

Hippocampal Contributions to Model-Based Planning and Spatial Memory

Highlights

- We tested planning and spatial memory in patients with hippocampal damage and controls
- Patients relied less on both model-based planning and allocentric spatial memory
- The planning impairment was related to the amount of damage to right hippocampus
- Planning and place memory covaried in controls but were less related in patients

Authors

Oliver M. Vikbladh, Michael R. Meager, John King, ..., Daphna Shohamy, Neil Burgess, Nathaniel D. Daw

Correspondence

omv208@nyu.edu (O.M.V.),
ndaw@princeton.edu (N.D.D.)

In Brief

Testing patients with hippocampal damage, Vikbladh et al. demonstrate that model-based planning and place memory rely on a common hippocampal substrate. The study bridges the reinforcement learning and spatial memory literatures to clarify the scope of hippocampal contributions to behavior.

Hippocampal Contributions to Model-Based Planning and Spatial Memory

Oliver M. Vikbladh,^{1,14,15,*} Michael R. Meager,^{2,3} John King,⁴ Karen Blackmon,⁵ Orrin Devinsky,^{3,6,7} Daphna Shohamy,^{8,9} Neil Burgess,^{10,11} and Nathaniel D. Daw^{12,13,*}

¹Center for Neural Science, New York University School of Arts and Science, New York, NY 10003, USA

²Department of Psychology, New York University School of Arts and Science, New York, NY 10003, USA

³Department of Neurology, New York University School of Medicine, New York, NY 10016, USA

⁴Division of Psychology & Language Sciences, Department of Clinical, Educational & Health Psychology, University College London, London WC1H 0AP, UK

⁵Department of Physiology, Neuroscience, and Behavioral Sciences, St. George's University School of Medicine, St. George, Grenada, West Indies

⁶Department of Neurosurgery, New York University School of Medicine, New York, NY 10016, USA

⁷Department of Psychiatry, New York University School of Medicine, New York, NY 10016, USA

⁸Department of Psychology, Columbia University, New York, NY 10027, USA

⁹Zuckerman Mind Brain Behavior Institute and Kavli Institute for Brain Science, Columbia University, New York, NY 10027, USA

¹⁰Institute of Cognitive Neuroscience, University College London, London WC1N 3AZ, UK

¹¹Institute of Neurology, University College London, London WC1N 3BG, UK

¹²Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08540, USA

¹³Department of Psychology, Princeton University, Princeton, NJ 08540, USA

¹⁴Twitter: @vikbladh

¹⁵Lead Contact

*Correspondence: omv208@nyu.edu (O.M.V.), ndaw@princeton.edu (N.D.D.)

<https://doi.org/10.1016/j.neuron.2019.02.014>

SUMMARY

Little is known about the neural mechanisms that allow humans and animals to plan actions using knowledge of task contingencies. Emerging theories hypothesize that it involves the same hippocampal mechanisms that support self-localization and memory for locations. Yet limited direct evidence supports the link between planning and the hippocampal place map. We addressed this by investigating model-based planning and place memory in healthy controls and epilepsy patients treated using unilateral anterior temporal lobectomy with hippocampal resection. Both functions were impaired in the patient group. Specifically, the planning impairment was related to right hippocampal lesion size, controlling for overall lesion size. Furthermore, although planning and boundary-driven place memory covaried in the control group, this relationship was attenuated in patients, consistent with both functions relying on the same structure in the healthy brain. These findings clarify both the neural mechanism of model-based planning and the scope of hippocampal contributions to behavior.

INTRODUCTION

Using knowledge of task contingencies, humans and other animals can plan novel courses of action, such as trajectories through a maze. Although the neural substrates for such

“model-based” planning are poorly understood, this ability is often viewed as similar to other functions supported by the hippocampus, like representing and remembering locations in space. Both model-based planning (Tolman, 1948) and place memory (O'Keefe and Nadel, 1978) are often described as requiring “cognitive maps” of the environment or task structure and are contrasted against habitual response-based behaviors that depend on the basal ganglia. Still, despite the commonalities, these functions are distinct in principle, and it is unclear whether they actually share a common neural mechanism and, if so, what that mechanism is.

Research into hippocampal spatial cognition most clearly emphasizes localization: determining one's position in allocentric space. This function is most famously exemplified by location-selective neural responses in the hippocampus (O'Keefe and Nadel, 1978) and behaviorally operationalized using spatial tasks, such as the Morris water maze (MWM) (Morris et al., 1982), where rodents have to find and remember the location of a hidden platform in an open arena. This type of “place memory,” keyed to allocentric configurations of cues like boundaries, is distinguished from more landmark-based strategies, such as egocentric stimulus-response strategies (e.g., turn left or right), which rely more on the basal ganglia (McDonald and White, 1994; Packard and McGaugh, 1996; Pearce et al., 1998). An analogous dissociation between hippocampal and basal-ganglia-dependent memory has been demonstrated in humans using fMRI in virtual spatial tasks (Hartley et al., 2003; Iaria et al., 2003; Voermans et al., 2004; Doeller et al., 2008).

By contrast, research into planning investigates how organisms use knowledge of task contingencies, like action outcomes and state transitions, to evaluate actions by mental simulation. Experiments probing such functions, including reward

devaluation in operant lever pressing (Adams and Dickinson, 1981; Adams, 1982) and multi-step reinforcement learning (Gläscher et al., 2010; Daw et al., 2011), support a distinction between two classes of strategies—referred to as goal-directed or model-based planning versus habitual or model-free learning (Balleine and Dickinson, 1998; Daw et al., 2005). This distinction seems to parallel the place versus response memory dichotomy from spatial cognition (Poldrack and Packard, 2003; Kosaki et al., 2018) and the related declarative versus procedural memory distinction from the memory literature (Squire, 1992; Knowlton et al., 1996; Foerde and Shohamy, 2011; Shohamy and Daw, 2015). Indeed, model-free learning, like landmark-based stimulus response strategies in some spatial tasks, is well captured by theories of dopamine and the basal ganglia (Schultz et al., 1997; Bayer and Glimcher, 2005).

It is less clear what neural mechanisms are responsible for model-based planning. There are, however, a number of suggestive reasons to suspect it shares a common hippocampal substrate with place memory (Hirsh, 1974; Dickinson and Balleine, 1993; Eilan et al., 1993; Johnson and Redish, 2007; Shohamy and Daw, 2015; Kumaran et al., 2016). Hippocampal function, of course, extends beyond spatial cognition to support declarative memory and, notably, a role in encoding the relationships among environmental stimuli (Eichenbaum and Cohen, 2004; Davachi and Wagner, 2002; Kumaran et al., 2009; Schapiro et al., 2016; Boorman et al., 2016; Garvert et al., 2017). Knowing such relations is critical to building a model of task contingencies. Tests of relational encoding have even relied on tasks that are similar in logic to probes for model-based planning, like transitive inference or acquired equivalence (Dusek and Eichenbaum, 1997; Heckers et al., 2004; Shohamy and Wagner, 2008; Wimmer and Shohamy, 2012). Moreover, the hippocampus has been implicated in the ability to imagine or simulate future events, a function that may be critical to model-based planning (Hassabis et al., 2007; Addis et al., 2011). Spatial navigation studies have further demonstrated that the hippocampus and surrounding medial temporal areas, in addition to the current location, encode other variables that are relevant to planning, such as boundaries or the identity, direction, and distance to a goal (Spiers and Maguire, 2007; Viard et al., 2011; Chadwick et al., 2015; Wikenheiser and Redish, 2015; Brown et al., 2016; Kaplan et al., 2017). Non-local place-cell firing, such as preplay of locations ahead of the animal, has also been proposed to support planning by mental simulation of candidate routes, drawing on a cognitive model or map of the world (Johnson and Redish, 2007; Pfeiffer and Foster, 2013; Shohamy and Daw, 2015; Mattar and Daw, 2018).

At the same time, there is a surprising lack of direct evidence for hippocampal involvement in model-based planning. For predominant rodent models of model-based behavior, including outcome devaluation and contingency degradation in operant lever pressing, hippocampal lesions have negligible effects (Corbit and Balleine, 2000; Corbit et al., 2002). One exception is a recent rodent study in which hippocampal lesions impaired model-based planning in a multi-step decision task (Miller et al., 2017). However, the extensive training needed to teach animals such sequential decision tasks may elicit model-free strategies that only mimic the signatures of planning in more

lightly trained humans (Akam et al., 2015; Economides et al., 2015). Even seemingly model-based rodent behavior (Miller et al., 2017) could thus rely on the hippocampus for different reasons. Finally, with few exceptions (Simon and Daw, 2011), little evidence links hippocampal activity in human neuroimaging or rodent place cell preplay to planning in tasks specifically designed to identify choice strategies that require knowledge of task contingencies.

We therefore sought to directly test the hypotheses that model-based planning uses a hippocampal mechanism in humans and whether this substrate is shared with boundary-driven place memory. To this end, we studied the performance on a model-based planning task (Daw et al., 2011) and a spatial memory task (Doeller et al., 2008) in healthy controls and patients with medically intractable epilepsy, treated by unilateral anterior temporal lobectomy (ATL) with hippocampal resection. We investigated whether damage to the temporal lobe impaired model-based planning and boundary-driven place memory and how it affected the relationship between them. If the hippocampus is a common neural substrate for both functions, we expected hippocampal damage to impair performance on both tasks. Furthermore, a common substrate could lead to correlated performance across the tasks, but this correlation should itself be attenuated if that substrate is impaired by hippocampal damage. Finally, because the lesions also affected overlying cortex, we explored to what extent performance in either task was related specifically to the extent of damage to hippocampus on either side, controlling for the overall extent of the lesion.

RESULTS

Participant Characteristics

We recruited 19 epilepsy patients, treated with unilateral ATL, i.e., surgical removal of the anterior temporal lobe on one side, and 19 healthy controls (see STAR Methods). Patients and controls displayed no significant group differences in IQ ($t_{36} = 0.2200$; $p = 0.8271$), age ($t_{36} = -0.7760$; $p = 0.4428$), or number of males versus females ($z = -0.3261$; $p = 0.7444$). For 10 out of 19 patients, ATL was right lateralized. There were no significant differences between the right- and left-lateralized ATL groups in IQ ($t = -0.5295$; $p = 0.6033$), age ($t_{17} = 1.0876$; $p = 0.2919$), or number of males versus females ($z = -0.6752$; $p = 0.4995$).

Individual lesion masks, derived from post-surgical structural brain scans normalized to the 1-mm Montreal Neurological Institute (MNI) template (Figure 1; STAR Methods) were used to estimate the standardized total lesion size of each patient. Total lesion size in the right-lateralized ATL patient group (mean = 65,314 voxels; SE = 2,791 voxels) was found to be significantly larger ($t_{17} = -2.6103$; $p = 0.0183$) than in the left-lateralized ATL patient group (mean = 40,536 voxel; SE = 2,296 voxels).

Each patient's lesion mask was also compared to the Harvard-Oxford Brain Lexicon ($p > 0.5$) in order to estimate what percentage of the hippocampus had been resected (see STAR Methods). Mean hippocampal lesion size estimated in the right-lateralized ATL patient group (mean = 2,791 voxels; SE = 321 voxels) was not significantly larger ($t_{17} = -1.1124$; $p = 0.2814$) than in the left-lateralized ATL patient group (mean = 2,296 voxels; SE = 302 voxels). The mean hippocampal

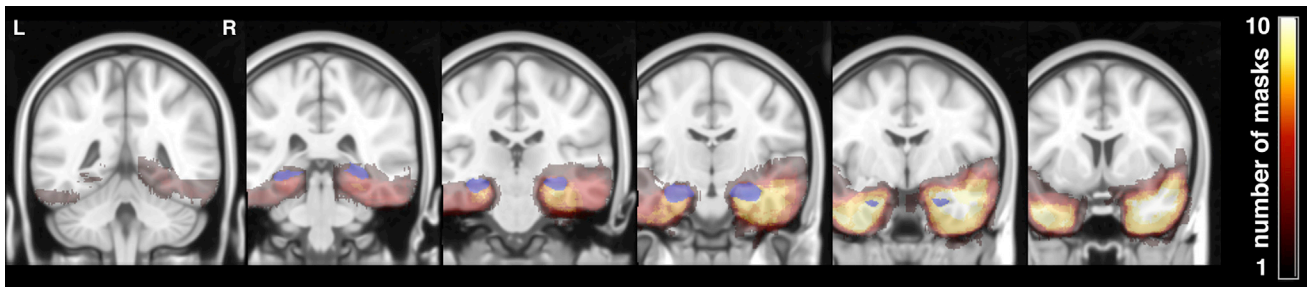


Figure 1. Patient Lesion Masks

Slices ($y = 82, 92, 102, 112, 122, \text{ and } 132$) showing all 19 hand-drawn patient ATL lesion masks normalized to the MNI template. Heatmaps indicate the number of masks overlapping at a given voxel. The hippocampus, as defined by Harvard-Oxford lexicon ($p > 0.5$), is shown in blue.

lesion sizes corresponded to 62.8% (SE = 7.2) and 53.7% (SE = 7.1) of the hippocampus resected for the right and left ATL groups, respectively.

Patients Display Shift from Model-Based to Model-free Strategy

Participants completed 200 trials of a two-step Markov decision task (Daw et al., 2011) designed to quantify the reliance on model-based and model-free strategies (see STAR Methods). The mean number of completed trials was 195.3 (SE = 2.1), with no difference between control and patient groups ($t_{38} = 1.2515$; $p = 0.2188$). The mean number of rewards received was 107.9 (SE = 2.6), also, with no significant difference between control and patient group ($t_{36} = -0.0399$; $p = 0.9684$). In general, rewards in this task are by design highly stochastic and not sensitive to differences in strategy.

On each trial, the participant first made a choice between two spaceships. One spaceship most commonly ($p = 0.7$) transitioned to the purple planet and otherwise made a rare transition ($p = 0.3$) to the red planet. For the other spaceship, probabilities were reversed. The participant then made a choice between two planet-specific aliens, each associated with a unique, slowly drifting probability of reward (Figure 2).

Figure 3 shows markers of both as a function of group, estimated from a factorial logistic regression (Table S1), which predicts choosing the same spaceship as on the previous trial. Model-free learning is signaled by a main effect of reward, i.e., a tendency to repeat choosing the spaceship that led to reward, whereas in model-based learning, choice of spaceship is mediated by expectations about the planets to which it leads, indicated by an interaction between reward and whether a rare or common transition occurred on the last trial. If, for instance, a reward is received but following a rare transition, a model-based agent should be less likely to repeat the choice of spaceship on the next trial. The difference in these effects measures the relative strength of model-based versus model-free choice. The regression also controls for additional nuisance explanatory factors, age and IQ.

The expression of model-based versus model-free strategies differed significantly by group, with controls showing a relatively even mixture of strategies (similar to previous reports using this task), but patients skewed away from model-based planning toward model-free learning. The relative reliance on model-

based over model-free strategies, calculated by taking the difference between these effects, differed significantly between groups, indicating a specific strategy change rather than a general impairment ($z = 2.028$; $p = 0.043$). This finding is consistent with our hypothesis that hippocampal damage in the ATL lesion group specifically affects the use of internal models or maps of task contingencies.

For simplicity, the above factorial analysis only considers the effect of the preceding trial's events on each trial's choice. To verify that our results were not dependent on this assumption, and in keeping with previous work (Daw et al., 2011), we repeated our analysis (Table S2) by fitting participants' choices with a full 6-parameter computational learning model (Daw et al., 2011; as modified by Gillan et al., 2016), which uses the full history of preceding rewards to predict each choice. The results recapitulate the findings from the regression: chiefly, a significant interaction of reinforcement learning strategy and experimental group such that patients are biased away from model-based and toward model-free strategies ($p = 0.036$). In addition, in this analysis (here going beyond the simpler regression analysis), the estimated strength of model-free learning is itself significantly higher in the patient group than the control group ($p = 0.043$). The remaining parameters of the computational model did not differ significantly between groups.

Patients Display Impaired Boundary-Driven Place Memory

Participants completed a spatial task (Doeller et al., 2008), where on each of 64 trials, they had to navigate in first person (Figure 4, left) to indicate from memory the correct location of one of four objects in a virtual arena (see STAR Methods). The mean number of completed trials was 61.3 (SE = 0.5) with no significant difference between control and patient group in number of completed trials ($t_{36} = -0.2882$; $p = 0.7749$).

For two of the objects, correct locations were defined in relation to distal boundary cues around the arena, and for the other two objects, correct locations were defined in relation to a landmark inside the arena (Figure 4, middle). Trials were presented in four blocks consisting of 16 trials each. Within the blocks, the landmark location was fixed with respect to the boundary cues, but between the blocks, the landmark moved (Figure 4, right). Participants were not instructed about the difference

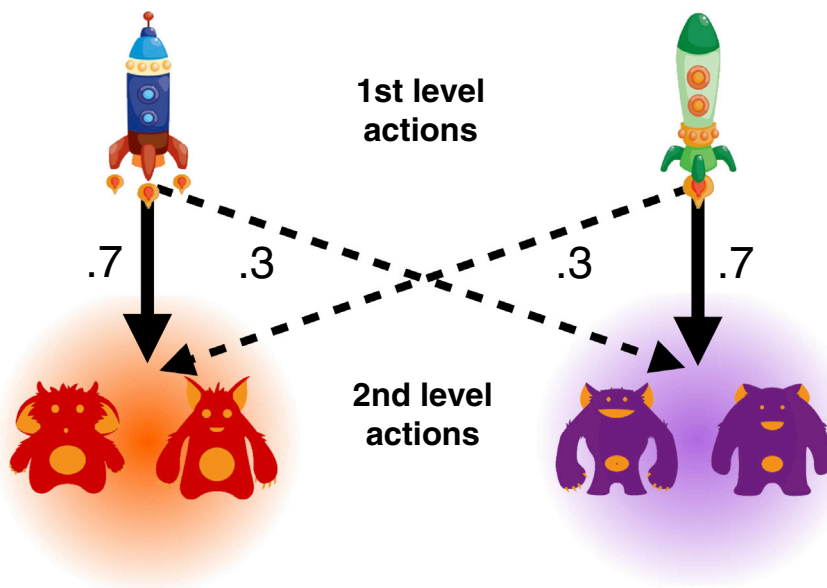


Figure 2. Two-Step Markov Decision Task

On each trial, the participants chose one of two first level actions (spaceships). One spaceship transitions the participant to red planet with $p = 0.7$, and the other spaceship transitions the participant to red planet with $p = 0.3$. Having transitioned to a second-level state, participants chose between two second-level actions (aliens) that were unique to each planet. Each alien was associated with a unique, slowly drifting probability of receiving reward.

distance errors (dB) and landmark distance errors (dL) between the chosen location and the correct locations as predicted by boundary cues and landmark cues, based on the previous block (Figure 4, right). dB and dL thus inversely reflect performance with respect to boundary and landmark cues. To assess group differences, we specified a regression model where the dependent variable, distance error (dB and dL for each trial) was regressed on the key explanatory variables

between landmark and boundary objects or about the block-wise landmark movements.

The movements of the landmarks with respect to the boundary cues serve to dissociate spatial memory performance based on either type of cue. Given previous results (Pearce et al., 1998; Doeller et al., 2008), we hypothesized that hippocampal damage would preferentially impair reliance on boundary cues. Following Doeller et al. (2008), for each object, we therefore focused our analyses on the first trial following each of the three landmark movements (average total 11.4 trials per participant, due to missed trials). This is because these trials cleanly dissociate performance based on recalling the object's location in the previous block relative to each type of cue (Figure 4, middle). Note that, because the boundary and landmark cues remain fixed with respect to each other within each block, performance on the remaining trials of each block are not as diagnostic of cue usage, because, to the extent behavior is based on recalling the object's most recent location with respect to either cue, this is equivalent for both cues (though see Figures S1 and S2 for analysis of remaining trials). Similarly, to avoid relying on the assumption that participants were able to learn to differentiate landmark from boundary objects (which is only possible following experience with at least one of the three landmark movements), we analyzed landmark and boundary error for all objects rather than differentiating by object type. We did, however, specify a regression model where we, for the critical trials following the landmark moves, additionally interacted group and distance error type by object type (Table S3); we did find a two-way interaction of object type by error type ($F_{1, 782.4} = 10.527$; $p < 0.001$), indicating that participants were ultimately able to treat the two object types differently. However, we found no significant three-way interaction of object type by error type by group ($F_{1, 782.1} = 2.6427$; $p = 0.1044$), indicating that we did not detect a group difference in this respect.

On the critical trials following movement of the landmark, we quantified reliance on either cue type by computing boundary

lesion group (control versus patient) and distance error type (dB or dL), while also controlling for additional nuisance explanatory factors, age and IQ.

We found a significant interaction of group by distance error type ($F_{1,97.58} = 5.5080$; $p = 0.021$), indicating a difference between groups in their relative reliance on the two cue types. This effect mainly reflected the finding that patient's dB was significantly higher, i.e., patients' performance was less driven by boundary cues ($F_{1,39.41} = 2.5102$; $p = 0.016$; Figure 5; Table S4).

In a follow-up analysis, aimed at simplifying the design for later elaboration by assessing relative reliance on boundary versus landmark cues using a single explanatory variable, we defined a relative measure of error: the ratio $dB/(dL+dB)$, which measures whether participants were relatively biased toward using boundary cues over landmark cues. Regressing it on group (controlling for nuisance variables age and IQ), we again found that patients were significantly biased toward relying on boundary cues ($F_{1,432} = 8.213$; $p = 0.004$; Table S5). All these results are consistent with our prediction that anterior temporal lobe structures like the hippocampus preferentially support boundary-over landmark-driven memory.

Relationship between Model-Based Planning and Boundary-Driven Place Memory

So far, we have shown impaired model-based planning and boundary-driven place memory in the patient group. Next, we examined the relationship between these two functions, first by investigating their baseline correlation in neurologically intact control participants. We did this by calculating the mean dB for each participant and using it as a covariate in the logistic regression model of our decision task. This approach is analogous to estimating participant-by-participant scores for model-based planning from the logistic model and then correlating

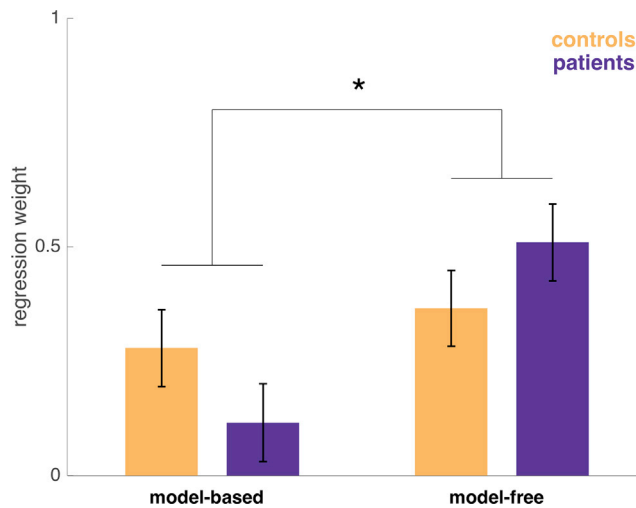


Figure 3. Model-Based and Model-free Regression Weights for Controls and Patients

Estimated with a logistic mixed-effects regression controlling for IQ and age. Error bars indicate SE. The interaction of strategy (model-free versus model-based) by group was significant ($z = 2.028$; $p = 0.043$).

those with dB in a second step but preferable because it takes account of statistical uncertainty about the participants' planning scores in computing their relationship to dB, which the naive correlation neglects. IQ was also included as a nuisance covariate to account for task-general variation.

Figure 6 displays the results of this regression (Table S6), broken down by group. Boundary-driven place memory significantly predicted a control participant's use of a model-based strategy ($z = 6.6455$; $p = 0.001$), consistent with the two measures sharing some underlying substrate.

Next, we engaged in a series of follow-up analyses to interrogate the specificity of this cross-task relationship. First, we wished to examine whether the relationship was specific to model-based and boundary-driven task strategies rather than general to both strategies tested in each task. We first verified that the increase in model-based planning associated with dB is significantly larger than any corresponding effect for model-free choice ($z = 2.137$; $p = 0.033$) or, in other words, that dB is associated with a relative increase in model-based relative to model-free choice. Next, we refined this analysis to also probe its specificity to boundary over landmark error. In a new regression model (Table S7), we replaced the boundary error dB with the relative error ratio $\text{dB}/(\text{dL}+\text{dB})$, which measures whether participants were relatively biased toward using boundary cues over landmark cues in the spatial task. In the control group, as with absolute boundary error, relative error was also significantly associated with relative increase in model-based, minus model-free, choice ($z = 2.069$; $p = 0.039$; Table S7). Thus, in healthy controls, there is a specific relationship between boundary-driven spatial memory and model-based choice, relative to their respective alternative strategies.

We next tested the specificity of the cross-task relationship to controls versus patients. We reasoned that, if the relationship in the control group depends on the intact hippocampus (e.g., if it

arose due to a common substrate located there), then over and above the effects on each task separately, their relationship would be affected by hippocampal and medial temporal lobe damage. Therefore, we tested the null hypothesis that the relationship between the tasks is unaffected by ATL damage, the rejection of which would support the alternative hypothesis that the ATL does affect their relationship. The relationship between boundary-driven memory and model-based planning was indeed significantly attenuated in the patient group ($z = 2.137$; $p = 0.032$). Reflecting this attenuation, the patient group, considered alone, did not display a significantly detectable relationship between the two functions ($z = 0.156$; $p = 0.875$). Critically, this null result does not imply that these functions are *unrelated* in the patient group.

Finally, we repeated this analysis using the full computational learning model in place of the simpler regression-based index of learning (Table S8). Again, while controlling for IQ, we observed a strong positive correlation between model-based planning and place memory in the control group ($p = 0.030$), but not the patient group ($p = 0.803$), although the group-wise interaction was merely trending in this version of the analysis ($p = 0.081$).

Deficits Are More Robust for Patients with Right-Lateralized ATL

Based on previous literature, we next sought to examine to what extent the reported effects might be preferentially associated with lesions lateralized to one side or the other. Breaking down the data this way requires examining small subgroups ($N = 9$ and 10), meaning that the key analyses comparing the two laterality groups against one another are underpowered relative to comparing either group to controls. Also, lesion laterality is correlated with overall lesion extent in our sample. Altogether, these analyses are fundamentally more exploratory and their results more tentative than those reported above.

With those caveats, we expected boundary-based memory, and spatial relations generally, to be more strongly associated with the right hippocampus (Burgess et al., 2002). For instance, boundary-memory-related hippocampal activity, in a previous fMRI study of the spatial memory task we used, was right lateralized (Doeller et al., 2008). It is less clear, *a priori*, how model-based planning might be lateralized.

Figure 7 shows decision task and spatial memory task performance with the patient data further subdivided by ATL laterality and the relationship between model-based planning and dB also broken down by laterality. In all three cases, the differences between patients and controls (Figure 7) appeared to be driven by the right lesion patients, with the left lesion patients more similar to controls. This impression is only partly borne out by statistics, however (Tables S9, S10, and S11). In particular, in all three cases, the right patients differ significantly from controls (model-based minus model-free: $z = 2.300$, $p = 0.022$; dB minus dL: $F_{1,92.22} = 4.464$, $p = 0.034$; across-task correlation between model-learning and boundary memory: $z = 2.550$, $p = 0.011$), whereas the left patient group did not differ from controls in any case (dB minus dL: $F_{1,100.7} = 2.644$, $p = 0.107$; model-based minus model-free: $z = 1.001$, $p = 0.317$; across-task correlation between model-learning and boundary memory: $z = 0.470$, $p = 0.639$). However, in no case were the lesion groups significantly different from one another (dB minus dL: $F_{1,97.8} = 0.150$,

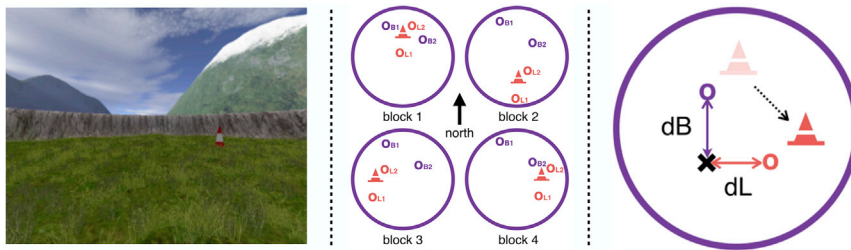


Figure 4. Spatial Memory Task

Left: virtual arena as seen from first-person perspective. Landmark (cone), arena boundary (wall), and distal cues (mountains) are visible.

Middle: spatial task block structure is shown, with virtual arena as seen as from above. Between blocks, the landmark (cone) moved in relation to the boundaries (large purple circle). Correct location of the two boundary objects (O_{B1} and O_{B2}) stayed constant with respect to boundaries across all blocks. Correct location of the two landmark objects (O_{L1} and O_{L2}) stayed constant with respect to the landmark across all blocks.

Right: measuring reliance on boundary and landmark cues is shown. The landmark (cone) moves (dotted line) in relation to boundaries (large purple circle) between blocks. In a trial preceding a landmark move, an example object's correct location (purple o), i.e., where the object appears during feedback, is in close proximity to the landmark (shaded cone). If participants remember this object location in relation to boundaries and distal cues, the predicted object location in the next block would also be indicated by the purple o. Conversely, if participants learned the object location in relation to the landmark, the predicted object location after the landmark moves to its new location (filled cone) would be the orange o. On the trials following movement of the landmark, we thus operationalize place memory by the dB between their response (cross) and the location predicted by boundaries and distal cues (purple o). Response memory is operationalized by the dL between their response (cross) and the location predicted by the landmark cue (orange o). Lower dB and dL thus means greater reliance on boundary and landmark cues, respectively.

$p = 0.699$; model-based minus model-free: $z = -1.149$, $p = 0.251$; across-task correlation between model-learning and boundary memory: $z = 1.552$, $p = 0.121$.

We also examined the breakdown, by lesion laterality, of the relationship between relative measures of planning and spatial memory to account for alternative strategies. In the regression model specified earlier that included the relative error ratio $dB/(dL+dB)$ as a covariate (Table S7), the association between relative preference for a model-based (minus model-free) strategy and the relative bias toward using boundary over landmark cues was larger in the left group than the right group ($z = 2.8082$; $p = 0.005$). An estimated effect in the same direction was also seen comparing the control with the right patient group, with the relationship being stronger in the control group, although it did not reach significance ($z = 1.567$; $p = 0.117$).

Thus, although noisy, there is a consistent suggestion across all three measures and different ways of examining the cross-task relationship that the results of this study were most robust in the right lesion group.

Lesion Size in the Right Hippocampus Predicts Model-Based Planning Deficits

One way to sharpen the foregoing analyses is to focus specifically on not just the side but the particular anatomical region hypothesized to underlie the effects: the hippocampus. Accordingly, we tested how performance on the tasks covaried with estimated lesion size on the right and left hippocampus, respectively. Hippocampal lesion size for each patient (Figure 1) was estimated by comparing the normalized anatomical masks to the Harvard-Oxford lexicon ($p > 0.5$; see STAR Methods). Importantly, the ATL procedure involves a pattern of damage to numerous temporal lobe structures in addition to hippocampus, which means one should be highly cautious interpreting these results with respect to any particular structure. Although it is not practical to control for damage to many different medial temporal lobe structures individually, we attempt to mitigate these concerns and focus on hippocampal lesion size by control-

ling for the overall lesion size as a nuisance effect. The regression analyses also controlled for age and IQ.

We found, as can be seen in Figure 8 (Table S12), that model-based planning was significantly worse for larger right hippocampal lesions ($z = 2.831$; $p = 0.005$). Conversely, planning was not significantly related to the amount of hippocampal damage on the left hippocampus ($z = 1.062$; $p = 0.288$), and this difference between right and left effects was itself significant ($z = 2.508$; $p = 0.012$). These results indicate that the amount of damage to the right, but not the left, hippocampus is related to model-based deficits. However, in order to further test specificity and ensure that the lesions in the right hippocampus are not simply causing general learning deficits, we also calculated the effect of each hemisphere's hippocampal lesion size on the difference between model-based and model-free learning, as estimated by the logistic regression. As predicted, we found that right hippocampal lesion size was significantly related to a shift away from model-based toward a model-free strategy ($z = 2.984$; $p = 0.003$) and that this effect was significantly larger for right compared to left hippocampal lesions ($z = -3.377$; $p < 0.001$).

For boundary memory, the effects of lesion size on performance were similar in magnitude and pattern, although not significant in either the right ($F_{1,12,16} = 2.6082$; $p = 0.1320$) or the left patient group ($F_{1,12,93} = 0.1204$; $p = 0.7340$; Table S13). It should be noted that this analysis is based on many fewer trials than the sequential decision task analysis.

DISCUSSION

Although extensive evidence indicates that the hippocampus supports localization in allocentric space, there is relatively little direct evidence for the hypothesis that the same mechanisms extend to model-based planning. We addressed this gap by testing model-based planning and place memory in patients with extensive hippocampal damage as a result of unilateral ATL lesions and matched, neurologically typical controls. Our results are consistent with the hypothesis that both

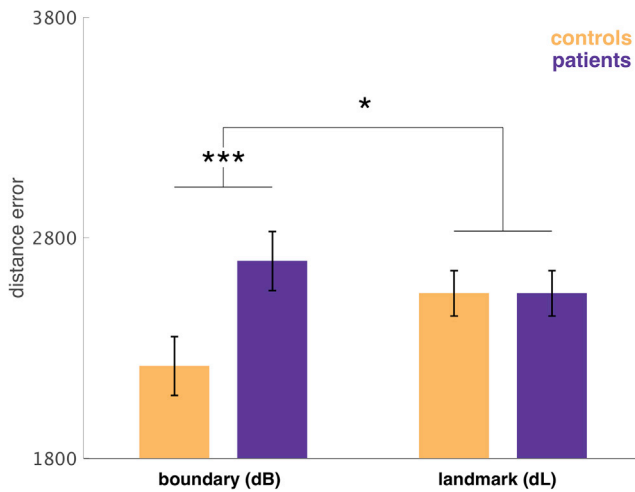


Figure 5. dB and dL (A.U.) for All Objects on Trials that Follow Movements of the Landmark

Estimated with a linear mixed-effects regression, controlling for IQ and age. Error bars indicate SE. There was a significant group difference in dB ($F_{1,39.41} = 2.510$; $p = 0.016$), but not dL ($F_{1,85.3} = 0$; $p = 0.9990$), with a significant interaction of group by cue type ($F_{1,97.58} = 5.508$; $p = 0.021$).

model-based planning and boundary-driven place memory share a common mechanism, which is affected in ATL patients and, more tentatively, associated with right hippocampus.

As predicted, ATL patients displayed significantly attenuated boundary-driven place memory in our MWM-like spatial memory task, alongside spared landmark-based memory. These results echo the dual systems view of navigation supported both by rodent lesion (O'Keefe and Nadel, 1978; McDonald and White, 1994; Packard and McGaugh, 1996; Pearce et al., 1998) and human neuroimaging experiments (Hartley et al., 2003; Iaria et al., 2003; Voermans et al., 2004; Doeller et al., 2008), whereby the hippocampus supports fast learning of allocentric spatial maps and the striatum facilitates slow, incremental associations between stimuli and responses. That said, one weakness of the current task in operationalizing the place versus response distinction is that, although use of boundary cues clearly exercises allocentric spatial localization, landmark usage imperfectly captures a striatal response system as classically envisioned (for instance, because some allocentric information is still needed to place objects correctly relative to the landmark). Nevertheless, the use of boundary cues to index hippocampal function (the measure most important to our results) is unambiguous and well validated (Pearce et al., 1998; Doeller et al., 2008), even if the landmark foil imperfectly captures a hypothetical striatal contribution. Also, our distance error measures assess the tendency of participants to use either sort of cue. It cannot distinguish deficits in boundary-driven place memory per se from performance deficits, such as reduced attention to these cues, or a greater belief that landmarks predict object locations. All these mechanisms, though, are consistent with the broader perspective that the hippocampus is ultimately involved in allocentric spatial localization based on configurations of cues.

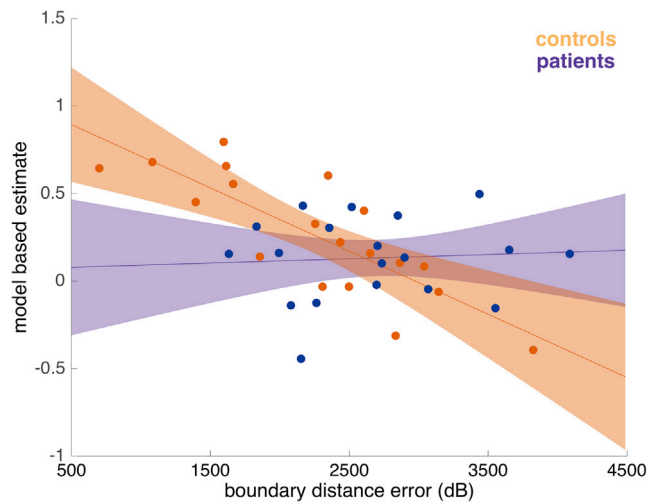


Figure 6. Relationship between Model-Based Planning and dB in Controls and Patients

Estimated with a logistic mixed-effects regression, controlling for IQ. Error bars indicate 80% confidence intervals. Individual place memory performance is reflected by mean dB from the spatial task. Dots indicate estimates for individual participants, calculated from the mixed-effects logistic regression. The trend was significant in the control group ($z = 6.6455$; $p = 0.001$), but not in the left patient group ($z = 0.156$; $p = 0.875$). The slope differed significantly between groups ($z = 2.137$; $p = 0.032$).

The patients were also significantly biased away from using model-based planning and toward model-free habitual strategies in the two-step Markov decision task. This result provides causal evidence for the inference that temporal lobe structures support model-based planning, over and above their role in place memory. The appearance of a compensatory shift toward improved model-free learning, which is rarely reported with this task (Frank et al., 2004), indicates that behavior in the ATL patients is not simply noisier and instead is consistent with models invoking multiple, potentially competing, reinforcement-learning systems in the human brain (Daw et al., 2005). Still, our results do not speak clearly to the perennial question whether the hippocampus plays a special role in such models for spatial versus more abstract relational tasks. This is because, although our planning task is structured like an abstract Markov decision process, its cover story, in terms of rocket trips to planets, might have elicited a spatial interpretation.

Our results also complement and extend previous research with rodents. Unit recording studies have shown results suggestive of hippocampal involvement in model-based planning, notably replay of forward trajectories in hippocampal place cells (Johnson and Redish, 2007; Pfeiffer and Foster, 2013). However, in contrast to our results, previous studies with place cell recordings have not yet shown behavioral evidence for a link between the hippocampus and the use of this knowledge in planning, nor do they provide evidence for a causal role of hippocampus in such a function. In these respects, our results more closely parallel a recent report of a related deficit in model-based learning in rodents during inactivation of the dorsal hippocampus, using an analogous two-stage Markov decision

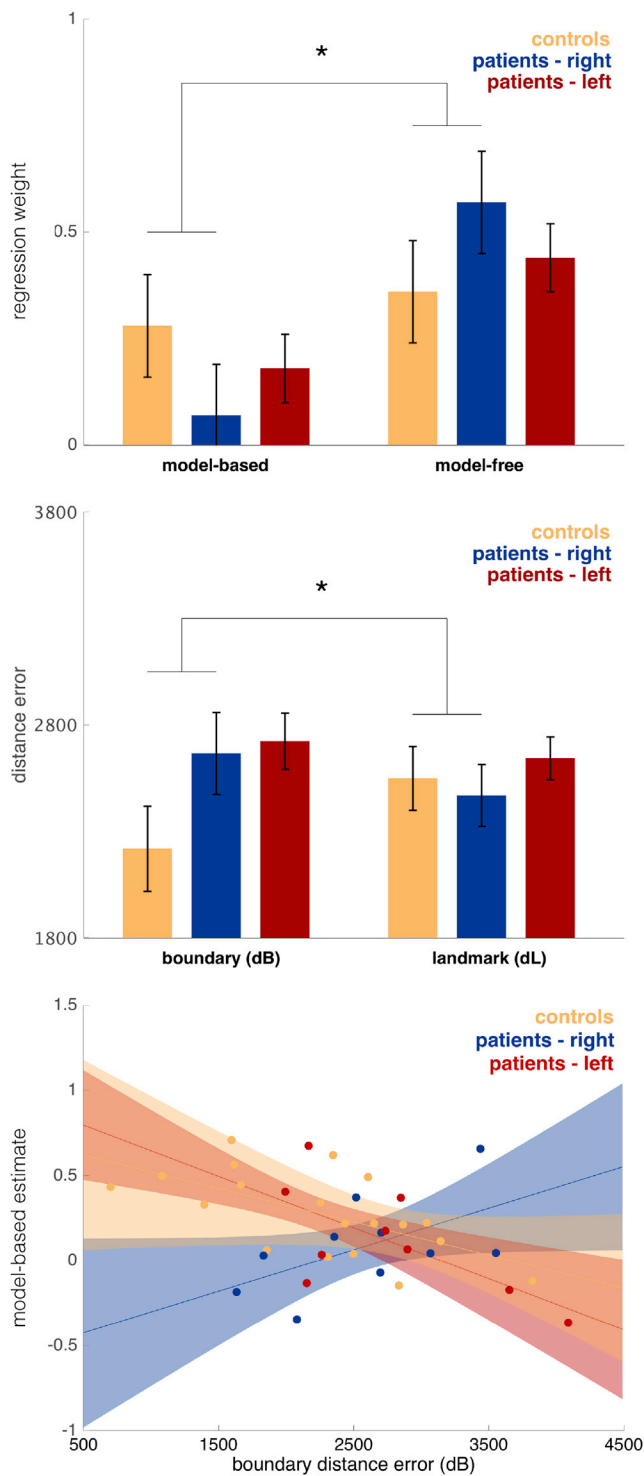


Figure 7. Effects for Both Tasks and Their Relationship, Broken Down by Lesion Laterality

Top: model-based and model-free regression weights for controls, right- and left-lateralized ATL patients. Estimated with a logistic mixed-effects regression, controlling for IQ and age. Error bars indicate SE. The difference in model free versus model based was significantly different between the control and right patient group ($z = 2.295$; $p = 0.022$).

task (Miller et al., 2017). The targeting of the inactivation to hippocampus in the rodent study sharpens the anatomical specificity of the effect. Conversely, our human result clarifies the contribution of the damaged structure, because we know more about the computations underlying model-based behavior on this task in humans. In particular, in humans, but not yet rodents, model-based choices have been explicitly linked to prospective neural activity at decision time (Doll et al., 2015). This helps to rule out other potentially confounding strategies, such as that the apparently model-based choices in rodents are instead produced by some learned-response-switching rule contingent on events spanning multiple trials. It has been suggested that such model-free strategies might arise following overtraining of the sort used to teach animals this task (Akam et al., 2015; Economides et al., 2015); this might also implicate hippocampus for other, confounding reasons, such as its involvement in trace conditioning and latent states (Solomon et al., 1986; Büchel et al., 1999; Gershman et al., 2010). Our findings of a similar result in humans, without extensive training, thus help corroborate the interpretation of the rodent study as well.

Comparing performance between our two tasks, we also found that, in healthy controls, boundary-driven place memory performance correlated with the extent of reliance on model-based planning strategies. Importantly, this relationship was also significantly attenuated in the patient group, a result that suggests the lesion affects some common substrate for the tasks that is otherwise provided by the temporal lobes in the healthy brain. Following damage to this structure, however, the two tasks may be at least partly supported by differential compensatory mechanisms, leading to their de-correlation. These findings are consistent with the hypothesis that both model-based planning and place memory share a common mechanism, which is impaired in ATL patients.

It is surprising and interesting that the effects we report emerge following damage to only one lobe, as unilateral temporal lobe damage is generally known to produce rather subtle effects on cognition in humans (Spiers et al., 2001) and animals (van Praag et al., 1998), relative to the famously dramatic effects of bilateral lesion (e.g., Scoville and Milner, 1957). This may relate to our use of two behavioral tasks that are well attuned to temporal lobe function. However, there exist inherent and important caveats in drawing conclusions about the neural bases of effects from a study of this sort. It is possible that the observed effects are caused, at least in part, by damage to the brain, including the hemisphere

Middle: boundary and landmark error dB and dL (A.U.) for controls and patients with right- and left-lateralized ATL are shown. Estimated with a linear mixed-effects regression, controlling for IQ and age. Error bars indicate SE. There was a significant difference between dB and dL when comparing the control and right patient group ($F_{1,92.22} = 4.463$; $p = 0.034$).

Bottom: relationship between model-based planning and dB (A.U.) for controls and patients with right- and left-lateralized ATL is shown. Estimated with a logistic mixed-effects regression, controlling for IQ. Error bars indicate 80% confidence intervals. Individual place memory performance is reflected by mean dB from the spatial task. Dots indicate estimates for individual participants, calculated from the mixed-effects logistic regression. The association differed significantly between the control and right patient group ($z = 2.5497$; $p = 0.011$).

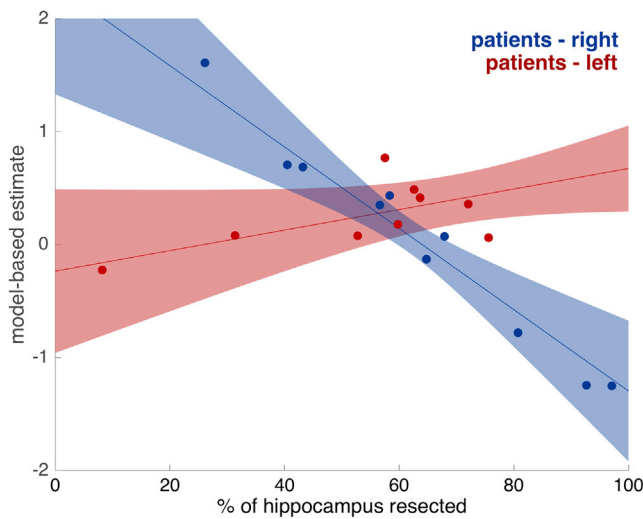


Figure 8. Relationships between Model-Based Planning and Hippocampal Lesion Size for Patients with Right- and Left-Lateralized ATL
 Estimated with mixed-effects regressions, controlling for IQ, age, and total size of lesion relative to the right-lateralized patient group. Error bars indicate 80% confidence intervals. Dots indicate estimates for individual participants, calculated from the mixed-effects logistic regression. There was a significant relationship in the right patient group ($z = 2.831$; $p = 0.005$), but not in the left patient group ($z = 1.062$; $p = 0.2882$), with a significant difference between the right and left patient groups ($z = 2.508$; $p = 0.0122$).

not surgically altered, as a result of the chronic epilepsy that prompted the surgery. Indeed, as with all studies of temporal lobe function in patients with epilepsy, the possibility that impaired behavior and cognition in patients is due to a history of epilepsy rather than the surgical intervention per se must be taken into account.

For this reason and others, we must also be cautious about associating the damage with individual structures. Our analyses indicate that the size of lesion to the right hippocampus is significantly related to model-based planning deficits. Still, ATL lesions additionally affect a number of other regions, including parahippocampal cortex, perirhinal cortex, and amygdala, which might also subserve these effects. Moreover, because the pattern of the lesions mainly varies in the extent that the temporal lobe has been removed in the dorsal direction, the patterns of damage to these structures tends to covary across individuals. Such collinearity makes it difficult to use variation across patients in damage to individual structures to fully disentangle their differential roles. We attempted to mitigate these issues by controlling for overall lesion size. Nevertheless, due to the very substantial analytic and interpretational issues, this anatomical specificity remains emphatically tentative.

These caveats aside, a final question posed by our results concerns how model-based planning and boundary-driven place memory actually relate to one another. In the spatial literature, the notion of a cognitive map primarily refers to place-selective hippocampal activity, which allows organisms to recognize and remember discrete locations in allocentric space. From the perspective of planning, a cognitive map goes beyond such a representation but is built upon it: the map captures the relation-

ships between locations, which can be used to evaluate candidate actions. This function fits well with the broader view of hippocampus supporting relational memory (Eichenbaum and Cohen, 2004; Davachi and Wagner, 2002; Kumaran et al., 2009; Schapiro et al., 2016; Boorman et al., 2016; Garvert et al., 2017), which indicates that the planning deficit in patients stems from hippocampal damage being accompanied by attenuation of the knowledge of relationships between actions and states. This view is also consistent with recent computational models describing how the hippocampus might serve model-based planning. In spatial tasks, sequential activations of place-selective cells are hypothesized to provide, not only a mnemonic function through supporting reactivation of previously traversed trajectories, but a planning function by generating novel place cell sequences, based on the learned contingencies between locations (Johnson and Redish, 2007; Pfeiffer and Foster, 2013; Mattar and Daw, 2018). The related successor representation model (Stachenfeld et al., 2014; Garvert et al., 2017) also focuses on learned relationships among locations by proposing that place selectivity itself is built from experience of state transitions to reflect expectations about future locations. A key challenge for future work addressing these ideas will be studying hippocampal activity in tasks, like the planning one used here, which manipulate animals' experience of environmental relationships, to reveal how they leverage this knowledge to guide choice.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - MRI Scanning
 - Task Order
 - Sequential-Decision Making Task
 - Spatial Memory Task
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - MRI Image Processing
 - Sequential-Decision Making Task – Regression Analysis
 - Sequential-Decision Making Task – Computational Model Fit
 - Group-level Modeling and Estimation for Computational Model Fit
 - Spatial Memory Task Analysis

SUPPLEMENTAL INFORMATION

Supplemental Information can be found with this article online at <https://doi.org/10.1016/j.neuron.2019.02.014>.

ACKNOWLEDGMENTS

We wish to thank Philip Parnamets and Aaron Bornstein for helpful input and Catherine Hartley for assistance with experimental stimuli. This project was

supported by NIH grant DA038891, part of the CRCNS program, and a gift from Google DeepMind.

AUTHOR CONTRIBUTIONS

Conceptualization, O.M.V. and N.D.D.; Methodology, O.M.V., N.D.D., and N.B.; Software, O.M.V., J.K., and N.D.D.; Formal Analysis, O.M.V.; Investigation, O.M.V., M.R.M., and K.B.; Resources, O.D., M.R.M., and K.B.; Data Curation, M.R.M. and K.B.; Writing – Original Draft, O.M.V. and N.D.D.; Writing – Review & Editing, O.M.V., N.D.D., D.S., N.B., O.D., and M.R.M.; Visualizations, O.M.V.; Supervision, N.D.D., N.B., and D.S.; Funding Acquisition, N.D.D. and D.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 29, 2018

Revised: December 21, 2018

Accepted: February 8, 2019

Published: March 11, 2019

REFERENCES

- Adams, C.D. (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q. J. Exp. Psychol.* *34*, 77–98.
- Adams, C.D., and Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Q. J. Exp. Psychol.* *33*, 109–121.
- Addis, D.R., Cheng, T., Roberts, R.P., and Schacter, D.L. (2011). Hippocampal contributions to the episodic simulation of specific and general future events. *Hippocampus* *21*, 1045–1052.
- Akam, T., Costa, R., and Dayan, P. (2015). Simple plans or sophisticated habits? State, transition and learning interactions in the two-step task. *PLoS Comput. Biol.* *11*, e1004648.
- Balleine, B.W., and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* *37*, 407–419.
- Bayer, H.M., and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* *47*, 129–141.
- Boorman, E.D., Rajendran, V.G., O'Reilly, J.X., and Behrens, T.E. (2016). Two anatomically and computationally distinct learning signals predict changes to stimulus-outcome associations in hippocampus. *Neuron* *89*, 1343–1354.
- Brown, T.I., Carr, V.A., LaRocque, K.F., Favila, S.E., Gordon, A.M., Bowles, B., Bailenson, J.N., and Wagner, A.D. (2016). Prospective representation of navigational goals in the human hippocampus. *Science* *352*, 1323–1326.
- Büchel, C., Dolan, R.J., Armony, J.L., and Friston, K.J. (1999). Amygdala-hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. *J. Neurosci.* *19*, 10869–10876.
- Burgess, N., Maguire, E.A., and O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron* *35*, 625–641.
- Chadwick, M.J., Jolly, A.E., Amos, D.P., Hassabis, D., and Spiers, H.J. (2015). A goal direction signal in the human entorhinal/subicular region. *Curr. Biol.* *25*, 87–92.
- Corbit, L.H., and Balleine, B.W. (2000). The role of the hippocampus in instrumental conditioning. *J. Neurosci.* *20*, 4233–4239.
- Corbit, L.H., Ostlund, S.B., and Balleine, B.W. (2002). Sensitivity to instrumental contingency degradation is mediated by the entorhinal cortex and its efferents via the dorsal hippocampus. *J. Neurosci.* *22*, 10976–10984.
- Davachi, L., and Wagner, A.D. (2002). Hippocampal contributions to episodic encoding: insights from relational and item-based learning. *J. Neurophysiol.* *88*, 982–990.
- Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* *8*, 1704–1711.
- Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., and Dolan, R.J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron* *69*, 1204–1215.
- Decker, J.H., Otto, A.R., Daw, N.D., and Hartley, C.A. (2016). From creatures of habit to goal-directed learners: Tracking the developmental emergence of model-based reinforcement learning. *Psychol. Sci.* *27*, 848–858.
- Dickinson, A., and Balleine, B. (1993). Actions and responses: the dual psychology of behaviour. In *Spatial Representation: Problems in Philosophy and Psychology*, N. Eilan, R.A. McCarthy, and B. Brewer, eds. (Blackwell), pp. 277–293.
- Doeller, C.F., King, J.A., and Burgess, N. (2008). Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. *Proc. Natl. Acad. Sci. USA* *105*, 5915–5920.
- Doll, B.B., Duncan, K.D., Simon, D.A., Shohamy, D., and Daw, N.D. (2015). Model-based choices involve prospective neural activity. *Nat. Neurosci.* *18*, 767–772.
- Dusek, J.A., and Eichenbaum, H. (1997). The hippocampus and memory for orderly stimulus relations. *Proc. Natl. Acad. Sci. USA* *94*, 7109–7114.
- Economides, M., Kurth-Nelson, Z., Lübbert, A., Guitart-Masip, M., and Dolan, R.J. (2015). Model-based reasoning in humans becomes automatic with training. *PLoS Comput. Biol.* *11*, e1004463.
- Eichenbaum, H., and Cohen, N.J. (2004). *From Conditioning to Conscious Recollection: Memory Systems of the Brain* (Oxford University).
- Eilan, N., McCarthy, R.A., and Brewer, B. (1993). *Spatial Representation: Problems in Philosophy and Psychology* (Blackwell).
- Foerde, K., and Shohamy, D. (2011). Feedback timing modulates brain systems for learning in humans. *J. Neurosci.* *31*, 13157–13167.
- Frank, M.J., Seeberger, L.C., and O'Reilly, R.C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* *306*, 1940–1943.
- Garvert, M.M., Dolan, R.J., and Behrens, T.E. (2017). A map of abstract relational knowledge in the human hippocampal-entorhinal cortex. *eLife* *6*, e17086.
- Gelman, A., and Rubin, D.B. (1992). Inference from iterative simulation using multiple sequences. *Stat. Sci.* *7*, 457–472.
- Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. (2003). *Bayesian Data Analysis* (Chapman and Hall/CRC Texts in Statistical Science).
- Gershman, S.J., Blei, D.M., and Niv, Y. (2010). Context, learning, and extinction. *Psychol. Rev.* *117*, 197–209.
- Gillan, C.M., Kosinski, M., Whelan, R., Phelps, E.A., and Daw, N.D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife* *5*, e11305.
- Gläscher, J., Daw, N., Dayan, P., and O'Doherty, J.P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* *66*, 585–595.
- Hartley, T., Maguire, E.A., Spiers, H.J., and 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.
- Hassabis, D., Kumaran, D., Vann, S.D., and Maguire, E.A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proc. Natl. Acad. Sci. USA* *104*, 1726–1731.
- Heckers, S., Zalesak, M., Weiss, A.P., Ditman, T., and Titone, D. (2004). Hippocampal activation during transitive inference in humans. *Hippocampus* *14*, 153–162.
- Hirsh, R. (1974). The hippocampus and contextual retrieval of information from memory: a theory. *Behav. Biol.* *12*, 421–444.
- Iaria, G., Petrides, M., Dagher, A., Pike, B., and Bohbot, V.D. (2003). Cognitive strategies dependent on the hippocampus and caudate nucleus in human navigation: variability and change with practice. *J. Neurosci.* *23*, 5945–5952.

- Jenkinson, M., and Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5, 143–156.
- Johnson, A., and Redish, A.D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J. Neurosci.* 27, 12176–12189.
- Kaplan, R., King, J., Koster, R., Penny, W.D., Burgess, N., and Friston, K.J. (2017). The neural representation of prospective choice during spatial planning and decisions. *PLoS Biol.* 15, e1002588.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., and Broussard, C. (2007). What's new in psychtoolbox-3. *Perception* 36, 1–16.
- Knowlton, B.J., Mangels, J.A., and Squire, L.R. (1996). A neostriatal habit learning system in humans. *Science* 273, 1399–1402.
- Kosaki, Y., Pearce, J.M., and McGregor, A. (2018). The response strategy and the place strategy in a plus-maze have different sensitivities to devaluation of expected outcome. *Hippocampus* 28, 484–496.
- Kruschke, J.K. (2010). Bayesian data analysis. Wiley Interdiscip. Rev. Cogn. Sci. 1, 658–676.
- Kumaran, D., Summerfield, J.J., Hassabis, D., and Maguire, E.A. (2009). Tracking the emergence of conceptual knowledge during human decision making. *Neuron* 63, 889–901.
- Kumaran, D., Hassabis, D., and McClelland, J.L. (2016). What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends Cogn. Sci.* 20, 512–534.
- Lewandowski, D., Kurowicka, D., and Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis* 100, 1989–2001.
- Mackey, W.E., Devinsky, O., Doyle, W.K., Meager, M.R., and Curtis, C.E. (2016). Human dorsolateral prefrontal cortex is not necessary for spatial working memory. *J. Neurosci.* 36, 2847–2856.
- Mattar, M.G., and Daw, N.D. (2018). Prioritized memory access explains planning and hippocampal replay. *Nat. Neurosci.* 21, 1609–1617.
- Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., et al. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 356, 1293–1322.
- McDonald, R.J., and White, N.M. (1994). Parallel information processing in the water maze: evidence for independent memory systems involving dorsal striatum and hippocampus. *Behav. Neural Biol.* 61, 260–270.
- Miller, K.J., Botvinick, M.M., and Brody, C.D. (2017). Dorsal hippocampus plays a causal role in model-based planning. *bioRxiv*. <https://doi.org/10.1101/096594>.
- Morris, R.G.M., Garrud, P., Rawlins, J.N., and O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature* 297, 681–683.
- O'Keefe, J., and Nadel, L. (1978). *The Hippocampus as a Cognitive Map* (Clarendon Press).
- Packard, M.G., and McGaugh, J.L. (1996). Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiol. Learn. Mem.* 65, 65–72.
- Pearce, J.M., Roberts, A.D., and Good, M. (1998). Hippocampal lesions disrupt navigation based on cognitive maps but not heading vectors. *Nature* 396, 75–77.
- Pfeiffer, B.E., and Foster, D.J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* 497, 74–79.
- Poldrack, R.A., and Packard, M.G. (2003). Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia* 41, 245–251.
- Schapiro, A.C., Turk-Browne, N.B., Norman, K.A., and Botvinick, M.M. (2016). Statistical learning of temporal community structure in the hippocampus. *Hippocampus* 26, 3–8.
- Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.
- Scoville, W.B., and Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry* 20, 11–21.
- Shohamy, D., and Daw, N.D. (2015). Integrating memories to guide decisions. *Curr. Opin. Behav. Sci.* 5, 85–90.
- Shohamy, D., and Wagner, A.D. (2008). Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron* 60, 378–389.
- Simon, D.A., and Daw, N.D. (2011). Neural correlates of forward planning in a spatial decision task in humans. *J. Neurosci.* 31, 5526–5539.
- Solomon, P.R., Vander Schaaf, E.R., Thompson, R.F., and Weisz, D.J. (1986). Hippocampus and trace conditioning of the rabbit's classically conditioned nictitating membrane response. *Behav. Neurosci.* 100, 729–744.
- Spiers, H.J., and Maguire, E.A. (2007). A navigational guidance system in the human brain. *Hippocampus* 17, 618–626.
- Spiers, H.J., Burgess, N., Maguire, E.A., Baxendale, S.A., Hartley, T., Thompson, P.J., and O'Keefe, J. (2001). Unilateral temporal lobectomy patients show lateralized topographical and episodic memory deficits in a virtual town. *Brain* 124, 2476–2489.
- Squire, L.R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.* 99, 195–231.
- Stachenfeld, K.L., Botvinick, M.M., and Gershman, S.J. (2014). Design principles of the hippocampal cognitive map. *NIPS'14 Proceedings of the 27th International Conference on Neural Information Processing Systems (MIT)*, pp. 2528–2536.
- Tolman, E.C. (1948). Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208.
- van Praag, H., Qu, P.M., Elliott, R.C., Wu, H., Dreyfus, C.F., and Black, I.B. (1998). Unilateral hippocampal lesions in newborn and adult rats: effects on spatial memory and BDNF gene expression. *Behav. Brain Res.* 92, 21–30.
- Viard, A., Doeller, C.F., Hartley, T., Bird, C.M., and Burgess, N. (2011). Anterior hippocampus and goal-directed spatial decision making. *J. Neurosci.* 31, 4613–4621.
- Voermans, N.C., Petersson, K.M., Daudey, L., Weber, B., Van Spaendonck, K.P., Kremer, H.P., and Fernández, G. (2004). Interaction between the human hippocampus and the caudate nucleus during route recognition. *Neuron* 43, 427–435.
- Wechsler, D. (2008). *Wechsler adult intelligence scale—Fourth Edition (WAIS-IV)*. NCS Pearson 22, 498.
- Wikenheiser, A.M., and Redish, A.D. (2015). Hippocampal theta sequences reflect current goals. *Nat. Neurosci.* 18, 289–294.
- Wimmer, G.E., and Shohamy, D. (2012). Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science* 338, 270–273.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
MATLAB	Mathworks	https://www.mathworks.com
Psychophysics Toolbox	Kleiner et al., 2007	http://psychtoolbox.org/
UnrealEngine2	Epic Games	http://api.unrealengine.com/udk/Two/WebHome.html
Stan Modeling Language	Stan Development Team	https://mc-stan.org/
FSL FLIRT	Jenkinson and Smith, 2001	https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/
MRICron	NITRC	https://www.mccauslandcenter.sc.edu/mricro/mricron/

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Oliver M. Vikbladh (omv208@nyu.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The NYU committee on activities involving human subjects approved this study, and all participants gave written informed consent before participation. 19 individuals, who had undergone unilateral anterior temporal lobectomy (ATL) for the treatment of intractable epilepsy, were recruited from the New York University (NYU) Patient Registry for the Study of Perception, Emotion and Cognition (PROSPEC). A clinical neuropsychologist (MRM or KB) conducted all standardized procedures for screening patients for inclusion into NYU PROSPEC. Patients were only selected for inclusion in the study if there was no evidence of global cognitive dysfunction as measured by a comprehensive neuropsychological evaluation, an FSIQ (Wechsler Adult Intelligence Scale-Fourth Edition; [Wechsler, 2008](#)) above 80, no evidence of diffuse atrophy on MRI (e.g., brain tumor or idiopathic epilepsy), or and no history of psychiatric or neurologic disease other than the primary etiology for the focal brain lesion. The patient participants had a mean age of 37.0 years (SE = 1.5), mean IQ score of 109.7 (SE = 2.8) and 10/9 male to female ratio. 19 healthy controls were also recruited from the local community through internet-based advertisement and gave consent to participate in the study. The control participants had a mean age of 39.3 years (SE = 3.6), mean IQ score of 108.7 (SE = 3.6) and 11/8 male to female ratio.

METHOD DETAILS

MRI Scanning

When post-surgical structural brain scans (T1 MP-RAGE) were not available from the referring center, the Department of Radiology at the NYU School of Medicine, patients were imaged at the NYU Center for Brain Imaging on a 3-Tesla Siemens Allegra head-only MR scanner. Medical Center scans were obtained using 1.5- or 3-Tesla Siemens full-body MR scanners. Image acquisitions included a conventional three-plane localizer and two T1-weighted gradient-echo sequence (MP-RAGE) volumes (TE = 3.25 ms, TR = 2530 ms, TI = 1.100 ms, flip angle = 7°, FOV = 256 mm, voxel size = 1 × 1 × 1.33 mm). Acquisition parameters were optimized for increased gray/white matter image contrast.

Task Order

On the day of testing, participants completed the two behavioral tasks separated by a short break. For all participants the sequential decision-making task was given first, followed by the spatial memory task. For the control participants the completion of the tasks was followed by the administration of the WAIS-IV. For the patient participants, the WAIS-IV had been completed during screening procedures for inclusion into PROSPEC.

Sequential-Decision Making Task

Participants completed 200 trials of a two-step Markov decision task designed to quantify the extent to which participants use a world model to prospectively evaluate actions ([Daw et al., 2011](#)). The task was framed as a game about mining for space treasure ([Decker et al., 2016](#)) and was presented with MATLAB, using the Psychophysics Toolbox extensions ([Kleiner et al., 2007](#)). Each trial involved two choices in succession, followed by reward ([Figure 2](#)). Participants first made a choice between two actions, depicted as spaceships, randomly presented on the left and right. The choice resulted in a transition to one of two second-stage states (depicted as a red or purple planet). One spaceship most commonly ($p = 0.7$) transitioned to the purple planet, and otherwise made a rare

transition ($p = 0.3$) to the red planet. For the other spaceship, probabilities were reversed. Participants were informed that each spaceship was more likely to go to a different planet but not which planet, nor the explicit transition probabilities.

Subsequently, participants made a choice between two actions depicted as a pair of aliens that were unique to the planet, randomly presented on the left and right. Each alien was associated with a probability of monetary reward (versus nothing) that slowly diffused over trials according to an independent random walk. Rewards were paid out at the end of the experiment at a rate of 15 cents per reward. The random change in the second-stage reward probabilities encouraged participants to adjust their choice preferences at both stages trial-by-trial, so as to maximize payoffs. For each choice, participants had 3 s to respond; or else the trial was aborted with a time-out message.

Prior to the experiment, participants completed an extensive instructional tutorial. The tutorial included a 20-trial practice run, using a different set of visual stimuli (planets, spaceships, and aliens) but otherwise identical.

Spatial Memory Task

Each participant completed 64 trials of a spatial memory task, identical (with one exception, see below) to the task used by [Doeller et al. \(2008\)](#). On each trial, participants navigated a virtual reality arena using keyboard presses. UnrealEngine2 Runtime software (Epic Games) was used to present a first-person perspective view of the arena. The virtual arena ([Figure 2 Left](#)) was bounded by a circular wall, contained a single intra-maze landmark in the form of a traffic cone, and was surrounded by distant cues (mountains, clouds, and the sun) projected at infinity. Both the boundary (wall) and landmark (cone) were rotationally symmetric, leaving the distal cues as the main source of orientation.

At the beginning of each trial, a picture of one of four objects was presented on a gray background for 2 s. Participants were then placed in a random position within the arena without any objects, one-fifth of the radius from the center of the arena and facing a random direction (note that in [Doeller et al., 2008](#) the starting radius was not restricted). Participants subsequently had 12 s to navigate to the correct location of the object as they remembered it from previous trials, and indicate that position by a button press. Following this button press, the object immediately appeared in its correct location. If no response had been made in 12 s, the object also appeared in its correct location automatically. Participants ended the trial by collecting the object in its correct location. A fixation cross was then presented for 2 s, before the start of the next trial.

The task consisted of 64 trials divided into 4 continuous blocks, each containing 4 pseudo-randomized presentations of each of the 4 objects. Between blocks, the landmark moved in relation to the boundaries, such that there were four arena configurations, with the landmark roughly in the middle of the north, south, west, and east sectors of the arena, as defined by the distal cues ([Figure 2 Middle](#)). The order of arena configurations over blocks was counterbalanced across participants and experimental groups. Participants were not informed of the landmark movements prior to the experiment.

During the first block, the correct location of all objects was in rough proximity of the landmark. Two of the objects were ‘boundary objects’, for which the correct locations were fixed relative to the environmental boundaries across the whole experiment. The other two objects (unannounced to the participants) were ‘landmark objects,’ for which correct locations were fixed at a constant distance and direction to the intra-maze landmark even as the landmark moved.

The task probed for memory of correct object locations within the arena. Critically, by manipulating the landmark location in relation to the boundary and distal cues, the task distinguished whether participants stored place memory of allocentric location in relation to the boundary and distal cues, or by egocentric response memory in relation to individual landmarks ([Figure 4 Right](#)). The original study, using the same procedure during fMRI in healthy participants, showed that boundary-related and landmark-related memory correlated with activity in the right posterior hippocampus and striatum, respectively ([Doeller et al., 2008](#)).

Participants practiced in an unrelated virtual environment with a different set of object stimuli before performing the experiment. Additionally, before the first trial, participants collected each of the objects once in their correct block 1 locations.

QUANTIFICATION AND STATISTICAL ANALYSIS

MRI Image Processing

The high-resolution structural images from each patient were normalized to Montreal Neurological Institute (MNI) standard space using FSL FLIRT (FMRIB’s Linear Image Registration Tool; <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>) ([Jenkinson and Smith, 2001](#)). This consisted of a two-step procedure: First, using MRICron (<https://www.mccauslandcenter.sc.edu/mricro/mricron/>), a mask was drawn over the lesion and any craniotomy defect to prevent bias in the transformation, then masked voxels were assigned a weight of “0” and ignored during a subsequent 12-parameter affine transformation of the lesioned brain to the standard MNI 1 mm reference volume ([Mackey et al., 2016](#)). The second step was manually tracing the lesions on individual slices of the patients’ brains overlaid on the standard MNI brain template, while crosschecking in all three planes. This tracing procedure produced a 3D mask with “1” indicating the presence of the lesion and “0” the presence of normal tissue. All patients had surgical lesions, which made the margins readily visible on the T1-weighted MRI images. In instances where there was uncertainty regarding the lesion margins, the treating neurosurgeon(s) and/or neuro-radiologists were consulted.

The lesion masks drawn in MNI space were subsequently overlaid on the Harvard-Oxford Structural Atlas ([Mazziotta et al., 2001](#)) to estimate the extent of damage to the hippocampus ([Figure 1](#)). Hippocampal lesion size was calculated as the voxel overlap between the individual lesion masks and the hippocampus as defined by the Atlas with $p > 0.5$.

Sequential-Decision Making Task – Regression Analysis

The logic of the task exploits the noisy coupling between spaceships and planets to measure model-free learning - directly learning the value of spaceship choices versus model-based planning - prospectively computing the value of the spaceship choices in terms of the planets they lead to.

For instance, consider on some trial choosing the spaceship that usually transitions to the purple planet, but instead being taken to the red planet (a “rare” transition). On the red planet your choice of alien is subsequently rewarded. In this situation model-free and model-based strategies make conflicting predictions about first-level choice behavior on the next trial. Participants using a model-free strategy will be more likely to choose the same spaceship on next trial, as it was rewarded. Conversely, participants using a model-based strategy will be more likely to switch and choose the other first-level action. This is because the model-based strategy computes the value of the spaceships using a cognitive map or model of their transition probabilities to the respective planets and the reward expected at the planets.

The goal of analysis was to estimate, for each participant, the extent to which they followed either strategy. Following previous work (Daw et al., 2011), we did this two ways, using a factorial logistic regression that captures the above qualitative logic, and fits of a more elaborate, but more assumption-laden, computational learning model.

We analyzed the first-level choices over spaceships using mixed-effects logistic regression (estimated using the *fitglm* function in MATLAB). For each trial, the dependent variable (coded as stay with the same spaceship or switch, relative to the previous trial) was explained in terms of two events from the previous trial: whether reward was received, whether the planet encountered was reached following a common or rare transition given the spaceship chosen, and the interaction of these two factors. Our measure of model-free choice was the main effect of reward; our measure of model-based choice was the interaction of reward by transition type (common versus rare). We further interacted the task factors with experimental group (lesion versus control) as well as with two nuisance covariates, IQ and age, which have both been shown to affect behavior on this task (Gillan et al., 2016). The intercept, and the regression coefficients for reward, transition, and their interaction were all taken as random effects (allowed to vary across participants).

To test our predictions about the relationship between reinforcement learning strategies employed in the decision-making task and place memory performance from the spatial memory task, we also specified a second regression model which interacted the task- and group-related factors (reward, common versus rare, group, and their interactions) with participant-specific average dB. IQ was also included as a nuisance variable. The interactions with dB (up to four-way) measure the extent to which the various effects in the decision task systematically vary, across participants, with their spatial memory performance; i.e., this is analogous to extracting per-participant effect sizes from the decision model and correlating them with dB, but by estimating that correlation as an effect within the regression defining those decision effects, accounts properly for uncertainty in the per-participant estimates. We also calculated a ratio $\text{dB}/(\text{dL}+\text{dB})$, where dL and dB were participant-wise means of landmark and response distance error. This ratio was also used in a separate model interacted with task- and group-related factors.

Sequential-Decision Making Task – Computational Model Fit

The logistic regression analysis considers only the previous trial’s experience in predicting each choice; this simplification is motivated by a limiting argument over the learning rate parameter in a more elaborate RL model of the data (Daw et al., 2011). In order to ensure that our results were not affected by neglecting the effect of earlier trials, we repeated our analyses fitting each participant’s trial-by-trial choices with a full RL model in which each choice depends on values learned from all previous rewards (based on Daw et al., 2011, but using the version from Gillan et al., 2016). To estimate the model we utilized Markov Chain Monte Carlo (MCMC) methods, implemented in the Stan modeling language (Stan Development Team). Given an arbitrary generative model for data dependent on free parameters, the method permits samples to be drawn from the posterior probability distribution of parameter values, conditional on the observed data. From the quantiles of these distributions, we constructed confidence intervals – technically, credible intervals – over the likely values of the free parameters (Kruschke, 2010). We also report the posterior likelihood that the credible region contains zero, as one minus the size of the largest symmetric credible interval that excludes zero, which is roughly comparable to a two-sided P value.

For each model, we produced 4 chains of 10,000 samples each. The first 2500 samples from each chain were discarded to allow for equilibration. We verified the convergence of the chains by visual inspection, and additionally by computing for each parameter the ‘potential scale reduction factor’ \hat{R} (Gelman and Rubin, 1992). For all parameters, we verified that $\hat{R} < 1.1$, a range consistent with convergence (Gelman et al., 2003).

We simultaneously estimated a model of all the data, incorporating individual parameters for each participant nested within a population-level model of the distribution of these parameters for each group.

At the participant level, the model is the same as the one used by Gillan et al. (2016), and full equations are presented there. In brief, the model learns from experience to predict values $Q(s, a)$ for the different actions a (rockets, aliens) in the different states (planets and the starting state). Different RL algorithms, model-based and model-free, produce different estimates Q at each step. First-level (spaceship) choices are determined by softmax choice, according to the weighted combination of model-based and model-free Q values, with weightings controlled by the free inverse temperature parameters β^{MB} and β^{MF} ; a third parameter β^{stick} captures any value-independent bias to stay or switch. Second-level (alien) choices are determined by a single set of Q values (since

model-based and model-free evaluation coincide for terminal actions), with inverse temperature β^{stage2} . The various Q values are updated according to delta rules with a free learning rate parameter α . Finally, the net model-free weighting β^{MF} is itself derived from the weighted combination of Q values learned by two variants of TD learning, TD(0) and TD(1), with weights β^{MF0} and β^{MF1} . (This is a minor change of variables with respect to the standard model-free TD(λ) algorithm used to hybridize these learning rules in Daw et al. (2011). Here the second temperature parameter replaces the eligibility trace parameter λ used in that model, which has the advantage of eliminating its 0,1 boundaries.) Following estimation, we reverse the change of variables by computing the net model-free weighting as $\beta^{MF} = \beta^{MF0} / \alpha + \beta^{MF1}$, where the α accounts for a difference in scaling between the two parameters (see Gillan et al., 2016). When making group comparisons, group estimates of β^{MF0} are scaled by the estimated α of the corresponding group.

The model thus estimates six free parameters per participant: α , β^{MB} , β^{MF0} , β^{MF1} , β^{stick} , and β^{stage2} , and our main hypotheses of interest concern group-wise differences in β^{MB} and the net β^{MF} .

Group-level Modeling and Estimation for Computational Model Fit

The model was specified hierarchically, so that the participant-specific parameter estimates were assumed to be drawn from a population-level distribution, separately for the patient and control groups. In particular, parameters $\vec{\theta}_s$ (a six-vector) for each participant s were modeled as drawn from a multivariate normal with mean $\vec{\mu}$ and covariance Σ . An additional vector $\vec{\mu}^{les}$ coded any difference in means for the lesion group (i.e., their mean was $\vec{\mu} + \vec{\mu}^{les}$, allowing us to test for group differences in each parameter by comparing the corresponding element of $\vec{\mu}^{les}$ to zero). For the parameter α (which is constrained to 0,1), the corresponding element of $\vec{\theta}_s$ (which has infinite support) was transformed through the CDF of the standard normal.

We jointly estimated the posterior distribution over the individual and group-level parameters using MCMC as described above, which required specifying prior distributions (“hyper-priors”) on the parameters of the group level distributions. In particular, priors for the elements of $\vec{\mu}$ and $\vec{\mu}^{les}$ were individually normal (mean = 0, SD = 2), which is uninformative within the relevant range. The covariance Σ was specified (as recommended in the Stan documentation) as the product of a correlation matrix Ω (which had an LKJ prior with shape $\nu = 2$; Lewandowski et al., 2009) scaled element wise by the outer product of a scale vector $\vec{\tau}$ (whose elements were again taken as normal, mean = 0, sd = 2) with itself. This model also included individual IQ scores and age as covariates.

In order to test the interaction between performances in the two tasks, we then expressed a new model with group-specific parameter estimates (priors were normal distributions with mean = 0 and sd = 1) that specified how individual z-scored participant-wise estimates of mean boundary distance error (dB) predicted the participant-specific parameter estimates. This model also included individual IQ scores as a covariate.

Spatial Memory Task Analysis

To measure memory of locations in relation to boundary and landmark cues, we focused our main analysis on the trials following the relative movement of the landmark in relation to the boundaries. Reliance on boundary cues was quantified by boundary distance error (dB), where dB was the distance from the response location to the correct location as defined by the boundary and distal cues in the previous block (Figure 4 Right). Reliance on landmark cues was quantified by landmark distance error (dL), where dL was the distance from the response location to the correct location as defined by the landmark, according to the position of the object relative to the landmark in the previous block, translated with respect to the landmark’s new position (Figure 4 Right). Low dB thus indicated greater reliance on boundary cues, which we interpret as ‘place memory’, and low dL indicated greater reliance on landmark cues, which we interpret as ‘response memory’.

To capture the repeated-measure structure of the data, all statistical analyses of performance in the task were done using mixed-effects linear regression, treating participant as a random factor. The models were estimated using the *fitlme* function in MATLAB, with standard errors computed using the Satterthwaite approximation to the degrees of freedom when the model was linear, and Wald (asymptotic Gaussian) test for logistic models. The dependent variable, distance error (dB and dL respectively for each trial) was regressed on the key explanatory variables lesion group, distance error type (dB or dL) and object type (boundary or landmark), while also controlling for additional nuisance explanatory factors, age and IQ.