The hippocampus is necessary for the consolidation of a task that does not require the hippocampus for initial learning

Running title: Hippocampal contributions to consolidation

Anna C. Schapiro*[1,2], Allison G. Reid[3], Alexandra Morgan[1], Dara S. Manoach[2,4,5], Mieke Verfaellie[3,6], Robert Stickgold[1,2]

[1]Department of Psychiatry, Beth Israel Deaconess Medical Center, Boston, MA
[2]Harvard Medical School, Boston, MA
[3]Memory Disorders Research Center, VA Boston Healthcare System, Boston, MA
[4]Department of Psychiatry, Massachusetts General Hospital, Charlestown, MA
[5]Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA
[6]Department of Psychiatry, Boston University School of Medicine, Boston, MA

*Corresponding author:
Anna Schapiro
aschapir@bidmc.harvard.edu
Department of Psychiatry
Beth Israel Deaconess Medical Center
330 Brookline Ave

Acknowledgments

Abstract

During sleep, the hippocampus plays an active role in consolidating memories that depend on it for initial encoding. There are hints in the literature that the hippocampus may have a broader influence, contributing to the consolidation of memories that may not initially require the area. We tested this possibility by evaluating learning and consolidation of the motor sequence task (MST) in hippocampal amnesics and demographically matched control participants. While the groups showed similar initial learning, only controls exhibited evidence of overnight consolidation. These results demonstrate that the hippocampus can be required for normal consolidation of a task without being required for its acquisition, suggesting that the area plays a broader role in coordinating memory consolidation than has previously been assumed.

Keywords

Introduction

        The hippocampus plays an important and active role in memory consolidation during sleep (Antony & Paller, 2017). It replays recent experiences during high frequency ripple oscillations that often co-occur with sleep spindles in neocortex, which are themselves associated with memory replay (Cairney, Guttesen, El Marj, & Staresina, 2018; Ji & Wilson, 2007; Nadasdy, Hirase, Czurko, Csicsvari, & Buzsaki, 1999; Peyrache, Khamassi, Benchenane, Wiener, & Battaglia, 2009; Siapas & Wilson, 1998; Sirota, Csicsvari, Buhl, & Buzsaki, 2003; Staresina et al., 2015). This hippocampal-cortical dialogue is thought to facilitate the transfer of new memories encoded in the hippocampus to long term neocortical stores (Siapas & Wilson, 1998). Accordingly, studies to date have focused on uncovering evidence for hippocampal involvement in sleep-dependent consolidation for types of memory that depend on the hippocampus for initial encoding (Antony & Paller, 2017).

        The hippocampus could conceivably play a broader role, however, by helping to reinstate extra-hippocampal memory traces for processing during sleep, an idea supported by recent findings in rodents (Sawangjit et al., 2018). There are hints that this may be the case in humans for a motor learning paradigm known as the motor sequence task (MST; Karni et al., 1995). In this task, participants type a 5-digit sequence (e.g. 4-1-3-2-4) as quickly and accurately as they can for a number of timed trials. The hippocampus seems unlikely to be necessary for performing this task, as the similar serial reaction time task (SRTT) does not require a healthy hippocampus (Nissen, Willingham, & Hartman, 1989; Reber & Squire, 1994) for learning simple sequential dependencies (Curran, 1997). At the same time, performance on this task has been consistently found to benefit from sleep (King, Hoedlmoser, Hirschauer, Dolfen, & Albouy, 2017) and there is strong evidence that the hippocampus is involved in this offline consolidation: Hippocampal activity and connectivity with other regions during initial learning is associated with performance improvement across sleep (Albouy et al., 2015; Albouy, King, Maquet, & Doyon, 2013; Albouy, Sterpenich, et al., 2013; Genzel et al., 2015), post-learning hippocampal activity during sleep is associated with improvement (Moroni et al., 2008) and, after sleep, there is increased activity in the hippocampus while performing the task (Albouy, Sterpenich, et al., 2013; King, Saucier, et al., 2017; Walker, Stickgold, Alsop, Gaab, & Schlaug, 2005).

        Further indirect evidence that the hippocampus is important for sleep-dependent MST consolidation comes from associations between MST improvement and sleep spindles. Sleep spindles and stage 2 sleep (a stage defined by spindle events) are associated with improvement on the MST (Albouy, Fogel, et al., 2013; Barakat et al., 2013; Barakat et al., 2011; Boutin et al., 2018; S. Fogel, Albouy, et al., 2017; S. Fogel, Vien, et al., 2017; S. M. Fogel et al., 2014; Laventure et al., 2016; Laventure et al., 2018; Manoach et al., 2010; Nishida & Walker, 2007; Walker, Brakefield, Morgan, Hobson, & Stickgold, 2002), and spindles are in turn often associated with hippocampal replay (Ji & Wilson, 2007; Peyrache et al., 2009; Siapas & Wilson, 1998; Sirota et al., 2003; Staresina et al., 2015). Consistent with the idea that spindles can provide an index of hippocampal involvement in consolidation, they have often been associated with improvement in tasks that are known to depend on the hippocampus (Saletin & Walker, 2012) (though not all

hippocampally dependent tasks show spindle correlations; Ackermann, Hartmann, Papassotiropoulos, de Quervain, & Rasch, 2015). In addition, patients with hippocampal sclerosis due to temporal lobe epilepsy and patients with amnestic mild cognitive impairment, which is associated with hippocampal dysfunction (Petersen et al., 2006), have fewer spindles than normal and deficits in consolidation of hippocampally dependent memory (Fuentemilla et al., 2013; Westerberg et al., 2012).

It is unclear from this collection of findings whether the hippocampus plays a *causal* role in MST learning and consolidation. First, is the hippocampus required for initial learning? There is extensive neuroimaging evidence that the hippocampus is engaged by performing the MST and other motor learning tasks like the SRTT (Albouy, King, et al., 2013; Fernandez-Seara, Aznarez-Sanado, Mengual, Loayza, & Pastor, 2009; Gheysen, Van Opstal, Roggeman, Van Waelvelde, & Fias, 2010; Harrison, Duggins, & Friston, 2006; Schendan, Searl, Melrose, & Stern, 2003), but we cannot determine from neuroimaging studies whether the hippocampus is necessary for normal MST performance or whether it is merely engaged by it. Amnesics' normal performance on the SRTT (Nissen et al., 1989; Reber & Squire, 1994) favors the latter possibility, though the MST may engage the hippocampus differently. Second, is the hippocampus necessary for normal consolidation of the task? The hippocampus is associated with offline gains in performance in neuroimaging studies, but, again, this may be epiphenomenal.

To answer these questions, we tested learning and consolidation of the MST in patients with severe amnesia due to hippocampal damage and demographically matched control participants. We found that amnesic patients performed similarly to control participants in their initial learning of the MST, indicating that, as in the SRTT, the hippocampus is not required for normal learning. In contrast, unlike controls, the patients exhibited no evidence of overnight consolidation. These findings are consistent with the prior literature but indicate that the previously observed hippocampal engagement in the initial learning of the MST may be primarily setting the stage for later offline involvement. Our results demonstrate that the hippocampus is necessary for the consolidation of a form of memory that does not require the hippocampus for acquisition, suggesting that the hippocampus plays a broader role in memory consolidation than was previously understood.

Methods

Participants. Eight patients with medial temporal lobe lesions (6 males, 7 right-handed) and 12 control participants (10 males, 8 right-handed) participated in the study. Etiology for the patients was hypoxic-ischemic injury secondary to cardiac or respiratory arrest (n=5), encephalitis (n=1), stroke (n=1), and status epileptics followed by left temporal lobectomy (n=1). All patients were in the chronic phase of illness, with time post injury ranging from 3.5 to 36.4 years (mean= 21.1).

Four patients and two controls did not meet the inclusion criterion on their first session, which required a minimum of 10 correct sequences on average over the last three trials of training. The average scores for the excluded controls were 8.3 and 9.1 sequences and for patients were 8.6, 3.7, 3.2, and 3.1 sequences. The patient with the highest score was tested on a second sequence, but again failed to meet threshold, with a

score of 9.3. The other three patients were not tested on a second sequence. We do not believe that these low scores reflect a sequence learning deficit, but rather a motor deficit. These four patients were also the slowest of all participants on the warmup task (described below), which does not require sequence learning[1]. Their average time between keypresses during the warmup was 679, 900, 965, and 586 ms, relative to a mean of 348 ms (SD = 180) for other patients and mean of 408 ms (SD = 153) for controls. The group of included patients was no different than the control subjects in warmup task speed (t[14]=0.38, p=0.71) whereas the excluded patients were much slower (t[14]=4.97, p=0.0002). We include analysis of the data for the excluded participants in the Supporting Information.

Demographic and neuropsychological characteristics for the four included patients are provided in Table 1. The neuropsychological profiles of each patient indicated severe episodic memory impairment (mean General Memory Index = 64.5), with otherwise preserved cognition (mean VIQ = 110.3; mean Working Memory index = 108.5). Lesions for three of the patients are shown in Figure 1. The remaining patient (P03) had suffered cardiac arrest, and could not be scanned due to medical contraindications. Medial temporal lobe pathology for this patient was inferred based on etiology and neuropsychological profile. Two patients (P02 and P04) had lesions restricted to the hippocampus, and one patient had volume loss extending outside of the hippocampus (P01). P04 has a general anxiety disorder, both pre- and post-amnesia onset. Otherwise, the patients have no history of or present psychiatric disorder.

The 10 control participants included in analyses were well matched to the included patients in terms of sex (8 males; all included patients male), handedness (9 right-handed; all included patients right-handed), age (mean=57.7; patients mean=58.0), years of education (mean=14.7; patients mean=16.8), and VIQ (mean=112.1; patients mean=110.3).

All participants provided informed consent in accordance with the Institutional Review Board of VA Boston Healthcare System.

Procedure. The task was presented to participants on a laptop using MATLAB with Psychophysics Toolbox (Brainard, 1997). Participants were instructed to rest four fingers of their left hand on a button box with buttons labeled 1, 2, 3, and 4. At the beginning of each session, participants completed a warmup task, where they were instructed to repeatedly type the sequence 1-2-3-4 when the screen turned from red to green and to "try to be as fast and accurate as you can." The sequence was always displayed on the screen, during both rest and typing periods. The number of seconds until the screen turned green was then displayed as spelled out numbers ("ten", "nine", "eight"…). The screen remained green for 30 seconds. With every key press, a new dot appeared in a horizontal line on the screen to provide feedback to the participant that their key press was registered. After the line reached the right side of the screen, the dots then disappeared one at a time with each additional key press. The experimenter wore one ear bud through which she would hear a beep whenever the participant pressed a button out

---

[1] Two of these patients had basal ganglia damage: one had extensive volume reduction in caudate, putamen, and pallidum bilaterally (z's< -2.29), and the other had reduction in left pallidum only (z=-2.58). The former patient additionally has bipolar disorder. It is possible that these factors contributed to the slower motor performance in these patients.

of sequence. This allowed her to provide rapid feedback if the participant was not pressing the buttons correctly.

After 30 seconds, the screen turned red, and participants were instructed to stop typing and take a 30 second break. A thirty second countdown immediately began with the number of seconds left again spelled out on the screen. At the end of the countdown, the screen turned green again for 30 seconds, and the participants again typed the warmup sequence. If the participant did not yet seem comfortable with the task, the experimenter had the option to initiate additional warmup trials.

Once the participant was accustomed to the task, the experimenter initiated the training phase. The training had the same structure as the warmup, with 30 seconds of typing interspersed with 30 seconds of rest. One of four sequences was used: 4-1-3-2-4, 1-4-2-3-1, 3-1-4-2-3, or 2-4-1-3-2. The sequence assignment was counterbalanced across subjects. In addition to the sequence being displayed continuously on the screen, an index card displaying the sequence was also placed next to the keypad so that participants did not have to look at the screen while typing. Participants completed 12 trials of training. Throughout the session, the experimenter reminded participants to start and stop typing as needed.

This first session took place at a time that was convenient for the participant. 24 hours later, participants completed another warmup and then the test, which consisted of 12 additional trials of the same sequence that they had typed the previous day. Because there were a limited number of patients, we aimed to obtain high fidelity estimates of each patient's performance. Patients who met the performance criterion (10 sequences correct across the last three trials of training) for their first sequence were therefore tested again several months later on a second sequence. The sequence assignment was again counterbalanced across participants. Patients completed the same two-day procedure with the second sequence.

On each day of training or testing, participants filled out a survey asking how well they slept the previous night, the duration of their sleep, and how alert they felt. On the morning of test days, participants also answered these questions using a paper survey filled out at home around the time of awakening.

Mixed effects model. To assess whether the patients and controls differed in their behavior across the training and test days, we fit a mixed effects model, with participant as random effect and group, day, trial, and sequence as fixed effects:

$$\text{performance} \sim (1 \mid \text{participant}) + \text{group} * \text{day} + \text{trial} + \text{sequence}$$

where day indicates training vs. test, and sequence indicates the first or second sequence used for patients. Significance of factors was assessed by removing the factor (or interaction between factors) from the model and assessing the difference between the original and modified models using the χ^2 statistic.

Trial outlier removal. As a preprocessing step for percent change analyses, where a small subset of trials was used (making the analyses potentially sensitive to trial outliers), we removed trials for each participant that fell far from an estimated learning curve, as follows. For each participant and each day, a power function ($y=b \cdot x^m$) was fit to performance across the 12 trials. The squared residuals for each trial were calculated, and trials falling more than 2 SD outside the distribution of squared residuals across all trials

and subjects were excluded. This resulted in exclusion of 19 trials out of 432 (4.4%). 12 of these trials came from patients and 7 from controls. We then averaged the data across the two sequences that each patient completed. When a trial was missing from one sequence and not the other, the non-missing data point was used. This resulted in just one missing trial across all the patient data. Figures 2B and 2C shows the data for individual controls and patients after this trial exclusion process (dashed blue line indicates the one missing trial for patient data). These preprocessing steps served to provide smoother estimates of patient performance given the small number of patients, minimizing the influence of outlier trials.

  Percent change calculation. Our estimate of the "initial" percent change from the end of training to the beginning of test was 100•(mean of first 3 test trials–mean of last 3 training trials)/mean of last 3 training trials. The "plateau" change was calculated as 100•(mean of last 6 test trials–mean of last 3 training trials)/mean of last 3 training trials. Percent improvement over the course of training was 100•(mean of last 3 training trials–first training trial)/first training trial. Two control subjects were not included in this analysis: for one, the first trial of training was excluded as an outlier, and for the second, a technical issue resulted in loss of data for the first three trials of training (this was the only data loss that occurred during the study). Differences between groups and differences of each group from zero were calculated using two-tailed t tests.

  Data Availability. Behavioral data are available as a supplemental file.

<center>Results</center>

  Sleep and alertness surveys. Controls and patients reported similar amounts of sleep both the night before training (controls = 7.2 ± 1.1 (S.D.) hours; patients = 8.0 ± 1.5; $t[12]=1.05$, $p=0.32$) and the night between training and test (controls = 7.3 ± 1.1 hours, patients = 7.8 ± 1.8; $t[12]=0.65$, $p=0.53$). They also reported similar sleep quality on both nights (1 = slept very poorly to 7 = very well; pre-training: controls = 5.6 ± 1.1, patients = 5.3 ± 0.3; $t[12]=0.63$, $p=0.54$; post-training: controls = 5.3 ± 1.1; patients = 5.8 ± 0.6; $t[12]=0.78$, $p=0.45$) and similar alertness (1 = may fall asleep to 7 = wide awake) both at the time of training (controls = 6.2 ± 0.8, patients = 6.1 ± 0.9; $t[12]=0.16$, $p=0.88$) and at test (controls = 5.9 ± 1.2; patients = 6.3 ± 1.5; $t[12]=0.46$, $p=0.65$).
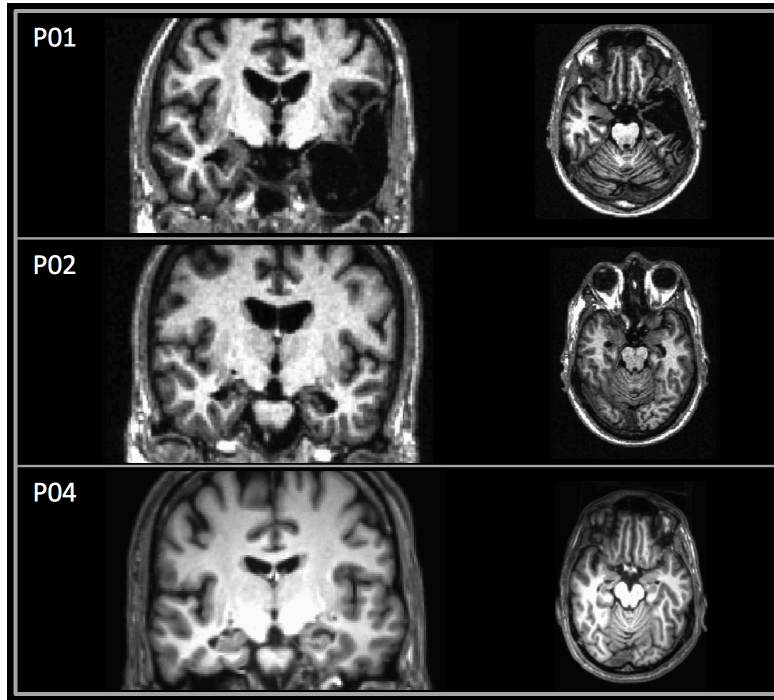
Figure 1. Coronal and axial T1-weighted magnetic resonance images depict lesions for patients P01, P02, and P04 (no scans were available for P03). The left side of the brain is displayed on the right side of the image.

Table 1. Demographic and neuropsychological characteristics of amnesic patients

| | Etiology | Age | Edu | WAIS III | | WMS III | | | Years since onset | % VL in bilateral hippocampus | % VL in subhippocampal cortex |
				VIQ	WMI	GMI	VD	AD			
P01	Status epilepticus + left temp. lobectomy	53	16	93	94	49	53	52	27.3	63%	60%[a]
P02	Hypoxic-ischemic	61	14	106	115	59	72	52	24.2	22%	---
P03	Hypoxic-ischemic	65	17	131	126	86	78	86	15.0	unknown	unknown
P04	Stroke	53	20	111	99	60	65	58	3.45	43%	---

Note: Age = age in years at time of first training session; Edu = education in years; WAIS-III = Wechsler Adult Intelligence Scale-III; WMS-III = Wechsler Memory Scale-III; VIQ = verbal IQ; WMI = working memory index; GMI = general memory index; VD = visual delayed; AD = auditory delayed; VL = volume loss. [a]VL in left anterior parahippocampal gyrus (i.e., entorhinal cortex, medial portion of the temporal pole, and the medial portion of perirhinal cortex). See Kan, Giovanello, Schnyer, Makris, and Verfaellie (2007) for methodology.

Learning time course. Control and patient learning curves were remarkably similar throughout the course of training, whereas the two groups separated in the test phase, with patients performing worse than controls (Figure 2a). Both groups displayed a drop in performance from the end of training to the beginning of test and then a quick rise to a stable performance level that was maintained for the rest of test. This pattern has been observed before in older adult participants performing the MST and may reflect a need for older participants to re-establish a task set before true performance levels can be expressed (Tucker, McKinley, & Stickgold, 2011). A similar lag in test performance has also been seen in younger subjects performing a 9-digit bimanual version of the task (Kuriyama, Stickgold, & Walker, 2004).
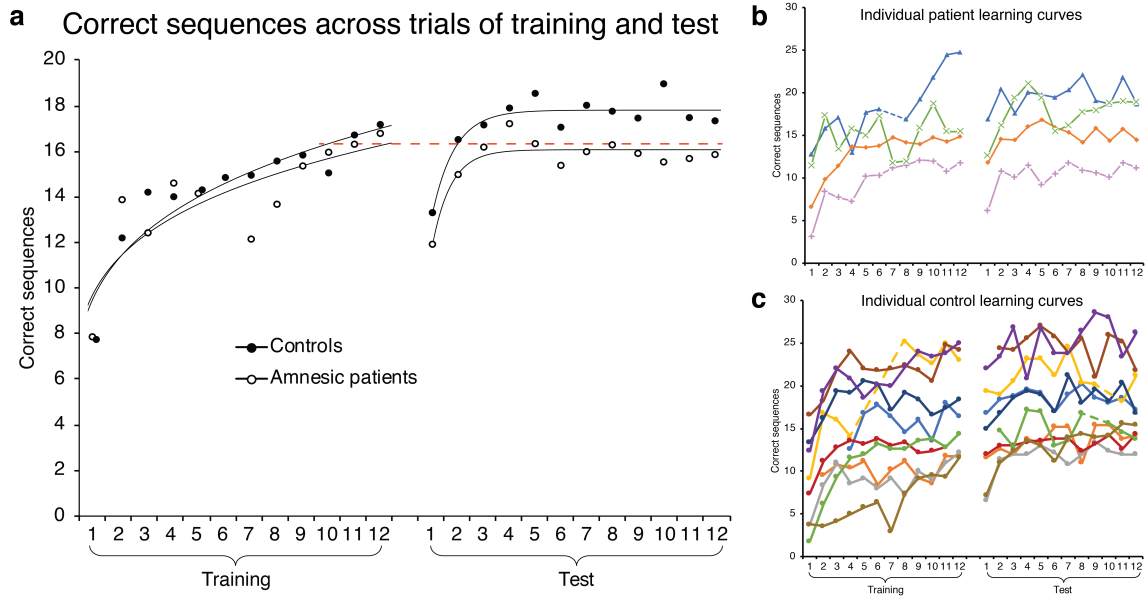

Figure 2. Performance across training and test for controls and amnesic patients. (a) Curves fit to average performance are power functions for training and exponential functions for test. Dashed red line set at average performance across last three trials of Training, which was the same for controls and patients (16.3 sequences correct). Correct sequences across trials of training and test (b) for individual patients and (c) for individual control participants, after removing outlier trials. Patient curves are averaged across the two sequences that each patient completed. Dashed lines connect across removed outlier trials that occurred in the middle of training or test. Each patient is plotted with a unique marker: P01 is green "x"; P02 is orange diamond; P03 is pink plus; P04 is blue triangle.

To test whether group differences were statistically reliable, we ran a mixed effects model. The model revealed reliable effects of day ($\chi^2(1)=13.65$, $p=0.0002$), trial ($\chi^2(1)=78.46$, $p<0.0001$), and, critically, a group by day interaction ($\chi^2(1)=5.63$, $p=0.018$), with the difference between controls and patients larger at test than during training. This indicates that the amnesic patients showed less offline improvement on the MST than controls. There was no effect of first vs. second sequence ($\chi^2(1)=0.06$, $p=0.80$), indicating that patients did not perform differently on their first vs. second set of two-day sessions. The number of correct sequences was very similar for these two sets of sessions: Mean performance across training trials was 13.5 in the first set of sessions and 14.4 in the

second set, and mean performance across test trials was 15.9 in the first set of sessions and 15.2 in the second set.

   Change from end of training to test. Following the analysis strategy of a prior paper assessing MST consolidation in a patient group (Manoach et al., 2010), we calculated for each participant the percent change from the last three training trials to both the first three (initial change) and last six (plateau change) test trials. For the initial change, patients exhibited a near-significant decline in performance (Figure 3; mean= –13.6%, $t[3]$=2.68, p=0.075, *Cohen's d*=1.34) whereas at plateau performance they almost fully recovered to the level of performance achieved at the end of training (mean= –1.9%, $t[3]$=0.36, p=0.74, d=0.18). Controls, in contrast, performed at the same level at the beginning of test as at the end of training (mean = –0.05%, $t[9]$=0.02, p=0.99, d=0.01) and showed reliable improvement at plateau (mean = 12.4%, $t[9]$=2.54, p=0.032, d=0.80). The difference between groups was reliable for the initial change ($t[12]$=2.51, p=0.027, d=1.49) but not for plateau change: $t[12]$=1.68, p=0.118, d=1.0).

   We verified that these results were qualitatively similar (albeit noisier) when considering only the first of the two-day sessions completed by the patients. For the initial change, patients exhibited a numerical decline in performance of the same magnitude as seen for the combined visits (mean=13.3%, $t[3]$=1.60, p=0.21) and a return to training performance at plateau (mean=0.0%). Patients performed marginally worse than controls for initial change ($t[12]$=2.00, p=0.07) and numerically worse than controls at plateau ($t[12]$=1.54, p=0.15).
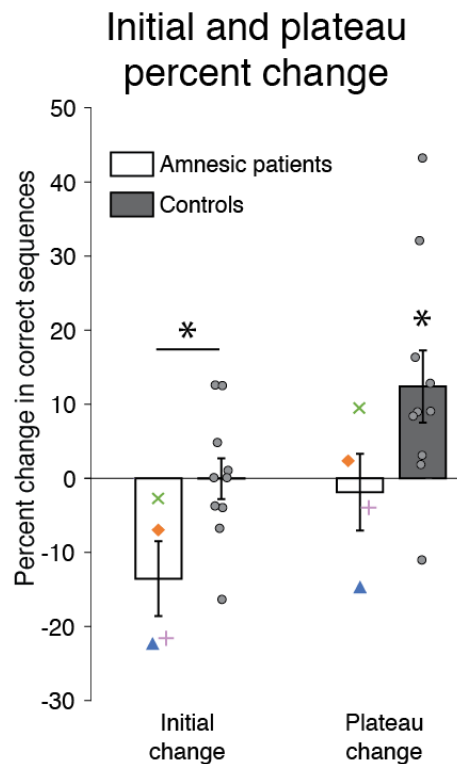

Figure 3. Percent change in performance from training to test. Individual controls are plotted as gray circles. Asterisk above horizontal line denotes significant difference between groups;

asterisk without line indicates where condition differs from zero. Error bars denote ± 1 SEM. *
p<0.05, t-test.

Training performance. To verify that there was no difference in initial learning between the two groups, we assessed the percent change from the first training trial to the last three training trials. Percent improvement across training averaged 131% for patients and 176% for controls, levels which were not significantly different between the groups ($t[10]=0.41, p=0.69, d=0.25$).

Discussion

These findings demonstrate a critical role for the hippocampus in the consolidation of a task that does not require the hippocampus for initial learning. Performance on the MST was assessed in two sessions separated by 24 hours in patients with amnesia due to hippocampal damage and in matched control participants. Patients and controls performed similarly on the first day during training, but unlike controls, patients did not improve overnight; they performed significantly worse than controls on the second relative to first day. The patients retained much of their learning across sessions, exhibiting savings in performing the motor sequence across days, as has been well-documented in related tasks in amnesics since patient H.M. (Corkin, 1968). But their poor performance at test relative to controls indicates a selective deficit in offline consolidation. Indeed, even at the end of test, patients showed no improvement compared to the end of training. Based on these findings, we conclude that even when it is not necessary for initial learning, the hippocampus can play a causal role in consolidation.

How might the hippocampus be involved in consolidation of the MST? There is evidence that the hippocampus is engaged during initial learning of the task, and that this activity predicts sleep-dependent consolidation (Albouy, King, et al., 2013). There are several possibilities for what this activity might reflect. One possibility is that the hippocampus is learning a representation of the sequence in parallel with, and perhaps in interaction with, the striatum (Albouy, Sterpenich, et al., 2013). There is extensive evidence that the hippocampus is involved in learning sequential content in both motor (Fernandez-Seara et al., 2009; Gheysen et al., 2010; Harrison et al., 2006; Schendan et al., 2003) and non-motor domains (Rose, Haider, Salari, & Buchel, 2011; Schapiro, Kustner, & Turk-Browne, 2012; Schapiro, Turk-Browne, Norman, & Botvinick, 2016) and that it is necessary for normal sequential learning in non-motor domains (Covington, Brown-Schmidt, & Duff, 2018; Schapiro, Gregory, Landau, McCloskey, & Turk-Browne, 2014) (perhaps because there is less parallel learning occurring in the striatum in non-motor tasks). The hippocampus may thus benefit consolidation by replaying the content of the sequence during sleep in the same way it would for a paradigm that required the hippocampus for acquisition. If the hippocampus is learning the sequential content alongside the striatum, it is likely that this representation would take a different form than the striatal representation. Indeed, there is evidence that the striatum learns the motor contingencies of the MST (an "egocentric" representation) while the hippocampus learns a more abstract ("allocentric") representation of effector-independent spatial contingencies, and that it is this hippocampal version of the representation that undergoes

sleep-dependent improvement (Albouy et al., 2015; Albouy, Fogel, et al., 2013; King, Hoedlmoser, et al., 2017).

Another possible role for the hippocampus during initial learning is that instead of learning the content of the sequence per se, it is tagging striatal or motor cortical memories for later offline processing (Albouy, King, et al., 2013; Albouy et al., 2008). One version of this possibility is that the hippocampus binds incidentally encoded contextual information with sequential representations stored in other areas. This contextual information could then be revisited during sleep, in turn helping to reinstate the sequential representations stored in the other areas. In other words, the hippocampus could be reminding the sleeping brain that it did a sequence learning experiment earlier that day. Future work will be needed to investigate and adjudicate between these possibilities.

These possibilities assume that patients and controls employed the same learning mechanisms during acquisition, as suggested by their similar learning curves. However, it is worth noting the possibility that patients might have used compensatory mechanisms that are not representative of typical function, for example using their striatum while controls relied more on the hippocampus. This could be assessed in future work by testing whether patients were learning egocentric or allocentric representations of the task.

We based our protocol on a prior MST study in older adults (ages 60–79) who completed a training session and then a test 24 hours later, including a night of sleep, as in the current study, or a training session in the morning and a test 12 hours later, with no intervening sleep (Tucker et al., 2011). The participants in the sleep condition showed a similar pattern to our control subjects, with no initial improvement but reliable plateau improvement. A direct statistical comparison between the control group in the current study and the sleep group from the Tucker et al. (2001) study found no difference for immediate or plateau percent change (ps>0.69). The results for participants in the Tucker et al. (2001) wake group were remarkably similar to those for our amnesic patients, with worse initial performance and plateau performance at the same level as performance at the end of training (testing for differences between these two groups: ps>0.82). The similarity between the amnesic patients and the healthy older adults in the wake group suggests that the impact of hippocampal damage on sleep-dependent memory processing is similar to that of not having any post-training sleep.

This correspondence between the two studies helps to rule out an alternative explanation for our results, which is that the hippocampus is not needed for consolidation but instead for reinstating context of the task from the previous day (Palombo, Di Lascio, Howard, & Verfaellie, 2018). If improved performance at test in healthy subjects is due to the hippocampus reinstating context, then the Tucker et al. (2001) study should have found improved performance at test in both the wake and sleep group, as both groups had functioning hippocampi. Thus, hippocampal context reinstatement does not appear to explain MST improvement, and therefore, in turn, is not likely to explain the difference between patients and controls in the current study. Taking the results from the two studies together, it seems likely that hippocampal damage impairs sleep-dependent consolidation of the MST.

Due to constraints on patient availability, we were limited to a 24-hour design for this study, where the time between training and test included long periods of both sleep

12

and wake. This design does not permit direct claims about the sleep-dependence of the observed effects, separate from effects of elapsed time. However, prior studies have shown that change in MST performance is similar across 12 hours of sleep and across 24 hours including both wake and sleep (Manoach et al., 2004; Manoach et al., 2010; Walker et al., 2002). We therefore expect that the behavior we observed would be similar in a 12 hour sleep design and that the effects we observed are very likely to be sleep-dependent.

Older adults do not show the same overall typing speed nor the same sleep-dependent benefit in initial improvement as younger adults on the MST (S. M. Fogel et al., 2014; Gudberg, Wulff, & Johansen-Berg, 2015; King, Fogel, Albouy, & Doyon, 2013; Tucker et al., 2011). Younger adults tend to show more robust consolidation and a boost in performance in the first three trials of the test phase after sleep (Walker et al., 2002; Walker et al., 2003). We followed the same protocol as the study described above that found no improvement in initial test performance in older adults but an improvement in plateau performance (Tucker et al., 2011), and we replicated those findings here in our control participants. Other studies have looked only at initial performance and found results consistent with ours (Gudberg et al., 2015; King, Saucier, et al., 2017). One additional study found no benefit of a nap for the MST in older adults (S. M. Fogel et al., 2014), which may have been related to poor performance in the training phase (King, Saucier, et al., 2017). An intriguing possibility is that the difference between younger and older adults on this task may be functionally similar to the difference we observed between patients and controls. There is a known reduction in hippocampal function with healthy aging (Golomb et al., 1993) as well as a reduction in sleep spindles with age (Nicolas, Petit, Rompre, & Montplaisir, 2001), and the reduction in sleep spindles in older adults has been related to MST consolidation (S. M. Fogel et al., 2014). Thus, young adults, older adults, and patients with amnesia may fall on a spectrum of decreasing contribution of the hippocampus and spindles to consolidation.

There is debate in the literature as to whether the benefit of sleep to the MST for young adults is to boost performance or simply to stabilize it (Nettersheim, Hallschmid, Born, & Diekelmann, 2015; Pan & Rickard, 2015), alternatives which tend to support either an active role for sleep or a passive period of rest and reduced interference. Though the present findings do not directly speak to this debate or hinge on it, we believe they are easier to explain from the perspective of an active role for sleep: If a brain region is not critical for initial performance of a task but becomes critical offline, it seems likely that the region is playing an active role during that offline period.

A recent study assessed MST performance in patients experiencing transient global amnesia, a form of hippocampal amnesia lasting less than 24 hours (Dohring et al., 2017). During training, these patients typed fewer sequences overall than controls but exhibited similar percent improvement across training. Patients were tested on the same sequence again two days later, when they were no longer experiencing amnestic symptoms, and they performed in the same range as control subjects who were learning a new sequence. Improvement across days was larger in patients than controls, which the authors interpreted as evidence that consolidation boosted performance more in patients than controls. However, the findings can also be interpreted simply as better overall performance outside of an acute amnestic state. We therefore do not view these findings as inconsistent with ours. Another recent study assessed learning of the MST in patients

with hippocampal dysfunction associated with medial temporal lobe epilepsy (Long, Feng, Liao, Zhou, & Urbin, 2018). During training, patients initially performed at the same level as controls, but their percent improvement across training was reliably lower than controls. This is a counterintuitive finding, as the disruption to hippocampal function in these patients is likely much less than that in our patients, who have substantial hippocampal lesions. One possibility is that hippocampal dysfunction in epilepsy patients may disrupt hippocampal-striatal interactions during learning (Albouy, King, et al., 2013), whereas larger hippocampal lesions may leave the striatum to act more functionally and independently. This is an interesting possibility to explore in future work. For now, our sample of patients demonstrates that it is possible to exhibit normal MST learning despite extensive hippocampal damage.

While we have demonstrated that the hippocampus can be involved in the consolidation of a task that does not require the hippocampus for initial learning, we are not claiming that the hippocampus is necessarily involved in the offline processing of all such tasks. It is possible that the hippocampus is involved in consolidation of the MST specifically because it is attuned to sequential information, for example (Fortin, Agster, & Eichenbaum, 2002). The classic literature looking at retention of nondeclarative memory across days in amnesics may seem to support this idea, as the general claim has been that amnesics show impressive retention of hippocampally independent memory across days (Cohen & Squire, 1980; Corkin, 1968; Gabrieli, Corkin, Mickel, & Growdon, 1993; Nissen et al., 1989; Warrington & Weiskrantz, 1968). However, i) the evidence is mixed or absent as to whether the forms of memory tested in these studies have consolidation components that are sleep dependent, ii) initial task learning is typically not comparable between amnesics and controls in these studies (or there are no controls), making it difficult to assess whether there are deficits in offline changes, and iii) there often do appear to be deficits in performance change from the end of one day of training to the beginning of the next for amnesics in these studies. We thus view this literature as consistent with the possibility that the hippocampus plays a general enough role in encoding experience and its context that it will be important for a broad range of tasks that undergo sleep-dependent consolidation. A recent study in rats lends direct support to this claim, showing that the hippocampus is critical for the sleep-dependent consolidation of a hippocampus-independent novel object recognition task (Sawangjit et al., 2018). More work is still needed to determine the precise scope and nature of hippocampal involvement, but these findings open the door to new possibilities for the offline functions of the hippocampus by providing a proof of concept that the hippocampus can be critical for offline consolidation a task even when it is not necessary during initial learning.

References

Ackermann, S., Hartmann, F., Papassotiropoulos, A., de Quervain, D. J., & Rasch, B. (2015). No Associations between Interindividual Differences in Sleep Parameters and Episodic Memory Consolidation. *Sleep, 38*(6), 951-959. doi: 10.5665/sleep.4748

Albouy, G., Fogel, S., King, B. R., Laventure, S., Benali, H., Karni, A., . . . Doyon, J. (2015). Maintaining vs. enhancing motor sequence memories: respective roles of

striatal and hippocampal systems. *Neuroimage, 108*, 423-434. doi: 10.1016/j.neuroimage.2014.12.049

Albouy, G., Fogel, S., Pottiez, H., Nguyen, V. A., Ray, L., Lungu, O., . . . Doyon, J. (2013). Daytime sleep enhances consolidation of the spatial but not motoric representation of motor sequence memory. *PLoS One, 8*(1), e52805. doi: 10.1371/journal.pone.0052805

Albouy, G., King, B. R., Maquet, P., & Doyon, J. (2013). Hippocampus and striatum: dynamics and interaction during acquisition and sleep-related motor sequence memory consolidation. *Hippocampus, 23*(11), 985-1004. doi: 10.1002/hipo.22183

Albouy, G., Sterpenich, V., Balteau, E., Vandewalle, G., Desseilles, M., Dang-Vu, T., . . . Maquet, P. (2008). Both the hippocampus and striatum are involved in consolidation of motor sequence memory. *Neuron, 58*(2), 261-272. doi: 10.1016/j.neuron.2008.02.008

Albouy, G., Sterpenich, V., Vandewalle, G., Darsaud, A., Gais, S., Rauchs, G., . . . Maquet, P. (2013). Interaction between hippocampal and striatal systems predicts subsequent consolidation of motor sequence memory. *PLoS One, 8*(3), e59490. doi: 10.1371/journal.pone.0059490

Antony, J. W., & Paller, K. A. (2017). Hippocampal Contributions to Declarative Memory Consolidation During Sleep. In D. E. Hannula & M. C. Duff (Eds.), *The Hippocampus from Cells to Systems*: Springer International Publishing.

Barakat, M., Carrier, J., Debas, K., Lungu, O., Fogel, S., Vandewalle, G., . . . Doyon, J. (2013). Sleep spindles predict neural and behavioral changes in motor sequence consolidation. *Hum Brain Mapp, 34*(11), 2918-2928. doi: 10.1002/hbm.22116

Barakat, M., Doyon, J., Debas, K., Vandewalle, G., Morin, A., Poirier, G., . . . Carrier, J. (2011). Fast and slow spindle involvement in the consolidation of a new motor sequence. *Behav Brain Res, 217*(1), 117-121. doi: 10.1016/j.bbr.2010.10.019

Boutin, A., Pinsard, B., Bore, A., Carrier, J., Fogel, S. M., & Doyon, J. (2018). Transient synchronization of hippocampo-striato-thalamo-cortical networks during sleep spindle oscillations induces motor memory consolidation. *Neuroimage, 169*, 419-430. doi: 10.1016/j.neuroimage.2017.12.066

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis, 10*(4), 433-436.

Cairney, S. A., Guttesen, A. A. V., El Marj, N., & Staresina, B. P. (2018). Memory consolidation is linked to spindle-mediated information processing during sleep. *Curr Biol, 28*(6), 948-954 e944. doi: 10.1016/j.cub.2018.01.087

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that. *Science, 210*(4466), 207-210.

Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia, 6*, 255-265.

Covington, N. V., Brown-Schmidt, S., & Duff, M. C. (2018). The necessity of the hippocampus for statistical learning. *J Cogn Neurosci, 30*(5), 680-697. doi: 10.1162/jocn_a_01228

Curran, T. (1997). Higher-order associative learning in amnesia: evidence from the serial reaction time task. *J Cogn Neurosci, 9*(4), 522-533. doi: 10.1162/jocn.1997.9.4.522

Dohring, J., Stoldt, A., Witt, K., Schonfeld, R., Deuschl, G., Born, J., & Bartsch, T. (2017). Motor skill learning and offline-changes in TGA patients with acute hippocampal CA1 lesions. *Cortex, 89*, 156-168. doi: 10.1016/j.cortex.2016.10.009

Fernandez-Seara, M. A., Aznarez-Sanado, M., Mengual, E., Loayza, F. R., & Pastor, M. A. (2009). Continuous performance of a novel motor sequence leads to highly correlated striatal and hippocampal perfusion increases. *Neuroimage, 47*(4), 1797-1808. doi: 10.1016/j.neuroimage.2009.05.061

Fogel, S., Albouy, G., King, B. R., Lungu, O., Vien, C., Bore, A., . . . Doyon, J. (2017). Reactivation or transformation? Motor memory consolidation associated with cerebral activation time-locked to sleep spindles. *PLoS One, 12*(4), e0174755. doi: 10.1371/journal.pone.0174755

Fogel, S., Vien, C., Karni, A., Benali, H., Carrier, J., & Doyon, J. (2017). Sleep spindles: a physiological marker of age-related changes in gray matter in brain regions supporting motor skill memory consolidation. *Neurobiol Aging, 49*, 154-164. doi: 10.1016/j.neurobiolaging.2016.10.009

Fogel, S. M., Albouy, G., Vien, C., Popovicci, R., King, B. R., Hoge, R., . . . Doyon, J. (2014). fMRI and sleep correlates of the age-related impairment in motor memory consolidation. *Hum Brain Mapp, 35*(8), 3625-3645. doi: 10.1002/hbm.22426

Fortin, N. J., Agster, K. L., & Eichenbaum, H. B. (2002). Critical role of the hippocampus in memory for sequences of events. *Nat Neurosci, 5*(5), 458-462.

Fuentemilla, L., Miro, J., Ripolles, P., Vila-Ballo, A., Juncadella, M., Castaner, S., . . . Rodriguez-Fornells, A. (2013). Hippocampus-dependent strengthening of targeted memories via reactivation during sleep in humans. *Curr Biol, 23*(18), 1769-1775. doi: 10.1016/j.cub.2013.07.006

Gabrieli, J. D., Corkin, S., Mickel, S. F., & Growdon, J. H. (1993). Intact acquisition and long-term retention of mirror-tracing skill in Alzheimer's disease and in global amnesia. *Behav Neurosci, 107*(6), 899-910.

Genzel, L., Dresler, M., Cornu, M., Jager, E., Konrad, B., Adamczyk, M., . . . Goya-Maldonado, R. (2015). Medial prefrontal-hippocampal connectivity and motor memory consolidation in depression and schizophrenia. *Biol Psychiatry, 77*(2), 177-186. doi: 10.1016/j.biopsych.2014.06.004

Gheysen, F., Van Opstal, F., Roggeman, C., Van Waelvelde, H., & Fias, W. (2010). Hippocampal contribution to early and later stages of implicit motor sequence learning. *Exp Brain Res, 202*(4), 795-807. doi: 10.1007/s00221-010-2186-6

Golomb, J., de Leon, M. J., Kluger, A., George, A. E., Tarshish, C., & Ferris, S. H. (1993). Hippocampal atrophy in normal aging. An association with recent memory impairment. *Arch Neurol, 50*(9), 967-973.

Gudberg, C., Wulff, K., & Johansen-Berg, H. (2015). Sleep-dependent motor memory consolidation in older adults depends on task demands. *Neurobiol Aging, 36*(3), 1409-1416. doi: 10.1016/j.neurobiolaging.2014.12.014

Harrison, L. M., Duggins, A., & Friston, K. J. (2006). Encoding uncertainty in the hippocampus. *Neural Netw, 19*(5), 535-546. doi: 10.1016/j.neunet.2005.11.002

Ji, D., & Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nat Neurosci, 10*(1), 100-107. doi: 10.1038/nn1825

Kan, I. P., Giovanello, K. S., Schnyer, D. M., Makris, N., & Verfaellie, M. (2007). Role of the medial temporal lobes in relational memory: neuropsychological evidence

from a cued recognition paradigm. *Neuropsychologia, 45*(11), 2589-2597. doi: 10.1016/j.neuropsychologia.2007.03.006

Karni, A., Meyer, G., Jezzard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature, 377*(6545), 155-158. doi: 10.1038/377155a0

King, B. R., Fogel, S. M., Albouy, G., & Doyon, J. (2013). Neural correlates of the age-related changes in motor sequence learning and motor adaptation in older adults. *Front Hum Neurosci, 7*, 142. doi: 10.3389/fnhum.2013.00142

King, B. R., Hoedlmoser, K., Hirschauer, F., Dolfen, N., & Albouy, G. (2017). Sleeping on the motor engram: The multifaceted nature of sleep-related motor memory consolidation. *Neurosci Biobehav Rev, 80*, 1-22. doi: 10.1016/j.neubiorev.2017.04.026

King, B. R., Saucier, P., Albouy, G., Fogel, S. M., Rumpf, J. J., Klann, J., . . . Doyon, J. (2017). Cerebral activation during initial motor learning forecasts subsequent sleep-facilitated memory consolidation in older adults. *Cereb Cortex, 27*(2), 1588-1601. doi: 10.1093/cercor/bhv347

Kuriyama, K., Stickgold, R., & Walker, M. P. (2004). Sleep-dependent learning and motor-skill complexity. *Learn Mem, 11*(6), 705-713. doi: 10.1101/lm.76304

Laventure, S., Fogel, S., Lungu, O., Albouy, G., Sevigny-Dupont, P., Vien, C., . . . Doyon, J. (2016). NREM2 and sleep spindles are instrumental to the consolidation of motor sequence memories. *PLoS Biol, 14*(3), e1002429. doi: 10.1371/journal.pbio.1002429

Laventure, S., Pinsard, B., Lungu, O., Carrier, J., Fogel, S., Benali, H., . . . Doyon, J. (2018). Beyond spindles: interactions between sleep spindles and boundary frequencies during cued reactivation of motor memory representations. *Sleep, 41*(9). doi: 10.1093/sleep/zsy142

Long, J., Feng, Y., Liao, H., Zhou, Q., & Urbin, M. A. (2018). Motor sequence learning is associated with hippocampal subfield volume in humans with medial temporal lobe epilepsy. *Frontiers in Human Neuroscience, 12*(367).

Manoach, D. S., Cain, M. S., Vangel, M. G., Khurana, A., Goff, D. C., & Stickgold, R. (2004). A failure of sleep-dependent procedural learning in chronic, medicated schizophrenia. *Biol Psychiatry, 56*(12), 951-956. doi: 10.1016/j.biopsych.2004.09.012

Manoach, D. S., Thakkar, K. N., Stroynowski, E., Ely, A., McKinley, S. K., Wamsley, E., . . . Stickgold, R. (2010). Reduced overnight consolidation of procedural learning in chronic medicated schizophrenia is related to specific sleep stages. *J Psychiatr Res, 44*(2), 112-120. doi: 10.1016/j.jpsychires.2009.06.011

Moroni, F., Nobili, L., Curcio, G., De Carli, F., Tempesta, D., Marzano, C., . . . Ferrara, M. (2008). Procedural learning and sleep hippocampal low frequencies in humans. *Neuroimage, 42*(2), 911-918. doi: 10.1016/j.neuroimage.2008.05.027

Nadasdy, Z., Hirase, H., Czurko, A., Csicsvari, J., & Buzsaki, G. (1999). Replay and time compression of recurring spike sequences in the hippocampus. *J Neurosci, 19*(21), 9497-9507.

Nettersheim, A., Hallschmid, M., Born, J., & Diekelmann, S. (2015). The role of sleep in motor sequence consolidation: stabilization rather than enhancement. *J Neurosci, 35*(17), 6696-6702. doi: 10.1523/JNEUROSCI.1236-14.2015

Nicolas, A., Petit, D., Rompre, S., & Montplaisir, J. (2001). Sleep spindle characteristics in healthy subjects of different age groups. *Clin Neurophysiol, 112*(3), 521-527.

Nishida, M., & Walker, M. P. (2007). Daytime naps, motor memory consolidation and regionally specific sleep spindles. *PLoS One, 2*(4), e341. doi: 10.1371/journal.pone.0000341

Nissen, M. J., Willingham, D., & Hartman, M. (1989). Explicit and implicit remembering: when is learning preserved in amnesia? *Neuropsychologia, 27*(3), 341-352.

Palombo, D. J., Di Lascio, J. M., Howard, M. W., & Verfaellie, M. (2018). Medial temporal lobe amnesia is associated with a deficit in recovering temporal context. *J Cogn Neurosci*, 1-13. doi: 10.1162/jocn_a_01344

Pan, S. C., & Rickard, T. C. (2015). Sleep and motor learning: Is there room for consolidation? *Psychol Bull, 141*(4), 812-834. doi: 10.1037/bul0000009

Petersen, R. C., Parisi, J. E., Dickson, D. W., Johnson, K. A., Knopman, D. S., Boeve, B. F., . . . Kokmen, E. (2006). Neuropathologic features of amnestic mild cognitive impairment. *Arch Neurol, 63*(5), 665-672. doi: 10.1001/archneur.63.5.665

Peyrache, A., Khamassi, M., Benchenane, K., Wiener, S. I., & Battaglia, F. P. (2009). Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. *Nat Neurosci, 12*(7), 919-926. doi: 10.1038/nn.2337

Reber, P. J., & Squire, L. R. (1994). Parallel brain systems for learning with and without awareness. *Learn Mem, 1*(4), 217-229.

Rose, M., Haider, H., Salari, N., & Buchel, C. (2011). Functional dissociation of hippocampal mechanism during implicit learning based on the domain of associations. *J Neurosci, 31*(39), 13739-13745. doi: 10.1523/JNEUROSCI.3020-11.2011

Saletin, J. M., & Walker, M. P. (2012). Nocturnal mnemonics: sleep and hippocampal memory processing. *Front Neurol, 3*, 59. doi: 10.3389/fneur.2012.00059

Sawangjit, A., Oyanedel, C. N., Niethard, N., Salazar, C., Born, J., & Inostroza, M. (2018). The hippocampus is crucial for forming non-hippocampal long-term memory during sleep. *Nature, 564*(7734), 109-113. doi: 10.1038/s41586-018-0716-8

Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The necessity of the medial temporal lobe for statistical learning. *J Cogn Neurosci, 26*(8), 1736-1747. doi: 10.1162/jocn_a_00578

Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Curr Biol, 22*(17), 1622-1627. doi: 10.1016/j.cub.2012.06.056

Schapiro, A. C., Turk-Browne, N. B., Norman, K. A., & Botvinick, M. M. (2016). Statistical learning of temporal community structure in the hippocampus. *Hippocampus, 26*(1), 3-8. doi: 10.1002/hipo.22523

Schendan, H. E., Searl, M. M., Melrose, R. J., & Stern, C. E. (2003). An FMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron, 37*(6), 1013-1025. doi: S0896627303001235 [pii]

Siapas, A. G., & Wilson, M. A. (1998). Coordinated interactions between hippocampal ripples and cortical spindles during slow-wave sleep. *Neuron, 21*(5), 1123-1128.

Sirota, A., Csicsvari, J., Buhl, D., & Buzsaki, G. (2003). Communication between neocortex and hippocampus during sleep in rodents. *Proc Natl Acad Sci U S A, 100*(4), 2065-2069. doi: 10.1073/pnas.0437938100

Staresina, B. P., Bergmann, T. O., Bonnefond, M., van der Meij, R., Jensen, O., Deuker, L., . . . Fell, J. (2015). Hierarchical nesting of slow oscillations, spindles and ripples in the human hippocampus during sleep. *Nat Neurosci, 18*(11), 1679-1686. doi: 10.1038/nn.4119

Tucker, M., McKinley, S., & Stickgold, R. (2011). Sleep optimizes motor skill in older adults. *J Am Geriatr Soc, 59*(4), 603-609. doi: 10.1111/j.1532-5415.2011.03324.x

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

Walker, M. P., Brakefield, T., Seidman, J., Morgan, A., Hobson, J. A., & Stickgold, R. (2003). Sleep and the time course of motor skill learning. *Learn Mem, 10*(4), 275-284. doi: 10.1101/lm.58503

Walker, M. P., Stickgold, R., Alsop, D., Gaab, N., & Schlaug, G. (2005). Sleep-dependent motor memory plasticity in the human brain. *Neuroscience, 133*(4), 911-917. doi: 10.1016/j.neuroscience.2005.04.007

Warrington, E. K., & Weiskrantz, L. (1968). New method of testing long-term retention with special reference to amnesic patients. *Nature, 217*(5132), 972-974.

Westerberg, C. E., Mander, B. A., Florczak, S. M., Weintraub, S., Mesulam, M. M., Zee, P. C., & Paller, K. A. (2012). Concurrent impairments in sleep and memory in amnestic mild cognitive impairment. *J Int Neuropsychol Soc, 18*(3), 490-500. doi: 10.1017/S135561771200001X

Supporting Information for:

The hippocampus is necessary for the consolidation of a task that does not require the hippocampus for initial learning

Anna C. Schapiro, Allison G. Reid, Alexandra Morgan, Dara S. Manoach, Mieke Verfaellie, and Robert Stickgold

Analysis of excluded participants. Four out of eight patients were excluded from the main analyses for failing to meet the inclusion threshold of 10 correctly typed sequences on average across the last three trials of training. Figure S1 shows the learning time courses for the excluded patients, with the included patients for comparison. Three of these patients performed far below threshold (3.7, 3.2, and 3.1 sequences; yellow, purple, and red curves in Fig. S1) whereas one patient was close to the threshold (8.6). We completed the protocol with this latter patient, testing them with a new sequence on a second set of two-day sessions, where performance was again near but below threshold (9.3). Across the two sequences (light blue curve in Fig. S1), this patient had a 29.4% decline in performance for initial change (last three training trials to first three test trials) and an 8.6% improvement for plateau change (last three training trials to last six test trials). These results indicate that this patient showed the same pattern as the included patients: a decrement in initial performance at test without much evidence for improvement by the end of test.

We performed an analysis adding this patient as well as the two excluded control subjects (who also performed near criterion, with 8.3 and 9.1 typed sequences) to the participants who passed the inclusion threshold. For initial change, patients exhibited a significant decline in performance ($t[4]=3.32$, $p=0.029$, $d=1.48$), and they showed no difference between the end of training and the test plateau ($t[4]=0.05$, $p=0.964$, $d=0.02$). The difference between patients and controls was marginally significant for initial change ($t[15]=1.83$, $p=0.087$, $d=0.98$) and not reliably different for plateau change ($t[15]=1.43$, $p=0.173$, $d=0.76$). Thus, relaxing the inclusion threshold yielded results very similar to those reported in the main text.

The three patients who performed far below the inclusion threshold were performing virtually at floor during training, making it difficult to test for any decrement in performance at test. They showed no evidence of improvement across training: From the first three trials to the last three trials of training, they typed on average 0.2 *fewer* sequences. For the data collected in these patients (one set of two-day sessions), they exhibited a 13.9% improvement on average in initial change and a 23.8% improvement for plateau change. By the last three trials of test, they were typing an average of only 4 correct sequences per trial. While this is a different pattern than that seen in the other patients, we do not view these results as informative, as having some evidence of memory formed during training is a prerequisite for meaningfully asking how that memory changes at test.
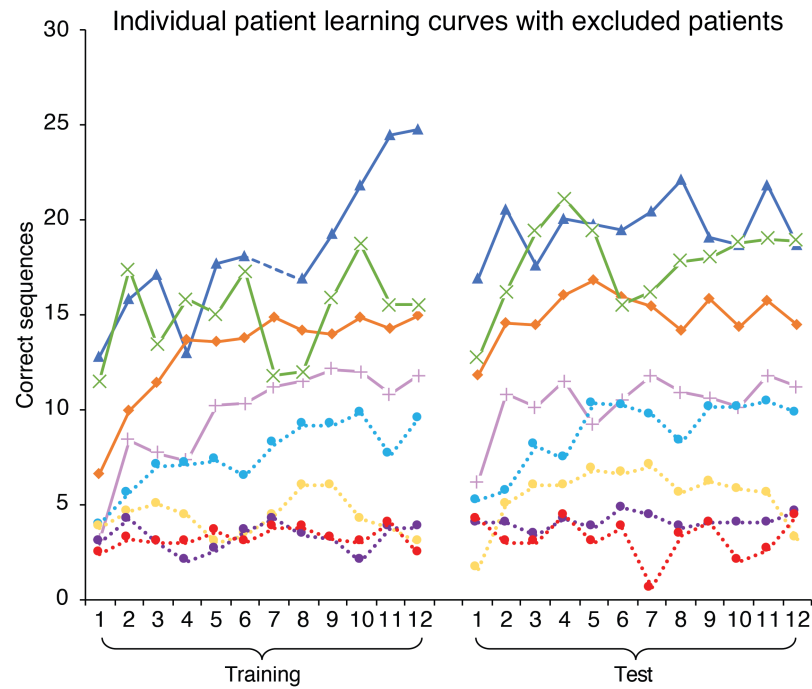
Figure S1. Correct sequences across trials of training and test for the four excluded patients, shown with dotted lines, as well as the four included patients (as in Fig. 2b). Curves are averaged across two sequences for the patients that completed two sequences (the top five performing patients).