Hippocampal Contributions to
Model-Based Planning and Spatial Memory

Authors: Oliver Vikbladh (Lead Contact: omv208@nyu.edu), Michael R. Meager, John King, Karen E. Blackmon, Orrin Devinsky, Daphna Shohamy, Neil Burgess, Nathaniel D. Daw

Abstract

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, there is limited direct evidence for the link between model-based 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. We found that 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, while planning and boundary-driven place memory covaried with one another, this relationship was attenuated in the patient group, which is 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

functional magnetic resonance imaging (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 vs. habitual or model-free learning (Balleine and Dickinson, 1998, Daw et al., 2005). This distinction seems to parallel the place vs. response memory dichotomy from spatial cognition (Poldrack and Packard, 2003, Kosaki et al., 2018) and the related declarative vs. procedural memory distinction from the memory literature (Squire, 1992, Knowlton et al., 1996, Foerde and Shohamy, 2011, Daw and Shohamy, 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 (Hirsch, 1974, Dickinson and Balleine, 1993, Eilan et al., 1993, Johnson and Redish, 2007, Daw and Shohamy, 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, 2001, 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 which 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 (Hassibis 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, Daw and Shohamy, 2015, Mattar and Daw, 2017).

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 anterior temporal lobectomy (ATL) i.e. surgical removal of the anterior temporal lobe on one side (Table 1), and 19 healthy controls (see 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 vs. females (z=-0.3261, p=0.7444).

|	**Patients**	**Controls**
Number of participants	19	19
Mean IQ (SE)	109.7 (2.8)	108.7 (3.6)
Mean age (SE)	37.0 years (1.9)	39.3 years (2.3)
Males/Females	*10/9*	*11/8*

Table 1: Characteristics of patients and controls.

For 10 out of 19 patients ATL was right lateralized (Table 2). 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 vs. females (z=-0.6752, p=0.4995). Participant-wise lesion masks were normalized to the MNI template (see Figure 1) and compared to the Harvard-Oxford Lexicon (p>.5) in order to estimate size of lesion to the hippocampus (See Methods). There was no group difference in estimated size of hippocampal resection (t_{17}= -

1.1124, p= 0.2814), however, total lesion size was significantly larger in the right lateralized group (t_{17}=-2.6103, p=0.0183).

	Right lateralized ATL	Left lateralized ATL
Number of patients	10	9
Mean IQ (SE)	*111.1 (4.8)*	*108.1 (2.5)*
Mean age (SE)	35.0 years (2.8)	39.2 years (2.7)
Males/Females	6/4	4/5
Mean total lesion (SE)	65314 Voxels (2791)	*40536 Voxels (2296)*
Mean hipp. lesion (SE)	*2791 Voxels (321)*	2296 Voxels (302)

***Table 2: Characteristics for patients with right and left lateralized ATL.***
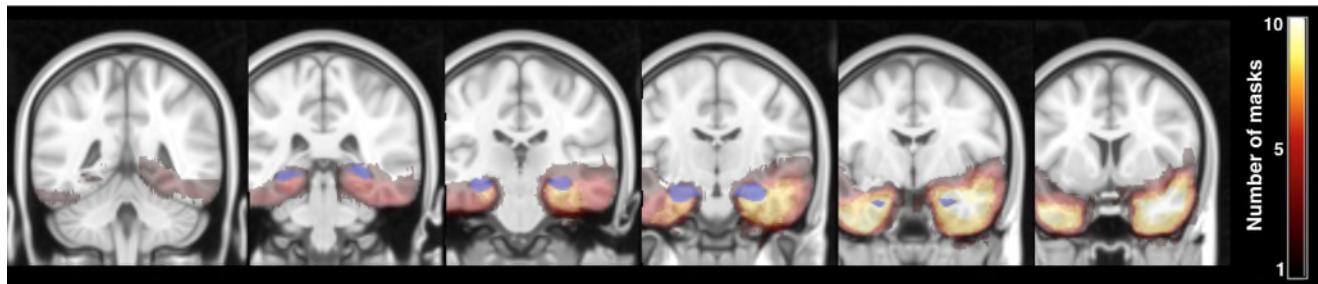
Figure 1: Patient lesion masks. *Slices (y=82, 92, 102, 112, 122, 132) showing all 19 hand-drawn patient ATL lesion masks normalized to the MNI template. Heat maps indicate the number of masks overlapping at a given voxel. The hippocampus, as defined by Harvard-Oxford Lexicon (p>.5), is shown in blue.*

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 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 or 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=.7) transitioned to the purple planet, and otherwise made a rare transition (p=.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 (see Figure 2).
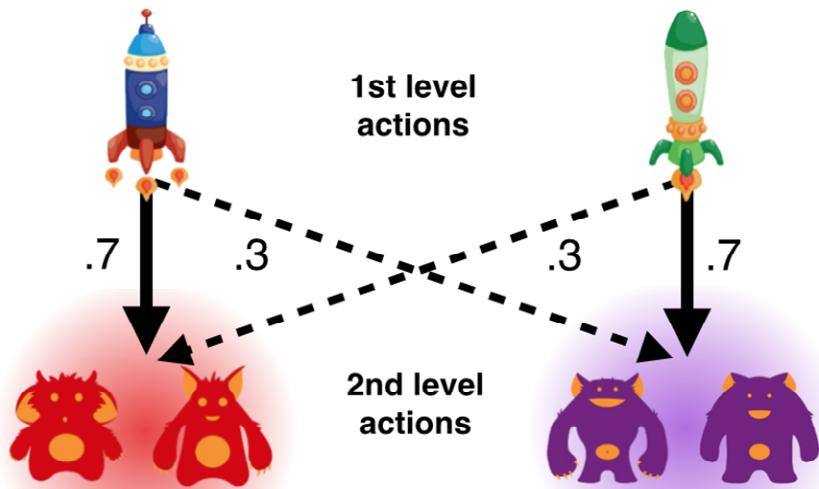
Figure 2: Two-step Markov Decision-Task*. On each trial the participants chose one of two first level actions (spaceships). One space ship transitions the participant to red planet with p=.7 while the other space ship transitions the participant to red planet with p=.3. Having transitioned to a second 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.*

Figure 3 shows markers of both as a function of group, estimated from a factorial logistic regression (see Supplementary Table 2), 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 vs model-free choice. The regression also controls for additional nuisance explanatory factors, age and IQ.

The expression of model-based vs 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=2.0278, 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.
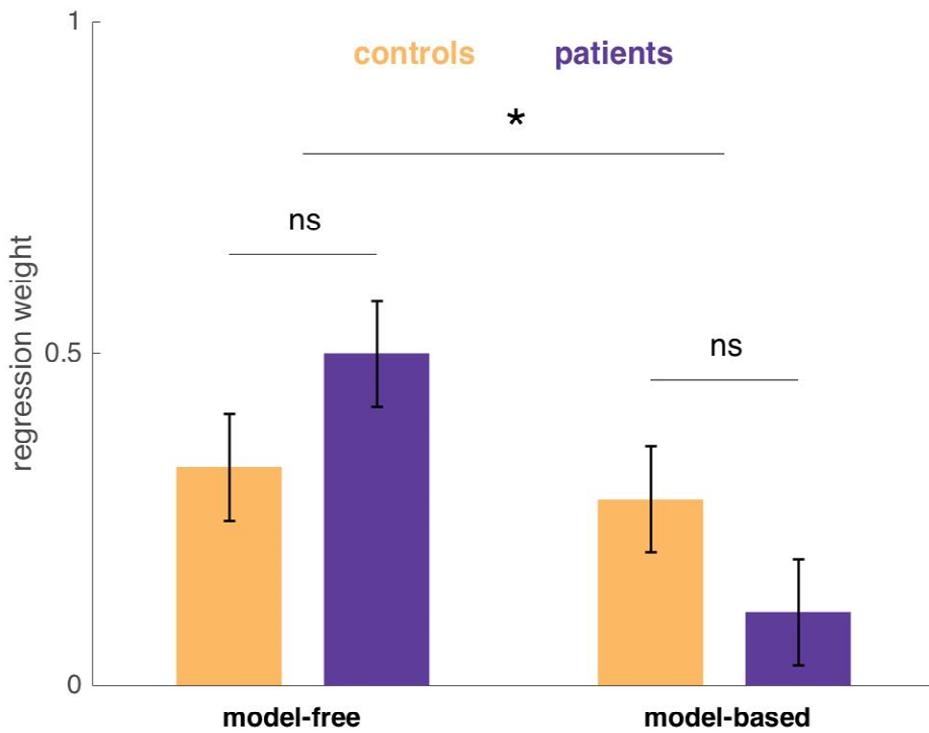
Figure 3: Model-free and model-based regression weights for controls and patients. *Estimated with a logistic mixed-effects regression controlling for IQ and age. Error bars indicate standard error. The interaction of strategy (model-free vs. model-based) by group was significant (z=2.0278, p=0.043).*

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 (see Supplementary Table 13) 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 RL strategy and experimental group such that patients are biased away from model-based and towards 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 indicate, from memory, the correct location of one of four objects in a virtual arena (see 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).
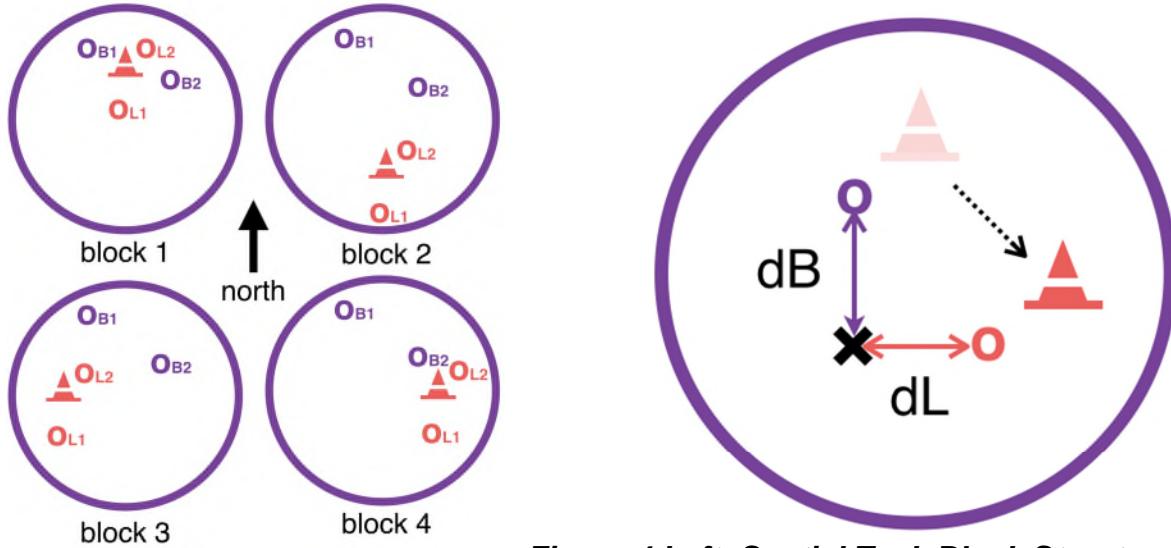
Figure 4 Left: Spatial Task Block Structure. *Virtual arena as seen as from above. Between blocks the landmark (cone) moved in relation to the boundaries (large purple circle). Correct location of the two boundary objects (o_{B1} and o_{B2}) stayed constant with respect to boundaries across all blocks. 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.*** *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 coned) would be the orange o. On the trials following movement of the landmark we thus operationalize place memory by the boundary distance error (dB) between their response (cross) and the location predicted by boundaries and distal cues (purple o). Response memory is operationalized by the landmark distance error (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.*

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 (see Figure 2 Left). 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 (see Figure 2 Left). Participants were not instructed about the difference 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 boundary cue usage. 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, right). Note that since the boundary and landmark cues remain fixed with respect to each other within each block, performance on the remaining trials of each block is not as diagnostic of cue usage, since 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 Supplementary Figure 2 and Supplementary Table 11 for results on these trials). Similarly, to avoid relying on the assumption that subjects 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 (For results broken down by object type, see Supplemental Results and Supplementary Table 12).

On the critical trials, those following movement of the landmark, we quantified reliance on either cue type by computing distance errors dB and dL, respectively, between the chosen location and the correct locations as predicted by boundary cues and landmark cues, based on the previous block (see 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 lesion group (control vs. 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) (see Figure 5 and Supplementary Table 1).

In a follow-up analysis, aimed at simplifying the design for later elaboration by assessing relative reliance on boundary vs. 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 towards relying on boundary cues ($F_{1,432}$=8.213, p=0.004) (see Supplementary Table 9). All these results are consistent with our prediction that anterior temporal lobe structures like the hippocampus preferentially support boundary- over landmark-driven memory.
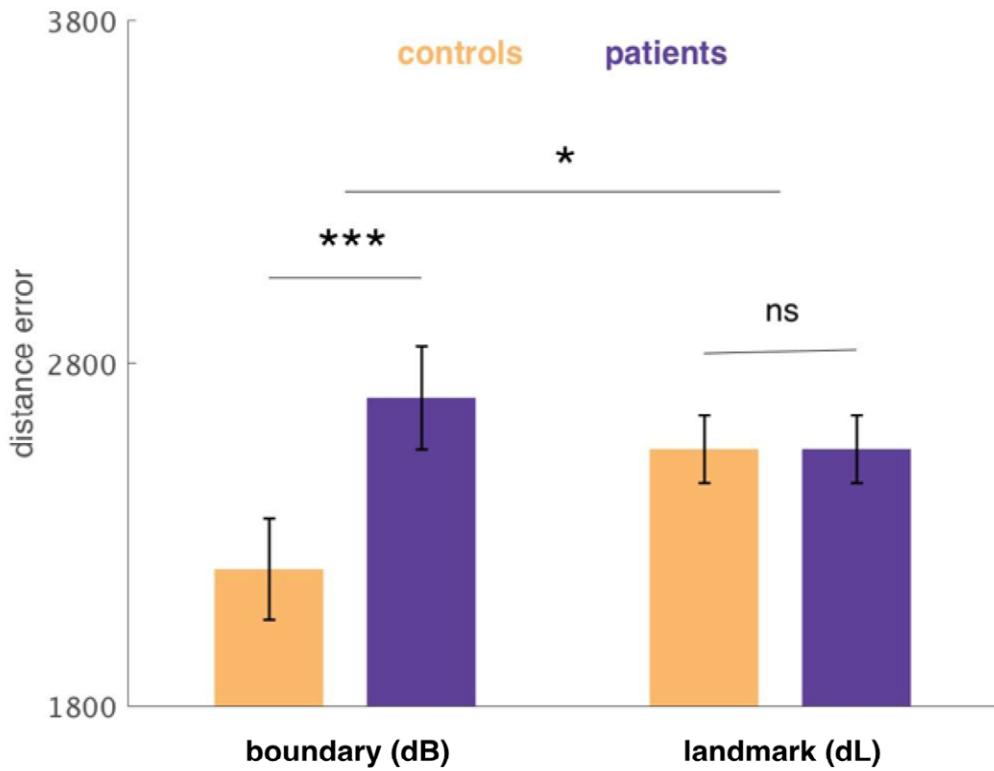
Figure 5:
Boundary (dB) and response (dL) distance error (arbitrary units) for all objects on trials
that follow movements of the landmark. *Estimated with a linear mixed-effects regression,*
controlling for IQ and age. Error bars indicate standard error. There was a significant group
difference in boundary distance error ($F_{1,39.41}=2.5102, p=0.016$), but not landmark distance
error ($F_{1,85.3}=0$, $p=0.9990$) with a significant interaction of group by cue type ($F_{1,97.58}=5.5080$,
$p=0.0210$)

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 boundary distance error 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,
then correlating 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 naïve correlation neglects. IQ was also included as a nuisance covariate to
account for task-general variation.

Figure 6 displays the results of this regression (also see Supplementary Table 3), 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=2.0473, p=0.0326), 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 (see Supplementary Table 14), we replaced the boundary error dB with the relative error ratio dB/(dL+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=2.0578, p=0.039) (see Supplementary Table 10). 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 vs 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/MTL 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 (See Supplementary Table 14). 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).
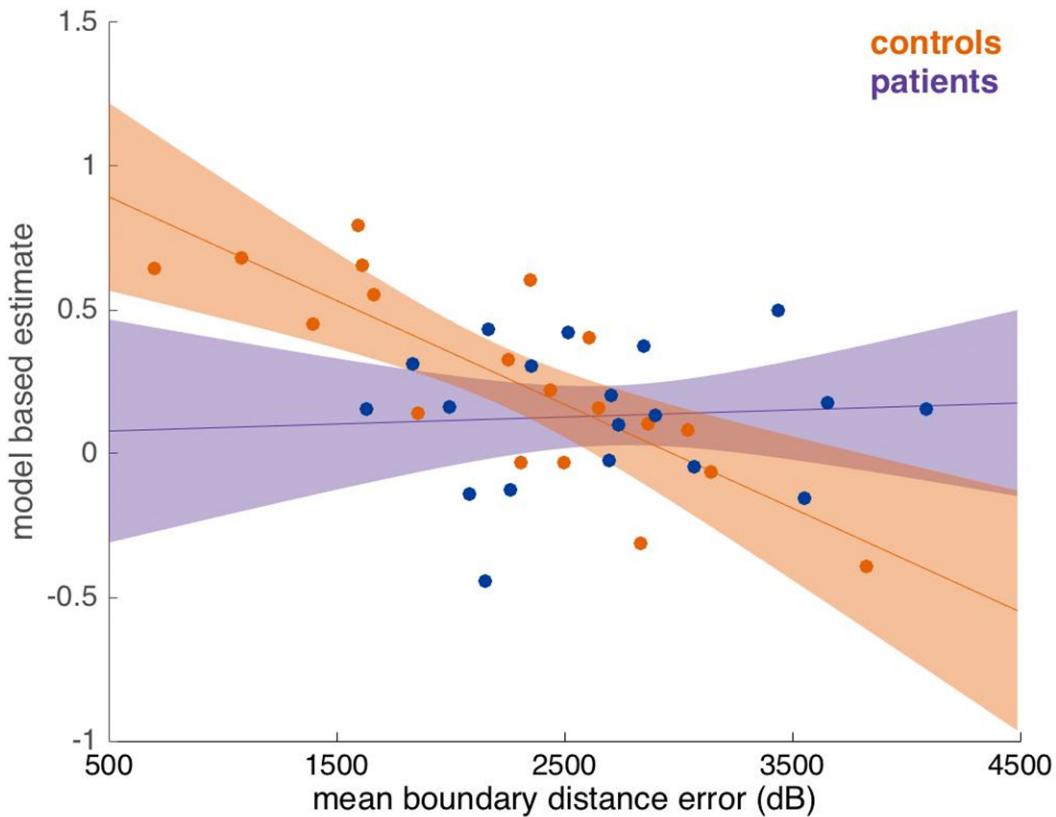
Figure 6: Relationship between model-based planning and boundary distance error (arbitrary unit) in controls and patients*. Estimated with a logistic mixed-effects regression, controlling for IQ. Error bars indicate 80% confidence intervals. Individual place memory performance is reflected by mean boundary distance error (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)*.*

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 (Table 2). 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 the 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 Figure 8 shows the relationship between model-based planning and dB also broken down by laterality. In all three cases the differences between patients and controls (from Figures 4-6) 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 (see Supplementary Tables 4, 5 and 6). In particular, in all three cases the right patients differ significantly from controls (dB minus dL: $F_{1,92.22}=4.4635$, p=0.034; model-based minus model-free: z=2.2950, p=0.022; across-task correlation between model-learning and boundary memory: z=2.5497, 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=1.0008, p=0.317; across-task correlation between model-learning and boundary memory: z=0.4696, p=0.639). However, in no case were the lesion groups significantly different from one another (dB minus dL: $F_{1,97.8}=0.1503$, p=0.69903; model-based minus model-free: z=-1.1490, p=0.251; across-task correlation between model-learning and boundary memory: z=1.5524, 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 (see Supplementary Table 10), the association between relative preference for model-based (minus model-free) planning and the relative bias toward using boundary cues was larger in the left group than the right group (z=3.5278, 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=1.7278, p=0.1170).
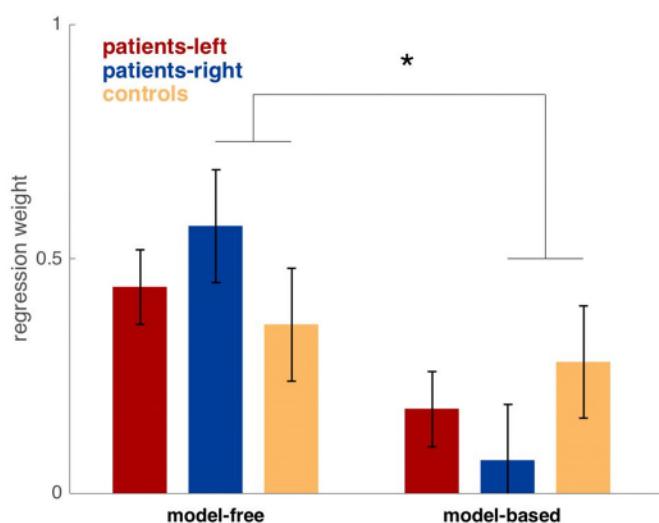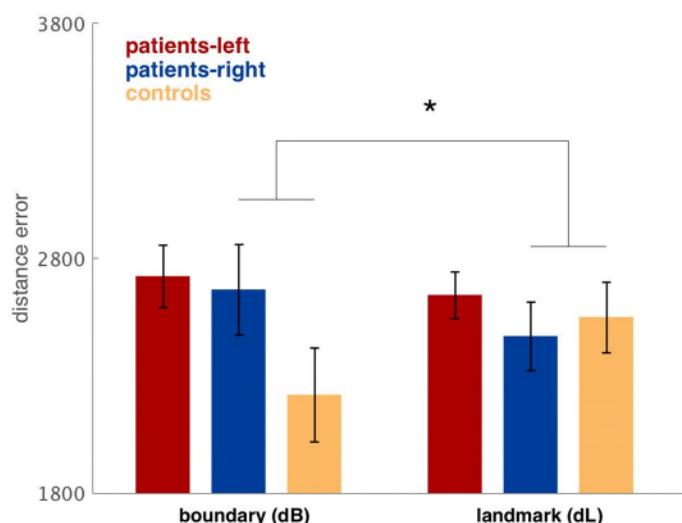
Figure 7 Top: **Boundary (dB)**
and response (dL) distance error (arbitrary units) for controls and patients with right
and left lateralized ATL. *Estimated with a linear mixed-effects regression, controlling for IQ*
and age. Error bars indicate standard error. There was a significant difference between dB and
dL when comparing the control and right patient group ($F_{1,92.22}=4.4635$ p=0.034). **Bottom:**
Model-free and model-based regression weights controls, right and left lateralized ATL
patients. *Estimated with a logistic mixed-effects regression, controlling for IQ and age. Error*
bars indicate standard error. The difference in model-free vs. model-based was significantly
different between the control and right patient group (z=2.295 p=0.022).

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
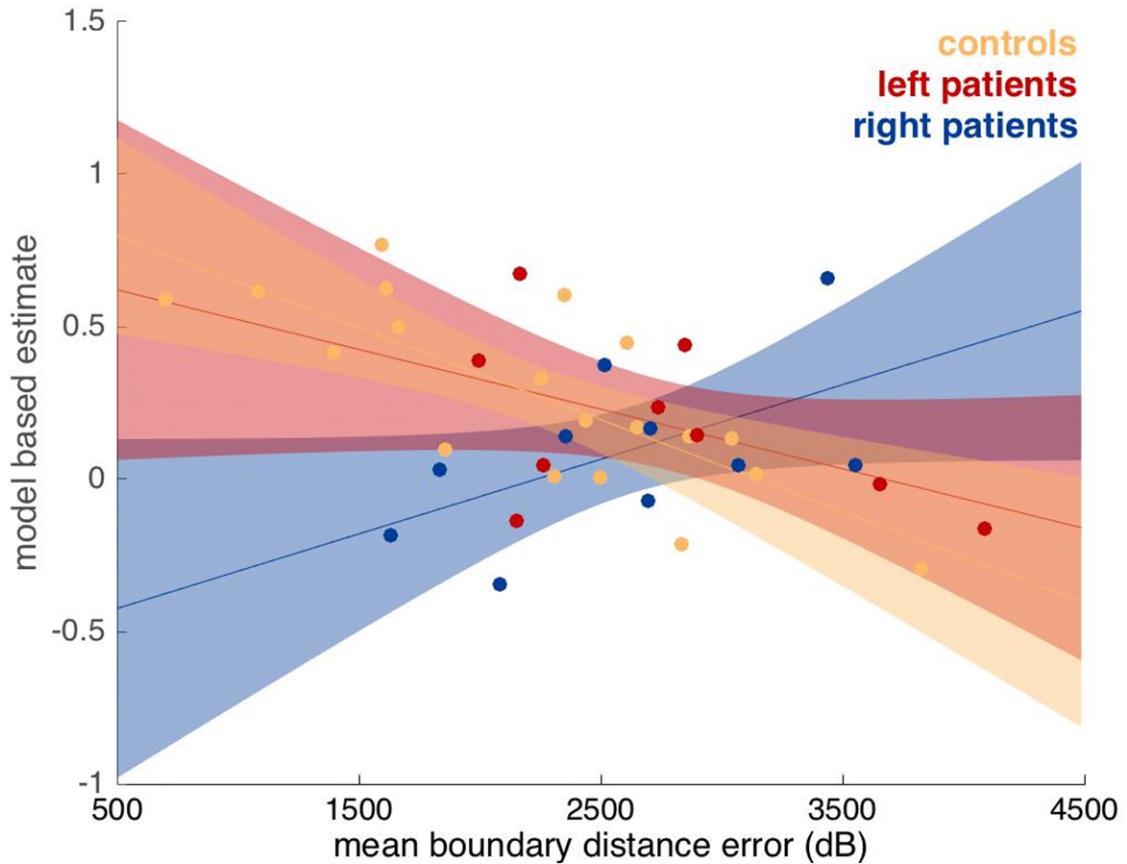
Figu
re 8: Relationship between model-based planning and boundary distance error
(arbitrary unit) for controls and patients with right and left lateralized ATL*. Estimated with*
a logistic mixed-effects regression, controlling for IQ. Error bars indicate 80% confidence
intervals. Individual place memory performance is reflected by mean boundary distance error
(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).

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 co-varied with estimated lesion size on
the right and left hippocampus respectively. Hippocampal lesion size for each patient (see
Figure 1) was estimated by comparing the normalized anatomical masks to the Harvard-
Oxford Lexicon (p>.5) (see 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 results with respect to any particular structure. Although
it is not practical to control for damage to many different MTL structures individually, we
attempt to mitigate these concerns and focus on hippocampal lesion size by controlling 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 9 (see Supplementary Table 7), that model-based planning was significantly worse for larger right hippocampal lesions ($z=2.8305$, $p=0.005$). Conversely, planning was not significantly related to the amount of hippocampal damage on the left hippocampus ($z=1.0622$, $p=0.2882$) and this difference between right and left effects was itself significant ($z=2.5076$, $p=0.0122$). 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 towards a model-free strategy ($z=4.0262$, $p=0.0028$) and that this effect was significantly larger for right compared to left hippocampal lesions ($z=4.5278$, $p=0.0001$).

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 ($F_{1,12.93}=0.1204$, $p=0.7340$) (see Supplementary Table 8). It should be noted that this analysis is based on many fewer trials than the sequential decision task analysis.
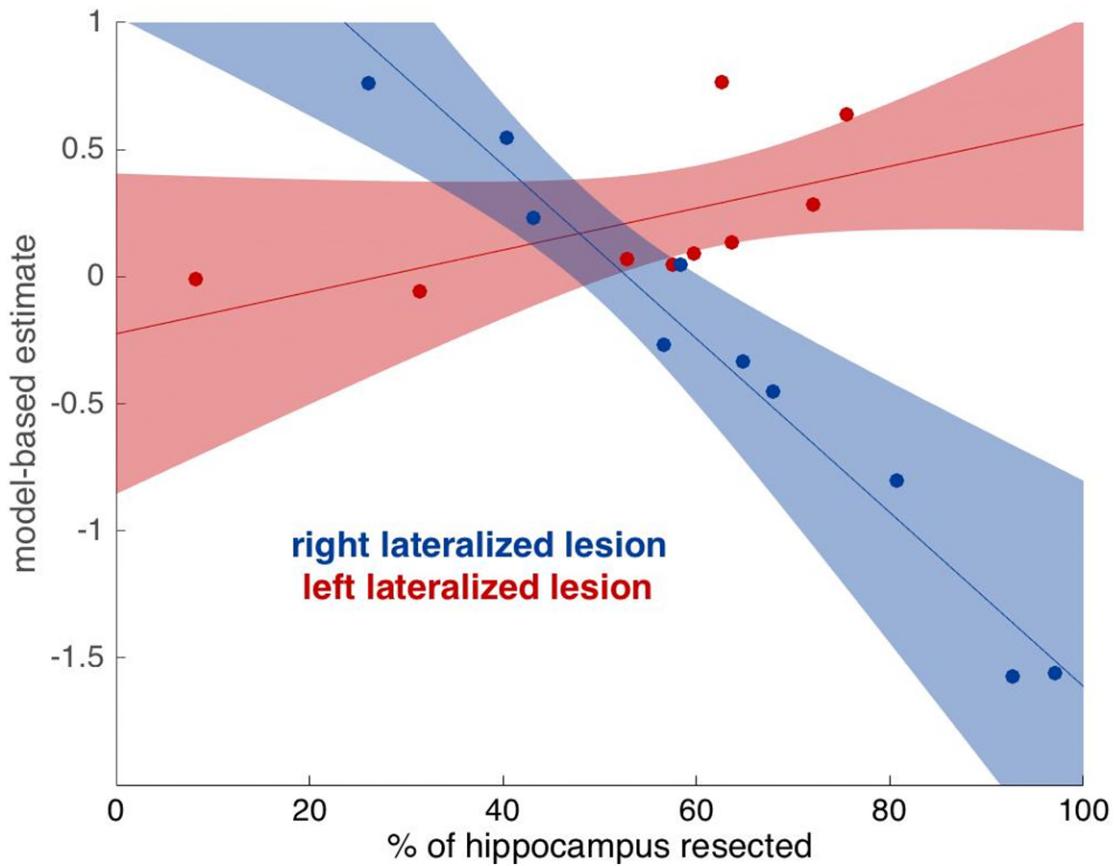
***Figure 9: Relationships between model-based planning and hippocampal lesion size.,
for patients with right and left lateralized ATL**. Estimated with a mixed-effects regressions,
controlling for IQ, age, and total size of lesion. Error bars indicate 80% confidence intervals.
There was a significant relationship in the right patient group (z=2.8305, p=0.005) but not in
the left patient group (z=1.0622, p=0.2882), with a significant difference between the right and
left patient groups (z=2.5076, p=0.0122).*

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 and 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 vs. 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 anterior temporal lobe is ultimately involved in allocentric spatial localization based on configurations of cues.

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 vs. 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 inactivations of the dorsal hippocampus, using an analogous two-stage Markov decision 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 decorrelation. 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 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, since 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 relationships 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, 2001, Davachi and Wagner, 2002, Kumaran et al., 2009, Schapiro et al., 2016, Boorman et al., 2016, Garvert et al., 2017) which indicate 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., 2017, 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.

Acknowledgements

References:

Adams, C. D. (1982). Variations in the sensitivity of instrumental responding to reinforcer

devaluation. *The Quarterly Journal of Experimental Psychology*, *34*(2), 77–98.

Adams, C. D., and Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *The Quarterly Journal of Experimental Psychology*, *33*(2), 109–121.

Addis, D. R., Cheng, T., P Roberts, R., & Schacter, D. L. (2011). Hippocampal contributions to the episodic simulation of specific and general future events. *Hippocampus*, *21*(10), 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 Computational Biology*, *11*(12), e1004648.

Balleine, B. W., and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, *37*(4), 407–419.

Bayer, H. M., and Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*(1), 129–141.

Boorman, E. D., Rajendran, V. G., O'Reilly, J. X., & Behrens, T. E. (2016). Two anatomically and computationally distinct learning signals predict changes to stimulus-outcome associations in hippocampus. *Neuron*, *89*(6), 1343-1354.

Brown, T. I., Carr, V. A., LaRocque, K. F., Favila, S. E., Gordon, A. M., Bowles, B., ... and Wagner, A. D. (2016). Prospective representation of navigational goals in the human hippocampus. *Science*, *352*(6291), 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. *Journal of Neuroscience*, *19*(24), 10869–10876.

Burgess, N., Maguire, E. A., and O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*(4), 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. *Current Biology*, *25*(1), 87–92.

Corbit, L. H., and Balleine, B. W. (2000). The role of the hippocampus in instrumental conditioning. *Journal of Neuroscience*, *20*(11), 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. *Journal of Neuroscience*, *22*(24), 10976–10984.

Davachi, L., and Wagner, A. D. (2002). Hippocampal contributions to episodic encoding: insights from relational and item-based learning. *Journal of Neurophysiology*, *88*(2), 982–990.

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*(6), 1204–1215.

Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704.

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. *Psychological Science*, *27*(6), 848–858.

Dickinson, A., and Balleine, B. (1993). Actions and responses: The dual psychology of behaviour.

Doeller, C. F., King, J. A., and Burgess, N. (2008). Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. *Proceedings of the National Academy of Sciences*, *105*(15), 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. *Nature Neuroscience*, *18*(5), 767–772.

Dusek, J. A., and Eichenbaum, H. (1997). The hippocampus and memory for orderly stimulus relations. *Proceedings of the National Academy of Sciences*, *94*(13), 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 Computational Biology*, *11*(9), e1004463.

Eichenbaum, H., and Cohen, N. J. (2004). *From conditioning to conscious recollection: Memory systems of the brain* (No. 35). Oxford University Press on Demand.

Eilan, N., McCarthy, R. A., and Brewer, B. (1993). Spatial representation: Problems in philosophy and psychology.

Foerde, K., and Shohamy, D. (2011). Feedback timing modulates brain systems for learning in humans. *Journal of Neuroscience*, *31*(37), 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*(5703), 1940-1943.

Garvert, M. M., Dolan, R. J., & Behrens, T. E. (2017). A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *Elife*, *6*.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003). Bayesian Data Analysis, (Chapman and Hall/CRC Texts in Statistical Science). Retrieved from http://www.citeulike.org/group/302/article/105949

Gelman, A., and Rubin, D. B. (1992). Inference from iterative simulation using multiple

sequences. *Statistical Science*, 457–472.

Gershman, S. J., Blei, D. M., and Niv, Y. (2010). Context, learning, and extinction.

Psychological Review, *117*(1), 197.

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*(4), 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*(5), 877–888.

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. (2007). Patients with hippocampal

amnesia cannot imagine new experiences. *Proceedings of the National Academy of

Sciences*, *104*(5), 1726-1731.

Heckers, S., Zalesak, M., Weiss, A. P., Ditman, T., and Titone, D. (2004). Hippocampal

activation during transitive inference in humans. *Hippocampus*, *14*(2), 153–162.

Hirsh, R. (1974). The hippocampus and contextual retrieval of information from memory: A

theory. *Behavioral biology*, *12*(4), 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. *Journal of Neuroscience*, *23*(13), 5945-5952.

Jenkinson, M., and Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143–156.

Johnson, A., and Redish, A. D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *Journal of Neuroscience*, *27*(45), 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 Biology*, *15*(1), e1002588.

Knowlton, B. J., Mangels, J. A., and Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science*, *273*(5280), 1399-1402.

Kruschke, J. K. (2010). Bayesian data analysis. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(5), 658–676.

Kumaran, D., Summerfield, J. J., Hassabis, D., & Maguire, E. A. (2009). Tracking the emergence of conceptual knowledge during human decision making. *Neuron*, *63*(6), 889-901.

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends in cognitive sciences*, *20*(7), 512-534.

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. *Journal of Neuroscience*, *36*(10), 2847–2856.

Mattar, M. G., and Daw, N. D. (2018). Prioritized memory access explains planning and hippocampal replay. *BioRxiv*, 225664. https://doi.org/10.1101/225664

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., … others. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *356*(1412), 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. *Behavioral and neural biology*, *61*(3), 260-270.

Miller, K. J., Botvinick, M. M., and Brody, C. D. (2017). Dorsal hippocampus plays a causal role in model-based planning. *BioRxiv*, 096594.

Morris, R. G. M., Garrud, P., Rawlins, J. A., and O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature*, *297*(5868), 681.

O'keefe, J., and Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: 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. *Neurobiology of Learning and Memory*, *65*(1), 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*(6706), 75.

Pfeiffer, B. E., and Foster, D. J. (2013). Hippocampal place cell sequences depict future paths to remembered goals. *Nature*, *497*(7447), 74.

Poldrack, R. A., and Packard, M. G. (2003). Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia*, *41*(3), 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*(1), 3–8.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599.

Scoville, W. B., and Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, *20*(1), 11.

Shohamy, D., and Wagner, A. D. (2008). Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron*, *60*(2), 378–389.

Shohamy, D., and Daw, N. D. (2015). Integrating memories to guide decisions. *Current Opinion in Behavioral Sciences*, *5*, 85-90.

Simon, D. A., and Daw, N. D. (2011). Neural correlates of forward planning in a spatial decision task in humans. *Journal of Neuroscience*, *31*(14), 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. *Behavioral Neuroscience*, *100*(5), 729.

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*(12), 2476–2489.

Spiers, H. J., and Maguire, E. A. (2007). A navigational guidance system in the human brain. *Hippocampus*, *17*(8), 618–626.

Squire, L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological review*, *99*(2), 195.

Stachenfeld, K. L., Botvinick, M., and Gershman, S. J. (2014). Design principles of the hippocampal cognitive map. In *Advances in neural information processing systems* (pp. 2528-2536).

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189.

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. *Behavioural Brain Research*, *92*(1), 21–30.

Viard, A., Doeller, C. F., Hartley, T., Bird, C. M., and Burgess, N. (2011). Anterior hippocampus and goal-directed spatial decision making. *Journal of Neuroscience*, *31*(12), 4613–4621.

Wimmer, G. E., and Shohamy, D. (2012). Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science*, *338*(6104), 270-273.

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*(3), 427-435.

Wikenheiser, A. M., and Redish, A. D. (2015). Hippocampal theta sequences reflect current goals. *Nature neuroscience*, *18*(2), 289.

Methods

Participants

The patient sample consisted of 19 individuals who had undergone unilateral anterior temporal lobectomy (ATL) for the treatment of intractable epilepsy. Patients were recruited from the New York University (NYU) Patient Registry for the Study of Perception, Emotion and Cognition (PROSPEC). The control sample consisted of 19 healthy controls (HC), who were recruited from the local community through internet-based advertisement and gave consent to participate in the study.

A clinical neuropsychologist (MRM or KB) conducted all standardized procedures for screening patients for inclusion into NYU PROSPEC. Patients were only selected for inclusion 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.

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.

MRI Scanning and Image Processing

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.

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; http://fsl.fmrib.ox.ac.uk/fsl) (Jenkinson and Smith, 2001). This consisted of a two-step procedure: First, using MRIcron (http://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 neuroradiologists were consulted.

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 (see 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 $p>.5$.

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 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 grey 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 seconds 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 seconds, 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 (See Figure 2 Left). 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 (See Figure 4 Right). The original study, using the same procedure during fMRI in healthy participants, showed that place and response 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.

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 (See 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 (See 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.

Sequential-Decision Making Task

Participants also 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). Each trial involved two choices in succession, followed by reward (See 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=.7) transitioned to the purple planet, and otherwise made a rare transition (p=.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 (vs 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 seconds 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.

Sequential-Decision Making Task 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 vs. 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 *fitglme* 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 vs. rare). We further interacted the task factors with experimental group (lesion vs. 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 vs rare, group, and their interactions) with participant-specific average boundary distance error (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 regressing defining those decision effects, accounts properly for uncertainty in the per-participant estimates. We also calculated a ratio dB/(dL+dB), where dL and dB where 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.

Supplementary Methods for 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 250 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, Carlin, Stern, and Rubin, 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 $\beta^{MF} = \frac{\beta^{MF0}}{\alpha} + \beta^{MF1}$, 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: $\alpha, \beta^{MB}, \beta^{MF0}, \beta^{MF1}, \beta^{stick}$, 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

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 ("hyperpriors") 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 performance 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 estimates of place memory (dB predicted the participant-specific parameter estimates. dB was extracted identically as previously. This model also included individual IQ scores as a covariate.

Supplementary Results from Spatial Task:

We compared the behavior of patients and controls in the spatial memory task, which probes the extent to which they can learn locations via boundary or landmark cues. During the first block the landmark has not yet moved and it is therefore not possible to distinguish which cue type is used (See Supplementary Figure 1). This is because boundary cues and landmark cues (fixed to one another and to each object) predict the same position, so distance error with respect to either cue type is the same.

Mean distance error across the whole first block was not quite significantly greater for patients during this first block ($F_{1,38}$= 4.0279, p=0.0519). Note that the low distance error already from object presentation 1 reflects the fact that some learning had already taken place, since the location of each object was shown once to participants before they made their first response.
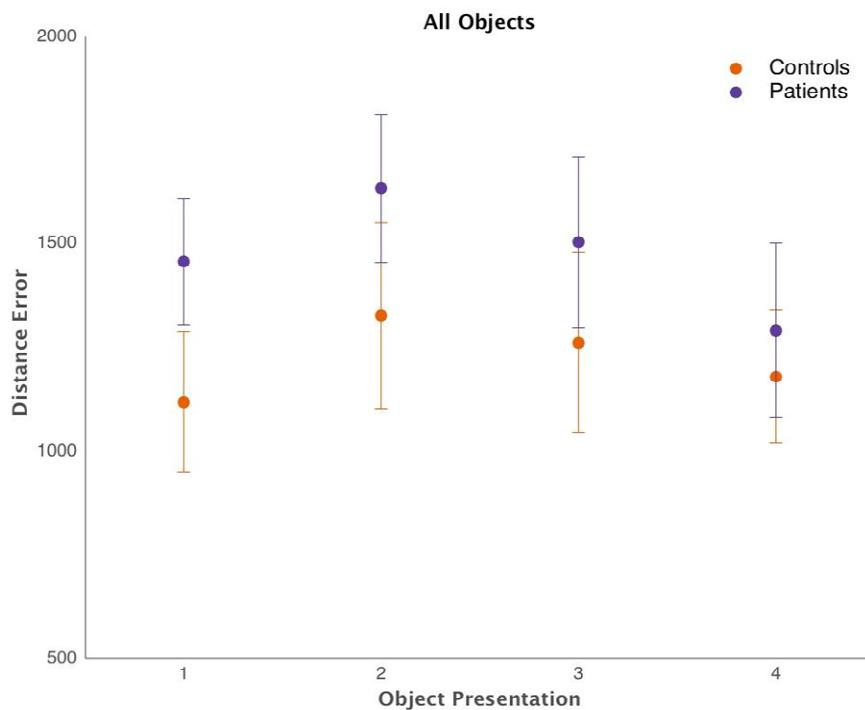
Supplementary Figure 1: Distance error (arbitrary units) for all objects during the first 4 presentations that constitute block. *Error bars indicate standard error.*

Distance error for all trials in blocks 2-4 can be seen in Supplementary Figure 2, now broken down with respect to error relative to either cue type's position in the previous block. Following object presentation 4, 8 and 12 (i.e. after block 1,2 and 3) the landmark moved in relation to boundaries of the arena. This movement allows us to establish to what extent object locations

were learnt in relation to boundary cues or landmark cues on the most recent block. The distinguishability should be the greatest on the trials immediately following the movement of the landmark. Performance differences are not necessarily expected on the remaining trials, as performance based on the object's most recent location, with respect to either cue, would there be equivalent. Even so, when calculating the boundary and landmark distance error for each object on all trials in block 2-4, based on location in the most recent block, we still see group by distance error type interaction on mean distance error ($F_{1, 37.7} = 4.1524$, $p = 0.0486$) (see Supplementary Table 11).

Similarly, we expected that attempting to break down performance according to the true object type (landmark or boundary) would limit our ability to see effects because subjects could only have understood this distinction, if at all, in the latter portion of the experiment. Accordingly, in a regression model where we, for the critical trials following the landmark moves, additionally interacted group and distance error type by object type (see Supplementary Table 12) we did find a two-way interaction of object type by eror type (F, p), indicating that subjects 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 ($F_{1, 782.1} = 0. 2.6427$, $p = 0.1044$),indicating that we did not detect no group differences in this respect.
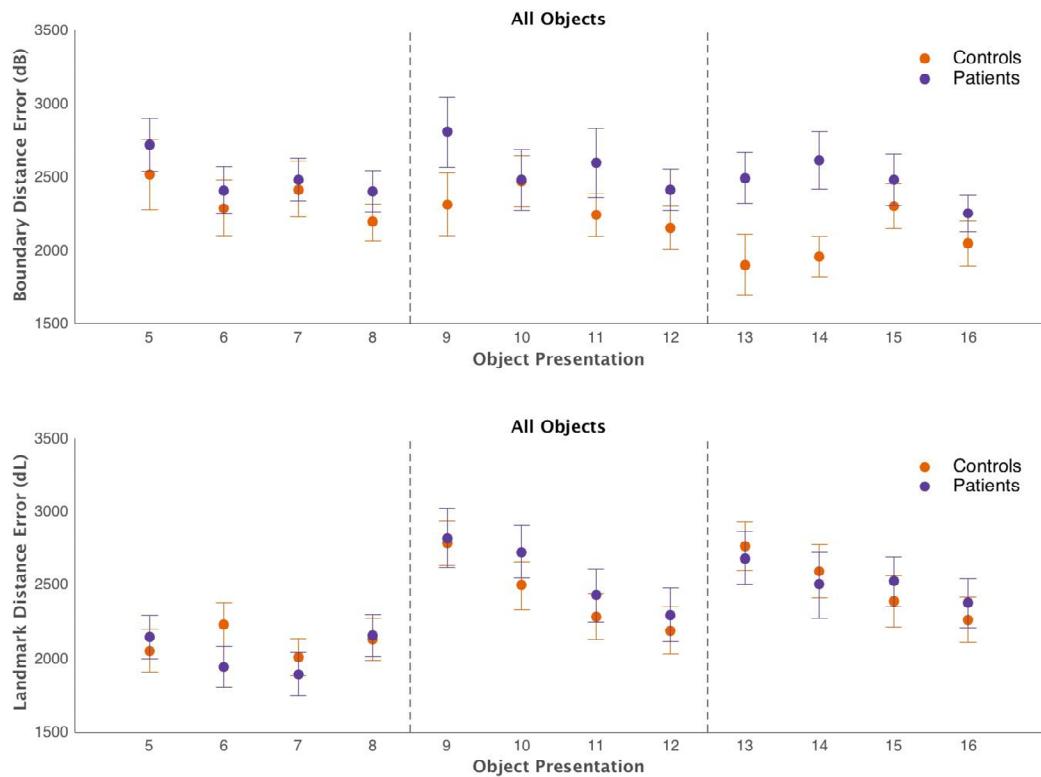
Supplementary Figure 2: Boundary (dB) and response (dL) distance error (arbitrary units) for all objects on trials across all object presentations in blocks 2-3. *Dotted lines separate blocks. Error bars indicate standard error.*

Note that our main result it Figure 5 were calculated as the mean of distance error from object presentations 5, 9 and 13 in Supplementary Figure 2. These trials were selected as in Doeller et al. (2008) since these trials cleanly dissociate performance based on what type of cue

Supplementary Result Tables:

Term	F	DF1	DF2	pValue	Estimate	SE
'zIQ:dB'	15.29	1	39.42	0	-369.1	94.42
'zAge:dB'	3.455	1	39.47	0.071	176.5	94.97
'Controls:dB'	275.3	1	39.48	0	2219	133.7
'Patients:dB'	406.2	1	39.51	0	2695	133.7
'zIQ:dL'	1.425	1	85.45	0.236	-86.58	72.52
'zAge:dL'	1.755	1	85.55	0.189	-96.65	72.95
'Controls:dL'	615.8	1	85.56	0	2549	102.7
'Patients:dL'	615.8	1	85.6	0	2549	102.7

Supplementary Table 1: Results from linear mixed regression model, with participant as random factor, implemented using Matlab function fitglme. Trials included are all critical trials following the movement of the landmark. Dependent variable is Distance Error. Age and IQ are z-scored. Effects seen for Patient group and Control group separately.

Term	z	pValue	Estimate	SE
'Patients'	7.57	0	1.635	0.2161
'Controls'	7.13	0	1.533	0.2152
'Reward:zAge'	1.55	0.121	0.0778	0.0502
'Rare:zAge'	0.96	0.34	0.0515	0.0539
'Reward:zIQ'	0.45	0.655	0.0234	0.0523
'Rare:zIQ'	0.16	0.872	-0.0089	0.0555
'Reward:Patients'	6.05	0	0.5114	0.0845
'Rare:Patients'	1.01	0.313	-0.079	0.0783
'Reward:Controls'	4.41	0	0.3664	0.0831
'Rare:Controls'	0.55	0.579	-0.0426	0.0767
'Reward:Rare:zAge'	1.45	0.147	0.0828	0.0571
'Reward:Rare:zIQ'	0.99	0.322	-0.0582	0.0587
'Reward:Rare:Patients'	1.35	0.176	-0.1158	0.0856
'Reward:Rare:Controls'	3.32	0.001	-0.2794	0.0841

Supplementary Table 2: Results from binomial mixed regression model, with participant as random factor, implemented using Matlab function fitglme. Dependent variable is whether participant stay with the first same first level action as on the last trial. Rare transition on last trial coded as 1, common as -1. Rewards on last trials coded as 1, no reward coded as -1. Age and IQ are z-scored. Effects seen for Patient group and Control group separately.

Term	z	pValue	Estimate	SE
'Controls'	6.74	0	1.505	0.2234
'Reward:zIQ'	1.04	0.296	0.065	0.0623
'Rare:zIQ'	0.62	0.533	0.042	0.0675
' dB:Patients'	0.57	0.569	-0.1413	0.2482
'Reward:Patients'	5.09	0	0.4601	0.0905
'Rare:Patients'	0.71	0.479	-0.0579	0.0817
' dB:Controls'	0.62	0.538	-0.1324	0.2151
'Reward:Controls'	4.44	0	0.3964	0.0892
'Rare:Controls'	0.13	0.898	0.0104	0.0809
'Reward:Rare:zIQ'	0.58	0.564	0.0409	0.0708
' dB:Reward:Patients'	1.29	0.196	0.132	0.1021
' dB:Rare:Patients'	1.05	0.293	-0.098	0.0933
'Reward:Rare:Patients'	1.46	0.145	-0.1271	0.0872
' dB:Reward:Controls'	0.19	0.848	0.0195	0.1017
' dB:Rare:Controls'	1.56	0.119	0.1516	0.0973
'Reward:Rare:Controls'	2.12	0.034	-0.1834	0.0866
'dB:Reward:Rare:Patients'	0.18	0.859	-0.0176	0.0991
'dB:Reward:Rare:Controls'	2.66	0.008	0.2752	0.1036

Supplementary Table 3: Results from binomial mixed regression model, with participant as random factor, implemented using Matlab function fitglme. Dependent variable is whether participant stay with the first same first level action as on the last trial. Rare transition on last trial coded as 1, common as -1. Rewards on last trials coded as 1, no reward coded as -1. Controlling for z-scored IQ. dB is z-scored participant wise place memory performance. Effects seen for Patient group and Control group separately.

Term	F	DF1	DF2	pValue	Estimate	SE
'zIQ:dB'	15.21	1	39.36	0	-368.4	94.44
'zAge:dB'	3.232	1	39.57	0.08	173.2	96.36
'Controls:dB	275.7	1	39.4	0	2219	133.7
'Left:dB'	192.7	1	41.36	0	2724	196.2
'Right:dB'	208.3	1	38.11	0	2668	184.9
'zIQ:dL'	1.35	1	88.62	0.248	-83.97	72.27
'zAge:dL'	2.114	1	89.27	0.15	-107.3	73.8
'Controls:dL'	621.7	1	88.67	0	2550	102.3
'Left:dL'	304	1	95.42	0	2640	151.4
'Right:dL'	308.5	1	84.24	0	2471	140.7

Supplementary Table 4: Results from linear mixed regression model, with participant as random factor, implemented using Matlab function fitglme. Dependent variable is Distance Error. Age and IQ are z-scored. Effects seen for Left patient group, Right patient group and Control group separately, each interacted with distance error types dL and dB seperately.

Term	z	pValue	Estimate	SE
'Controls'	7.22	0	1.5300	0.2120
'Left'	4.77	0	1.4760	0.3095
'Right'	6.04	0	1.7720	0.2935
'Reward:zAge'	1.58	0.1140	0.0802	0.0507
'Rare:zAge'	1.19	0.2340	0.0622	0.0522
'Reward:zIQ'	0.48	0.6300	0.0254	0.0526
'Rare:zIQ'	0.15	0.8840	-0.0079	0.0539
'Reward:Controls'	4.43	0	0.3634	0.0821
'Rare:Controls'	0.66	0.5070	-0.0490	0.0739
'Reward:Left'	3.69	0	0.4409	0.1193
'Rare:Left'	1.96	0.0500	-0.2102	0.1072
'Reward:Right'	4.84	0	0.5682	0.1175
'Rare:Right'	0.31	0.7530	0.0336	0.1069
'Reward:Rare:zAge'	1.59	0.1120	0.0892	0.0562
'Reward:Rare:zIQ'	0.96	0.3360	-0.0554	0.0576
'Reward:Rare:Controls'	3.38	0.001	-0.2810	0.0832
'Reward:Rare:Left'	1.48	0.1400	-0.1789	0.1211
'Reward:Rare:Right'	0.55	0.5840	-0.0653	0.1192

Supplementary Table 5: Results from binomial mixed regression model, with participant as random factor, implemented using Matlab function fitglme. Dependent variable is whether participant stay with the first same first level action as on the last trial. Rare transition on last trial coded as 1, common as -1. Rewards on last trials coded as 1, no reward coded as -1. Age and IQ are z-scored. Effects seen for Left Patient group, Right Patient group and Control group separately.

Term	z	pValue	Estimate	SE
'Controls'	6.81	0	1.4980	0.2201

Term	z	pValue	Estimate	SE
'Left'	4.4	0	1.5000	0.3405
'Right'	6.02	0	1.8080	0.3002
'Reward:zIQ'	0.42	0.6720	0.0276	0.0651
'Rare:zIQ'	0.29	0.7730	0.0201	0.0697
'Reward:Controls'	4.38	0	0.3763	0.0860
'Rare:Controls'	0.04	0.9660	-0.0033	0.0778
'Reward:Left'	2.41	0.0160	0.3154	0.1309
'Rare:Left'	1.69	0.0900	-0.1993	0.1177
'Reward:Right'	4.67	0	0.5586	0.1196
'Rare:Right'	0.41	0.6810	0.0446	0.1086
'Controls:zdB'	0.66	0.5090	-0.1400	0.2119
'Left:zdB'	0.15	0.8790	-0.0523	0.3448
'Right:zdB'	0.46	0.6470	-0.1631	0.3559
'Reward:Rare:zIQ'	0.03	0.9760	-0.0022	0.0728
'Reward:Rare:Controls'	2.42	0.0160	-0.2034	0.0841
'Reward:Rare:Left'	1.84	0.0660	-0.2340	0.1275
'Reward:Rare:Right'	0.5	0.6160	-0.0587	0.1168
'Reward:Controls:zdB'	0.18	0.8560	-0.0182	0.0997
'Rare:Controls:zdB'	1.32	0.1870	0.1258	0.0954
'Reward:Left:zdB'	2.03	0.0420	0.2899	0.1429
'Rare:Left:zdB'	0.01	0.9890	0.0018	0.1304
'Reward:Right:zdB'	0.02	0.9810	-0.0036	0.1468
'Rare:Right:zdB'	1.11	0.2680	-0.1500	0.1355
'Reward:Rare:Controls:zdB'	2.29	0.0220	0.2332	0.1019
'Reward:Rare:Left:zdB'	1.01	0.3150	0.1402	0.1394
'Reward:Rare:Right:zdB'	1.17	0.2430	-0.1698	0.1453

Supplementary Table 6: Results from binomial mixed regression model, with participant as random factor, implemented using Matlab function fitglme. Dependent variable is whether participant stay with the first same first level action as on the last trial. Rare transition on last trial coded as 1, common as -1. Rewards on last trials coded as 1, no reward coded as -1. IQ and participant-wise dB scores are z-scored. Effects seen for Left Patient group, Right Patient group and Control group separately.

Term	z	pValue	Estimate	SE
'Left'	5.0003	0	1.5818	0.3163
'Right'	5.7589	0	1.9119	0.332

Term		F		pValue	Estimate	SE
'Reward:zAge'		2.7688		0.006	0.3278	0.1184
'Rare:zAge'		0.5229		0.601	-0.068	0.1301
'Reward:zIQ'		0.03		0.976	0.0026	0.0844
'Rare:zIQ'		1.8214		0.069	0.1629	0.0894
'Reward:Left'		3.8632		0	0.4548	0.1177
'Rare:Left'		1.9511		0.051	-0.2516	0.1289
'Reward:Right'		5.5357		0	0.9101	0.1644
'Rare:Right'		0.4984		0.618	0.0886	0.1778
'Left:HippLesion'		1.3754		0.169	0.4727	0.3437
'Right:HippLesion'		1.4965		0.135	-0.6947	0.4642
'Left:TotalLesion'		0.0424		0.967	0.0151	0.3584
'Right:TotalLesion'		0.3306		0.741	0.1903	0.5757
'Reward:Rare:zAge'		1.7328		0.083	0.2477	0.143
'Reward:Rare:zIQ'		1.2396		0.215	0.125	0.1008
'Reward:Rare:Left'		1.9357		0.053	-0.2573	0.1329
'Reward:Rare:Right'		1.0946		0.274	0.2051	0.1874
'Reward:Left:HippLesion'		1.8877		0.059	-0.2651	0.1404
'Rare:Left:HippLesion'		1.4744		0.14	0.2301	0.156
'Reward:Right:HippLesion'		0.4876		0.626	0.1115	0.2287
'Rare:Right:HippLesion'		0.6427		0.52	0.1597	0.2485
'Reward:Left:TotalLesion'		1.5294		0.126	0.235	0.1536
'Rare:Left:TotalLesion'		1.7828		0.075	-0.3011	0.1689
'Reward:Right:TotalLesion'		1.6206		0.105	-0.4798	0.2961

Supplementary Table 7: Results from binomial mixed regression model, with participant as random factor, implemented using Matlab function fitglme. Dependent variable is whether participant stay with the first same first level action as on the last trial. Rare transition on last trial coded as 1, common as -1. Rewards on last trials coded as 1, no reward coded as -1. Age, IQ Total Lesion size and Hippocampal Lesion size are z-scored. Effects seen for Left Patient group and Right Patient group separately.

Term	F	DF1	DF2	pValue	Estimate	SE
'zIQ'	2.7923	1	216	0.096	0.1305	0.0781
'zAge'	0.7262	1	216	0.395	-0.0918	0.1077
'Left'	1.5152	1	216	0.220	0.1429	0.1161
'Right'	0.0236	1	216	0.878	-0.0212	0.1380

'Left:HippLesion'	0.0888	1	216	0.766	0.0468	0.1572
'Right:HippLesion'	1.7341	1	216	0.189	-0.2611	0.1983
'Left:TotalLesion'	4.3513	1	216	0.038	0.3157	0.1514
'Right:TotalLesion'	0.8002	1	216	0.372	0.2263	0.2530

Supplementary Table 8: Results from linear mixed regression model, with participant as random factor, implemented using Matlab function fitglme. Dependent variable is Boundary Distance Error. Age, IQ Total Lesion size and Hippocampal Lesion size are z-scored. Effects seen for Left Patient group and Right Patient group separately.

Term	F	DF1	DF2	pValue	Estimate	SE
'(Intercept)'	939.9385	1	432	0	0.4586	0.0150
'zIQ'	0.5339	1	432	0.4650	-0.0128	0.0175
'zAge'	1.7860	1	432	0.1820	0.0218	0.0163
'Patients'	8.2130	1	432	0.0040	0.0608	0.0212

Supplementary Table 9: Results from linear mixed regression model, with participant as random factor, implemented using Matlab function fitglme. Trials included are all critical trials following the movement of the landmark. Dependent variable is ratio dB/(dB+dL). Age and IQ are z-scored. Effect of Patient is the difference to control group.

Term	FStat	pValue	Estimate	SE
'Controls'	47.411	5.75E-12	1.4836	0.2155
'Left'	15.792	7.07E-05	1.3798	0.3472
'Right'	30.443	3.44E-08	1.7744	0.3216
'Reward:zIQ'	1.2543	0.26273	0.0692	0.0618
'Rare:zIQ'	3.20E-05	0.99549	0.0004	0.0638
'ratio:Controls'	0.72676	0.39394	1.8831	2.2089
'Reward:Controls'	21.275	3.98E-06	0.3864	0.0838

Term	F	pValue	Estimate	SE
'Rare:Controls'	0.20918	0.64741	-0.0347	0.0759
'ratio:Left'	0.4084	0.52278	-4.1843	6.5476
'Reward:Left'	4.716	0.029883	0.2910	0.1340
'Rare:Left'	4.3656	0.036671	-0.2524	0.1208
'ratio:Right'	0.001657	0.96753	0.2433	5.9776
'Reward:Right'	18.587	1.62E-05	0.5550	0.1287
'Rare:Right'	0.14136	0.70693	0.0441	0.1173
'Reward:Rare:zIQ'	0.00029568	0.98628	0.0011	0.0642
'ratio:Reward:Controls'	0.45995	0.49765	-0.6826	1.0065
'ratio:Rare:Controls'	0.099326	0.75264	-0.2978	0.9448
'Reward:Rare:Controls'	8.2962	0.0039729	-0.2320	0.0806
'ratio:Reward:Left'	4.7003	0.030158	-6.1804	2.8507
'ratio:Rare:Left'	0.77308	0.37927	-2.3035	2.6198
'Reward:Rare:Left'	5.3569	0.02064	-0.2967	0.1282
'ratio:Reward:Right'	0.076529	0.78206	0.6662	2.4083
'ratio:Rare:Right'	0.37175	0.54205	1.3451	2.2061
'Reward:Rare:Right'	0.203	0.65231	-0.0560	0.1243
'ratio:Reward:Rare:Controls'	3.3529	0.067086	-1.8130	0.9901
'ratio:Reward:Rare:Left'	3.2995	0.0693	-5.0096	2.7579
'ratio:Reward:Rare:Right'	0.6052	0.4366	1.8149	2.3329

Supplementary Table 10: Results from binomial mixed regression model, with participant as random factor, implemented using Matlab function fitglme. Dependent variable is whether participant stay with the first same first level action as on the last trial. Rare transition on last trial coded as 1, common as -1. Rewards on last trials coded as 1, no reward coded as -1. These are interacted by the participant specific estimate of dB/(dB+dL). Effects seen for Left Patient group, Right Patient group and Control group separately.

Term	F	DF1	DF2	pValue	Estimate	SE
'zIQ:dB'	1.6794	1	37.5803	0.203	-162.8533	125.6656
'zAge:dB'	0.0082	1	37.9909	0.928	10.466	115.5669
'Patients:dB'	560.5218	1	37.7903	0	2516.7	106.3016
'Controls:dB'	450.8038	1	37.8373	0	2236.8	105.348
'zIQ:dL'	0.1111	1	37.473	0.741	26.8459	80.545
'zAge:dL'	0.8551	1	38.3598	0.361	-68.7394	74.3367
'Patients:dL'	1204.1	1	37.9252	0	2368	68.2585

'Controls:dL'	1206.7	1	38.0391	0	2350.9	67.676

Supplementary Table 11: Results from linear mixed regression model, with participant as random factor, implemented using Matlab function fitglme. Trials included are all trials in blocks 2-4. Dependent variable is Distance Error. Age and IQ are z-scored. Effects seen for Patient group and Control group, each interacted with distance error types dL and dB separately

Term	F	DF1	DF2	pValue	Estimate	SE
'zIQ:dL'	1.1611	1	104.5756	0.2840	-74.4111	69.0562
'zAge:dL'	1.8066	1	105.4405	0.1820	-93.2646	69.3873
'Controls:dL'	416.4943	1	75.2530	0	2.9607e+03	145.0740
'Patients:dL'	425.7442	1	75.3192	0	2.9934e+03	145.0733
'zIQ:dB'	13.0348	1	45.6346	1.0000e-03	-332.5745	92.1164
'zAge:dB'	4.0539	1	45.5598	0.0500	186.2976	92.5279
'Controls:dB'	152.4128	1	39.9651	0	2.2115e+03	179.1342
'Patients:dB'	183.1393	1	39.9833	0	2.4239e+03	179.1084
'Controls:dL:L_Object'	18.7613	1	156.0378	0	-817.8130	188.8092
'Patients:dL:L_Object'	21.8280	1	156.0260	0	-882.1264	188.8096
'Controls:dB:L_Object'	0.0039	1	173.4617	0.9500	11.6564	186.8515
'Patients:dB:L_Object'	8.2023	1	173.6667	0.0050	535.1180	186.8446

Supplementary Table 12: Results from linear mixed regression model, with participant as random factor, implemented using Matlab function fitglme. Trials included are all trials in blocks critiacal trials following landmark movements. Dependent variable is Distance Error. Age and IQ are z-scored. Effects seen for Patient group and Control group, each interacted with distance error types dL and dB separately. There variables are additionally interacted with a variable L_object which is coded 1 for landmark objects and 0 for boundary objects.

	mean	2.5%	97.5%
β^{MB}_{cont}	0.33	0.1	0.55
β^{MF0}_{cont}	-0.06	-0.14	0.01
β^{MF1}_{cont}	0.30	0.15	0.46
β^{stick}_{cont}	1.16	0.78	1.54
β^{2stage}_{cont}	1.17	0.85	1.5
α_{cont}	0.26	0.18	0.38
$\beta^{MB}_{pat\_diff}$	-0.2	-0.52	0.13
$\beta^{MF0}_{pat\_diff}$	0.05	-0.05	0.15
$\beta^{MF1}_{pat\_diff}$	0.2	-0.02	0.42
$\beta^{stick}_{pat\_diff}$	0.25	-0.29	0.79
$\beta^{2stage}_{pat\_diff}$	-0.04	-0.49	0.41

α_{pat_diff} 0.07   -0.06   0.22

Supplementary Table 13: Estimates (mean, 2.5% confidence interval and 97.5% confidence interval) from the full RL model which includes age and IQ as covariates but not individual place-memory performance estimates. Cont indicated free-parameter estimates of the control group. Pat_diff indicates how patient group free-parameter estimates differ from control group.

	mean	2.5%	97.5%
$\beta^{MB\text{-}dB}_{cont}$	-0.27	-0.52	-0.02
$\beta^{MF0\text{-}dB}_{cont}$	0.00	-0.08	0.07
$\beta^{MF1\text{-}dB}_{cont}$	0.12	-0.07	0.31
$\beta^{stick\text{-}dB}_{cont}$	-0.13	-0.59	0.32
$\beta^{2stage\text{-}dB}_{cont}$	-0.38	-0.73	-0.03
$\alpha\text{-}dB_{cont}$	0.24	0.15	0.34
$\beta^{MB\text{-}dB}_{pat\_diff}$	0.29	-0.05	0.63
$\beta^{MF0\text{-}dB}_{pat\_diff}$	-0.03	-0.15	0.09
$\beta^{MF1\text{-}dB}_{pat\_diff}$	-0.05	-0.31	0.21
$\beta^{stick\text{-}dB}_{pat\_diff}$	0.06	-0.54	0.66
$\beta^{2stage\text{-}dB}_{pat\_diff}$	0.51	0.03	0.97
$\alpha\text{-}dB_{pat\_diff}$	0.06	-0.06	0.22

Supplementary Table 14: Estimates (mean, 2.5% confidence interval and 97.5% confidence interval) from the full RL model which includes IQ and individual place memory performance estimates as covariates. Cont indicates effect of place-memory (dB) on free-parameter estimates of the control group. Pat_diff indicates how patient group effect of place-memory on free parameter estimates differ from control group.