgrunewald@u.northwestern.edu.

References

- Belal, S., Cousins, J., El-Derey, W., Parkes, L., Schneider, J., Tsujimura, H., . . . Lewis, P. (2018). Identification of memory reactivation during sleep by EEG classification. *NeuroImage*, *176*, 203–214. doi:10.1016/j.neuroimage.2018.04.029
- Bendor, D., & Wilson, M. A. (2012). Biasing the content of hippocampal replay during sleep. *Nature Neuroscience*, *15*, 1439–1444. doi:10.1038/nn.3203
- Brodts, S., Pöhlchen, D., Täumer, E., Gais, S., & Schönauer, M. (2018). Incubation, not sleep, aids problem-solving. *Sleep*, *41*(10), Article zsy155. doi:10.1093/sleep/zsy155
- Cai, D. J., Mednick, S. A., Harrison, E. M., Kanady, J. C., & Mednick, S. C. (2009). REM, not incubation, improves creativity by priming associative networks. *Proceedings of the National Academy of Sciences, USA*, *106*, 10130–10134. doi:10.1073/pnas.0900271106
- Chein, J. M., Weisberg, R. W., Streeter, N. L., & Kwok, S. (2010). Working memory and insight in the nine-dot problem. *Memory & Cognition*, *38*, 883–892. doi:10.3758/MC.38.7.883
- Cousins, J. N., El-Derey, W., Parkes, L. M., Hennies, N., & Lewis, P. A. (2014). Cued memory reactivation during slow-wave sleep promotes explicit knowledge of a motor sequence. *The Journal of Neuroscience*, *34*, 15870–15876. doi:10.1523/JNEUROSCI.1011-14.2014
- Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. *Behavioural Brain Research*, *208*, 425–429. doi:10.1016/j.bbr.2009.12.021
- Ellenbogen, J. M., Hu, P. T., Payne, J. D., Titone, D., & Walker, M. P. (2007). Human relational memory requires time and sleep. *Proceedings of the National Academy of Sciences, USA*, *104*, 7723–7728. doi:10.1073/pnas.0700094104
- Giuditta, A., Ambrosini, M. V., Montagnese, P., Mandile, P., Cotugno, M., Zucconi, G. G., & Vescia, S. (1995). The sequential hypothesis of the function of sleep. *Behavioural Brain Research*, *69*, 157–166. doi:10.1016/0166-4328(95)00012-1
- Hu, X., Cheng, L., Chiu, M. H., & Paller, K. A. (2019). *Promoting memory consolidation during sleep: A systematic review and meta-analysis of targeted memory reactivation*. Unpublished manuscript.
- Lewis, P. A., Knoblich, G., & Poe, G. (2018). How memory replay in sleep boosts creative problem-solving. *Trends in Cognitive Sciences*, *22*, 491–503. doi:10.1016/j.tics.2018.03.009
- MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (2001). Information processing and insight: A process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 176–201. doi:10.1037/0278-7393.27.1.176
- Mazzarello, P. (2000). What dreams may come? *Nature*, *408*, 523. doi:10.1038/35046170
- Paller, K. A., Mayes, A. R., Antony, J. W., & Norman, K. A. (in press). Replay-based consolidation governs enduring memory storage. In M. S. Gazzaniga, G. R. Mangun, & D. Poeppel (Eds.), *The cognitive neurosciences* (6th ed.). Cambridge, MA: MIT Press.
- Payne, J. D. (2011). Learning, memory, and sleep in humans. *Sleep Medicine Clinics*, *6*, 15–30. doi:10.1016/j.jsmc.2010.12.005
- Payne, J. D., Schacter, D. L., Propper, R. E., Huang, L.-W., Wamsley, E. J., Tucker, M. A., . . . Stickgold, R. (2009). The role of sleep in false memory formation. *Neurobiology of Learning and Memory*, *92*, 327–334. doi:10.1016/j.nlm.2009.03.007

- Peigneux, P., Laureys, S., Fuchs, S., Collette, F., Perrin, F., Reggers, J., . . . Maquet, P. (2004). Are spatial memories strengthened in the human hippocampus during slow wave sleep? *Neuron*, *44*, 535–545. doi:10.1016/j.neuron.2004.10.007
- Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological Reviews*, *93*, 681–766. doi:10.1152/physrev.00032.2012
- Rasch, B., Büchel, C., Gais, S., & Born, J. (2007). Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science*, *315*, 1426–1429. doi:10.1126/science.1138581
- Ritter, S. M., Strick, M., Bos, M. W., Van Baaren, R. B., & Dijksterhuis, A. P. (2012). Good morning creativity: Task reactivation during sleep enhances beneficial effect of sleep on creative performance. *Journal of Sleep Research*, *21*, 643–647. doi:10.1111/j.1365-2869.2012.01006.x
- Rudoy, J. D., Voss, J. L., Westerberg, C. E., & Paller, K. A. (2009). Strengthening individual memories by reactivating them during sleep. *Science*, *326*, 1079. doi:10.1126/science.1179013
- Schönauer, M., Alizadeh, S., Jamalabadi, H., Abraham, A., Pawlizki, A., & Gais, S. (2017). Decoding material-specific memory reprocessing during sleep in humans. *Nature Communications*, *8*, Article 15404. doi:10.1038/ncomms15404
- Schönauer, M., Brodt, S., Pöhlchen, D., Breßmer, A., Danek, A. H., & Gais, S. (2018). Sleep does not promote solving classical insight problems and magic tricks. *Frontiers in Human Neuroscience*, *12*, Article 72. doi:10.3389/fnhum.2018.00072
- Schooler, J. W., & Melcher, J. (1995). The ineffability of insight. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 97–133). Cambridge, MA: MIT Press.
- Shambroom, J. R., Fábregas, S. E., & Johnstone, J. (2012). Validation of an automated wireless system to monitor sleep in healthy adults. *Journal of Sleep Research*, *21*, 221–230. doi:10.1111/j.1365-2869.2011.00944.x
- Sio, U. N., Monaghan, P., & Ormerod, T. (2013). Sleep on it, but only if it is difficult: Effects of sleep on problem solving. *Memory & Cognition*, *41*, 159–166.
- Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, *135*, 94–120. doi:10.1037/a0014212
- Stickgold, R., Scott, L., Rittenhouse, C., & Hobson, J. A. (1999). Sleep-induced changes in associative memory. *Journal of Cognitive Neuroscience*, *11*, 182–193. doi:10.1162/089892999563319
- Stickgold, R., & Walker, M. P. (2013). Sleep-dependent memory triage: Evolving generalization through selective processing. *Nature Neuroscience*, *16*, 139–145. doi:10.1038/nn.3303
- Tambini, A., Berners-Lee, A., & Davachi, L. (2017). Brief targeted memory reactivation during the awake state enhances memory stability and benefits the weakest memories. *Scientific Reports*, *7*, Article 15325. doi:10.1038/s41598-017-15608-x
- Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature*, *427*, 352–355. doi:10.1038/nature02223
- Walker, M. P., Liston, C., Hobson, J. A., & Stickgold, R. (2002). Cognitive flexibility across the sleep–wake cycle: REM-sleep enhancement of anagram problem solving. *Cognitive Brain Research*, *14*, 317–324. doi:10.1016/S0926-6410(02)00134-9
- Wang, B., Antony, J. W., Lurie, S., Brooks, P. P., Paller, K. A., & Norman, K. A. (2019). Targeted memory reactivation during sleep elicits neural signals related to learning content. *The Journal of Neuroscience*, *39*, 6728–6736. doi:10.1523/JNEUROSCI.2798-18.2019
- Wilson, M. A., & McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, *265*, 676–679. doi:10.1126/science.8036517

Supplemental Material for

Targeted Memory Reactivation during Sleep Improves Next-Day Problem Solving

Study Design Considerations

A recent meta-analytic review found that incubation effects are moderated by design factors such as problem type, length of incubation period, and type of incubation task (Sio & Ormerod, 2009). Therefore, we designed the current study to surmount challenges that may have produced inconsistencies in the sleep, incubation, and problem-solving literatures. We selected puzzles likely to require restructuring, broadly construed, as is common in incubation studies. Solving these puzzles often depends on a person ignoring initially prepotent ideas and integrating more distantly related ideas, or utilizing information that was initially ignored. Some extant lab studies presented a single problem, which could lack both sensitivity and generalizability. Other studies briefly presented many short problems of a single type, which could lead participants to confuse problem elements and, therefore, decrease the usefulness of incubation. To avoid both problems, for each participant, we collected data from 12 dissimilar puzzles they were unable to solve on the first attempt and randomly paired each puzzle with a unique sound. Additionally, the data collection was spread over two days (two evenings and two mornings), so that on any given night participants were not attempting too many puzzles or hearing too many sounds, which could lead to participants' confusion of the sound-puzzle pairings and dilute potential effects. Furthermore, the impact of sleep processing on problem solving may take time and multiple cycles of both slow-wave sleep (SWS) and rapid eye movement (REM) sleep to develop. Whereas some sleep studies with Targeted Memory Reactivation (TMR) have utilized a short nap, we presented sound cues over a full night of sleep. Finally, to increase power and control of participant-related and situational variables, we tested a relatively large number of participants in a within-participant design.

Sleep Monitoring and Cueing System Validation

The Sleep-Monitoring and Cueing System (SMCS) in our study employed an automated wireless sleep-stage monitoring device (WS, Zeo Inc). We modified the WS to use wet electrodes (silver/silver-chloride electrodes filled with electrolyte gel), which generally should improve signal quality compared to the dry electrodes used in validation studies. We targeted SWS for two reasons. First, this strategy follows the TMR literature, which largely presents cues during SWS. Second, we wanted to present cues without waking participants, which is less-likely in SWS. However, our hypotheses and outcomes are not critically tied to SWS, as being able to produce a cueing effect by presenting puzzle-associated cues in any sleep stage would be meaningful. Nonetheless, the device we utilized has a good record for targeting SWS.

In prior validation studies of the WS, agreement between the WS and human scorers on epochs of sleep versus wakefulness was 93.6%, compared to 95.8% between two human raters (Shambroom, Fábregas, & Johnstone, 2012) although compared to another automated scoring system the WS showed

only 80.9% agreement (Griessenberger, Heib, Kunz, Hoedlmoser, & Schabus, 2013). In addition, the WS has high sensitivity to detect sleep (97.6%) but lower specificity (56.1%), suggesting the WS underestimates wake (Tonetti et al., 2013). As mentioned in the manuscript, in our study we observed general agreement between participants' self-reported and SMCS-reported sleep (395 minutes self-reported sleep, 95% CI = [380, 410] vs. 393 minutes SMCS-recorded sleep, 95% CI = [377, 409]) and fairly few reports of awakening during the night and hearing the puzzle-associated sounds (3.5% of trials), suggesting that the device did not regularly awaken participants.

Regarding specific sleep stages, in one study the WS agreed with a consensus score from two human PSG scorers 81.1% of the time while the two scorers agreed with each other 83.2% of the time (Cohen's kappa of .70 and .74 respectively) and showed moderate to high agreement between the scorers and the WS with positive predictive values of 85.6%, 74.4%, and 69.1% for light, REM, and deep sleep respectively (Shambroom et al., 2012). Two other validation studies also showed moderate to high agreement with expert scorers with Cohen's kappas of .56, .67, and .70 in an overnight study (Tonetti et al., 2013) and .49, .28, .72 in a nap study (Cellini, McDevitt, Ricker, Rowe, & Mednick, 2015) for light, REM, and deep sleep respectively. Together, these validation studies suggest that, despite occasional errors, overall the sleep stages indicated by the device are similar to those from human PSG scorers. The SMCS, with the improvement of wet electrodes, has also been successfully used in a prior study to present sound cues during sleep and show behavioral consequences the next day (Honma et al., 2016).



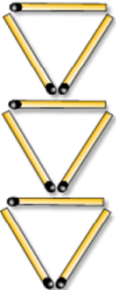
A. Spatial	B. Rebus	C. Verbal
<p>The triangle shown below points to the top of the page. Show how you can move three circles to create an identical triangle that points to the bottom of the page.</p> 		<p>On a wall outside a closet door are 3 standard on/off switches. One (and only one) controls a light bulb inside the light-tight, well-insulated closet. The other two switches do nothing. You can only open the closet door once, and cannot change any switches after the door is open (or re-closed, for that matter). Damaging or disassembling the door, walls, or switches is against the rules.</p> <p>Within these constraints, how can you determine with certainty which switch controls the light bulb?</p>
D. Matchstick		
<p>Move exactly 3 matches to make 4 equilateral triangles (equilateral means having three sides of the same length). All matchsticks must be part of an enclosed shape (no loose ends).</p>		

Fig. S1.

Example puzzles from each of the four puzzle types. (A) Spatial puzzles required participants to manipulate a spatial configuration to attain a goal state. (B) Rebus puzzles required participants to identify the word or phrase that the image depicted. (C) Verbal puzzles required participants to complete a specified objective based on a scenario. (D) Matchstick puzzles required participants to add, move, or remove matches to attain a configuration or goal. Answers: (A): move the top circle to the bottom center and the two bottom corner circles to be in line with the second row. (B): “scrambled eggs.” (C): turn on switch 1 and leave it on for a couple minutes, then turn it off, turn on switch 2 and open the door. If the lightbulb is on, switch 2 controls it, if it is off and warm, switch 1 controls it, and if it is off and cold, switch 3 controls it. (D): move the bottom triangle so that it sits to the right of the two above triangles forming a small triangle in the space where the three matchstick triangles meet. Each triangle is equilateral, although the newly created one is smaller than the other three.

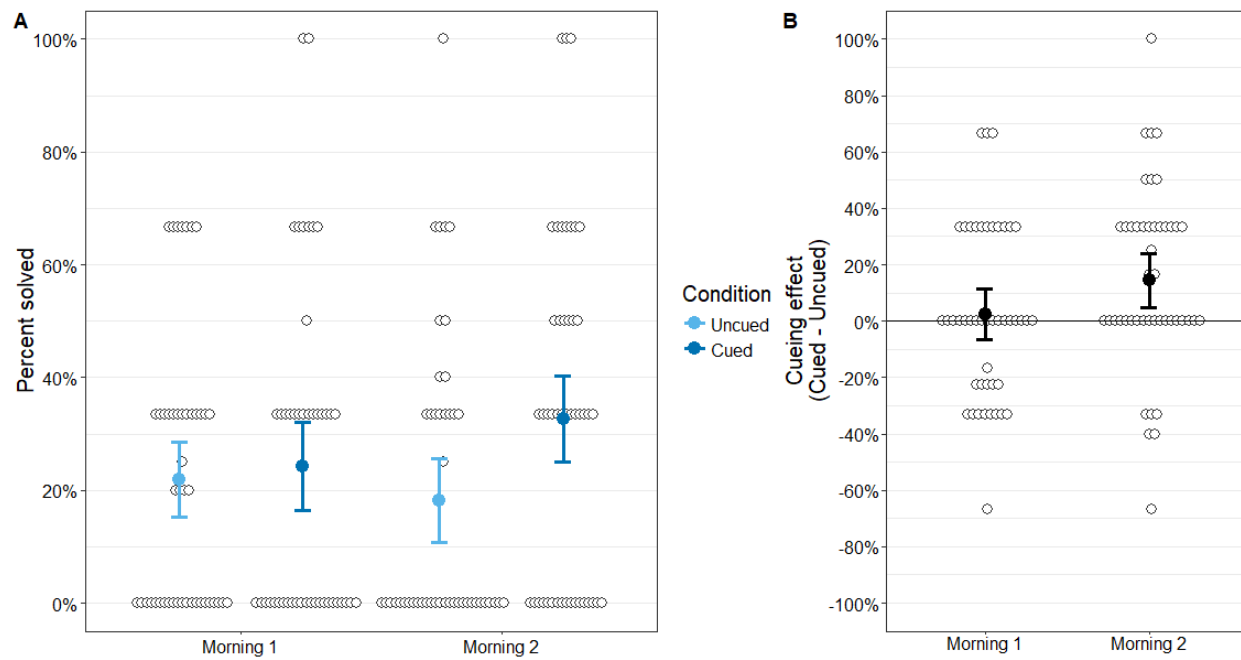


Fig. S2.

A. Solving rates for Uncued vs. Cued puzzles separately for each of the morning sessions. Error bars represent 95% within-subject confidence intervals. **B.** Within-participant cueing effect (percent Cued solving minus percent Uncued solving). Error bars represent between-subject 95% confidence intervals. A 2 (Cueing) x 2 (Session) within-subjects ANOVA shows a main effect of Cueing ($F(1, 45) = 7.55, \eta_p^2 = .14, p = .009$), no main effect of Session ($F(1, 45) = 0.45, \eta_p^2 = .01, p = .505$), and no interaction ($F(1, 45) = 3.01, \eta_p^2 = .06, p = .090$). Note that only 46 participants are included in this analysis due to missing data.

Table S1*Problem solving cueing effect by exclusion criteria*

	Exclusion Criteria	Cued Mean	Cued SD	Uncued Mean	Uncued SD	t	df	p
1.	Exclude heard sounds and nights with no sounds (9.6% of data)	31.7%	28.3%	20.5%	18.9%	3.22	56	.002
2.	None	30.0%	25.5%	20.1%	18.6%	3.29	56	.002
3.	Exclude only heard sounds (3.6% of data)	31.7%	28.3%	20.1%	18.6%	3.52	56	.001
4.	Exclude only nights with no sounds (6.0% of data)	30.0%	25.5%	20.5%	18.9%	2.97	56	.004

Note. Line 1 includes data as described in the text, excluding any data that could affect results. Line 2 includes all possible data; Lines 3 and 4 include all data except for the specific criteria noted. Regardless of exclusion criteria, participants solved more Cued than Uncued puzzles.

Table S2

Overnight sleep stage duration (minutes and percent of total sleep period) and correlations with solving and memory cueing effects

	Mean (minutes)	SD (minutes)	Mean (percent)	SD (percent)	Correlation with Solving Cueing Effect	<i>p</i>	Correlation with Memory Cueing Effect	<i>p</i>
Light Sleep	205.01	46.73	52.2%	9.3%	.05	.70	.15	.29
SWS	61.00	33.30	16.0%	9.1%	.16	.24	-.12	.37
REM	100.73	35.87	25.2%	6.6%	-.02	.87	-.18	.19
Wake	17.95	17.87	4.5%	4.2%	.04	.75	.32	.02
Missing	6.16	17.87	1.5%	3.9%	.11	.43	.13	.33
Total Sleep	393.20	59.37	----	----	.17	.22	.07	.63

Note. Correlation values are for minutes in the sleep stage and the respective cueing effects. *P* values are uncorrected for multiple comparisons. Sound cues were presented during all instances of recorded SWS.

Table S3*Problem solving cueing effect by puzzle type (matchstick, rebus, spatial, and verbal).*

	<i>N</i>	Uncued Percent Solved	Uncued <i>SD</i>	Cued Percent Solved	Cued <i>SD</i>	Cued - Uncued Difference	<i>t</i>	<i>p</i>
Matchstick	8	28.1	21.9	39.4	19.6	11.3	2.17	0.07
Rebus	11	25.2	18.5	36.7	27.4	11.5	1.91	0.09
Spatial	11	23.4	27.9	40.1	22.5	16.7	2.35	0.04
Verbal	11	18.3	14.0	23.1	22.4	4.8	0.59	0.57

Note. *P* values are uncorrected for multiple comparisons. One matchstick puzzle is excluded from analysis because it was often solved in the evening (so it was excluded from morning datasets) and the few times it was presented in the morning, it was excluded for other reasons (e.g., the sound was heard during the night).

References for Supplemental Materials

- Cellini, N., McDevitt, E. A., Ricker, A. A., Rowe, K. M., & Mednick, S. C. (2015). Validation of an automated wireless system for sleep monitoring during daytime naps. *Behavioral Sleep Medicine, 13*(2), 157–168. <https://doi.org/10.1080/15402002.2013.845782>
- Griessenberger, H., Heib, D. P. J., Kunz, A. B., Hoedlmoser, K., & Schabus, M. (2013). Assessment of a wireless headband for automatic sleep scoring. *Sleep and Breathing, 17*(2), 747–752. <https://doi.org/10.1007/s11325-012-0757-4>
- Honma, M., Plass, J., Brang, D., Florczak, S. M., Grabowecky, M., & Paller, K. A. (2016). Sleeping on the rubber-hand illusion: memory reactivation during sleep facilitates multisensory recalibration. *Neuroscience of Consciousness, 2016*(1), niw020. <https://doi.org/10.1093/nc/niw020>
- Shambroom, J. R., Fábregas, S. E., & Johnstone, J. (2012). Validation of an automated wireless system to monitor sleep in healthy adults. *Journal of Sleep Research, 21*(2), 221–230. <https://doi.org/10.1111/j.1365-2869.2011.00944.x>
- Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin, 135*(1), 94–120. <https://doi.org/10.1037/a0014212>
- Tonetti, L., Cellini, N., de Zambotti, M., Fabbri, M., Martoni, M., Fábregas, S. E., ... Natale, V. (2013). Polysomnographic validation of a wireless dry headband technology for sleep monitoring in healthy young adults. *Physiology & Behavior, 118*, 185–188. <https://doi.org/10.1016/j.physbeh.2013.05.036>