

# Sleep-dependent reductions in reality-monitoring errors arise from more conservative decision criteria

Carmen E. Westerberg, Christopher A. Hawkins, and Lauren Rendon

*Psychology Department, Texas State University, San Marcos, Texas 78666, USA*

Reality-monitoring errors occur when internally generated thoughts are remembered as external occurrences. We hypothesized that sleep-dependent memory consolidation could reduce them by strengthening connections between items and their contexts during an afternoon nap. Participants viewed words and imagined their referents. Pictures of the referents also accompanied half of the words. After a 2-h break filled with sleep ( $n = 31$ ) or wakefulness ( $n = 32$ ), participants indicated if they previously viewed a picture of each word. Nap participants made fewer reality-monitoring errors than wake participants by adopting more stringent response criteria, suggesting that sleep reduces reality-monitoring errors primarily by influencing post-retrieval decision processes.

New memories are strengthened and stabilized for long-term storage through consolidation, which optimally progresses during sleep (Diekelmann and Born 2010). Sleep-dependent gains have been observed for both procedural and declarative memories, and are thought to involve the redistribution of memories from temporary to long-term storage sites (Diekelmann and Born 2010). Sleep is thought to aid in the redistribution of declarative memories from hippocampal to neocortical sites through the slow-oscillatory activity present during slow-wave sleep (SWS), as the synchronized neuronal firing promotes crosstalk between different brain areas (Crunelli and Hughes 2010). Stage-2 sleep spindles have also been implicated in promoting regional information exchange (Siapas and Wilson 1998).

A common division in declarative memory concerns item memory and source memory (Johnson et al. 1993). Item memory is memory for an individual item at the focus of attention, and source memory refers to memory for the contextual details associated with an item (background, spatial location, color, etc.). Theoretically, sleep may be especially relevant for source-memory consolidation, as the hippocampus is necessary to connect items and their contexts during memory acquisition, but is not always required to support item memories (Eichenbaum et al. 2007). Supporting this assertion, van der Helm et al. (2011) found that sleep improves memory for the learning context of a word list but has no effect on memories for the words themselves. Furthermore, paradigms which require participants to indicate the spatial location in which an object was learned typically show sleep-related benefits, whereas item memories seldom show benefits from sleep (Diekelmann et al. 2009).

Errors of source memory are common (Johnson et al. 1993). One type, known as reality-monitoring errors, occur when individuals mistake a memory as externally generated when it was really internally generated (Johnson and Raye 1981). For example, remembering that you paid your bills when, in reality, you only imagined paying them. The impact of sleep on reality-monitoring errors is unclear, with evidence primarily coming from the Deese–Roediger–McDermott (DRM) paradigm (Deese 1959; Roediger and McDermott 1995). In this paradigm, participants typically exhibit false memories for critical words (e.g., *sweet*) that represent the “themes” of encoded word lists (e.g., *sour, candy, sugar, etc.*), but

were never actually studied. One explanation for the high level of false memories in DRM experiments is that as each list word is presented, activation spreads to related words. Because critical words are highly related to each presented word, they accrue a large amount of spreading activation. At test, the internal spreading activation received by critical words is confused with activation from actual presentation, leading to false memories for critical words (Roediger and McDermott 1995; Roediger et al. 2001).

Fenn et al. (2009) demonstrated that in the DRM paradigm, participants exhibited fewer false memories following sleep compared with wakefulness, and attributed this result to improved reality-monitoring following sleep. However, Payne et al. (2009) found increased false memories for critical words after sleep, whereas Diekelmann et al. (2008, 2010) found no effect of sleep on false memories. These different results could be due to differences in sleep architecture across samples, the use of recall versus recognition tests, or other factors specific to the DRM paradigm. In particular, there is a high degree of semantic similarity across the studied words, which could result in the strengthening of the gist of word lists relative to item-specific information during consolidation (Stickgold and Walker 2013). Thus, the effects of sleep-dependent consolidation on reality-monitoring errors not involving highly related stimuli may differ from those observed with the DRM paradigm.

With regard to sleep architecture, some DRM experiments have found an association between SWS and increased false memories, but this is typically attributed to processing the semantic similarity shared by list words (Diekelmann et al. 2008, 2010; Payne et al. 2009; Pardilla-Delgado and Payne 2017). Using unrelated word lists, van der Helm et al. (2011) found that memory for list-learning context was associated with stage-2 sleep time and fast (13–15 Hz) spindle activity. Therefore, these aspects of sleep might be most relevant for reducing reality-monitoring errors when semantic similarity is not a factor.

The goal of the current experiment was to investigate the impact of sleep on reality-monitoring errors using a procedure wherein high levels of semantic similarity across stimuli do not

**Corresponding author:** cw54@txstate.edu

Article is online at <http://www.learnmem.org/cgi/doi/10.1101/lm.046342.118>.

© 2018 Westerberg et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first 12 months after the full-issue publication date (see <http://learnmem.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

complicate interpretations. Using a reality-monitoring paradigm (Gonsalves and Paller 2000), participants viewed 250 words one at a time for 1000 msec each and were asked to imagine the object that corresponded to each word (Figure 1). For half of the words, a picture of the object described by the word was shown immediately after for 500 msec (picture trials). For the other half, a blank rectangle appeared on the screen for 500 msec (imagine-only trials). Immediately after, participants completed a page of basic math problems to clear their short-term memories. Later, participants completed a recognition test wherein 375 spoken words were presented one at a time through speakers at a rate of one word every 2 sec. In particular, 125 of the words were from picture trials during encoding, 125 were from imagine-only trials, and 125 were new words, randomly intermixed. For each test word, participants were instructed to indicate either “yes,” they saw a picture of the word earlier in the experiment (pressed “1” key), or ‘no,’ they did not (pressed “2” key). The words used for picture, imagine-only, and new test trials were matched in frequency and concreteness and counterbalanced across participants. Pictures were 266 × 281 pixel color images depicting a typical referent of each word against a white background.

Participants aged 18–29 years with no preexisting sleep disorders were randomly assigned to one of two groups. Participants in the wake group ( $n = 32$ ) studied the words, completed the math problems, and were instructed not to sleep and to return to the laboratory 2 h later. Participants in the nap group ( $n = 32$ ; 15 others excluded for failure to sleep) were first prepped for polysomnography (PSG) and then studied the words, completed the math problems, and given 90 min to nap in a bed in a dark, quiet room. PSG included electroencephalography recorded from electrodes at seven standard sites (C3, C4, Fpz, F3, F4, O1, O2), each referenced to the mastoid average. Electrocardiography, electromyography, and electrooculography were also obtained through electrodes placed on the chest, chin, and near the eyes. Data were recorded and am-

plified with a 250 Hz sampling rate and band-pass filtered between 0.25 and 100 Hz. After waking, they were given 30 min to clean up and recover from any sleep inertia. One nap participant was excluded from analyses for poor memory ( $>2.5$  SD below the mean percent correct score on the recognition test). PSG data from the remaining 31 participants were scored using standard sleep-staging criteria (Table 1). Artifacts were rejected from PSG data based on visual inspection, and fast (13–15 Hz) spindles were identified for C3 and C4 electrodes and slow (11–13 Hz) spindles were identified for Fpz, F3, and F4 electrodes, using an automated algorithm in PRANA (PhiTools; duration: 0.5–3.0 sec, amplitude fluctuations  $>2.5$  SD above mean amplitude). Electrodes selected for spindle analyses were based on known topographical differences between the two spindle types (Schabus et al. 2007). Density was computed as the average number of spindles per minute during stage-2 sleep (Table 2).

Prior to studying the objects, all participants completed the Karolinska Sleep Log (Akerstedt et al. 1994) to get estimates of sleep from the night preceding the experiment. No differences in sleep time [ $t_{(61)} = 1.3$ ,  $P = 0.2$ ], sleep latency [ $t_{(61)} = 1.6$ ,  $P = 0.1$ ], or number of awakenings [ $t_{(61)} = 1.0$ ,  $P = 0.3$ ] were present across groups. Participants in the nap group also completed the Pittsburgh Sleep Quality Index (Buysse et al. 1989) to assess sleep quality within the past month. On average, participants had good (score  $<5$ ) sleep quality (mean = 3.5, SD = 1.3). Immediately before the recognition test, all participants completed the positive and negative affectivity scale (PANAS) (Watson et al. 1988). Two participants (one nap, one wake) failed to properly follow instructions for completing it and were excluded from analyses of PANAS data. No group differences in positive [ $t_{(59)} = 0.1$ ,  $P = 0.9$ ] or negative affect [ $t_{(59)} = 1.7$ ,  $P = 0.09$ ] were observed.

For the recognition test, “yes” responses to imagine-only words were considered reality-monitoring error hits, “yes” responses to picture words were considered picture recognition hits, and “yes” responses to new words were considered false alarms. These values were used to calculate Pr and Br, measures of recognition sensitivity and response bias, respectively, for reality-monitoring errors and picture recognition (Snodgrass and Corwin 1988). Sensitivity reflects the ability to discriminate between old and new items. Thus, if sleep-dependent consolidation strengthened connections between items and their contexts, nap participants would be less likely than wake participants to mistake a picture they imagined for one they actually viewed, and this would be reflected in Pr. Response bias reflects the general tendency to give “yes” responses. Br values  $<0.5$  reflect a conservative bias and  $>0.5$  reflect a liberal bias (more likely to respond “yes”). If sleep influenced post-retrieval decision processes, nap participants would be more or less likely to endorse all words as having been accompanied by a picture during study compared with the wake group, and Br would differ between groups.

To test these possibilities, a  $2 \times 2$  ANOVA was conducted on Pr with group (wake, nap) and item type (reality-monitoring errors, picture recognition) as independent variables. There was no main effect of group,  $F_{(1,61)} = 0.6$ ,  $P = 0.5$  and no group  $\times$  item type interaction,  $F_{(1,61)} = 2.6$ ,  $P = 0.1$ , indicating that sleep did not influence sensitivity. As expected, Pr was significantly higher for picture recognition than for reality-monitoring errors,  $F_{(1,61)} = 327.1$ ,  $P = 3.4 \times 10^{-26}$ , Cohen’s  $d = 3.0$ .

Another  $2 \times 2$  ANOVA examined the effects of group and item type on Br. Although an afternoon nap did not influence sensitivity, it did affect bias (Figure 2). The nap group was more conservative in giving “yes” responses (0.09) than the wake group (0.15),  $F_{(1,61)} = 6.7$ ,  $P = 0.01$ ,  $d = 0.7$ , and there was no item type  $\times$  group interaction,  $F_{(1,61)} = 3.1$ ,  $P = 0.08$ , indicating that the nap group engaged in more conservative responding for both picture recognition and reality-monitoring errors. All participants were more

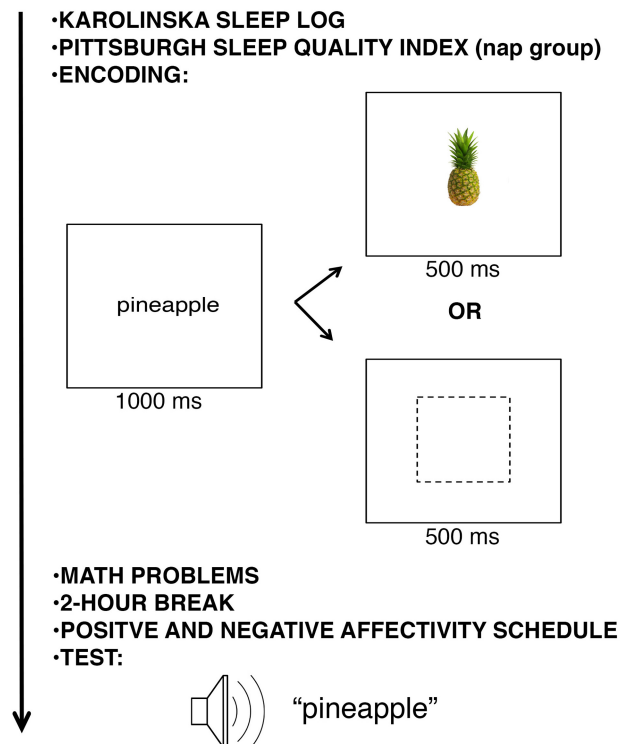


Figure 1. Experimental procedure.

**Table 1.** Average time (min) until sleep onset (latency), and average time (min) spent in each sleep stage and awake after sleep onset for participants in the nap group (standard error in parentheses)

Latency	Stage 1	Stage 2	SWS	REM <sup>a</sup>	Wake	TST
15.3 (1.9)	12.3 (1.3)	34.4 (1.8)	12.4 (2.1)	16.6 (2.6)	11.2 (1.7)	66.6 (2.1)

<sup>a</sup>Only includes participants who achieved REM sleep ( $n = 14$ ).

SWS, slow-wave sleep; REM, rapid eye movement; TST, total sleep time.

conservative in giving “yes” responses for reality-monitoring errors compared with picture recognition,  $F_{(1,61)} = 97.2$ ,  $P = 3.0 \times 10^{-14}$ ,  $d = 0.5$ .

To determine whether the more conservative bias in the nap group was systematically related to specific aspects of sleep, Br measures (reality-monitoring error, picture recognition) were separately correlated with total sleep time, stage-2 sleep time, SWS time, and stage-2 spindle density, as these measures have previously been associated with changes in memory (van der Helm et al. 2011; Pardilla-Delgado and Payne 2017). No significant relationships were observed (see Table 3 for correlation coefficients and  $p$  values for each comparison). Although sleep did not significantly affect Pr, certain aspects of sleep may still be systematically related to sensitivity. Pearson’s correlations revealed that picture recognition Pr was not related to any sleep measures, and reality-monitoring error Pr was not associated with total sleep time, SWS time, or fast-spindle density (Table 3). However, reality-monitoring Pr was negatively correlated with stage-2 sleep time ( $r = -0.35$ ,  $P = 0.05$ ) and with stage-2 slow-spindle density at Fpz ( $r = -0.38$ ,  $P = 0.04$ ), indicating that as stage-2 time and slow-spindle activity increases, reality-monitoring error sensitivity decreases.

The current results indicate that an afternoon nap reduces reality-monitoring errors primarily through the adoption of a more conservative response bias following sleep. The only other studies to examine the influence of sleep on reality-monitoring errors used the DRM paradigm and the results were mixed (Diekelmann et al. 2008; Fenn et al. 2009; Payne et al. 2009; Pardilla-Delgado and Payne 2017). The present experiment provides a more straightforward test than DRM experiments, as studied items did not share high semantic similarity, which may influence sleep-dependent memory processes in idiosyncratic ways (Stickgold and Walker 2013).

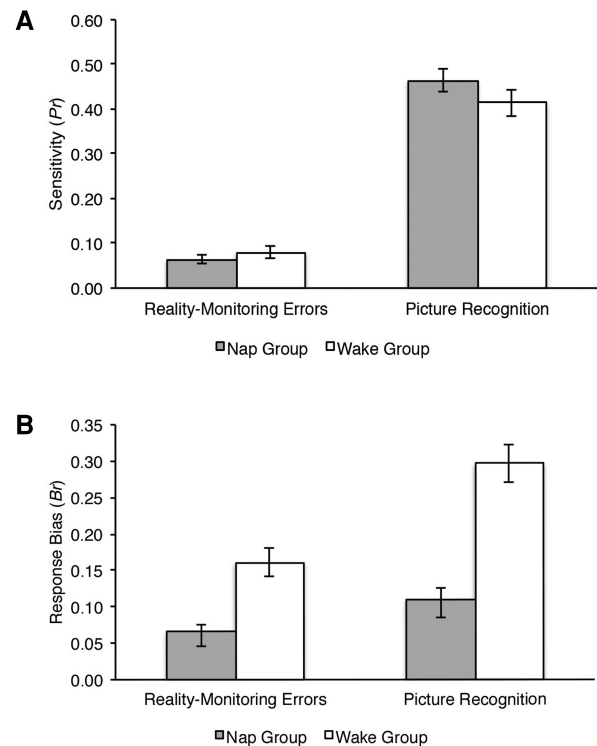
The more conservative response bias in the nap group suggests that after sleep, participants are less likely to endorse all test items as having been accompanied by a picture during study. To date, the influence of sleep on response bias has received little attention and results have been mixed, with some reports of more liberal (Pardilla-Delgado and Payne 2017) and some of more conservative (Hu et al. 2006) responding after sleep. Hu et al. (2006) observed a more conservative bias after sleep for neutral and emotional picture memories, and suggested that high levels of hippocampal acetylcholine during rapid eye movement (REM) sleep could underlie the conservative shift. This explanation cannot account for the current findings, as many nap participants did not achieve any REM sleep (17 of 31). Another possibility is that the

**Table 2.** Average fast (13–15 Hz) and slow (11–13 Hz) spindle densities (number/min) at individual electrodes during stage-2 sleep (standard error in parentheses)

	C3	C4	Fpz	F3	F4
Fast spindle density	3.6 (0.3)	3.5 (0.3)			
Slow spindle density			3.2 (0.3)	3.0 (0.3)	2.9 (0.3)

conservative shift could be a consequence of sleep-dependent changes in memory strength. Hirshman (1995) demonstrated that the perceived mean memory strength of studied items is a powerful cue for criterion placement at test initiation, and weak mean memory strength results in a more liberal bias than strong. Here, it is possible that sleep-dependent consolidation strengthened memories for studied items, which led to the assessment of stronger mean memory strength at test initiation and thus a more conservative criterion placement in the nap group compared with the wake group.

If memories were strengthened during sleep, however, changes in sensitivity may also be expected. Although sensitivity changes were not observed, it is possible that an afternoon nap caused subtle changes in memory strength that were enough to influence criterion placement but not sensitivity. In the nap versus the wake group, sensitivity was numerically lower for reality-monitoring errors (0.06 versus 0.08, respectively) and numerically higher for picture recognition (0.46 versus 0.41, respectively); the expected pattern if connections between items and their contexts were strengthened during sleep. Furthermore, Hu et al. (2006) also did not find sensitivity differences despite a more conservative bias after sleep compared with wake, although they did observe increased

**Figure 2.** (A) Recognition sensitivity (Pr) and (B) response bias (Br), for reality-monitoring errors and picture recognition for the nap and wake groups. Bars indicate standard error of the mean.

**Table 3.** Pearson's correlation coefficients (*r*) and associated *P* values for correlations between memory (Pr and Br) measures and sleep measures

	Reality-monitoring Pr		Picture recognition Pr		Reality-monitoring Br		Picture recognition Br	
	<i>r</i>	<i>P</i>	<i>r</i>	<i>P</i>	<i>r</i>	<i>P</i>	<i>r</i>	<i>P</i>
Total sleep time (min)	-0.26	0.16	-0.07	0.71	0.13	0.50	0.12	0.53
Stage 2 (min)	-0.35	0.05*	0.15	0.44	-0.13	0.50	-0.19	0.31
Slow-wave sleep (min)	0.06	0.77	-0.17	0.35	0.11	0.57	0.06	0.75
C3 fast spindles (number/min)	-0.24	0.19	-0.14	0.45	-0.10	0.59	-0.11	0.54
C4 fast spindles (number/min)	-0.14	0.44	-0.16	0.39	0.01	0.94	0.01	0.97
Fpz slow spindles (number/min)	-0.38	0.04*	-0.29	0.12	-0.11	0.56	-0.17	0.35
F3 slow spindles (number/min)	-0.20	0.29	-0.29	0.12	0.06	0.74	0.00	0.99
F4 slow spindles (number/min)	-0.20	0.27	-0.33	0.07	0.08	0.69	0.01	0.97

\**P* value ≤ 0.05.

confidence in memories following sleep. Finally, in the nap group, as time spent in stage-2 sleep increased and as slow-spindle density increased, reality-monitoring error sensitivity decreased. These findings are consistent with results indicating that stage 2 and spindle activity during this stage are associated with context memory (van der Helm et al. 2011). Stage 2 and associated spindle activity may strengthen item-context bindings, which would reduce reality-monitoring errors. It is possible that if participants spent more time in stage 2, through a longer nap or overnight sleep, sensitivity differences may emerge. Of note is that fast and slow spindles appear to play complementary roles in memory consolidation (Möller et al. 2011). Van der Helm et al. (2011) implicated fast spindles, whereas here, slow spindles were implicated. Methodological differences may have influenced the relevance of fast versus slow spindles across studies, and future research is necessary to test this possibility.

These results suggest that sleep reduces reality-monitoring errors through the adoption of more conservative response criteria following sleep. Clearly, sleep impacted post-retrieval decision processes, and it is possible that these decision processes were influenced by subtle changes in memory strength resulting from sleep-dependent consolidation. To date, few experiments have examined the influence of sleep on response bias. Future experiments aimed at understanding how sleep may influence post-retrieval decisions will ultimately provide a more comprehensive account of how reality-monitoring and other memory errors evolve over time.

## Acknowledgments

This research was supported by a Research Enhancement Program grant at Texas State University. We thank Mary Aiello, Joni Fields, and David Lee for assistance with data collection.

## REFERENCES

Akerstedt T, Hume K, Minors D, Waterhouse J. 1994. The subjective meaning of good sleep, an intraindividual approach using the Karolinska Sleep Diary. *Percept Mot Skills* **79**: 287–296.

Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. 1989. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res* **28**: 193–213.

Crunelli V, Hughes SW. 2010. The slow (<1 Hz) rhythm of non-REM sleep: a dialogue between three cardinal oscillators. *Nat Neurosci* **13**: 9–17.

Deese J. 1959. On the prediction of occurrence of particular verbal intrusions in immediate recall. *J Exp Psychol* **58**: 17–22.

Diekelmann S, Born J. 2010. The memory function of sleep. *Nat Rev Neurosci* **11**: 114–126.

Diekelmann S, Born J, Wagner U. 2010. Sleep enhances false memories depending on general memory performance. *Behav Brain Res* **208**: 425–429.

Diekelmann S, Landolt HP, Lahl O, Born J, Wagner U. 2008. Sleep loss produces false memories. *PLoS One* **3**: e3512.

Diekelmann S, Wilhelm I, Born J. 2009. The whats and whens of sleep-dependent memory consolidation. *Sleep Med Rev* **13**: 309–321.

Eichenbaum H, Yonelinas AP, Ranganath C. 2007. The medial temporal lobe and recognition memory. *Annu Rev Neurosci* **30**: 123–152.

Fenn KM, Gallo DA, Margoliash D, Roediger HL 3rd, Nusbaum HC. 2009. Reduced false memory after sleep. *Learn Mem* **16**: 509–513.

Gonsalves B, Paller KA. 2000. Neural events that underlie remembering something that never happened. *Nat Neurosci* **3**: 1316–1321.

Hirshman E. 1995. Decision processes in recognition memory: criterion shifts and the list-strength paradigm. *J Exp Psychol: Learn Mem Cogn* **21**: 302–313.

Hu P, Stylos-Allan M, Walker MP. 2006. Sleep facilitates consolidation of emotionally arousing declarative memory. *Psychol Sci* **17**: 891–898.

Johnson MK, Hashtroudi S, Lindsay DS. 1993. Source monitoring. *Psychol Bull* **114**: 3–28.

Johnson MK, Raye CL. 1981. Reality monitoring. *Psychol Rev* **88**: 67–85.

Möller M, Bergmann TO, Marshall L, Born J. 2011. Fast and slow spindles during the sleep slow oscillation: disparate coalescence and engagement in memory processing. *Sleep* **34**: 1411–1421.

Pardilla-Delgado E, Payne JD. 2017. The impact of sleep on true and false memory across long delays. *Neurobiol Learn Mem* **137**: 123–133.

Payne JD, Schacter DL, Propper R, Huang L-W, Wamsley E, Tucker MA, Walker MP, Stickgold R. 2009. The role of sleep in false memory formation. *Neurobiol Learn Mem* **92**: 327–334.

Roediger HL, Balota DA, Watson JM. 2001. Spreading activation and arousal of false memories. In *The nature of remembering: essays in honor of Robert G. Crowder* (eds. Roediger HL, Nairne JS, Neath I, Surprenant AM), pp. 95–115. American Psychological Association Press, Washington, D.C.

Roediger HL III, McDermott KB. 1995. Creating false memories: remembering words not presented in lists. *J Exp Psychol Learn Mem Cogn* **21**: 803–814.

Schabus M, Dang-Vu TT, Albouy G, Baletu E, Boly M, Carrier J, Darsaud A, Degueldre C, Desseilles M, Gais S, et al. 2007. Hemodynamic cerebral correlates of sleep spindles during human non-rapid eye movement sleep. *Proc Natl Acad Sci* **104**: 13164–13169.

Siapas AG, Wilson MA. 1998. Coordinated interactions between hippocampal ripples and cortical spindles during slow-wave sleep. *Neuron* **21**: 1123–1128.

Snodgrass JG, Corwin J. 1988. Pragmatics of measuring recognition memory: applications to dementia and amnesia. *J Exp Psychol Gen* **117**: 34–50.

Stickgold R, Walker MP. 2013. Sleep-dependent memory triage: evolving generalization through selective processing. *Nat Neurosci* **16**: 139–145.

van der Helm E, Gujar N, Nishida M, Walker MP. 2011. Sleep-dependent facilitation of episodic memory details. *PLoS One* **6**: e27421.

Watson D, Clark LA, Tellegen A. 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol* **54**: 1063–1070.

Received August 10, 2017; accepted in revised form November 28, 2017.