

The Role of Sleep in Forming a Memory Representation of a Two-Dimensional Space

Marc N. Coutanche,^{1*} Carol A. Gianessi,¹ Avi J.H. Chanales,²
Kate W. Willison,² and Sharon L. Thompson-Schill¹

ABSTRACT: There is ample evidence from human and animal models that sleep contributes to the consolidation of newly learned information. The precise role of sleep for integrating information into interconnected memory representations is less well understood. Building on prior findings that following sleep (as compared to wakefulness) people are better able to draw inferences across learned associations in a simple hierarchy, we ask how sleep helps consolidate relationships in a more complex representational space. We taught 60 subjects spatial relationships between pairs of buildings, which (unknown to participants) formed a two-dimensional grid. Critically, participants were only taught a subset of the many possible spatial relations, which allowed them to potentially infer the remainder. After a 12 h period that either did or did not include a normal period of sleep, participants returned to the lab. We examined the quality of each participant's map of the two-dimensional space, and their knowledge of relative distances between buildings. After 12 h with sleep, subjects could more accurately map the full space than subjects who experienced only wakefulness. The incorporation of untaught, but inferable, associations was particularly improved. We further found that participants' distance judgment performance related to self-reported navigational style, but only after sleep. These findings demonstrate that consolidation over a night of sleep begins to integrate relations into an interconnected complex representation, in a way that supports spatial relational inference. © 2013 Wiley Periodicals, Inc.

KEY WORDS: learning; inferences; relational; cognitive map; spatial

INTRODUCTION

One of the most important characteristics of the human memory system is its integrated nature. Far from being a series of isolated packets of knowledge, the material we learn becomes integrated into highly interconnected relational networks. In its most basic form, a relational memory connects overlapping pieces of information, so that all parts become connected. A prototypical example is a “transitive inference,” where a series of overlapping pairings ($A > B$ and $B > C$) can be combined to construct a hierarchy ($A > B > C$), and inferences ($A > C$)

become at least as accessible to a person as the original constituent pairings. A range of studies into this form of learning point to the hippocampus as playing a central role in encoding, and later retrieving, such inferences (Dusek and Eichenbaum, 1997; Zeithamova et al., 2012). One recent neuroimaging study provided evidence that hippocampal neural activity tracks the linear structure of hierarchies (Kumaran et al., 2012), adding to evidence that suggests, “a core mnemonic function of the hippocampus is the capacity to bridge representational gaps between elements of our experiences” (Staresina and Davachi, 2009, p. 271). Spatial learning—another process linked to the hippocampus (O'Keefe and Nadel, 1978)—has been positioned as both a subset of relational memory (e.g., Konkel and Cohen, 2009) and as a competing hypothesis for hippocampal function (e.g., Kumaran and Maguire, 2005), although the hippocampus plays a key role in learning spatial relations under both accounts (Bird and Burgess, 2008).

Sleep has been found to help consolidate relational memories. Ellenbogen et al. (2007) found that, following sleep, people are better able to access distant transitive inferences to correctly select between items in a hierarchy, despite never having explicitly learned these relationships. Importantly, sleep did not improve performance for the set of underlying taught pairs, supporting the idea that sleep helps integrate, rather than just strengthen, individual memory traces. It has been hypothesized that overlapping learned material is consolidated through relevant neural activity being “replayed” during slow-wave sleep (Lewis and Durrant, 2011). The replay of neuronal patterns may be particularly important for relational memories, which depend on drawing connections between memories. For example, the improved access to inferences reported in Ellenbogen et al. (2007) may result from multiple overlapping memories being bound together by synchronized neuronal reactivation.

Sleep also plays a role in consolidating spatial information. After freely navigating through a virtual space (while being encouraged to form a map), people who have slept can navigate to landmarks faster than those who have spent time awake (Ferrara et al., 2008; Wamsley et al., 2010). Hippocampal reactivation during slow-wave sleep can predict postsleep improvements in virtual route retrieval (Peigneux et al., 2004), suggesting that

¹ Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania; ² Institute for Research in Cognitive Science, University of Pennsylvania, Philadelphia, Pennsylvania

Grant sponsor: National Science Foundation; Grant number: 0910247.

*Correspondence to: Marc Coutanche, Department of Psychology, University of Pennsylvania, 3720 Walnut Street, Philadelphia, PA 19104, USA. E-mail: coumarc@psych.upenn.edu

Accepted for publication 29 May 2013.

DOI 10.1002/hipo.22157

Published online 00 Month 2013 in Wiley Online Library (wileyonlinelibrary.com).

navigational improvements are also linked to neural reactivation (Wilson and McNaughton, 1994; Peigneux et al., 2004).

Learned items that follow a hierarchy have just one-dimension. This creates a constraint that the number of steps between each item and their distance apart in the representational space are always equal (e.g., for $A > B > C > D > E > F$: the distance between B and D is twice that of between B and C in connecting steps, and also in distance), but not all inferences follow this simple pattern. In representations with more than one dimension, such as spatial layouts, items can be separated by a large number of learned connections, yet be close together in the space. For example, after traveling one block North, one block East and then one block South, a person's destination can be just one block (West) from their starting position in terms of Euclidian distance, but three blocks away in connecting steps. This study focuses on the effects of sleep on knowledge spaces in which the numbers of learned connections, and Euclidian distances, can diverge.

In this study, we employ a two-dimensional spatial layout to examine whether sleep facilitates the transformation of unstructured propositional information about a space into a more coherent integrated network. Participants learned a subset of possible relationships between landmarks (that formed an untaught spatial grid). Later, after a delay that either did or did not include a period of sleep, we asked subjects to recreate a map of the representational space they had learned, and to make distance comparisons between landmarks in that space. For the latter, we were particularly interested in circumstances in which the number of connections and distance between items were discordant: Success on these "discordant" questions would represent a consolidated understanding of the space beyond the explicitly taught pairings. We hypothesized that participants who slept between training and testing would construct maps that more accurately represented relationships between landmarks, than subjects who had not slept. We further hypothesized that sleep would particularly improve discordant distance judgments.

Finally, we also examined whether participants' navigational strategies would interact with the effects of sleep. Individuals differ in how they process spatial information during navigation (O'Keefe and Nadel, 1978; Pazzaglia and Taylor, 2007). One well-replicated individual difference is preference (or "cognitive style") for using a "survey" spatial approach, which involves forming a global map-like representation (Pazzaglia and Taylor, 2007). Individuals also vary in the degree to which they focus on their particular route, such as their particular movements (in a navigation task) or the sequence of items they encounter (a "route" preference/cognitive style). These approaches to spatial tasks have differing neural bases. Spontaneously using distant landmarks to build-up a spatial representation (analogous to a survey approach) is accompanied by greater task activity in the hippocampus, whereas counting or assigning numbers to possible routes is not (Iaria et al., 2003; also see Hartley et al., 2003). Drawing on the discussed evidence that hippocampal reactivation is important for consolidation, and on the relationship between hippocampal activity and

navigational style, we hypothesized that participants with a strong survey preference would particularly benefit from sleep, especially with discordant relationships, where understanding distances across the space is crucial.

MATERIALS AND METHODS

Subjects

Participants were recruited through a University of Pennsylvania subject recruitment website. All participants were native English speakers, under the age of 36, free from circadian, psychiatric or neurological disorders, and not taking psychoactive medications. We excluded participants who reported sleeping fewer than 6 h each night in an initial screening assessment of 112 subjects. Of these, 72 met the inclusion criteria and returned for the experiment [mean (M) age: 21.3, standard deviation (sd): 2.47]; we additionally excluded 10 subjects who did not reach training criteria within 10 blocks (M blocks for included subjects = 4.0), one subject with implausibly fast reaction times, and one wake-group participant who reported napping between sessions, resulting in a final sample size of 60 subjects. Informed consent was obtained from all participants each time they entered the lab. Participants completed the experiment for payment or class credit. All procedures were approved by the University of Pennsylvania Institutional Review Board.

Experimental Groups

We randomly assigned subjects to one of four groups (see Fig. 1): 12 h wake (20 subjects, nine male), 12 h sleep (20 subjects, seven male), 24 h morning (10 subjects, one male), 24 h evening (10 subjects, eight male). Subjects came to the lab twice after the screening session. The 12 h wake group trained between 9 and 10 am, and returned 12 h later on the same day. The 12 h sleep group trained between 9 and 10 pm, and returned 12 h later on the next morning, following a normal night of sleep. The 24 h morning group's first session began between 9 and 10 am, with a testing session 24 h later on the following morning. The 24 h evening group trained between 9 and 10 pm, and returned 24 h later on the next evening. The 24 h groups were included for two purposes. First, they allowed us to protect against the possibility that group differences could result from time-of-day effects on training and testing (i.e., evening vs. morning). If such effects did influence this paradigm, we would see differences between the two 24 h groups. Second, the 24 h conditions (which contain both sleep and wakefulness) can give insight into whether effects are susceptible to daytime interference. One 24 h morning participant was removed from all analyses after reporting that they rehearsed.

Procedure

Prior to the first session, subjects completed a battery of self-report measures to assess eligibility and characterize individual

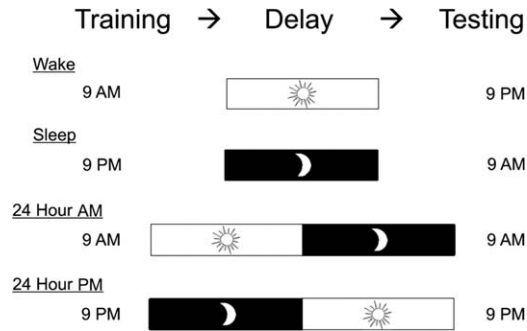


FIGURE 1. Group design. Subjects were randomly allocated to a 12 h group with wakefulness or sleep between training and testing, or to a 24 h group with both training and testing in the morning or evening.

differences in spatial strategies. The Questionnaire on Spatial Representation (QSR; Pazzaglia and Taylor, 2007) and Santa Barbara Sense of Direction Scale (SBSOD; Hegarty et al., 2002) were administered to query spatial strategies and spatial abilities, respectively. Questions from the QSR that reflected survey (e.g., “Would you describe yourself as a person who tries to create a mental map of the environment?”) or route (e.g., “Would you describe yourself as a person who orients him/herself by remembering routes connecting one place to another?”) navigational strategies were identified before data collection. The Pittsburgh Sleep Quality Index (PSQI; Buysse et al., 1989), Epworth Sleepiness Scale (Johns, 1991) and Morningness-Eveningness Questionnaire (MEQ; Horne and Ostberg, 1976) assessed participants’ sleep patterns and allowed

us to check that the groups did not differ in circadian individual differences.

Experimental procedure

Subjects were taught and tested on their knowledge of the relative locations of nine novel landmarks (e.g., a coffee shop). Rather than allowing subjects to freely navigate this environment, we taught them a small number of the many possible relations among the landmarks. This limited set of relations was sufficient for subjects to (potentially) infer a full map of the space. The stimuli were all color images of storefronts collected from online searches of typical shops that could be found in any town or city (see Fig. 2). The images were large, high resolution, and included a store sign written in English. Chain stores were excluded to avoid interference from existing knowledge. Storefronts with any reference to a particular geographic location were also excluded. The allocations of stimuli to the nine map positions were randomly selected from one of four random arrangements. Stimuli presentation and data collection were managed through E-Prime 2.0 Professional Edition (Psychology Software Tools).

Training session

During training, subjects were taught (through feedback) spatial relations between eight pairs of landmarks. On each trial, one pair of storefronts was presented on-screen: one storefront (the target) appeared in the center, and a second storefront (the test item) appeared in two positions (above and below, or left and right)



FIGURE 2. The set of stimuli. Storefronts were randomly assigned to different map locations. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

around the target image. Having two directly opposing options like this connected the paradigm to one-dimensional transitive inference investigations, in which participants select items from just one dimension. The subjects responded with an arrow key to indicate which of the relationships was correct [e.g., to respond that the bakery is West (left arrow) or East (right arrow) of the coffee shop target]. Feedback based on their response was then displayed for 5 s with “Correct!” in green or “Incorrect.” in red. The incorrect choice then disappeared, leaving the target and correct relationship on-screen for 5 s. The training task is shown in Figure 3. The eight relationships included six East–West pairs (e.g., store 1 is West of store 2) and two North–South pairs (e.g., store 3 is North of store 6), which together formed a three-by-three grid (Fig. 4), although this fact was deliberately never told to participants. Presenting more East–West pairs helped maximize the number of untaught inferences that we could later test.

The eight taught relations were each presented four times per block (following two practice trials with different stimuli). The trial order was pseudo-randomized so that pairs were not presented twice in a row and overlapping pairs were never presented consecutively (e.g., never: A is West of B followed by B is West of C). All subjects completed at least three blocks of 32 trials, after which they continued until reaching at least 75% accuracy for East–West pairs and 100% accuracy on the two North–South pairs. Using a below-ceiling criterion is consistent with prior investigations of sleep and transitive inference (e.g., 75% criterion in Ellenbogen et al., 2007) and has an advantage of not saturating participants with knowledge of the trained pairings (see Ellenbogen et al., 2007, for a discussion). We employed a criterion of 100% for North–South pairs, as this was required knowledge to be able to learn the untaught inferences. We could therefore be confident that by the end of the training session, every participant had the potential to consolidate material into a grid. Immediately after training, participants completed a block without feedback, to measure their post-training knowledge. Before leaving the lab, all subjects were asked to refrain from daytime napping and consuming caffeinated or alcoholic beverages until their return.

Testing session

After a delay (of 12 or 24 h depending on their group), subjects completed a block of the training procedure without feedback, to measure retention of the taught associations. In addition, we assessed inferred spatial relations using two measures: a distance judgment task and a map assembly task.

Distance judgment task. On each trial, subjects viewed a target building displayed above two building choices. Participants then selected the building that was closest in distance (as the crow flies) to the target. The correct choice was equally likely to be on the left or right of the screen. After each response, a 100 ms blank screen appeared, separating each trial. Each taught pair was included in a question four times, with each pair-member serving as the target twice. Intermixed with these trials were distance judgments of novel (“untaught”)



FIGURE 3. An example trial from the training task. Top: Two alternatives were presented on opposite sides of a central building—left and right (shown here) or above and below. Bottom: After each response, participants received feedback: “Correct!” in green or “Incorrect.” in red, with the correct arrangement displayed for 5 s (full details in text). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

combinations of buildings that had not been encountered before, but potentially could be answered using spatial inference. We separated the novel combinations into two categories: (i) inferences with a concordant number of connecting steps and separating distance (i.e., the correct choice was closer in distance and had fewer separating steps; e.g., store 2 in Figure 4 is closer in both distance and separating steps to store 6 than it is to store 9); (ii) inferences with a discordant number of connecting steps and separating distance (i.e., the correct choice is closer in distance but further in steps; e.g., store 2 in Figure 4 is closer in distance but not steps to store 5 than it is to store 6). Subjects who relied on choosing landmarks with the fewest separating learned associations would therefore perform poorly on this last set of inferences. On the other hand, we reasoned that correct answers to these “discordant” questions would suggest that a participant has access to direct distances between landmarks through spatial inference.

Map task. Finally, subjects performed a map assembly task. We gave subjects images of the nine landmarks and asked them to arrange the images into a map. No other instructions or guidance about this map was provided. The experimenter

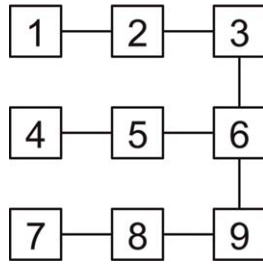


FIGURE 4. Taught pairings. Lines in the diagram reflect taught pairings between storefronts. Gaps (e.g., 1 North of 4) reflect associations that can be inferred from a consolidated map.

photographed the maps for later analysis. One participant's map scores were more than three standard deviations below the mean (from creating a completely reversed map—suggesting they learned the relationships but did not complete the task correctly), so were excluded from map analyses. Examples of maps created by participants are shown in Figure 5.

Two raters (blind to subject group) computed scores for each subject's map: (i) Overall accuracy, where each map was analyzed with a strict binary metric for whether all buildings were correctly placed according to the information presented to participants during training; (ii) Continuous assessments of spatial relations. The first metric measured whether storefronts were correctly placed in the same line—row, column, or diagonal—as others (irrespective of orientation). Each correct relationship (e.g., in Fig. 4, placing stores 1 and 4 in the same column), gave one point. Points were totaled for taught pairs (i.e., those in the three rows and right-most column) and also for untaught inferred pairs (i.e., the two left columns and the diagonals). The second metric assessed each participant's untaught pairings *relative* to their performance on the relevant underlying taught associations: For instance, underlying the building 1 to building 4 relationship: (i) buildings 1 and 3 sharing a row, (ii) buildings 3 and 6 sharing a column, and (iii) buildings 6 and 4 sharing a row. The inferred associations were totaled relative to whether the underlying buildings were correctly placed in the same taught row or column. For example, a participant with three inferences, where 100% of underlying pairings are present for two of these inferences, while

33% are present for the third, would score 129% ($3/2.333$). On the other hand, a participant that has two inferences, but sufficient pairings in taught rows/column for 3.333 inferences, would score 60% ($2/3.333$).

RESULTS

Map Knowledge

We first examined participants' maps for overall accuracy. Participants who experienced sleep in the 12 h between training and testing were more likely to create a map that was correct based on the directional information they were exposed to during training, as compared to participants who were awake during the interval ($\chi^2(1) = 7.65$, $P = 0.006$), with 15 of 19 sleep participants creating maps that were fully correct based on their training, compared to seven of 20 in the wake group. Using a detailed continuous measure of how well relationships were integrated into the map's structure—whether landmarks were correctly placed in the same column, row, or diagonal—the 12 h wake participants performed worse for inferred (55%) than for the taught (73%) relationships ($P < 0.001$ in a Wilcoxon signed rank test). In contrast, the sleep group had high performance on both (taught: 84%, inferred: 79%; $P = 0.25$ in Wilcoxon signed rank test), with a significant interaction between group and relationship-type ($P = 0.01$ in Wilcoxon rank sum test), indicating that sleep-benefits were disproportionately found for the inferred relationships (see Fig. 6). We also examined how well each participant incorporated inferred columns and diagonals *relative* to underlying associations in the taught rows and column (see Methods section for details). This relative inference performance revealed that the 12 h sleep group ($M = 114\%$) had higher scores ($P = 0.04$ in Wilcoxon rank sum test) than the wake group ($M = 77\%$), suggesting that the wake group was less able to translate associations in the taught rows and column into inferences.

To check that any group differences were not due to circadian differences (e.g., morning testing for the 12 h sleep group vs. evening testing for the 12 h wake group), we also ran

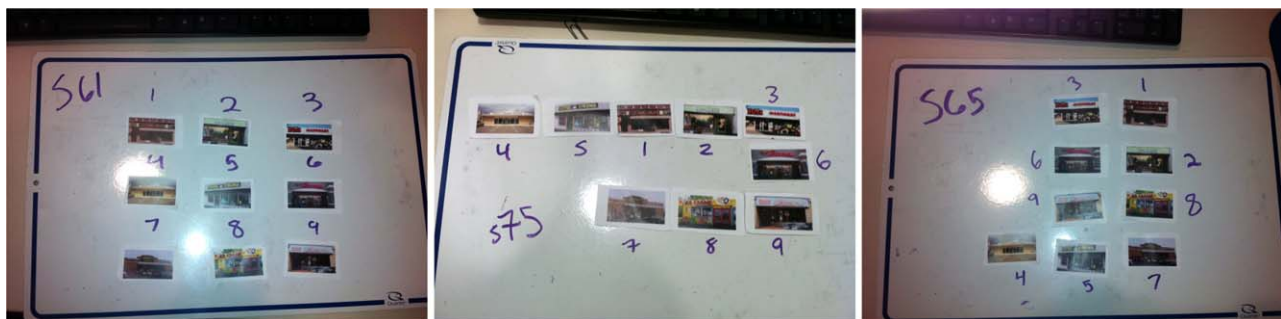


FIGURE 5. Example maps constructed by subjects. Left: A completely correct map. Center: A map with correctly orientated taught associations (e.g., storefront 2 is West of storefront 3), but without the two untaught columns. Right: An incorrect map. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

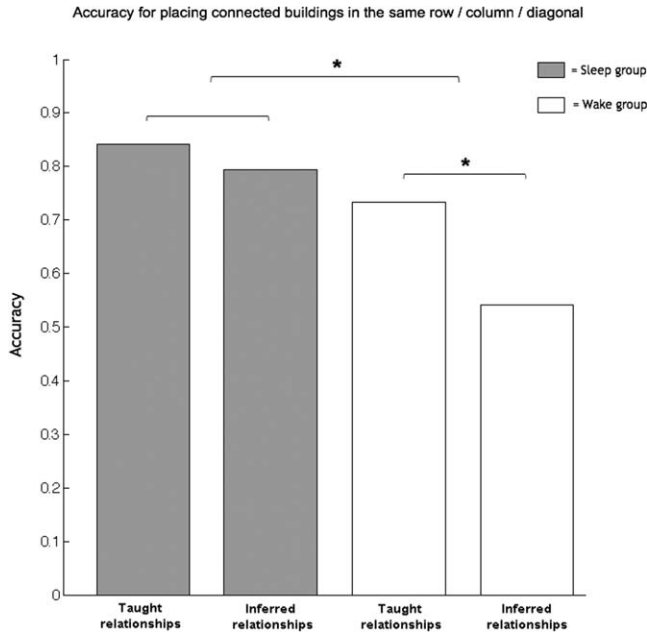


FIGURE 6. Accuracies for placing pairs of buildings in the same row, column or diagonal. Asterisks reflect significant differences at $P < 0.05$. The top asterisk reflects a significant interaction between group and relationship-type. Error bars are not included due to the nonparametric tests performed.

participants in two 24 h conditions: morning-to-morning and evening-to-evening. These groups had both sleep and wakefulness before testing but differed in their training and testing times, like the main 12 h sleep and wake groups. If the main group findings are explainable by time-of-day effects, we would expect to see differences between the 24 h groups. Instead, they had almost identical results for overall map accuracies (evening group: 6 of 10; morning group: 5 of 9; $\chi^2(1) = 0.07$, $P = 0.79$ with Yates' correction for low frequencies) and continuous map performance (untaught: evening group: 64%; morning group: 68%; $P = 0.82$ in Wilcoxon rank sum test). As the two 24 h groups did not differ in these time-of-day comparisons, they were collapsed for subsequent analyses. Also arguing against circadian effects, the 12 h sleep and wake groups did not differ in the numbers of training blocks required to reach criterion ($P = 0.51$ in Wilcoxon rank sum test) or in their MEQ scores ($t(1,17) = 0.65$, $P = 0.52$), reducing the possibility that the above effects were due to one group being more alert at a certain time-of-day.

The longer retention period of the 24 h participants included both sleep and wakefulness. Although we expected the 24 h participants' longer retention interval to slightly lower their behavioral performance (as often found: Ellenbogen et al., 2006, 2007), we hypothesized that the presence of sleep (rather than addition of interference) would significantly influence performance, particularly on the continuous map metric, which separated-out measures that were, and were not, hypothesized to show a sleep effect. The results followed these predictions. Overall map accuracy in the 24 h participants (11 of 19) was numerically between the two 12 h groups. This measure

included material that was believed to be both affected and unaffected by sleep, so we next examined the continuous map metric. We probed the specific role of sleep in the training-testing interval by conducting orthogonal planned contrasts on the full set of data (including 24 h participants [M taught = 75%; M inferred = 66%]). The 12 h sleep and 24 h participants all experienced sleep in the training-testing interval, while the 12 h wake participants did not. The specific inclusion of sleep in the interval affected whether inference performance was similar to, or less than, taught performance: $F(2,55) = -2.23$, $P = 0.03$. In contrast, adding 12 h of wakefulness to a period of sleep did not reduce the effect ($F(2,55) = -0.74$, $P = 0.46$).

Retention of Taught Associations

The sleep and wake groups did not differ in retention-loss of taught pairs over the training-testing interval (Wake M loss = 6.2% of post-training knowledge; Sleep M loss = 1.7%; $P = 0.17$ in Wilcoxon rank sum test). As there was no sleep effect, there was no reason to compare these results to the 24 h participants. The two different 24 h groups did not differ in retention rates ($t(17) = 0.69$, $P = 0.50$), again arguing against time-of-day effects.

The immediate post-training test (i.e., the training task without feedback) gave an opportunity to repeat the prior assessment of map inferences in relation to each participant's knowledge of underlying taught associations (but now using the immediate-test as the indicator of taught-pair knowledge). In line with the above results, participants with a 12 h interval that included sleep ($M = 82\%$) had more associations in the inference columns and diagonals relative to their own knowledge of taught steps (at a trend-level: $P = 0.09$ in Wilcoxon rank sum test), as compared to participants experiencing just wakefulness ($M = 61\%$).

Distance Judgments

Two participants in the 12 h wake and sleep groups failed to give accurate distance judgments about taught pairings (<54% as compared to the next lowest subject at 72%) despite strong performance in the training task with the same associations. This suggested they did not transition well to the new task, so their data were removed from subsequent analyses. In the distance judgments, participants selected which of two landmarks was closer to a target building. The sleep ($M = 84\%$) and wake ($M = 80\%$) groups did not differ in performance on the distance judgments involving buildings in explicitly taught pairings ($t(1,34) = -1.05$, $P = 0.30$). We separated the distance judgments involving nontaught relationships into two categories based on whether they relied on: (i) inferences with concordant numbers of separating taught connections and distance, (ii) inferences with discordant steps and distance. Participants in the wake and sleep groups performed significantly above chance for both question types (Fig. 7), although unexpectedly, the groups did not differ in accuracy ($P > 0.80$ for pairwise comparisons).

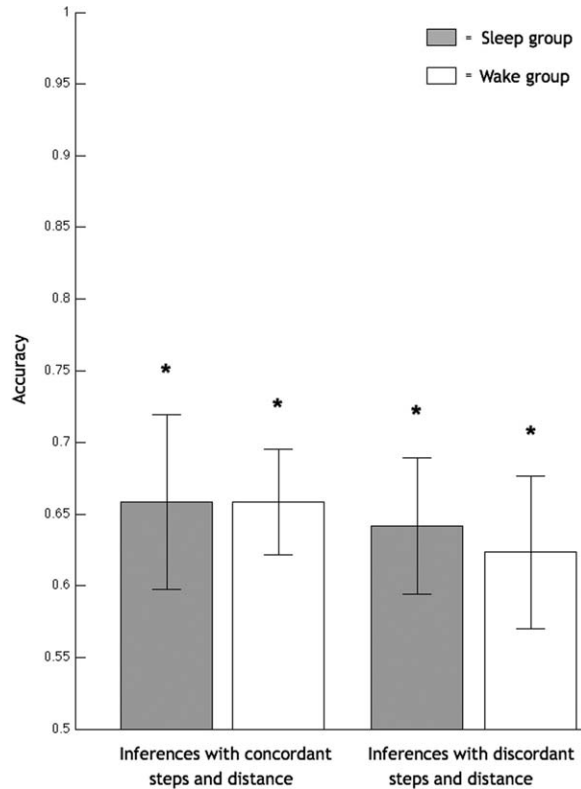


FIGURE 7. Accuracies for distance judgments. Mean group performance for questions involving inferences with concordant steps and distance (left), and questions with discordant steps and distance (right). Error bars reflect the standard error of the mean. Asterisks indicate significantly above-chance performance. The y-axis begins at the level of chance.

We next examined whether performance on the distance judgments that specifically required knowledge of distance rather than numbers of connecting steps (i.e., where distance and steps were discordant) was related to the extent of participants' survey or route navigational styles. Sleep-group participants with greater accuracies on discordant questions had higher QSR survey ($r(17) = 0.49$, $P = 0.04$), but not route ($r(17) = 0.02$, $P = 0.94$) scores (Hotelling/Williams test for different dependent correlations: $t(15) = 1.92$, $P = 0.07$; see Fig. 8). In contrast, the wake group's accuracy was not related to either survey ($r(17) = 0.36$, $P = 0.15$) or route ($r(17) = 0.12$, $P = 0.63$) scores (Hotelling/Williams: $t(15) = 0.79$, $P = 0.44$). The above effects were specific to discordant questions (requiring knowledge of distance), as the sleep group's survey scores did not significantly correlate with concordant question accuracies ($r(17) = 0.13$, $P = 0.61$; Hotelling/Williams: difference from discordant relationship: $t = 2.01$, $P = 0.06$). The wake group had similar survey-score relationships with both discordant and concordant accuracies (Hotelling/Williams: $t = -0.26$, $P = 0.80$). General navigation ability (independent of navigation style) measured through the SBSOD scale, was not related to discordant performance in the wake ($r(17) = 0.30$, $P = 0.22$) or sleep ($r(17) = 0.12$, $P = 0.62$) groups.

DISCUSSION

In this study, we taught participants a subset of spatial relations that formed a two-dimensional space, and measured the effects of sleep on their knowledge of taught and inferred (untaught) relationships that spanned the learned space. After a 12 h period including sleep, participants produced more accurate maps of the space. A detailed analysis of map errors revealed that untaught inferred associations were particularly benefited by sleep. In contrast, the explicitly taught associations were successfully incorporated into the map after both wakefulness and sleep, suggesting the taught information did not require sleep to become incorporated. An analysis of inference performance relative to taught associations showed that participants more readily translated their learned material into map inferences following sleep. A group-level analysis did not show any differences in distance-judgment accuracy; however, an analysis of individual differences revealed that after 12 h including sleep (but not wakefulness), performance could be predicted by self-reported survey (but not route) navigational styles.

Analyses of two 24 h groups that were trained and tested at different times argue that these effects were unlikely to be due to time-of-day differences. An examination of the specific influence of sleep (which was possible from including 24 h participants) showed that the presence of sleep affected participants' knowledge of inferred relationships. In contrast, this was not affected by whether (potentially interfering) wakefulness was added to sleep in the training-testing interval. These results are consistent with an effect of sleep, rather than an absence of daytime interference, on consolidation.

The one-dimensional hierarchy typically learned in transitive inference experiments does not reveal whether the representations formed by participants have a spatial component (akin to one dimension, such as East to West) or if they are less spatial and more abstract. Our design used a task with a spatial framing. Our reported relationship between "discordant" comparisons (where knowledge of Euclidian distance is required over the numbers of steps) and participants' navigational strategy indicates a connection between spatial processing and relational memory in this context. Future work may wish to address whether spatial strategies are deployed in similar inference tasks, or if this is specific to the characteristics of this study.

Our finding that judgments requiring knowledge of Euclidian distance (and not just connecting steps) related to individuals' survey navigational style in the sleep, but not wake, group raises the intriguing possibility that some behavioral consequences of a survey strategy may only become accessible after spatial knowledge is consolidated. In a prior investigation of the survey style, Pazzaglia and Taylor (2007) speculated that high-survey participants might have navigated well in a virtual environment by mentally reconstructing a previously shown map, or perhaps by implementing a different navigational strategy to low-survey individuals. As our study did not show participants a map of the space (nor mention that they would later construct one), and also did not test participants in a virtual environment, our results

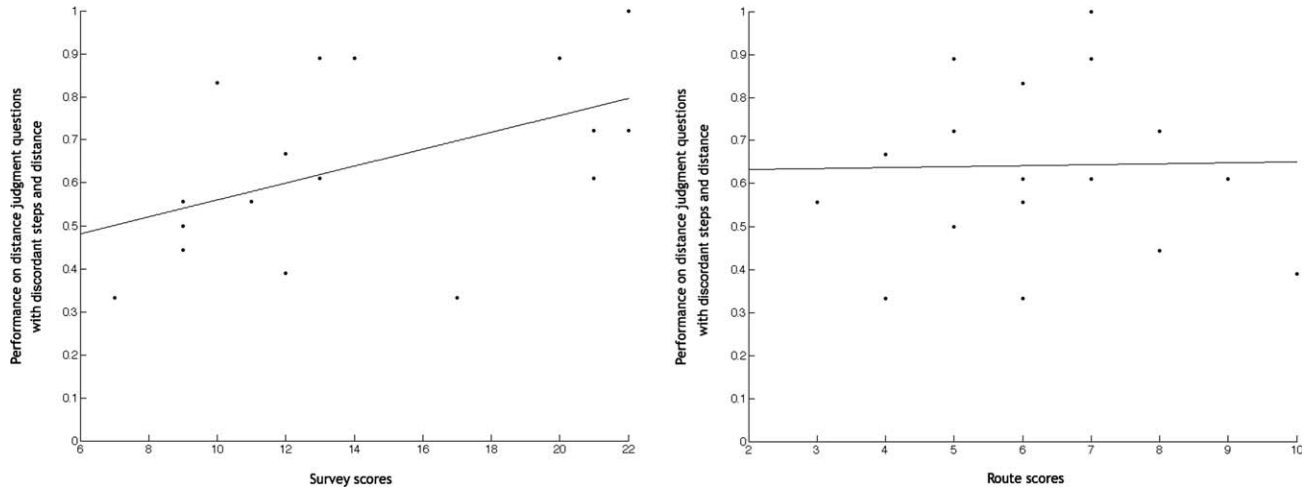


FIGURE 8. Relationships between navigation styles and distance judgment performance in the sleep group. The *x*-axes show survey (left) and route (right) navigation strategy scores for participants in the 12 h sleep group; the *y*-axis shows performance on the subset of distance judgments requiring spatial inference (i.e., discordant steps and distance; see text for details). The strength of the survey style was significantly correlated with distance task performance in the sleep group.

raise the interesting possibility that following sleep, high-survey individuals may not need to have seen a map to fully realize their spatial knowledge.

What mechanism might be enhancing participants' retrieval of inferred relationships? The most compelling possibility is the consolidation of a cognitive map, which then allows participants to move beyond taught pairings and access inferred spatial relationships. This mechanism fits well with the discovered relationship between the survey approach and performance on distance judgments that require knowledge of true Euclidian distance in the space (i.e., “discordant” questions). This result—namely, that participants who navigate in a way conducive to forming cognitive maps do particularly well on these questions after sleep—provides convincing evidence that consolidation processes transform learned information to a format similar to a cognitive map. One alternative possibility, however, is that taught material is consolidated in a manner that allows it to be recombined “online” during testing. As performance on explicitly taught associations did not significantly improve after sleep (also found in Ellenbogen et al., 2007), this alternative would have to allow flexible use without a change in explicit knowledge. A third possible consolidation effect is for spatial inferences becoming more accessible from taught information, without being part of a cognitive map *per se*.

Prior evidence for hippocampal involvement in building relational memories—for example, hippocampal lesions reducing rats' abilities to correctly respond to inferences, but not underlying pairings (Dusek and Eichenbaum, 1997)—is suggestive that the hippocampus does more than strengthen taught pairs. The “cognitive map” and “spatial inference” possibilities are both consistent with theories that sleep reorganizes learned material, allowing insight into a “hidden structure” (Born and Wilhelm, 2012, p. 200), which could be considered the grid in this study.

Hippocampal activation has been particularly associated with the survey spatial approach (but not other strategies: Hartley et al., 2003; Iaria et al., 2003), so what does the selective relationship between survey style and discordant performance say about how memory networks change with consolidation? There are several possibilities. First, individual differences in the survey style may be associated with encoding different components of the presented information. If certain information is not encoded, it will not be reactivated during consolidation, which might then lower inference performance if that information is crucial for forming inferences. Second, some models of hippocampal consolidation suggest that neurons associated with strong overlapping memories are more likely to be reactivated (than weaker memories) during sleep (Meeter and Murre, 2005; Lewis and Durrant, 2011). If differences in spatial approaches moderate memory strengths of various components of encountered information (such as the link between associations and the map structure), individuals will likely vary in the consolidated memories available for making distance judgments. This idea, that survey-based individuals may consolidate differently, parallels findings that CA1 pyramidal cells with overlapping place fields (representing adjacent spatial locations) are more likely to fire together in postexploration sleep than neurons for distant locations (Wilson and McNaughton, 1994; O'Neill et al., 2010). If survey-based participants do construct a cognitive map during consolidation, locations that are close only through untaught inferred connections (e.g., storefronts 1 and 4 in Fig. 4) might become linked in an analogous fashion to CA1 cells with overlapping place fields. It is worth noting that a route-based approach may also be associated with sleep benefits. The survey cognitive style was particularly sensitive to judgments that evaluated an understanding of the overall space, but greater route-based processing may relate to performance in other tasks after sleep.

In our study, a flexible map-creation task was more sensitive to sleep-induced consolidation effects than a distance judgment task. Recreating a map notably involves combining all forms of spatial information into an interconnected representation of space. A key virtue of this study's design (and map test) is that subjects did not navigate the entire space, but instead received limited access to underlying relations, allowing us to later compare taught and inferred aspects within participants' maps. We do note, however, that the map task's flexibility also brings a limitation that subjects can complete a map through a number of different approaches. The particular construction strategy that a subject employs before and after sleep would be a useful focus for future studies. Tools such as eye-tracking and video recordings could help elucidate the proximate cause of sleep participants' superior map performance. That said, it is worth noting that the learning task we employed—designed specifically to build on one-dimensional transitive inference procedures—does not mimic how individuals typically experience spatial environments. Real-world learning of spatial relationships is often unsupervised and involves additional factors such as self-motion. Accordingly, the use of this paradigm will depend on an investigator's hypotheses.

To conclude, we have found that sleeping in the 12 h after learning a subset of associations that form a two-dimensional space, allows participants to retrieve a more accurate map, and better integrate untaught inferred relationships into its structure. Participants with a stronger survey-based navigational style can make more accurate distance judgments requiring spatial inferences, but only after sleeping. Together, these results point to an important role for sleep in consolidating simple relationships into a complex representational space.

Acknowledgments

The authors thank members of the Thompson-Schill lab for helpful discussions about this work and Victor Schinazi and Steven Marchette for comments on earlier versions of the manuscript. Marc Coutanche was funded by a fellowship from the Howard Hughes Medical Institute.

REFERENCES

- Bird CM, Burgess N. 2008. The hippocampus and memory: Insights from spatial processing. *Nat Rev Neurosci* 9:182–194.
- Born J, Wilhelm I. 2012. System consolidation of memory during sleep. *Psychol Res* 76:192–203.
- Buyse 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.
- Dusek JA, Eichenbaum H. 1997. The hippocampus and memory for orderly stimulus relations. *Proc Natl Acad Sci U S A* 94:7109–7114.
- Ellenbogen J, Hu P, Payne J, Titone D, Walker M. 2007. Human relational memory requires time and sleep. *Proc Natl Acad Sci U S A* 104:7723–7728.
- Ellenbogen JM, Hulbert JC, Stickgold R, Dinges DF, Thompson-Schill SL. 2006. Interfering with theories of sleep and memory: sleep, declarative memory, and associative interference. *Curr Biol* 16:1290–1294.
- Ferrara M, Iaria G, Tempesta D, Curcio G, Moroni F, Marzano C, De Gennaro L, Pacitti C. 2008. Sleep to find your way: The role of sleep in the consolidation of memory for navigation in humans. *Hippocampus* 18:844–851.
- Hartley T, Maguire EA, Spiers HJ, Burgess N. 2003. The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron* 37:877–888.
- Hegarty M, Richardson AE, Montello DR, Lovelace K, Subbiah I. 2002. Development of a self-report measure of environmental spatial ability. *Intelligence* 30:425–447.
- Horne JA, Ostberg O. 1976. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *Int J Chronobiol* 4:97.
- Iaria G, Petrides M, Dagher A, Pike B, Bohbot VD. 2003. Cognitive strategies dependent on the hippocampus and caudate nucleus in human navigation: Variability and change with practice. *J Neurosci* 23:5945–5952.
- Johns MW. 1991. A new method for measuring daytime sleepiness: The Epworth sleepiness scale. *Sleep* 14:540–545.
- Konkel A, Cohen NJ. 2009. Relational memory and the hippocampus: Representations and methods. *Front Neurosci* 3:166–174.
- Kumaran D, Maguire E. 2005. The human hippocampus: Cognitive maps or relational memory? *J Neurosci* 25:1103–1105.
- Kumaran D, Melo HL, Duzel E. 2012. The emergence and representation of knowledge about social and nonsocial hierarchies. *Neuron* 76:653–666.
- Lewis PA, Durrant SJ. 2011. Overlapping memory replay during sleep builds cognitive schemata. *Trends Cogn Sci* 15:343–351.
- Meeter M, Murre JMJ. 2005. Tracelink: A model of consolidation and amnesia. *Cogn Neuropsychol* 22:559–587.
- O'Keefe J, Nadel L. 1978. The hippocampus as a cognitive map. *Behav Brain Sci* 2:487–494.
- O'Neill J, Pleydell-Bouverie B, Dupret D, Csicsvari J. 2010. Play it again: Reactivation of waking experience and memory. *Trends Neurosci* 33:220–229.
- Pazzaglia F, Taylor HA. 2007. Perspective, instruction, and cognitive style in spatial representation of a virtual environment. *Spat Cogn Comput* 7:349–364.
- Peigneux P, Laureys S, Fuchs S, Collette F, Perrin F, Reggers J, Phillips C, Degueldre C, Del Fiore G, Aerts J, Luxen A, Maquet P. 2004. Are spatial memories strengthened in the human hippocampus during slow wave sleep? *Neuron* 44:535–545.
- Staresina BP, Davachi L. 2009. Mind the gap: binding experiences across space and time in the human hippocampus. *Neuron* 63: 267–276.
- Wamsley EJ, Tucker MA, Payne JD, Stickgold R. 2010. A brief nap is beneficial for human route-learning: The role of navigation experience and EEG spectral power. *Learn Memory* 17:332–336.
- Wilson MA, McNaughton BL. 1994. Reactivation of hippocampal ensemble memories during sleep. *Science (NY)* 265:676–679.
- Zeithamova D, Schlichting ML, Preston AR. 2012. The hippocampus and inferential reasoning: Building memories to navigate future decisions. *Front Human Neurosci* 6:70.