

# Unraveling the Role of the Hippocampus in Reversal Learning

Adrià Vilà-Balló,<sup>1,2,3\*</sup> Ernest Mas-Herrero,<sup>1,2,4\*</sup> Pablo Ripollés,<sup>1,2,5</sup> Marta Simó,<sup>1,6</sup> Júlia Miró,<sup>1,7</sup> David Cucurell,<sup>1</sup> Diana López-Barroso,<sup>1</sup> Montserrat Juncadella,<sup>7</sup>  Josep Marco-Pallarés,<sup>1,2</sup> Mercè Falip,<sup>7</sup> and  Antoni Rodríguez-Fornells<sup>1,2,8</sup>

<sup>1</sup>Cognition and Brain Plasticity Group (Bellvitge Biomedical Research Institute-IDIBELL) and <sup>2</sup>Department of Basic Psychology, Campus Bellvitge, University of Barcelona, L'Hospitalet de Llobregat, Barcelona 08907, Spain, <sup>3</sup>Department of Psychology, Faculty of Education and Psychology, University of Girona, Girona 17071, Spain, <sup>4</sup>Montreal Neurological Institute, McGill University, Montreal, Quebec H3A 2B4, Canada, <sup>5</sup>Poeppel Laboratory, Department of Psychology, New York University, New York, New York 10003, <sup>6</sup>Neuro-Oncology Unit, Hospital Universitari de Bellvitge-Hospital Duran i Reynals (Institut Català d'Oncologia), L'Hospitalet de Llobregat, Barcelona 08907, Spain, <sup>7</sup>Epilepsy Unit, Neurological Service, Hospital Universitari de Bellvitge, L'Hospitalet de Llobregat, Barcelona 08907, Spain, and <sup>8</sup>Catalan Institution for Research and Advanced Studies, Barcelona 08907, Spain

Research in reversal learning has mainly focused on the functional role of dopamine and striatal structures in driving behavior on the basis of classic reinforcement learning mechanisms. However, recent evidence indicates that, beyond classic reinforcement learning adaptations, individuals may also learn the inherent task structure and anticipate the occurrence of reversals. A candidate structure to support such task representation is the hippocampus, which might create a flexible representation of the environment that can be adaptively applied to goal-directed behavior. To investigate the functional role of the hippocampus in the implementation of anticipatory strategies in reversal learning, we first studied, in 20 healthy individuals (11 women), whether the gray matter anatomy and volume of the hippocampus were related to anticipatory strategies in a reversal learning task. Second, we tested 20 refractory temporal lobe epileptic patients (11 women) with unilateral hippocampal sclerosis, who served as a hippocampal lesion model. Our results indicate that healthy participants were able to learn the task structure and use it to guide their behavior and optimize their performance. Participants' ability to adopt anticipatory strategies correlated with the gray matter volume of the hippocampus. In contrast, hippocampal patients were unable to grasp the higher-order structure of the task with the same success than controls. Present results indicate that the hippocampus is necessary to respond in an appropriately flexible manner to high-order environments, and disruptions in this structure can render behavior habitual and inflexible.

**Key words:** epilepsy; flexible behavior; hippocampus; model-based strategies; reinforcement learning; reversal learning

## Significance Statement

Understanding the neural substrates involved in reversal learning has provoked a great deal of interest in the last years. Studies with nonhuman primates have shown that, through repetition, individuals are able to anticipate the occurrence of reversals and, thus, adjust their behavior accordingly. The present investigation is devoted to know the role of the hippocampus in such strategies. Importantly, our findings evidence that the hippocampus is necessary to anticipate the occurrence of reversals, and disruptions in this structure can render behavior habitual and inflexible.

## Introduction

To accomplish their goals, humans adapt their behavior on the basis of learning stimulus–reward contingencies. A powerful ex-

planation for reward-guided learning is provided by reinforcement learning (RL) theories (e.g., Sutton and Barto, 1998; Daw et al., 2005). RL models achieve behavioral adaptation by reinforcing successful actions and devaluing erroneous or nonrewarded actions through the computation of prediction errors (Watkins

Received Oct. 15, 2016; revised April 25, 2017; accepted May 14, 2017.

Author contributions: D.C., J.M.-P., and A.R.-F. designed research; A.V.-B., P.R., J.M., D.L.-B., M.J., and M.F. performed research; D.C. and J.M.-P. contributed unpublished reagents/analytic tools; A.V.-B., E.M.-H., P.R., and M.S. analyzed data; A.V.-B., E.M.-H., P.R., and A.R.-F. wrote the paper.

This work was supported by Spanish Government Grants PSI2011-29219 and PSI2015-69178-P to A.R.-F. and Catalan Government (Generalitat de Catalunya) Grant 2009 SGR 93. A.V.-B. was supported by Predoctoral IDIBELL Grant 06/IDB-001. E.M.-H. was supported by FPI Program BES-2010-032702 and Montreal Neurological Institute Jeanne Timmins Costello Fellowship. P.R. was supported by FPU program AP2010-4179. We thank all the patients and control participants for great collaboration in the present project; and A. Suades for help with data collection.

The authors declare no competing financial interests.

\*A.V.-B. and E.M.-H. contributed equally to this work.

Correspondence should be addressed to either Dr. Adrià Vilà-Balló or Dr. Antoni Rodríguez-Fornells, Cognition and Brain Plasticity Unit (Campus Bellvitge), Department of Basic Psychology, University of Barcelona, L'Hospitalet de Llobregat, Feixa Larga, Barcelona 08907, Spain. E-mail: adriavilaballo@gmail.com or arfornells@gmail.com.

DOI:10.1523/JNEUROSCI.3212-16.2017

Copyright © 2017 the authors 0270-6474/17/376686-12\$15.00/0

and Dayan, 1992). However, the main limitation of these classic RL models is that they work without using any representation of the task structure, such as rules describing the relationship between actions, rewards, and external variables (e.g., time) (Sutton and Barto, 1998; Daw et al., 2011). Indeed, in real life, reward contingencies often change in a predictable manner. For instance, in natural settings, some types of food are only available in a particular period of the year. It would be advantageous to take this into account to anticipate the occurrence of such periods and then initiate and exploit foraging behavior only in this particular temporal context.

Previous studies using the reversal learning task have shown that humans (Hampton et al., 2006) and other animals (Huh et al., 2009; Costa et al., 2015; Jang et al., 2015) can learn such high-order structure and use this knowledge to improve their decision-making. During this type of task, participants are requested to learn specific stimulus–response associations to attain certain rewards (e.g., money). After a variable number of trials (e.g., 16–24), the stimulus–reward associative structure changes. Although participants cannot fully predict the occurrence of a reversal, there is an underlying structure that can be learned. Specifically, the rule never changes during the first trials after a reversal (in the example given, the first 15 trials), but the reversal likelihood increases after that. Studies with nonhuman primates have shown that, through repetition, individuals are able to identify these two temporal contexts, anticipating the occurrence of reversals and, thus, adjust their behavior accordingly (Costa et al., 2015; Jang et al., 2015). Importantly, this high-order structure of the environment is not captured by classic RL models and is acquired slower than stimulus–reward mappings, as it involves creating an internal model of the task structure (Costa et al., 2015).

One brain region that may be especially involved in learning such structures is the hippocampus. Existing evidence suggests that the hippocampus supports rapid encoding of multiple items, sequences, and events in relation to the temporal context (Allen et al., 2016) by creating a flexible representation of the environment (Howard and Kahana, 2002; Konkel and Cohen, 2009; Hsieh et al., 2014), which can be later used to guide performance (Eichenbaum, 2000; Squire et al., 2004; Aggleton et al., 2007). Thus, the hippocampus might be critical for maintaining the representations of the contingency structure of a task given a particular context, which can be then adaptively applied to anticipate potential changes in the environment (Bornstein and Daw, 2013).

To unravel the functional role of the hippocampus in the implementation of anticipatory strategies in reversal learning, we first studied whether the gray matter (GM) anatomy and volume of the hippocampus were related to participants' ability to implement strategies based on the predictability of reversals. Additionally, we included a group of refractory temporal lobe epileptic patients with unilateral hippocampal sclerosis (who served as a hippocampal lesion model), to test whether the use of these strategies was hippocampus-dependent. In that case, we hypothesized that lesions within the hippocampus should impair the use of strategies determined by the temporal context, leading behavior to be dependent on classic RL mechanisms.

## Materials and Methods

**Participants.** Initially, 20 healthy subjects were recruited for the first part of the study (11 women, 55 ± 51.0%; age: 42.1 ± 14.3 years; education: 11.6 ± 4.2 years). However, the final sample was reduced to 19 participants (10 women, 53 ± 51.3%; age: 40.7 ± 13.3 years; education: 11.8 ±

4.1 years). One subject was excluded because of poor performance (he never changed his choice during the reversal learning task). The temporal lobe epileptic group with unilateral sclerotic hippocampus (TLE-UHS) consisted of 20 patients (11 women, 55 ± 51.0%; age: 41.8 ± 12.3 years; education: 11.9 ± 4.0 years), 10 suffering from left and 10 from right medically refractory TLE. All patients were recruited after a presurgical evaluation at the Hospital of Bellvitge. Patient diagnoses were established according to clinical EEG and MRI data (Cendes, 2005; Malmgren and Thom, 2012; Tatum, 2012). All of the patients underwent a standardized neurological and neuropsychological examination, prolonged interictal and ictal video-EEG monitoring, and brain MRI assessed by an expert neurologist and a neuroradiologist. All patients were evaluated before and 3 months after an anterior temporal lobe resection for the relief of medically intractable TLE. The surgery, performed by the same neurosurgeon, consisted of *en bloc* resection of the mesiotemporal structures. The resection line extended 4.0–5.0 cm from the temporal pole in the nondominant hemisphere and 3.5–4.0 cm in the dominant hemisphere. Hippocampal sclerosis was confirmed in all patients with an anatomopathological study by the same pathologist. None of the patients suffered a seizure during the experimental task or 24 h before it, and all of the patients were on habitual antiepileptic drug regimens. In the current study, the chronic temporal lobe epileptic patients were matched in gender ( $\chi^2(1, N = 39) = 0.02, p = 0.88$ ), age ( $t_{(37)} = -0.25, p = 0.81$ ), and years of education ( $t_{(37)} = -0.01, p = 0.99$ ) with the healthy control group (demographic data are summarized in Tables 1, 2). The study was approved by the Ethical Committee of Bellvitge University Hospital. Written informed consent was obtained from all of the patients and controls. The experiment was performed in accordance with the guidelines of the Declaration of Helsinki and approved by the Ethical Committee of University Hospital of Bellvitge, Spain.

**Procedure.** Regarding behavioral acquisition, an initial reversal learning session was performed for each group. The neuropsychological battery was performed only one time for each group, and 1 d before the initial reversal learning session. The MRI acquisition session was held before the reversal learning session. To test the effects of removing unilateral medial-temporal lobe (MTL) on reversal learning, patients performed an additional session. This session (postsession) was undertaken a minimum of 6 months after the initial session, to avoid learning effects and at least 3 months after the surgery.

**Experimental design.** We used a modified version of the probabilistic reversal learning task of Jocham et al. (2009) (Fig. 1, task illustration). The experiment was divided into 63 blocks (pseudorandom order), with 16–24 trials each, which resulted in 1260 trials. In every trial, two gray squares, located in the middle of the screen on either side of a central fixation point, were presented over a black background for 1000 ms. All participants had to select one of the two gray squares by pressing the right or left button of the mouse. A feedback stimulus indicating a win (happy face) or loss (sad face) of 0.06€ appeared 700 ms after the response. It was presented in the middle of the screen for 800 ms. To avoid automatic responses, the intertrial stimulus was randomly set to 500–900 ms. If participants did not respond in the established 1000 ms after the presentation of the stimulus, a question mark appeared in the middle of the screen. Every 3 blocks, self-paced resting breaks were given between the seventh and 10th trial of each block to impede guessing of the rule change. During these breaks, the information about the amount of money accumulated to date was provided.

Participants were informed that the probability of winning (75% to win on one side, and 25% to win on the other) remained constant during some trials, but after an unspecified period, the probability reversed. On each trial, participants had to choose the stimulus leading to the reward outcome. Reward contingencies reversed (rule reversal) after a randomly jittered block length of 16–24 trials (rule change occurred always between the 16th and 24th trial), and participants had to switch their selection to the new rewarded alternative. The responses to the side with the highest probability of winning were considered as correct choices; in like manner, responses to the side with the lowest probability of losing were considered as failures. In the present study, the accuracy rate was calculated as the sum of correct responses divided by the total number of responses. Because of the established probabilities, participants had to

**Table 1. Demographic data for temporal lobe epilepsy (left and right) patients and controls included in the study**

Code	Group	Age (yr)	Gender	Education (yr)	Age at epilepsy onset	Type	Frequency	No. of epileptic drugs	Benzodiazepine and barbiturates
T01	TLE-R	47	M	16	10 mo	cpc	1–2/mo	4	No
T02	TLE-L	39	F	8	14 mo	cpc	1–2/mo	3	Clobazam 10 mg/d
T03	TLE-R	51	M	11	11 yr	cpc	1/mo	3	Clobazam 20 mg/d
T04	TLE-R	36	M	11	18 yr	cpc	4–6/mo	3	Phenobarbital 0–0–2
T05	TLE-R	50	F	8	18 yr	cpc	6–8/mo	3	Phenobarbital 1/–0–1
T06	TLE-L	65	F	0	4 yr	cpc	4–5/mo	2	Phenobarbital 0–0–2
T07	TLE-R	66	M	8	41 yr	cpc	4/mo	2	No
T08	TLE-L	33	F	11	16 yr	cpc	30–35/mo	3	Clobazam 10 mg/d
T09	TLE-L	34	M	16	23 yr	cpc	8–10/mo	2	No
T10	TLE-R	21	M	16	17 yr	cpc	5/mo	2	No
T11	TLE-L	37	F	12	13 mo	cpc	7–9/mo	2	No
T12	TLE-L	41	F	14	12 mo	cpc	5–6/mo	4	Phenobarbital 1–1–1
T13	TLE-R	34	F	14	10 yr	cpc	3/mo	2	No
T14	TLE-D	61	M	12	31 yr	cpc	4–5/mo	3	No
T15	TLE-L	29	M	14	15 yr	cpc	3–4/mo	3	Phenobarbital 1–1–2
T16	TLE-L	25	M	9	13 yr	cpc	1/mo	2	No
T17	TLE-L	48	F	14	32 yr	cpc	9–10/mo	3	Clobazam 25 mg/d
T18	TLE-R	34	F	12	21 yr	cpc	2/mo	2	No
T19	TLE-L	43	F	14	8 yr	cpc	18–20/mo	2	No
T20	TLE-L	35	F	17	2 yr	cpc	4–6/mo	2	No
C01	Control	49	M	15	—	—	—	—	—
C0 <sub>2</sub>	Control	42	F	10	—	—	—	—	—
C03	Control	48	M	11	—	—	—	—	—
C04	Control	39	M	10	—	—	—	—	—
C05	Control	28	F	11	—	—	—	—	—
C06	Control	35	F	16	—	—	—	—	—
C07	Control	53	F	8	—	—	—	—	—
C08 <sup>a</sup>	Control	68	F	6	—	—	—	—	—
C09	Control	71	M	0	—	—	—	—	—
C10	Control	25	F	17	—	—	—	—	—
C11	Control	35	M	10	—	—	—	—	—
C12	Control	25	M	17	—	—	—	—	—
C13	Control	43	F	10	—	—	—	—	—
C14	Control	43	F	10	—	—	—	—	—
C15	Control	29	F	18	—	—	—	—	—
C16	Control	61	M	10	—	—	—	—	—
C17	Control	28	M	12	—	—	—	—	—
C18	Control	21	M	13	—	—	—	—	—
C19	Control	51	F	12	—	—	—	—	—
C20	Control	48	F	15	—	—	—	—	—

<sup>a</sup>Control C08 was not included in the final analyses because of poor performance.

monitor the meaning of the different feedbacks and thus recognize when a correct choice was nonrewarded or when the negative feedback signaled the need for behavioral change because the rule had reversed (Fig. 1).

Importantly, participants were instructed to change the response pattern only when they were completely sure that the rule had changed. A brief training session was provided at the beginning of the encounter to ensure comprehension of the task. Participants were encouraged to win as much money as possible during the task.

**RL model.** A Q-learning model used by Watkins and Dayan (1992) was implemented. The model uses reward prediction error to update the weights associated with each stimulus and probabilistically chooses the stimulus with the higher weight. The weight is then updated using the following algorithm:

$$W(t) = W(t-1) + \alpha \times \delta \quad (1)$$

where  $\alpha$  is the learning rate and  $\delta$  represents the prediction errors, calculated as the difference between the outcome and the expectancy or weight of the selected figure. Softmax action selection was used to compute the probability of choosing one of the following two options:

$$P_A(t) = \frac{e^{\gamma W_A(t)}}{e^{\gamma W_A(t)} + e^{\gamma W_B(t)}} \quad (2)$$

where  $\gamma$  is the inverse of the temperature parameter.  $\gamma$  determines the degree to which differences in reward values between two potential actions are translated into a more deterministic choice.

The model was run 10 times, using random initial values for each subject by maximizing the log likelihood estimate with the `fmincon` function of MATLAB R2008 (The MathWorks). The parameters  $\alpha$  and  $\gamma$  with the best log likelihood estimate were selected. Once  $\alpha$  and  $\gamma$  were individually calculated, model predictions could be determined on a trial-by-trial basis.

**MRI data acquisition.** Controls and TLE-UHS patients underwent MRI whole-brain structural scans, including T1 weighted and FLAIR acquisition protocols. The high-resolution T1 weighted image (slice thickness = 1 mm; no gap; number of slices = 240; TR = 2300 ms; TE = 3 ms; matrix = 256 × 256; FOV = 244 mm; voxel size = 1 × 1 × 1 mm) was acquired with a 3.0 tesla Siemens Trio MRI system at the Hospital Clinic of Barcelona. The FLAIR image (slice thickness = 5.2 mm; no gap; number of slices = 19; TR = 7295 ms; TE = 12 ms; matrix = 256 × 256; FOV = 230 mm; voxel size = 0.89 × 0.89 × 5.2 mm) was acquired on a 1.5 Philips Intera scan at the University Hospital of Bellvitge. Both structural scans were assessed by an expert neurologist who found no structural abnormalities in addition to unilateral hippocampal sclerosis in TLE-UHS patients.

**Table 2. Demographic data for controls and TLE+UHS patients included in the study<sup>a</sup>**

	Controls		TLE+UHS		<i>t</i>	<i>p</i>
	Mean	SD	Mean	SD		
Age (yr)	40.7	13.3	41.8	12.3	−0.2	0.806
Education (yr)	11.8	4.1	11.9	4.0	0.0	0.995
Logical Memory I	35.2	13.7	28.8	8.8	1.7	0.097
Visual Reproduction I	89.6	19.0	75.4	19.3	2.3	0.026
Logical Memory II	21.7	10.3	15.6	7.2	2.2	0.036
Visual Reproduction II	78.4	22.3	52.0	24.4	3.5	0.001*
Digits Span	15.7	4.5	13.1	6.2	1.5	0.138
Letters and numbers	10.2	2.9	8.4	2.7	2.0	0.056
RAVLT total learning at trial 5	13.3	1.8	11.5	2.5	2.5	0.017
RAVLT immediate recall	12.1	3.0	8.9	3.2	3.2	0.003
RAVLT delayed recall	12.0	3.1	8.7	3.5	3.2	0.003
RAVLT recognition	13.8	2.2	12.8	2.4	1.4	0.168
Trial Making Test A	46.3	29.6	51.1	22.9	−0.6	0.579
Trial Making Test B	88.2	58.9	142.9	131.5	−1.6	0.112
Vocabulary	44.7	10.5	35.8	8.5	2.9	0.006*
Boston Naming Test	52.5	8.3	48.9	6.4	1.5	0.133
Semantic Fluency	20.7	5.5	17.3	4.3	2.2	0.035
Phonemic Fluency	16.2	5.7	13.8	7.5	1.1	0.259
Rey-Osterrieth Complex Figure copy	37.4	11.7	30.5	6.9	2.2	0.033
RCF copy time	148.4	63.0	187.7	68.7	−1.8	0.074
RCF immediate recall	23.1	6.9	14.8	7.7	3.5	0.001*

<sup>a</sup>Mean (SD) scores of neuropsychological data prescores for controls and TLE+UHS patients included in the study. *t* test compares the control group versus the TLE+UHS. Significant FDR-corrected differences are indicated.

\**p* < 0.05 FDR-corrected.

**Voxel-based morphometry analyses.** To evaluate the possible neural correlates between “anticipatory switching” and the hippocampus in healthy participants, regions of interest (ROIs) for the left and right hippocampus were defined based on the Anatomical Automatic Labeling Atlas in MNI space using the WFU pickatlas tool (Tzourio-Mazoyer et al., 2002; Maldjian et al., 2003; Maldjian et al., 2004). Two extra ROIs covering the caudate and the ventral striatum were also created. These regions were selected given the strong connections between the striatum and the hippocampus, and their shared role in learning and memory (Lisman and Grace, 2005; Haber and Knutson, 2010). The caudate ROI was built using the toolbox WFU pickatlas and the Automated Anatomical Labeling Atlas, as with the hippocampal ROI. The ventral striatum ROI was created using the results from an independent monetary gambling task from a previous experiment, which was conceived to functionally localize this structure (contrast: gains > losses, thresholded at FWE-corrected *p* < 0.001) (Ripollés et al., 2014). These ROIs served as control regions to further assess the specificity of our findings (as in Ripollés et al., 2016). Voxel-based morphometry within these ROIs (VBM) (Ashburner and Friston, 2000) was performed using Statistical Parametric Mapping software (SPM8; Wellcome Department of Imaging Neuroscience, University College, London; www.fil.ion.ucl.ac.uk/spm). Specifically, New Segment (Ashburner and Friston, 2005) was applied to the structural T1 weighted images of each subject. The resulting GM tissue probability maps were imported and fed into Diffeomorphic Anatomical Registration using Exponentiated Lie algebra (DARTEL) (Ashburner, 2007) to achieve spatial normalization in MNI space (using “modulation” to compensate for the effect of spatial normalization). Normalized images were smoothed using an isotropic spatial filter (FWHM = 8 mm) to reduce residual interindividual variability. The individual smoothed GM volume images were entered into a second-level analysis using a one-sample *t* test and were correlated with the behavioral measures of anticipatory switching (deviation from RL in “anticipatory switching,” see Behavioral results).

VBM was also used to calculate the differences in GM volume between patients and controls within the hippocampus. Patients’ T1 weighted images (before surgery) underwent the same processing pipeline described above. Individual smoothed GM volume images for controls and left TLE-UHS patients (LTLE-UHS) were entered into a two-sample *t* test, and two contrasts of interest were calculated: Controls > LTLE-

UHS and LTLE-UHS > Controls. The same analysis was repeated using the controls and the right TLE-UHS patients (RTLE-UHS).

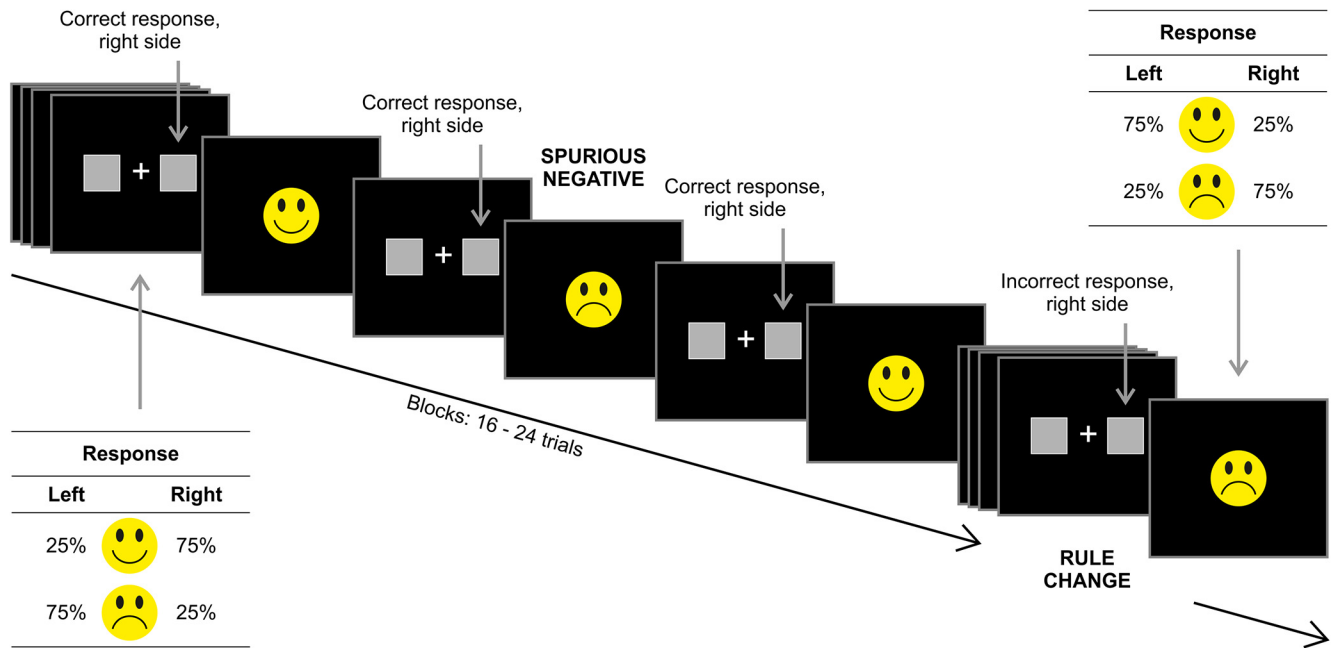
Contrasts were thresholded at a *p* < 0.005 uncorrected threshold at the voxel level with a cluster extent of >50 contiguous voxels (Lieberman and Cunningham, 2009). Corrected peaks are reported using a *p* < 0.05 FWE small volume correction (SVC) threshold.

**Volumetric analyses of the hippocampus.** Following previous studies showing that hippocampal volume is a good predictor of behavior in TLE patients (Fuentemilla et al., 2013) and in other types of patients with hippocampal damage (Horner et al., 2012), we used Freesurfer (<http://surfer.nmr.mgh.harvard.edu>, version 5.0.0) to compute the volume of the hippocampus in both patients and controls. Freesurfer volume estimates of the hippocampus have been previously validated both in controls (Morey et al., 2009) and in TLE patients with hippocampal sclerosis (Pardoe et al., 2009). The Freesurfer pipeline includes removal of non-brain tissue (Ségonne et al., 2004), automated Talairach transformation, intensity normalization (Sled et al., 1998), segmentation of the subcortical white matter and deep GM (Fischl et al., 2002, 2004), tessellation of the GM/white matter boundary, automated topology correction (Fischl et al., 2001; Ségonne et al., 2007), and surface deformation to detect GM/white matter and GM/CSF boundaries (Fischl and Dale, 2000). A deformable procedure was applied to parcellate the cerebral cortex into different regions according to gyral and sulcal structure information (Desikan et al., 2006).

Volumes for the hippocampus in both hemispheres were extracted and corrected for total intracranial volume (TIV) (Horner et al., 2012; Fuentemilla et al., 2013). First, a linear regression with the extracted volumes as the dependent variable and TIV as the independent was performed. Using the unstandardized  $\beta$  weights of this correlation, the volumes were corrected according to the equation ( $V_{ic} = V_i + (\text{mean-TIV} - \text{TIV}_i) \times \beta$ ), where  $V_{ic}$  is the corrected volume for a particular structure and subject,  $V_i$  is the raw volume as calculated by Freesurfer, meanTIV is the mean TIV for a particular group,  $\text{TIV}_i$  is the specific TIV of the subject, and  $\beta$  is the  $\beta$  weight calculated for this subject during the linear regression. To perform this analysis, we used the sum of the left and the right hippocampal corrected volumes, given that half of the patients have damage to one hemisphere specifically (similar to Fuentemilla et al., 2013). First, using an independent-samples *t* test, hippocampal volumes were compared between patients and controls. Then, hippocampal volume was correlated with deviation from RL in “anticipatory switching” by using Pearson’s *r*.

**Neuropsychological scoring.** Neuropsychological data for all participants were obtained only for the “pre” sessions (presession for the control group, and presurgery for the patient group). The neuropsychological data are summarized in Table 2. All participants (patients and controls) completed the Logical Memory I (immediate verbal memory) and Logical Memory II (delayed verbal memory), the Visual Reproduction I (immediate visual memory) and Visual Reproduction II (delayed visual memory), the Digits Span, and the Letters and Numbers subtests of the Wechsler Memory Scale III (Wechsler, 2004); the Vocabulary subtest of the Wechsler Adult Intelligence Scale (Wechsler, 1999); the Rey Auditory Verbal Learning Test (Rey, 1941; Schmidt, 1996); the Trail Making Test (Reitan, 1955; Davies, 1968); the Semantic Fluency and the Phonemic Fluency subtests of the Barcelona Test-R (Peña-Casanova, 2005); and the Rey-Osterrieth Complex Figure (copy and memory) (Rey, 1941; Osterrieth, 1944; Peña-Casanova, 2005). For each neuropsychological subtest, we used independent-sample *t* tests to elucidate the differences between groups of participants (Controls, TLE-UHS). The mean scores and statistical analyses of the neuropsychological test for patients and controls are summarized in Table 2. Even though most of the verbal and visuospatial functions were in some regards preserved before the surgery in the TLE-UHS group (e.g., correct performance on Rey Auditory Verbal Learning Test, in Logical Memory I and II, in Visual Reproduction I, and in copy Rey-Osterrieth Complex Figure), some effects were observed, as previously reported (Hermann et al., 1995, 1997; Helmstaedter, 2002; Pereira et al., 2010), in figural memory (low scores in Visual Reproduction II and in Rey-Osterrieth Complex Figure memory) and in verbal aspects (low scores on the Vocabulary subtest).





**Figure 1.** Schematic illustration of the task design. For each trial, participants had to choose one of the two squares placed on each side of the screen by pressing the corresponding mouse button. After a delay period of 700 ms, feedback was presented. An example of a sequence of trials in the probabilistic reversal learning task is given. The probability of obtaining positive or negative feedback on each side is stated for each sequence. Note the differences between the two negative feedbacks displayed. In the first, the participant responded correctly (choosing the most rewarding action) but got a spurious negative feedback. In contrast, in the second, the participant made an incorrect choice as the rule changed and then kept selecting the previous correct choice.

*Statistical analysis.* An analysis comparing the behavioral data from the left and the right TLE-UHS patients was performed to control for any effect of the laterality of the lesion or surgery. As no differences were found in any behavioral measure regarding the side of the lesion, we decided to collapse the results of both the right and the left TLE-UHS patients. Therefore, in the subsequent analysis, patients are treated as a single TLE-UHS group. Statistical effects regarding comparison within groups were obtained using univariate or repeated-measures ANOVA with age as a covariate. The same procedure was used to study differences between groups, adding group as a between-subject factor. To obtain an “anticipatory switching” measure, we used a principal components analysis. Factor coefficient scores for individual subjects were estimated using the Anderson-Rubin method (Anderson and Rubin, 1956).

For all statistical effects involving two or more degrees of freedom in the numerator, the Greenhouse-Geisser  $\epsilon$  was used to correct possible violations of the sphericity assumption. *p* values after this correction are reported.

## Results

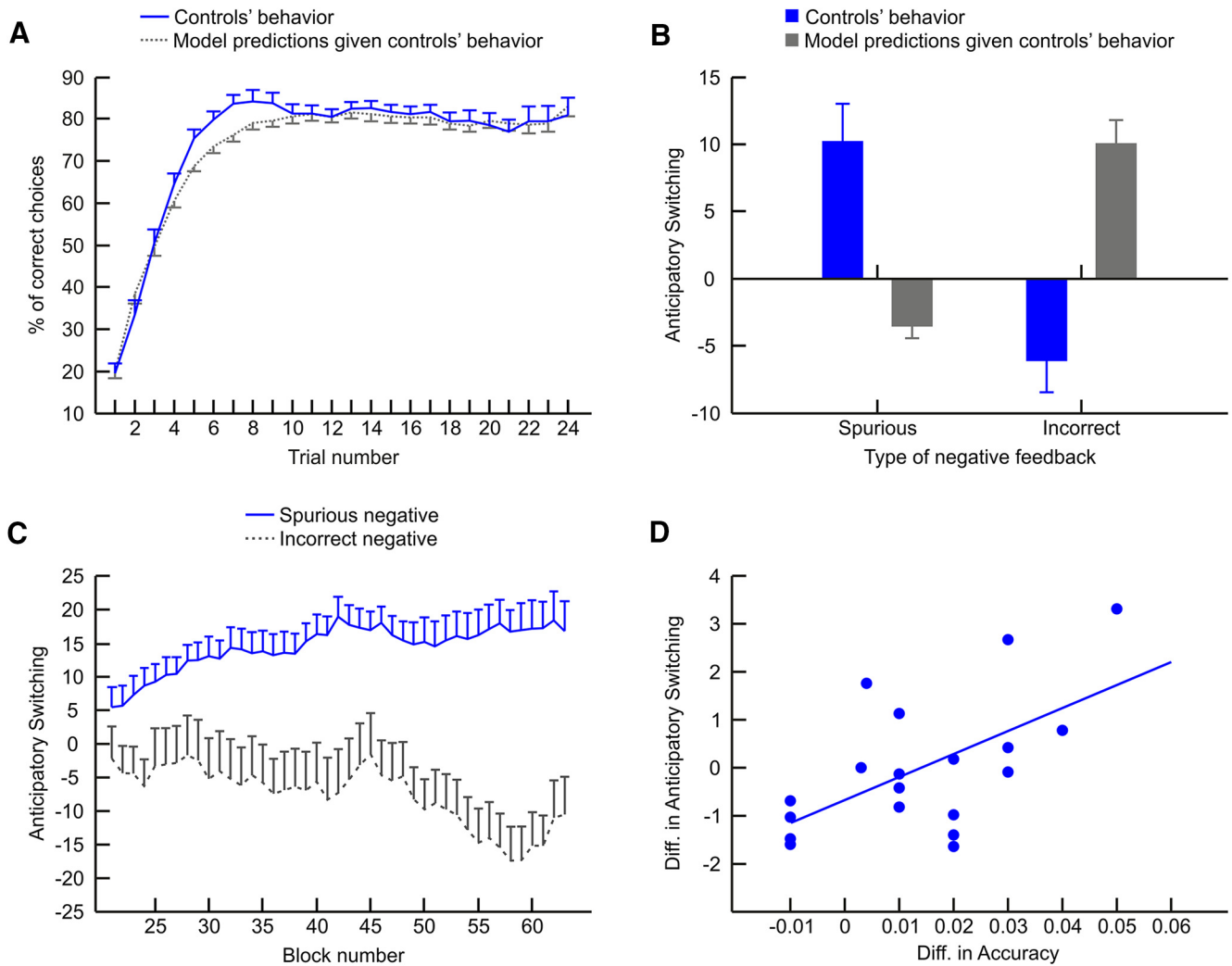
### Behavioral analyses in healthy subjects

Healthy subjects presented, on average, an accuracy rate (proportion of correct responses divided by the total number of responses) of 73.0% (SD = 5.3%) and earned 18.6€ (SD = 5.1€). To examine whether participant performance was captured by classic RL assumptions, a standard RL model, following a Q-learning model procedure used by Watkins and Dayan (1992), was fitted to participants’ behavioral performance (pseudo- $R^2 = 0.49$ , SD = 0.13). Participants had a mean learning rate of 0.71 (SD = 0.14) and an inverse temperature value of 2.12 (SD = 0.73).

In Figure 2A, we show the performance of healthy subjects across trials and the performance predicted by the RL model. Although the model predicts most of the participants’ performance, during early trials participants show faster adaptation than what the model predicts. In particular, the predictions of the model differed from the performance observed in healthy subjects in the first trials of each block (from trial 4 to 9, all

*p* values < 0.05, false discovery rate [FDR] corrected). A similar pattern of behavior has been previously reported in other studies (Mas-Herrero and Marco-Pallarés, 2014). These differences could be driven by participants’ ability to predict the occurrence of a reversal, which in turn may allow them to adapt faster. If so, participants should be more likely to switch from the current correct response to the current incorrect response following a negative feedback by the end of the block, when reversals are more likely. In contrast, once in the current incorrect response, they should be less likely to switch to the current correct response following a negative feedback by the end of the block than at the beginning.

Thus, we analyzed the probability of a behavioral switch following spurious negative feedback (negative feedback with current correct action) and incorrect negative feedback (negative feedback with current incorrect action) during the first (from 1 to 12) and last trials of each block (from 13 to 24), where the occurrence of reversal is more likely. As expected, individuals switched more following a spurious negative feedback at the end than at the beginning of the block ( $t_{(18)} = -3.66$ ,  $p = 0.002$ ), whereas they stayed more in the current incorrect target following an incorrect negative feedback in the last trials of the block ( $t_{(18)} = 2.59$ ,  $p = 0.019$ ). That is, participants were anticipating the occurrence of reversals, as shown in the increase of the probability of selecting the current incorrect target along the block according to the increasing probability of the appearance of a reversal trial. Furthermore, we computed for each individual, the mean difference in the percentage of switching between early and late trials of each block following spurious and incorrect negative feedback and compared the resulting values with those predicted by the model. The probabilities of being rewarded are the same during these trials, and thus the fact that the probability of choosing the current incorrect response following negative feedback increases across the block, cannot be accounted by a devaluation of the current correct action. Indeed, differences were found between



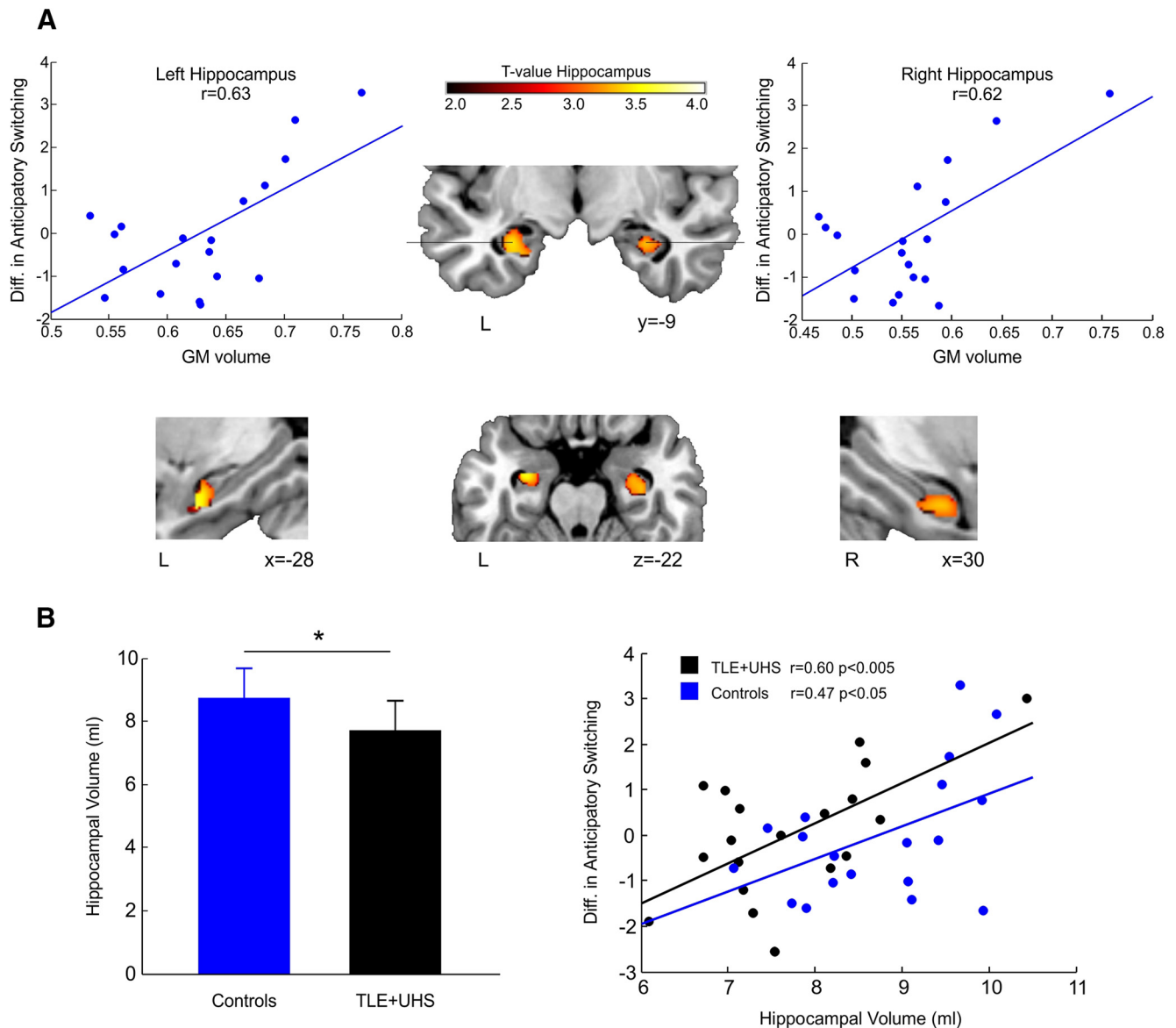
**Figure 2.** *A*, Learning curves during reversal learning. *y*-axis indicates the percentage of trials in which participants selected the rewarding choices. *x*-axis indicates the trial number. Blue lines indicate control participants' behavior. Gray-pointed line indicates the prediction of the model given participants' data. *B*, Mean percentage of switch following spurious negative feedback (negative feedback with current correct action) and incorrect negative feedback (negative feedback with current incorrect action). Blue bars represent control participants' behavior. Gray bars represent the prediction of the model given participants' data. *C*, Mean percentage of anticipatory switching following spurious negative and incorrect negative feedback across blocks. *D*, Correlation between deviation from RL model in "anticipatory switching" and "accuracy."

the model's predictions and the participants' behavior following both spurious ( $t_{(18)} = 4.57, p < 0.001$ ) and incorrect ( $t_{(18)} = -5.13, p < 0.001$ ) negative feedback (Fig. 2*B*).

We also examined to what extent this strategy was learned through the 63 blocks of the task. Figure 2*C* shows differences in switching between early and late trials of each block throughout the task for both kinds of feedback. As observed in the figure, difference in percentage switching between early and late trials increases through participants' experience following spurious (main effect of time,  $F_{(42,714)} = 1.65, p = 0.007$ ) and incorrect negative feedback (main effect of time,  $F_{(42,756)} = 1.94, p < 0.00$ ) but in opposite directions. Notably, difference in switching between early and late trials following spurious and incorrect negative feedback was not different from 0 during the first third of the task ( $t$  test against 0:  $p > 0.05$ , FDR corrected) but increased significantly during the second and final thirds ( $t$  test against 0:  $p < 0.05$ , FDR corrected). These results indicate that individuals learned to predict the occurrence of reversals through experience.

Notably, both measures (difference in percentage switching between early and late trials following spurious and incorrect negative feedback) were highly correlated ( $r_{(19)} = -0.59, p =$

0.008). This correlation further confirms that both patterns of behavior reflect the same strategy: anticipation of reversals. To have a single measure of this strategy, we performed a principal component analysis, including both measures. The resulting component was called "anticipatory switching." Similarly, we performed a second principal component analysis with the predictions driven by the RL model. Differences between the two resulting factors were computed for each individual to assess the deviation from the RL predictions. If anticipating the occurrence of reversals, as measured by the "anticipatory switching" factor, is an optimal strategy, we hypothesized that individuals showing higher "anticipatory switching" than that predicted by the model would show better accuracy. To assess the contribution of anticipatory switching in accuracy enhancement, we correlated the deviation from RL in anticipatory switching with deviation from RL in accuracy (the difference between the observed accuracy and the one predicted by the model; Fig. 2*D*). The correlational analysis revealed a significant positive relationship between both measures ( $r_{(19)} = 0.63, p = 0.004$ ), indicating that this strategy was optimal to perform the task beyond classic trial-and-error adaptations.



**Figure 3.** *A*, Correlates between GM volume in healthy participants and deviation from RL in “anticipatory switching.” Results for hippocampus (red–yellow) are shown at a  $p < 0.005$  uncorrected threshold at the voxel level (main peaks in the left cluster survived a FWE-corrected SVC,  $p < 0.05$  threshold). Neurological convention is used with MNI coordinates at the bottom right of each slice. Scatter plots represent the correlation between mean GM volume and the deviation from RL in “anticipatory switching.” *B*, Left, Difference between total hippocampal volume (mean  $\pm$  SD in milliliters) between controls (blue) and TLE+UHS patients (black). Right, Significant correlations between total hippocampal volume (in milliliters) and the deviation from RL in “anticipatory switching” in patients (black) and controls (blue).  $*p = 0.002$ . L, Left hemisphere; R, right hemisphere.

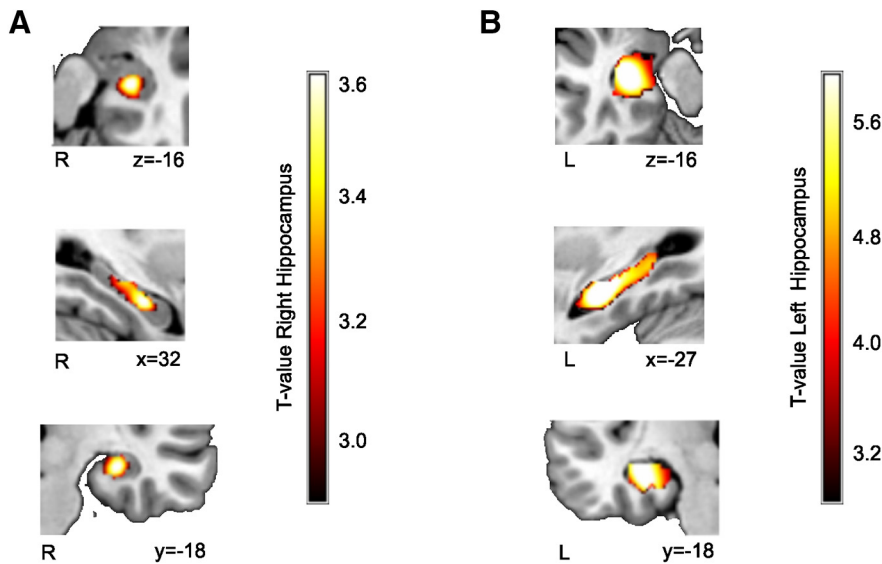
### Voxel-based morphometry analysis of the hippocampus and control regions in healthy subjects

Next, we used VBM to examine to what extent individual differences in GM volume in the hippocampus predicted anticipatory switching.”

In healthy controls, deviation from RL in “anticipatory switching” positively correlated with the hippocampus bilaterally (only the peak at the left hippocampus survived correction; Fig. 3*A*; 232 voxels at the left hippocampus, main peak at  $x = -30$ ,  $y = -3$ ,  $z = -29$ ,  $t_{(17)} = 3.92$ ,  $p < 0.05$  FWE-corrected using SVC; 215 voxels at the right hippocampus, main peak at  $x = 30$ ,  $y = -10$ ,  $z = -21$ ,  $t_{(17)} = 3.33$ ,  $p = 0.1$  FWE-corrected SVC; Fig. 3*A*). These results suggest that individuals with greater hippocampal GM volume showed greater anticipatory switching. No significant correlation was found, as expected, for any voxel within the control ROIs for caudate and ventral striatum.

### Between-group VBM comparison of the hippocampus

To confirm the relationship between hippocampal volume and the capacity to anticipate predictable changes in the environment, we analyzed a group of 20 TLE patients with unilateral hippocampal sclerosis (10 suffering from left and 10 from right medically refractory TLE). To further confirm the hippocampal lesions in the TLE-UHS group, we calculated the differences in GM volume between patients and controls within these regions. As expected, the right TLE-UHS patients group showed lower GM volume in the right hippocampus (one cluster with 408 voxels;  $x = 32$ ,  $y = -16$ ,  $z = -17$ ;  $t_{(28)} = 3.57$ ,  $p < 0.039$  FWE-corrected using SVC; Fig. 4*A*) compared with healthy controls. Similarly, left TLE-UHS patients also showed lower GM volume in the left hippocampus compared with healthy participants (one cluster with 1621 voxels;  $x = -32$ ,  $y = -15$ ,  $z = -17$ ;  $t_{(28)} = 6.85$ ,  $p < 0.001$  FWE-corrected using SVC; Fig. 4*B*). These results fur-



**Figure 4.** Increased GM volume in healthy participants compared with (A) left TLE+UHS and (B) left TLE+UHS patients. Results for hippocampus (red–yellow) are shown at an auxiliary  $p < 0.005$  uncorrected threshold at the voxel level (main peaks in both clusters survived an FWE-corrected SVC,  $p < 0.05$  threshold). Neurological convention is used with MNI coordinates at the bottom right of each slice. L, Left hemisphere; R, right hemisphere.

ther confirm that the TLE-UHS group presented less hippocampal GM volume than controls.

#### Behavioral comparisons between left and right TLE-UHS

No behavioral differences in any measure were found between left and right TLE-UHS. The two groups presented, on average, similar accuracy rates ( $F_{(1,17)} = 1.50$ ,  $p = 0.24$ ) and accumulated similar amounts of money ( $F_{(1,17)} = 2.54$ ,  $p = 0.13$ ). RL parameters, such as the learning rate (learning rate:  $F_{(1,17)} = 0.70$ ,  $p = 0.41$ ), inverse of the temperature ( $F_{(1,17)} = 0.89$ ,  $p = 0.36$ ), and model fitting (pseudo- $R^2$ :  $F_{(1,17)} = 0.71$ ,  $p = 0.41$ ), were also similar for the two groups. Both groups also showed similar percentage of switching between early and late trials following both spurious ( $F_{(1,17)} = 0.01$ ,  $p = 0.92$ ) and incorrect negative feedback ( $F_{(1,17)} = 0.06$ ,  $p = 0.81$ ). Taking into account the lack of behavioral differences between left and right TLE-UHS groups, both were collapsed into one TLE group.

#### Behavioral comparisons between controls and TLE-UHS

As no behavioral differences in any measure were found between left and right TLE-UHS, all patients were included in one single TLE-UHS group presenting on average an accuracy rate of 69.8% (SD = 5.2%), who earned 15.8€ (SD = 4.7€). Patients were less accurate than Controls (accuracy:  $F_{(1,36)} = 4.46$ ,  $p = 0.042$ ) and consequently gained less money ( $F_{(1,36)} = 4.34$ ,  $p = 0.044$ ).

The RL model was fitted to patients' behavioral performance (pseudo- $R^2 = 0.43$ , SD = 0.19), leading to a mean learning rate of 0.77 (SD = 0.20) and a mean inverse temperature value of 1.89 (SD = 0.98). None of these measures differed between control participants and patients (all  $p$  values > 0.20).

In Figure 5A, we show the performance of TLE-UHS patients across trials and the performance predicted by the RL algorithm. In contrast to healthy subjects, no differences were found between the performance predicted by the model and the observed behavior in the TLE-UHS group (all  $p$  values > 0.15, FDR corrected). These results might suggest that patients from the TLE-UHS group were not able to take advantage of the predictability of the environment and could not anticipate potential reversals.

This also means that they might be guiding their choices only according to standard RL principles.

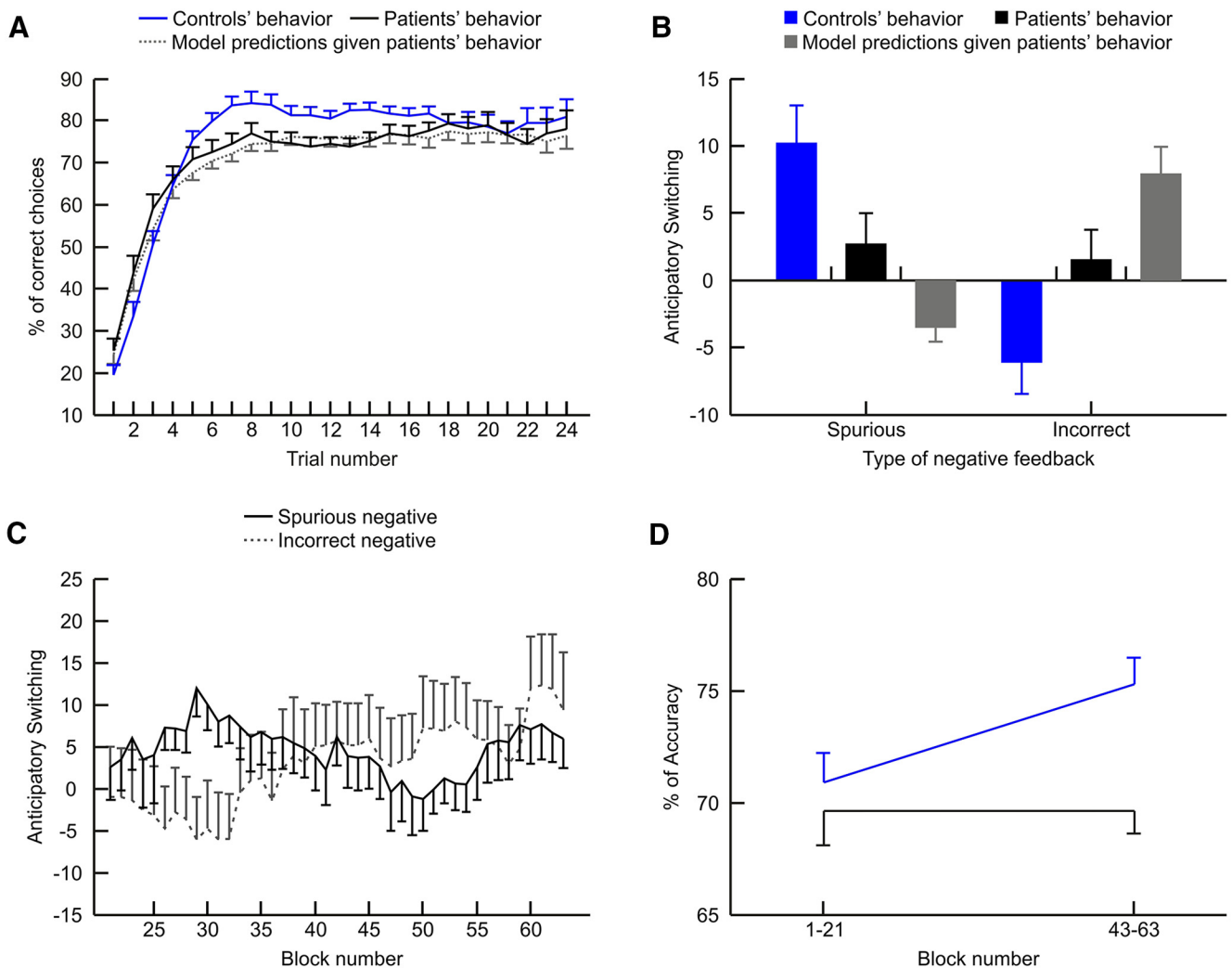
We then analyzed the pattern of switching developed by the TLE-UHS following spurious and incorrect negative feedback. Similar to controls, although less pronounced, the switching pattern of the TLE-UHS group between early and late trials of each block following either spurious ( $t_{(19)} = 2.35$ ,  $p = 0.03$ ) or incorrect negative feedback ( $t_{(19)} = -2.59$ ,  $p = 0.02$ ), also differed from the predictions of the RL model. Compared with the model, TLE-UHS patients switched more following a spurious negative feedback and stayed more following an incorrect negative feedback by the end each block, when reversals are more likely (Fig. 5B). However, controls developed greater “anticipatory switching” than patients following spurious ( $F_{(1,36)} = 4.42$ ,  $p = 0.04$ ) and incorrect negative feedback ( $F_{(1,36)} = 5.40$ ,  $p = 0.027$ ). In addition, the “anticipatory switching” strategy of the TLE-

UHS group (Fig. 5C) remained stable across blocks following both spurious (time effect:  $F_{(42,756)} = 0.87$ ,  $p = 0.71$ ) and incorrect negative feedback (time effect:  $F_{(42,714)} = 1.04$ ,  $p = 0.41$ ), in contrast to the control group (interaction group  $\times$  time following spurious:  $F_{(42,1512)} = 1.4$ ,  $p = 0.048$ ; and incorrect negative feedback:  $F_{(42,1386)} = 1.88$ ,  $p = 0.001$ ). These results further suggest that patients were not able to learn the high-order structure of the task through experience with the same success as healthy participants. Based on this, we also hypothesized that, whereas controls' performance should improve through experience, patients' performance should remain stable across the task. To test this, we compared the accuracy of both groups during the first and last third of the task. As observed in Fig. 5D, there was a clear interaction between time and group ( $F_{(1,36)} = 6.37$ ,  $p = 0.016$ ). The performance of controls improved between the first and the last third ( $F_{(1,5)} = 78.9$ ,  $p < 0.001$ ), something that did not occur in patients ( $F_{(1,4)} = 0.08$ ,  $p = 0.79$ ). But most importantly, controls and patients did not differ in their performance during the first third of the task ( $F_{(1,36)} = 0.3$ ,  $p = 0.59$ ), whereas controls surpassed patients in the last third ( $F_{(1,36)} = 18.2$ ,  $p < 0.001$ ). Finally, to study whether these changes in accuracy were driven by changes in learning rate across blocks rather than by the anticipation of reversals, we fitted the model separately for the first and the second half of the task. The resulting learning rate and inverse temperature parameter did not differ between first and second half of the task, either in control or patients (all  $p$  values > 0.15). No interaction or group effect was found either (all  $p$  values > 0.25). This lack of effect indicates that differences in accuracy and anticipatory switching found across blocks and between groups cannot be explained by changes in learning rate.

#### Volumetric analyses of the hippocampus

To confirm the relationship between the use of anticipatory strategies and intact hippocampal volume in the TLE-UHS group (as measured with Freesurfer; for a similar procedure, see Horner et al., 2012; Fuentemilla et al., 2013), we correlated the total hippocampal volume with the deviation from RL in “anticipatory switching” in both groups. First, as expected, patients showed





**Figure 5.** *A*, Learning curves during reversal learning. *y*-axis indicates the percentage of trials in which participants selected the rewarding choices. *x*-axis indicates the trial number. Blue lines indicate control participants' behavior. Gray-pointed line indicates the prediction of the model given TLE+UHS patients' data. Black lines indicate TLE+UHS patients' behavior. *B*, Mean percentage of switching between early and late trials following spurious negative feedback (negative feedback with current correct action) and incorrect negative feedback (negative feedback with current incorrect action). Black bars represent TLE+UHS patients' behavior. Blue bars represent control participants' behavior. Gray bars represent the prediction of the model given participants' data. *C*, Mean percentage of anticipatory switching following spurious negative and incorrect negative feedback across blocks for TLE+UHS (solid black line) patients and controls (solid blue line). *D*, Mean percentage of accuracy for the first third (1–21) and the last third of blocks of the task (43–63), for control participants (solid blue line) and TLE-UHS patients (solid black line).

significant lower total hippocampal volume than controls ( $7.68 \pm 0.97$  ml and  $8.72 \pm 0.93$  ml for patients and controls, respectively;  $t_{(37)} = 3.37$ ,  $p < 0.002$ ; Fig. 3*B*). Crucial for our hypothesis, hippocampal volume predicted anticipatory switching in both patients ( $r_{(20)} = 0.60$ ,  $p < 0.005$ ) and controls ( $r_{(19)} = 0.47$ ,  $p < 0.05$ ; Fig. 3*B*); note how the regression line is similar in both groups; Fig. 3*B*).

#### Behavioral comparisons between preoperative and postoperative sessions

The TLE-UHS group was also tested 3 months after an anterior temporal lobe resection. Patients showed similar performance after the surgery. They presented, on average, similar accuracy rates ( $F_{(1,18)} = 0.42$ ,  $p = 0.53$ ) and accumulated similar amounts of money ( $F_{(1,18)} = 0.76$ ,  $p = 0.39$ ), as during the first session. The value of the RL parameters derived from the RL model was also similar between sessions (learning rate:  $F_{(1,18)} = 0.001$ ,  $p = 0.98$ ; inverse temperature value:  $F_{(1,18)} = 0.15$ ,  $p = 0.70$ ; and model fitting:  $F_{(1,18)} = 0.008$ ,  $p = 0.93$ ). In both sessions, the TLE-UHS

presented similar behavioral switching patterns between early and late trials following spurious negative ( $t_{(19)} = 1.20$ ,  $p = 0.25$ ) and incorrect negative feedback ( $t_{(19)} = 1.08$ ,  $p = 0.30$ ). The lack of differences in any of the aforementioned measures implies that the damage induced by the surgery does not further affect the already diminished capacity of TLE-UHS patients to learn the high-order structure of the reversal learning task.

#### Discussion

In the present study, we investigated the role of the hippocampus in subserving anticipatory strategies, which rely on the correct identification of differential temporal contexts in reversal learning. To test this hypothesis, a probabilistic reversal learning task was administered to healthy participants and to a group of refractory temporal lobe epileptic patients with unilateral hippocampal sclerosis (who served as a hippocampal lesion model). Our results provide evidence in favor of a key role for the hippocampus in the creation of an internal model of the task structure and in the implementation of optimal strategies given this knowledge.

It has been proposed that two control systems might be used in reversal learning. The first involves learning which actions are more optimal through direct evaluation of their rewarding consequences using prediction error signals (Sutton and Barto, 1998). The second, which arises later in time after some training, represents the subject's belief that reversals or changes in reward contingency can occur. Thus, individuals generate a likelihood estimate that a reversal can occur, which in turn allow them to anticipate and rapidly adapt the rule change (Costa et al., 2015; Jang et al., 2015). This idea has been proved using the reversal learning task in extensively trained nonhuman primates. Using a Bayesian model, Costa et al. (2015) showed that animals were able to adjust their behavior according to the structure of the task and they also tended to switch before the reversal trial, especially in more uncertain conditions. The authors proposed that this behavior could be a combination of priors and evidence obtained during task performance. In agreement with this line of reasoning, we have shown that healthy controls overtook behavioral adaptations predicted by model-free learning models during the reversal learning task. Notably, our results indicate that participants could learn the task's structure and could anticipate the moment in time at which the rule was most likely to change, thus using this information to guide their choices (Fig. 2*B,C*). These findings are also consistent with a previous study that showed how healthy individuals may take advantage of the underlying structure of the reversal learning task to guide their choices, rather than only relying on the individual reward history of each action as captured by classic RL algorithms (Hampton et al., 2006). Additionally, we observed that participants with more anticipatory switching showed better accuracy during the task. These results converge with previous evidence suggesting that, in volatile environments, learning the sequential contingencies of events and actions (learning the task structure), and using them to guide decisions, improves individual performance (Acuña and Schrater, 2010) compared with standard RL strategies.

Importantly, in another study using the same Bayesian approach, Jang et al. (2015) demonstrated that nonhuman primates with lesions in the hippocampus (as well as orbitofrontal cortex, amygdala, and rhinal cortices) showed an impairment in a reversal learning task, which was explained by a reduction of the initial values of prior beliefs in reversals, rather than by an impairment in trial-and-error adaptations. Accordingly, early animal studies already indicated relevant participation of inferior temporal lobe and MTL regions (rhinal, parahippocampal, and hippocampal) in the modulation of reversal learning (Murray et al., 1998; Bussey et al., 2002; Browning et al., 2007; Morellini et al., 2010). In particular, previous findings indicate that lesions in the inferior temporal lobe and hippocampal regions disrupted reversal learning in monkeys (Murray et al., 1998; Bussey et al., 2002; Browning et al., 2007) and rats (for a review, see Whishaw, 1998). On the other hand, rats with a higher number of neurons in the dentate gyrus of the hippocampus showed faster adaptation to reversal learning (Morellini et al., 2010). In this vein, here we show a clear association between GM volume in the hippocampus and the use of anticipatory strategies relying on the temporal context, pointing up the key role of this region in goal-directed behaviors (Fig. 3*A,B*). Additionally, to further confirm the role of the hippocampus, we tested a group of refractory temporal lobe epileptic patients with unilateral hippocampal sclerosis, who served as a hippocampal lesion model. First and in convergence with the VBM analysis, volumetric measures of the hippocampus predicted anticipatory strategies in both patients and controls (Fig. 3*B*); as expected, patients had less hippocampal volume than

healthy individuals; Fig. 4). In addition, patients with hippocampal sclerosis were not able to anticipate the occurrence of the reversals in the same manner as controls, suggesting impairment in learning the structure of the task (Fig. 5). Indeed, patients' performance across trials fitted with the predictions generated by a model-free RL model, even at the beginning of the task when control behavior clearly differed from that of the model. These findings indicate that the behavior of TLE-UHS patients was not affected by the temporal context, relying mainly on RL strategies to complete the task. As expected, such behavioral pattern was less optimal (Acuña and Schrater, 2010), which was reflected by the fact that patients had lower accuracy rates and accumulated less money compared with the control group. Previous studies have already shown that MTL damage does not have a significant impact on individuals' abilities to learn reward contingencies, but rather in generalizing learned information. This deficit prevents patients from using previously learned knowledge when facing novel contexts with overlapping features of previous experiences (Meyers et al., 2003). We extend these results by showing that MTL damage also alters strategies that rely on the processing of contextual information.

Furthermore, converging evidence indicate that the hippocampus and surrounding medial temporal cortices subserve successful encoding (Davachi et al., 2003; Ranganath et al., 2004; Staresina and Davachi, 2006), retrieval (Cansino et al., 2002; Hannula and Ranganath, 2008; Diana et al., 2013), and representation (Bonnici et al., 2012) of contextual information, including temporal context (Jenkins and Ranganath, 2010; Ekstrom et al., 2011; Tubridy and Davachi, 2011). Such capacity might be related to the ability to distinguish between temporally distinct events that have overlapping elements (Hsieh et al., 2014). The capacity to use the representation of contextual information may be crucial to adaptively guide future actions and decisions in temporally structured sequences. Indeed, studies in rodents have shown that damage to the hippocampus results in impairments in the anticipatory selection of maze arms experienced in sequence (DeCoteau and Kesner, 2000). Hippocampal amnesic patients are impaired in tasks involving imagination of contextual information and representation of future events (Hassabis et al., 2007; Addis et al., 2009; Zeidman and Maguire, 2016). Similarly, the TLE-UHS group in the current study was not able to anticipate the occurrence of a reversal, leading to lower performance than the control group. In humans, there is no previous specific evidence regarding the role of the hippocampus in reversal learning. Here, we extend all these findings, by showing that changes in reversal learning performance may be driven by the role of the hippocampus in processing and representing contextual information to be used later for goal-directed behaviors.

Finally, the TLE-UHS group did not show differences before and after the resection of the hippocampus in their ability to solve the reversal learning task. We hypothesize that the damage present in the hippocampus before the surgery (reduced GM volume shown by patients in Figs. 3*B, 4*) was already impairing patient ability to use high-order strategies. Thus, the removal of part of the hippocampus might not further hinder their already disrupted system. Additionally, the fact that the surgery does not further disrupt their performance in reversal learning may contribute to improved clinical management.

## References

- Acuña DE, Schrater P (2010) Structure learning in human sequential decision-making. *PLoS Comput Biol* 6:e1001003. [CrossRef Medline](#)  
 Addis DR, Pan L, Vu MA, Laiser N, Schacter DL (2009) Constructive epi-

- sodic simulation of the future and the past: distinct subsystems of a core brain network mediate imagining and remembering. *Neuropsychologia* 47:2222–2238. [CrossRef Medline](#)
- Aggleton JP, Sanderson DJ, Pearce JM (2007) Structural learning and the hippocampus. *Hippocampus* 17:723–734. [CrossRef Medline](#)
- Allen TA, Salz DM, McKenzie S, Fortin NJ (2016) Nonspatial sequence coding in CA1 neurons. *J Neurosci* 36:1547–1563. [CrossRef Medline](#)
- Anderson TW, Rubin H (1956) Statistical inference in factor analysis. Regents of the University of California. <http://projecteuclid.org/euclid.bsmisp/1200511860>. Accessed April 24, 2017.
- Ashburner J (2007) A fast diffeomorphic image registration algorithm. *Neuroimage* 38:95–113. [CrossRef Medline](#)
- Ashburner J, Friston KJ (2000) Voxel-based morphometry: the methods. *Neuroimage* 11:805–821. [CrossRef Medline](#)
- Ashburner J, Friston KJ (2005) Unified segmentation. *Neuroimage* 26:839–851. [CrossRef Medline](#)
- Bonnici HM, Kumaran D, Chadwick MJ, Weiskopf N, Hassabis D, Maguire EA (2012) Decoding representations of scenes in the medial temporal lobes. *Hippocampus* 22:1143–1153. [CrossRef Medline](#)
- Bornstein AM, Daw ND (2013) Cortical and hippocampal correlates of de-liberation during model-based decisions for rewards in humans. *PLoS Comput Biol* 9:e1003387. [CrossRef Medline](#)
- Browning PG, Easton A, Gaffan D (2007) Frontal-temporal disconnection abolishes object discrimination learning set in macaque monkeys. *Cereb Cortex* 17:859–864. [CrossRef Medline](#)
- Bussey TJ, Wise SP, Murray EA (2002) Interaction of ventral and orbital prefrontal cortex with inferotemporal cortex in conditional visuomotor learning. *Behav Neurosci* 116:703–715. [CrossRef Medline](#)
- Cansino S, Maquet P, Dolan RJ, Rugg MD (2002) Brain activity underlying encoding and retrieval of source memory. *Cereb Cortex* 12:1048–1056. [CrossRef Medline](#)
- Cendes F (2005) Progressive hippocampal and extrahippocampal atrophy in drug resistant epilepsy. *Curr Opin Neurol* 18:173–177. [CrossRef Medline](#)
- Costa VD, Tran VL, Turchi J, Averbeck BB (2015) Reversal learning and dopamine: a bayesian perspective. *J Neurosci* 35:2407–2416. [CrossRef Medline](#)
- Davachi L, Mitchell JP, Wagner AD (2003) Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. *Proc Natl Acad Sci U S A* 100:2157–2162. [CrossRef Medline](#)
- Davies AD (1968) The influence of age on Trail Making Test performance. *J Clin Psychol* 24:96–98. [CrossRef Medline](#)
- Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 8:1704–1711. [CrossRef Medline](#)
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-based inferences on humans' choices and striatal prediction errors. *Neuron* 69:1204–1215. [CrossRef Medline](#)
- DeCoteau WE, Kesner RP (2000) A double dissociation between the rat hippocampus and medial caudoputamen in processing two forms of knowledge. *Behav Neurosci* 114:1096–1108. [CrossRef Medline](#)
- Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31:968–980. [CrossRef Medline](#)
- Diana RA, Yonelinas AP, Ranganath C (2013) Parahippocampal cortex activation during context reinstatement predicts item recollection. *J Exp Psychol Gen* 142:1287–1297. [CrossRef Medline](#)
- Eichenbaum H (2000) A cortical-hippocampal system for declarative memory. *Nat Rev Neurosci* 1:41–50. [CrossRef Medline](#)
- Ekstrom AD, Copara MS, Isham EA, Wang WC, Yonelinas AP (2011) Dissociable networks involved in spatial and temporal order source retrieval. *Neuroimage* 56:1803–1813. [CrossRef Medline](#)
- Fischl B, Dale AM (2000) Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A* 97:11050–11055. [CrossRef Medline](#)
- Fischl B, Liu A, Dale AM (2001) Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Trans Med Imaging* 20:70–80. [CrossRef Medline](#)
- Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM (2002) Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33:341–355. [CrossRef Medline](#)
- Fischl B, Salat DH, van der Kouwe AJ, Makris N, Ségonne F, Quinn BT, Dale AM (2004) Sequence-independent segmentation of magnetic resonance images. *Neuroimage* 23 [Suppl 1]:S69–S84.
- Fuentemilla L, Miró J, Ripollés P, Vilà-Balló A, Juncadella M, Castañer S, Salord N, Monasterio C, Falip M, Rodríguez-Fornells A (2013) Hippocampus-dependent strengthening of targeted memories via reactivation during sleep in humans. *Curr Biol* 23:1769–1775. [CrossRef Medline](#)
- Haber SN, Knutson B (2010) The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 35:4–26. [CrossRef Medline](#)
- Hampton AN, Bossaerts P, O'Doherty JP (2006) The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J Neurosci* 26:8360–8367. [CrossRef Medline](#)
- Hannula DE, Ranganath C (2008) Medial temporal lobe activity predicts successful relational memory binding. *J Neurosci* 28:116–124. [CrossRef Medline](#)
- Hassabis D, Kumaran D, Vann SD, Maguire EA (2007) Patients with hippocampal amnesia cannot imagine new experiences. *Proc Natl Acad Sci U S A* 104:1726–1731. [CrossRef Medline](#)
- Helmstaedter C (2002) Effects of chronic epilepsy on declarative memory systems. *Prog Brain Res* 135:439–453. [CrossRef Medline](#)
- Hermann BP, Gold J, Pusakulich R, Wyler AR, Randolph C, Rankin G, Hoy W (1995) Wechsler Adult Intelligence Scale-revised in the evaluation of anterior temporal lobectomy candidates. *Epilepsia* 36:480–487. [CrossRef Medline](#)
- Hermann BP, Seidenberg M, Schoenfeld J, Davies K (1997) Neuropsychological characteristics of the syndrome of mesial temporal lobe epilepsy. *Arch Neurol* 54:369–376. [CrossRef Medline](#)
- Horner AJ, Gadian DG, Fuentemilla L, Jentschke S, Vargha-Khadem F, Duzel E (2012) A rapid, hippocampus-dependent, item-memory signal that initiates context memory in humans. *Curr Biol* 22:2369–2374. [CrossRef Medline](#)
- Howard MW, Kahana MJ (2002) A distributed representation of temporal context. *J Math Psychol* 46:269–299. [CrossRef](#)
- Hsieh LT, Gruber MJ, Jenkins LJ, Ranganath C (2014) Hippocampal activity patterns carry information about objects in temporal context. *Neuron* 81:1165–1178. [CrossRef Medline](#)
- Huh N, Jo S, Kim H, Sul JH, Jung MW (2009) Model-based reinforcement learning under concurrent schedules of reinforcement in rodents. *Learn Mem* 16:315–323. [CrossRef Medline](#)
- Jang AI, Costa VD, Rudebeck PH, Chudasama Y, Murray EA, Averbeck BB (2015) The role of frontal cortical and medial-temporal lobe brain areas in learning a Bayesian prior belief on reversals. *J Neurosci* 35:11751–11760. [CrossRef Medline](#)
- Jenkins LJ, Ranganath C (2010) Prefrontal and medial temporal lobe activity at encoding predicts temporal context memory. *J Neurosci* 30:15558–15565. [CrossRef Medline](#)
- Jocham G, Klein TA, Neumann J, von Cramon DY, Reuter M, Ullsperger M (2009) Dopamine DRD2 polymorphism alters reversal learning and associated neural activity. *J Neurosci* 29:3695–3704. [CrossRef Medline](#)
- Konkel A, Cohen NJ (2009) Relational memory and the hippocampus: representations and methods. *Front Neurosci* 3:166–174. [CrossRef Medline](#)
- Lieberman MD, Cunningham WA (2009) Type I and Type II error concerns in fMRI research: re-balancing the scale. *Soc Cogn Affect Neurosci* 4:423–428. [CrossRef Medline](#)
- Lisman JE, Grace AA (2005) The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* 46:703–713. [CrossRef Medline](#)
- Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19:1233–1239. [CrossRef Medline](#)
- Maldjian JA, Laurienti PJ, Burdette JH (2004) Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *Neuroimage* 21:450–455. [CrossRef Medline](#)
- Malmgren K, Thom M (2012) Hippocampal sclerosis: origins and imaging. *Epilepsia* 53 [Suppl 4]:19–33.
- Mas-Herrero E, Marco-Pallarés J (2014) Frontal theta oscillatory activity is a common mechanism for the computation of unexpected outcomes and learning rate. *J Cogn Neurosci* 26:447–458. [CrossRef Medline](#)

- Meyers RA, Zavala AR, Neisewander JL (2003) Dorsal, but not ventral, hippocampal lesions disrupt cocaine place conditioning. *Neuroreport* 14:2127–2131. [CrossRef Medline](#)
- Morellini F, Sivukhina E, Stoenica L, Oulianova E, Bukalo O, Jakovcevski I, Dityatev A, Irintchev A, Schachner M (2010) Improved reversal learning and working memory and enhanced reactivity to novelty in mice with enhanced GABAergic innervation in the dentate gyrus. *Cereb Cortex* 20:2712–2727. [CrossRef Medline](#)
- Morey RA, Petty CM, Xu Y, Hayes JP, Wagner HR 2nd, Lewis DV, LaBar KS, Styner M, McCarthy G (2009) A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. *Neuroimage* 45:855–866. [CrossRef Medline](#)
- Murray EA, Baxter MG, Gaffan D (1998) Monkeys with rhinal cortex damage or neurotoxic hippocampal lesions are impaired on spatial scene learning and object reversals. *Behav Neurosci* 112:1291–1303. [CrossRef Medline](#)
- Osterrieth PA (1944) Le test de copie d'une figure complexe; contribution à l'étude de la perception et de la mémoire [Test of copying a complex figure; contribution to the study of perception and memory]. *Arch Psychol* 30:206–356.
- Pardoe HR, Pell GS, Abbott DF, Jackson GD (2009) Hippocampal volume assessment in temporal lobe epilepsy: How good is automated segmentation? *Epilepsia* 50:2586–2592. [CrossRef Medline](#)
- Peña-Casanova J (2005) Integrated Neuropsychological Exploration Program-Barcelona Test Revised. Barcelona: Masson.
- Pereira FR, Alessio A, Sercheli MS, Pedro T, Bilevicius E, Rondina JM, Ozelo HF, Castellano G, Covolan RJ, Damasceno BP, Cendes F (2010) Asymmetrical hippocampal connectivity in mesial temporal lobe epilepsy: evidence from resting state fMRI. *BMC Neurosci* 11:66. [CrossRef Medline](#)
- Ranganath C, Yonelinas AP, Cohen MX, Dy CJ, Tom SM, D'Esposito M (2004) Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia* 42:2–13. [CrossRef Medline](#)
- Reitan RM (1955) The relation of the Trail Making Test to organic brain damage. *J Consult Psychol* 19:393–394. [CrossRef Medline](#)
- Rey A (1941) L'examen psychologique dans les cas d'encéphalopathie traumatique (Les problèmes) [The psychological examination in cases of traumatic encephalopathy (problems).] *Arch Psychol* 28:215–285.
- Ripollés P, Marco-Pallarés J, Hielscher U, Mestres-Missé A, Tempelmann C, Heinze HJ, Rodríguez-Fornells A, Noesselt T (2014) The role of reward in word learning and its implications for language acquisition. *Curr Biol* 24:2606–2611. [CrossRef Medline](#)
- Ripollés P, Marco-Pallarés J, Alicart H, Tempelmann C, Rodríguez-Fornells A, Noesselt T (2016) Intrinsic monitoring of learning success facilitates memory encoding via the activation of the SN/VTA-Hippocampal loop. *eLife* 5:e17441. [CrossRef Medline](#)
- Schmidt M (1996) *Rey Auditory Verbal Learning Test (RAVLT): a handbook*. Torrance, CA: Western Psychological Services.
- Ségonne F, Dale AM, Busa E, Glessner M, Salat D, Hahn HK, Fischl B (2004) A hybrid approach to the skull stripping problem in MRI. *Neuroimage* 22:1060–1075. [CrossRef Medline](#)
- Ségonne F, Pacheco J, Fischl B (2007) Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *IEEE Trans Med Imaging* 26:518–529. [CrossRef Medline](#)
- Sled JG, Zijdenbos AP, Evans AC (1998) A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging* 17:87–97. [CrossRef Medline](#)
- Squire LR, Stark CE, Clark RE (2004) The medial temporal lobe. *Annu Rev Neurosci* 27:279–306. [CrossRef Medline](#)
- Staresina BP, Davachi L (2006) Differential encoding mechanisms for subsequent associative recognition and free recall. *J Neurosci* 26:9162–9172. [CrossRef Medline](#)
- Sutton RS, Barto AG (1998) *Reinforcement learning: an introduction*. Cambridge, MA: Massachusetts Institute of Technology.
- Tatum WO 4th (2012) Mesial temporal lobe epilepsy. *J Clin Neurophysiol* 29:356–365. [CrossRef Medline](#)
- Tubridy S, Davachi L (2011) Medial temporal lobe contributions to episodic sequence encoding. *Cereb Cortex* 21:272–280. [CrossRef Medline](#)
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289. [CrossRef Medline](#)
- Watkins CJ, Dayan P (1992) Technical note: Q-learning. *Mach Learn* 8:279–292. [CrossRef](#)
- Wechsler D (1999) *WAIS III Escala de Inteligencia de Wechsler para Adultos-III*. Madrid: TEA.
- Wechsler D (2004) *WMS-III: Escala de memoria de Wechsler-III*. Madrid: TEA.
- Whishaw IQ (1998) Place learning in hippocampal rats and the path integration hypothesis. *Neurosci Biobehav Rev* 22:209–220. [CrossRef Medline](#)
- Zeidman P, Maguire EA (2016) Anterior hippocampus: the anatomy of perception, imagination and episodic memory. *Nat Rev Neurosci* 17:173–182. [CrossRef Medline](#)