Prefrontal-hippocampal dynamics involved in learning regularities across episodes

Christian F. Doeller[1], Bertram Opitz[1], Christoph M. Krick[2], Axel Mecklinger[1] and Wolfgang Reith[2]

[1]Experimental Neuropsychology Unit, Department of Psychology, Saarland University, Saarbrücken, Germany and [2]Department of Neuroradiology, Saarland University Hospital, Homburg, Germany

Using functional magnetic resonance imaging, the neural correlates of context-specific memories and invariant memories about regularities across episodes were investigated. Volunteers had to learn conjunctions between objects and positions. In an invariant learning condition, positions were held constant, enabling subjects to learn regularities across trials. By contrast, in a context-specific condition object–position conjunctions were trial unique. Performance increase in the invariant learning condition was paralleled by a learning-related increase of inferior frontal gyrus activation and ventral striatal activation and a decrease of hippocampus activation. Conversely, in the context-specific condition hippocampal activation was constant across trials. We argue that the learning-related hippocampal activation pattern might be due to reduced relational binding requirements once regularities are extracted. Furthermore, we propose that the learning-related prefrontal modulation reflects the requirement to extract and maintain regularities across trials and the adjustment of object–position conjunctions on the basis of the extracted knowledge. Finally, our data suggest that the ventral striatum encodes the increased predictability of spatial features as a function of learning. Taken together, these results indicate a transition of the relative roles of distinct brain regions during learning regularities across multiple episodes: regularity learning is characterized by a shift from a hippocampal to a prefrontal–striatal brain system.

Keywords: fMRI, hippocampus, learning, prefrontal cortex, ventral striatum

Introduction

Episodic memory refers to the ability to remember specific events set in a spatio-temporal context (Tulving, 1983). A large number of lesion studies (Zola *et al.*, 2000) and single cell recordings in animals (Fortin *et al.*, 2002), as well as neuropsychological investigations (Yonelinas *et al.*, 2002), intracranial EEG data (Fernández *et al.*, 1999) and functional imaging data in humans (Davachi *et al.*, 2003), suggest that the hippocampus as part of the medial temporal lobe memory system plays a crucial role in encoding and retrieval of episodic memories. One influential view on hippocampal function suggests that this brain region is critically involved in binding of different features and events that compose episodic memories (Eichenbaum, 2000). For example, it has been shown that hippocampal cells in the rat bind singular conjunctions of episodic features, such as a specific odor occurring at a particular position (Wood *et al.*, 1999). Furthermore, the hippocampal binding mechanism has to prevent interference between distinct episodes. One seminal framework (O'Reilly and Norman, 2002; Norman and O'Reilly, 2003) proposes that the hippocampus assigns distinct (pattern-separated) representations to different episodes to minimize interference. In a similar vein, Shastri (2002) proposes two levels of hippocampal binding: (i) entities occurring in the event are bound to the roles they fill in the event and (ii) all role–entity bindings defining an event are grouped together in order to separate them from role–entity bindings of other events. To exemplify the first binding type, imagine the situation when a person paints a picture in a studio. In the representation of this specific episode the role 'painter' is bound to this particular person, the role 'location' is bound to studio and the role 'object' is bound to picture. Given the episode that this person watches television in his apartment, the second level of binding is required to distinguish both episodes.

How does the brain represent repeated overlapping features of episodes? In addition to context-specific memories about particular episodes, individuals are able to acquire knowledge about regularities across such specific episodes (Shanks and St John, 1994). Several recent models have addressed this issue. For instance, it has been assumed that this knowledge about regularities comprises a measure of the probability that certain types of entities are bound to a certain role (Shastri, 2002). Moreover, Eichenbaum (2000) assumes that particular hippocampal neurons are specialized to represent features that are common across many episodes. In contrast, O'Reilly and Norman (2002) suggest that the rhinal and parahippocampal cortices represent regularities in the environment by assigning overlapping distributed representations to similar stimuli.

Recent neuroscientific research has begun to study the neural correlates of learning regularities across specific episodes. For instance, studies using category learning tasks in humans (Reber *et al.*, 1998; Strange *et al.*, 2001) and monkeys (Freedman *et al.*, 2001) have identified the lateral prefrontal cortex (PFC) as a core structure in mediating this kind of learning. In a study by Strange *et al.* (2001) subjects had to learn abstract rules that define category membership of four-letter strings (e.g. 'The first two letters are always identical'). The lateral PFC was selectively engaged following rule change. Furthermore, Freedman *et al.* (2001) found category-sensitive neurons in the monkey's lateral PFC. After defining new categories based on the same stimulus set, the same neurons then adaptively represented the new categories, indicating that the lateral PFC is involved in the flexible detection of regularities in this task.

Another line of evidence for prefrontal involvement in learning regularities comes from artificial grammar learning research, where subjects acquire abstract knowledge about the rules of a finite-state grammar (Cleeremans *et al.*, 1998). Recent functional imaging studies have implicated the lateral PFC in artificial grammar learning (Fletcher *et al.*, 1999; Opitz and Friederici, 2003). The lateral PFC also supports sequence learning, as indicated by findings from several functional magnetic resonance imaging (fMRI) studies (Schendan *et al.*,

2003; Aizenstein *et al.*, 2004). In these studies, subjects acquire knowledge about a regular sequence of stimulus events. Moreover, the lateral PFC is associated with causal associative learning (Fletcher *et al.*, 2001) and with the detection of abstract sequence violations (Huettel *et al.*, 2002). It also shows learning-related activity during arbitrary rule learning (Toni *et al.*, 2001; Wallis *et al.*, 2001). Despite the different tasks used in the above-mentioned studies, taken as a whole, these latter findings underscore the importance of the lateral PFC in learning task-relevant regularities across different episodes, leading to the acquisition of abstract knowledge structures. However, it should be noted that the exact location of lateral PFC activation varies between studies, depending on stimulus properties and task requirements.

Using fMRI, we investigated the neural correlates of context-specific and invariant memories, i.e. memory for invariant features of episodes, thereby bringing together two research fields in cognitive neuroscience, episodic memory and rule learning. More specifically, we were interested in brain areas responsible for the acquisition of invariant memories by regularity learning. In each trial subjects were required to learn four sequential object–position conjunctions (see Materials and Methods and Fig. 1). Subsequently, they had to indicate whether or not a probe stimulus was identical to one of the four object–position conjunctions (i.e. whether or not there is an exact match of both stimulus features, the object and the position). In an invariant learning condition, the positions were invariant within a block. That is, different objects were presented at the same four positions in each trial of an experimental block, thereby enabling the extraction of spatial regularities and the formation of invariant memories. In a context-specific condition, by contrast, objects and positions were variable within experimental blocks, i.e. there were no spatial regularities across trials, requiring the processing of unique object–position combinations from trial to trial.

We expected enduring hippocampal activity in the context-specific condition, since variable objects have to be bound to variable positions in each trial of the blocks (constant binding requirements). In the invariant learning condition, by contrast, we hypothesized reduced hippocampal activation as a function of time within blocks. Here, variable objects have to be bound to invariant positions (reduced binding requirements). Moreover, we expected a hippocampal-prefrontal shift of activity in the invariant learning condition. More precisely, we assumed that the lateral PFC is involved in the extraction of invariant spatial features and the flexible adjustment of the extracted knowledge to the requirements of the ongoing task block. Thus, we expected increased PFC activity in the invariant learning condition.
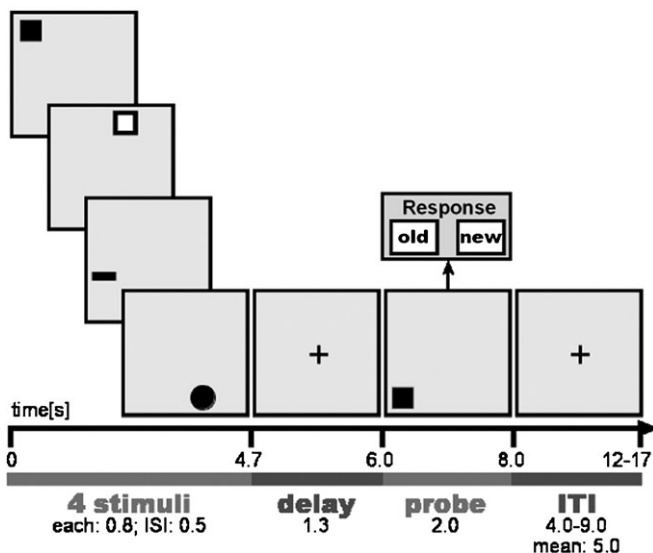
Materials and Methods

Subjects

Ten subjects participated in the study (aged 21-32 years, mean age = 27 years, six males). All subjects were right-handed with normal or corrected-to-normal vision and were paid for participating. Informed consent was obtained before scanning. All participants reported to be in good health with no history of neurological disease. One subject had to be excluded from further analysis due to technical problems during fMRI acquisition.

Stimuli, Task and Design

Sixteen stimuli, which consisted of basic shapes (e.g. square, circle) were used as stimulus materials and were presented within a 4 × 4 grid,
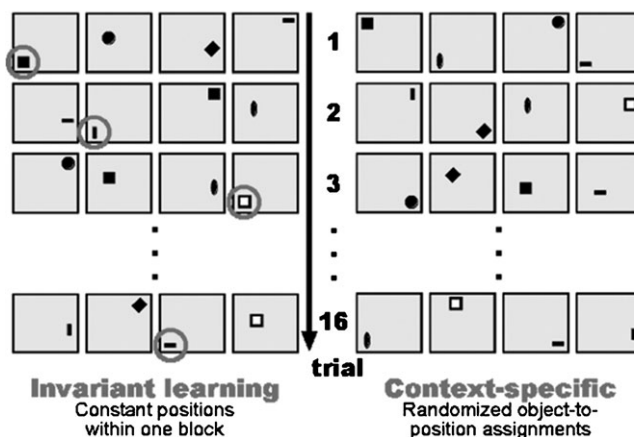

Figure 1. Trial and block structure of the experiment. (*A*) Trial structure of the experiment. Subjects learned four sequential relations among objects and positions within a 4 × 4 matrix (grid not shown). Each stimulus was presented for 800 ms at a particular position with an inter-stimulus interval (ISI) of 500 ms. Subsequently, subjects were probed and had 2000 ms to determine whether the probe stimulus was identical to one of the four object–position combinations (i.e. an exact match of both stimulus features, the object and the position). Visual feedback was given for 500 ms. In the present example, the probe comprised an old object at a new position, therefore requiring a 'new' response. (*B*) Block structure of the experiment. In the invariant learning condition, positions were invariant in each trial of an experimental block, whereas in the context-specific condition, objects and positions were variable in each trial. For illustration, one invariant position is highlighted by circles. One experimental block comprised 16 trials.

i.e. there were 16 possible positions. Stimuli were back projected onto a translucent screen that participants viewed through a mirror during fMRI acquisition.

Within one experimental trial four different stimuli were presented sequentially at four different positions. Each stimulus was delivered for 800 ms with a 500 ms inter-stimulus interval (Fig. 1*A*). After a short delay of 1300 ms a probe stimulus was presented for 2000 ms at a particular position within the grid. Participants had to indicate whether or not the current object–position conjunction (the probe stimulus) was identical to one of the four object–position combinations seen before. Responses were delivered by a button press with the right or left index finger (2000 ms response window) by means of MR-compatible response keys.

Response-to-hand mappings were counterbalanced across subjects. Visual feedback (500 ms) was given by means of the color of a fixation cross (green = correct, red = false, yellow = timeout), immediately after probe offset. Probes in each block (see below) comprised 50% old (old object at old position) and 50% new object–position conjunctions (three equally distributed categories: old object/new position, new object/old position and new object/new position). An exponentially distributed intertrial interval (ITI) of 4–9 s (mean = 5 s), which was varied in steps of 1 s was used in order to get an optimal tradeoff between detectability and estimation efficiency of the BOLD response (Hagberg *et al.*, 2001; Birn *et al.*, 2002).

After subjects had practiced 20 trials outside and ten trials in the scanner, they performed 224 trials during the whole experiment. One experimental session comprised seven blocks (16 trials each) in the context-specific condition and seven blocks in the invariant learning condition. Blocks of both conditions were presented in randomized order with the constraint that participants had to perform at most two blocks of the same condition in succession. Subjects were informed about the beginning of a new block. The sequence of blocks was balanced across subjects. In the invariant learning condition, four positions were held constant within one experimental block whereas in the context-specific condition, objects and positions were variable across trials (Fig. 1*B*). Within one invariant learning block the temporal order of the four positions was randomized across trials. The fixed configuration of four positions changed from block to block in the invariant learning condition. All other parameters were held constant across both conditions. Before scanning, participants were instructed that the detection and application of a rule would make the task more easy and that this rule could change when a new block starts. Participants were unaware of the experimental manipulation, i.e. the existence of two different conditions.

Behavioral Analysis
To get a first estimate for learning, mean *Pr* values (proportion of hits – proportion of false alarms) (Feenan and Snodgrass, 1990) were computed. Additional analyses were conducted for hits and false alarms. Furthermore, the false alarms were broken down by the type of false alarm committed. In all behavioral analyses, mean performance measures in the first and the last eight trials of each block (i.e. the two halves of the blocks) were averaged across all blocks separately for both conditions.

Imaging Parameters
A Siemens SONATA MR scanner (Erlangen, Germany) operating at 1.5 T with a standard circularly polarized whole head coil was employed to acquire both T1-weighted structural images and T2*-weighted BOLD-sensitive functional images. High-resolution (1 mm^3 voxel size) structural images were acquired using a 3-D MP RAGE sequence. Functional data were acquired using a gradient-echo EPI pulse sequence, with the following parameters: T_R = 1.8 s, T_E = 50 ms, flip angle = 85°, slice thickness = 4 mm, interslice gap = 1 mm, in-plane resolution = 3.5 × 3.5 mm, field of view = 224 mm, 20 axial slices parallel to anterior commissure–posterior commissure plane. The first four volumes were discarded to allow for T1 equilibration.

Data Preprocessing
FMRI time-series analysis was performed with Statistical Parametric Mapping (SPM2; Wellcome Department of Imaging Neuroscience, London, UK). To account for the different sampling time of the slices, voxel time series were corrected using sinc interpolation and resampled using the middle slice as a reference point. All functional volumes were motion corrected. Using a least squares approach and a six-parameter rigid body spatial transformation, realignment parameters were estimated. According to these determined parameters all volumes were resliced to the first volume using B-spline interpolation. After coregistering anatomical images to the mean functional image, the resulting images were normalized (Ashburner and Friston, 1999) to the standard T1 template based on the MNI reference brain (Cocosco *et al.*, 1997), using a 12-parameter affine transformation along with a nonlinear transformation (cosine basis functions). Based on the determined parameters the normalization algorithm was then applied to the func-

tional volumes. Finally, the normalized images were resampled into 2 mm isotropic voxels and spatially smoothed with an isotropic 8 mm full-width half-maximum Gaussian kernel.

Basic Statistical Analysis
Statistical analysis was performed in two stages in a mixed-effects model. For each subject, neural activity in each trial for both conditions was modeled by convolving a stimulus function, representing the onset of each sample phase with a canonical hemodynamic response function (HRF) (Friston *et al.*, 1998) (duration = 4.7 s), to cover the whole sample phase (epoch-related). The sample phase comprised the four sequential object–position conjunctions in each trial, including the inter-stimulus interval. To increase the power of our model, we additionally modeled neural activity for the probe stimulus in each trial for both conditions, using an event-related canonical HRF. For this purpose, an event train of δ-functions, time-locked to the onset of each probe stimulus was convolved with the canonical HRF. We assumed that learning-related neural activity should be optimally reflected in the sample phase, when subjects encode and maintain four object–position conjunctions. Thus, all analyses were restricted to the sample phase. The time series in each voxel were high-pass filtered to 1/128 Hz to remove low-frequency noise and proportionally scaled to a grand mean of 100 over voxels to remove effects due to global intensity fluctuations. Parameters for each covariate were estimated by a least-mean-squares fit of the model to the time series using a subject-specific fixed-effects model within the general linear model. During the estimation procedure serial correlations were estimated with a restricted maximum likelihood (ReML) algorithm using an AR(1) plus white noise model. In SPM2, the ReML estimates (hyperparameters) are then used to correct for non-sphericity (Friston *et al.*, 2002). Parameter estimates for the linear contrasts of interest (i.e. the direct contrasts of the invariant learning condition versus the context-specific condition and *vice versa*) entered into a second-level analysis treating subject as a random effect, using a one-sample *t*-test against a contrast value of zero at each voxel (Holmes and Friston, 1998). MNI-coordinates (Cocosco *et al.*, 1997) of all reported activations have been transformed to the canonical Talairach space (Talairach and Tournoux, 1988). In the direct contrasts between both conditions, statistical parametric maps (SPMs) were thresholded at $P < 0.0005$, uncorrected for multiple comparisons. In addition, an extent threshold of 5 voxels was used to emphasize coherent activation.

Learning-related Modeling
In a second analysis step, we conducted a condition × time interaction analysis and two parametric fMRI analyses to directly investigate learning-related modulations of brain activity. To get an initial estimate of a differential learning-related activation pattern in both conditions, we conducted a 2 × 2 analysis of variance (ANOVA) with the factors condition and time [mean blood oxygen level-dependent (BOLD) activity for trials 1–8 and 9–16 of the experimental blocks, separately for both conditions]. This analysis was restricted to the brain regions showing a main effect of condition in the basic statistical analysis ($P < 0.05$, small volume corrected; see below). Given our a-priori hypotheses with respect to the hippocampus, this analysis was also conducted for both hippocampi ($P < 0.05$, small volume corrected; see below).

In two subsequent parametric fMRI analyses, we examined learning-related activation pattern in more detail by incorporating subjects' individual performance and learning functions on a trial-by-trial level into the imaging analysis. In the first parametric analysis, we tested our a-priori hypotheses that (i) the hippocampus shows a learning-related decrease in activity in the invariant learning condition and (ii) constant activity in the context-specific condition. In this analysis subject-specific time series of each experimental block were weighted with subjects' behavioral performance measure. To increase the sampling rate of the learning functions, mean *Pr* values for four consecutive trials in each block (i.e. trials 1–4, 5–8, 9–12 and 13–16, respectively; see Fig. 4*D*, upper panel, for examples) were averaged across all blocks, separately for both conditions, resulting in four time bins for each subject and condition. The model functions were derived from each individual's learning curve by fitting a logarithmic function $y = a \times \ln(t_i) + b$ separately for both conditions (see Fig. 4*D*, lower panel for examples). The two original condition-specific stimulus functions

were multiplied by these parametric modulation functions, leading to additional regressors (Büchel *et al.*, 1998), either reflecting learning-related increase or decrease in the invariant learning condition or continuously enduring hemodynamic activity in the context-specific condition. Linear contrasts of the parameter estimates for each regressor were calculated for each subject and brought to the second-level random effects analysis. To test our specific hypotheses about the differential hippocampal involvement in both conditions, we performed a conjunction analysis (Friston *et al.*, 1999) across subjects. Here, we used linear contrasts of the parameter estimates for the respective regressors in both conditions, i.e. the regressor including a decreasing model function in the invariant learning condition and the regressor including a constant model function in the context-specific condition. This analysis was restricted to the hippocampus proper. Hippocampal regions of interest were determined a priori according to a detailed neuroanatomy atlas (Warner, 2001). An appropriate mask image was generated using the software package MRIcro. Statistical tests were corrected for multiple comparisons ($P < 0.05$, small volume corrected) (Worsley *et al.*, 1996).

In a second parametric analysis, we examined whether brain regions showing a main effect of condition in the direct comparisons also exhibit learning-related changes of activation in the invariant learning condition. Thus, learning-related increase and decrease of activity in this condition was modeled. In this analysis, SPMs from the invariant-learning versus context-specific contrast and the context-specific versus invariant-learning contrast served as mask images, respectively ($P < 0.05$, small volume corrected).

Results

Behavioral Results

Figure 2*A* shows mean *Pr* values (Feenan and Snodgrass, 1990), separately for the context-specific and invariant learning condition, collapsed across experimental blocks and subjects. In the invariant learning condition, *Pr* values increased within task blocks. This was not the case in the context-specific condition. This observation was confirmed by a two-way repeated-measures ANOVA with the factors condition (context-specific

versus invariant learning) and time (two levels: mean *Pr* values for eight consecutive trials each). This analysis revealed a main effect of time [$F(1,8) = 6.88$, $P < 0.05$] and a condition × time interaction [$F(1,8) = 24.66$, $P < 0.005$]. Additional one-way ANOVAs separately for both conditions showed a significant effect of time in the invariant learning condition [$F(1,8) = 20.58$, $P < 0.005$] but not in the context-specific condition [$F(1,8) = 5.27$, $P < 0.1$].

To elucidate in more detail the learning mechanisms and performance pattern in both condition, additional analyses were conducted separately for the mean hit and false alarm rates (Fig. 2*B*,*C*). There was a condition × time interaction for hits [$F(1,8) = 7.04$, $P < 0.05$] and for false alarms [$F(1,8) = 5.85$, $P < 0.05$]. Furthermore, a marginally significant increase of hits [$F(1,8) = 3.48$, $P < 0.1$] and a decrease of false alarms [$F(1,8) = 6.83$, $P < 0.05$] within blocks could be observed in the invariant learning condition. By contrast, in the context-specific condition, mean hit rate [$F(1,8) = 1.27$, $P > 0.2$] and mean false alarm rate [$F(1,8) < 1$] remained constant within blocks.

In our task, subjects can commit a false positive response by classifying as old an old object at a new position, a new object at an old position and a new object at a new position. Given this, we assumed that learning, i.e. the extraction of spatial regularities, should be most clearly revealed by a reduction of false alarms to trials that include new positions, since subjects could reject these probes solely on the basis of their knowledge about the invariant positions. To examine this, the false alarms in both conditions were broken down by the three types of errors described above (Fig. 2*D*–*F*). These supplementary analyses revealed a main effect of condition [$F(1,8) = 38.95$, $P < 0.001$] and a condition × time interaction [$F(1,8) = 7.69$, $P < 0.05$] for false alarms to old objects at new positions (Fig. 2*D*), but neither for false alarms to new objects at old positions [main effect: $F(1,8) = 1.31$, $P > 0.2$; interaction: $F(1,8) < 1$; Fig. 2*E*] nor for false alarms to new objects at new positions [main effect and
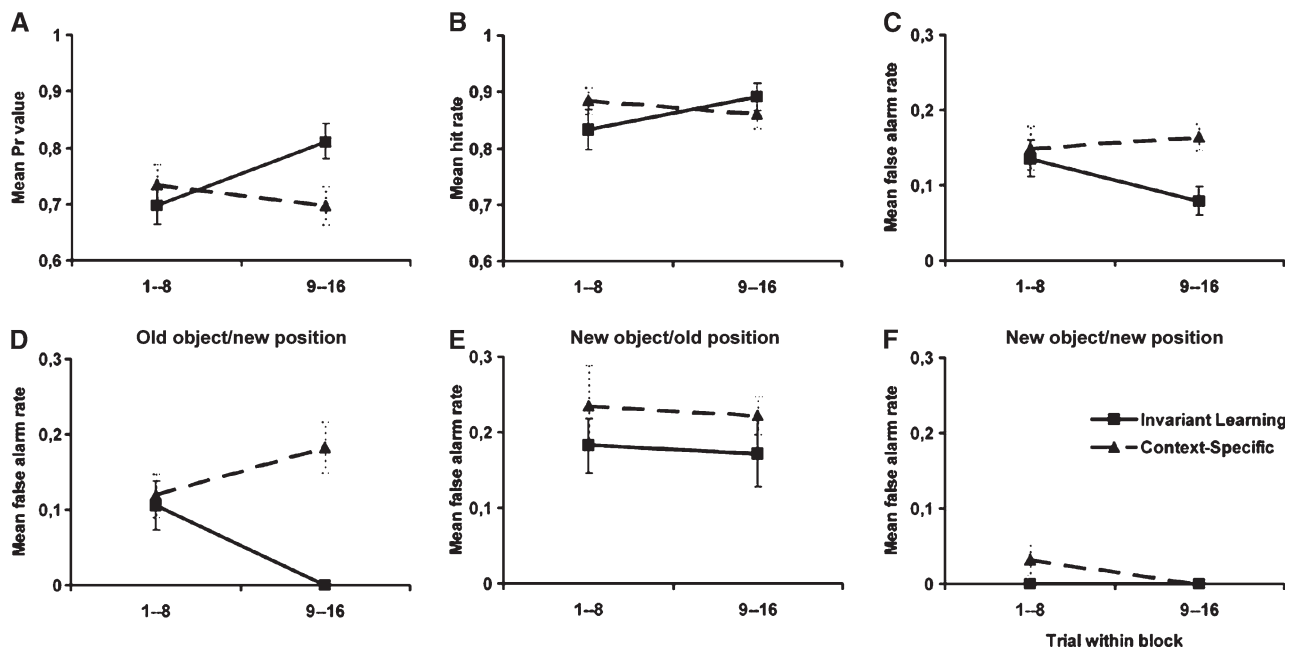

Figure 2. Behavioral results. Performance measures +— SEM) are plotted as a function of time within experimental blocks for the invariant learning (solid) and context-specific condition (dashed), collapsed across all blocks and subjects. (*A*) Mean *Pr* values, (*B*) mean hit rate, (*C*) mean false alarm rate. (*D*–*F*) The false alarms are further divided in three categories, depending on the specific probe type, i.e. the mean false alarm rate for probes including (*D*) an old object at a new position, (*E*) a new object at an old position and (*F*) a new object at a new position, respectively. The performance measure was averaged across trials 1–8 and 9–16, respectively.

interaction: $F(1,8) = 2.29$, $P > 0.1$; Fig. 2F]. Note that the latter type of errors was rarely committed, causing this effect to be non-significant (floor-effect). Moreover, an analysis separately for both conditions revealed that the false alarm rate to old objects at new positions decreased in the invariant learning condition [$F(1,8) = 10.36$, $P < 0.05$] but not in the context-specific condition [$F(1,8) = 2.72$, $P > 0.1$]. This differential modulation of the false alarm rates in the invariant learning condition confirms our initial hypothesis that learning takes place in the form of strengthening the representation of invariant positions within blocks.

Imaging Results: Basic Contrasts

In a first step, we calculated direct contrasts between the context-specific and the invariant learning condition (Fig. 3, Table 1). In the direct comparisons, only regions that survived a statistical threshold of $P < 0.0005$ (uncorrected) were considered significant. Regions that exhibited greater activation for invariant learning trials than for context-specific trials were considered to be sensitive for the formation of invariant memories. Several prefrontal regions, including the inferior portion of the left middle frontal gyrus [Brodmann area (BA) 46; peak Talairach coordinates x, y, z: −40, 30, 24], the opercular part of the right inferior frontal gyrus (BA 45; 40, 28, 19), the triangular part of the left inferior frontal gyrus (BA 45; −57, 20, 17) and the right inferior frontal sulcus (BA 9; 36, 21, 25) were activated in the invariant learning relative to the context-specific condition. In addition to this lateral prefrontal pattern, this contrast also revealed activation in the right ventral striatum (26, 2, −7), the right inferior parietal lobule (BA 40; 51, −42, 44 / 42, −56, 45) and the right lingual gyrus (BA 30; 22, −47, 2). The comparison between the context-specific and the invariant learning condition revealed enhanced activity in the left posterior cingulate cortex (BA 31; −20, −33, 48) and along the right superior occipital sulcus (BA 19; center activity: 34, −74, 28). All experimental trials were included in this analysis. When solely considering correct responses in the analysis, a highly similar pattern for both contrasts could be observed, with the exception that the left inferior parietal lobule additionally showed a significant activation in the contrast between the invariant learning condition and the context-specific condition (see Table 2 for details).

Imaging Results: Learning-Related Activity

In accordance with the behavioral analysis, we conducted a 2 × 2 ANOVA with the factors condition and time (BOLD activity for trials 1–8 and trials 9–16 of the experimental blocks, separately for both conditions) to get a first estimate of differential learning-related activation pattern in both conditions. This analysis was restricted to the hippocampus and to the activation foci found in the direct comparisons between conditions (see Materials and Methods). A subset of the regions showing a main effect of condition in the direct comparisons also exhibited a significant condition × time interaction, including the triangular part of the left inferior frontal gyrus (−57, 20, 17; z_{max} = 1.75), the left posterior cingulate cortex (−16, −29, 46; z_{max} = 2.70), the right inferior parietal lobule (48, −46, 48; z_{max} = 3.12) and the right lingual gyrus (24, −47, 1; z_{max} = 2.48). Moreover, this analysis revealed a significant interaction between condition and time in the right (24, −8, −10; z_{max} = 2.70 / 38, −20, −14;
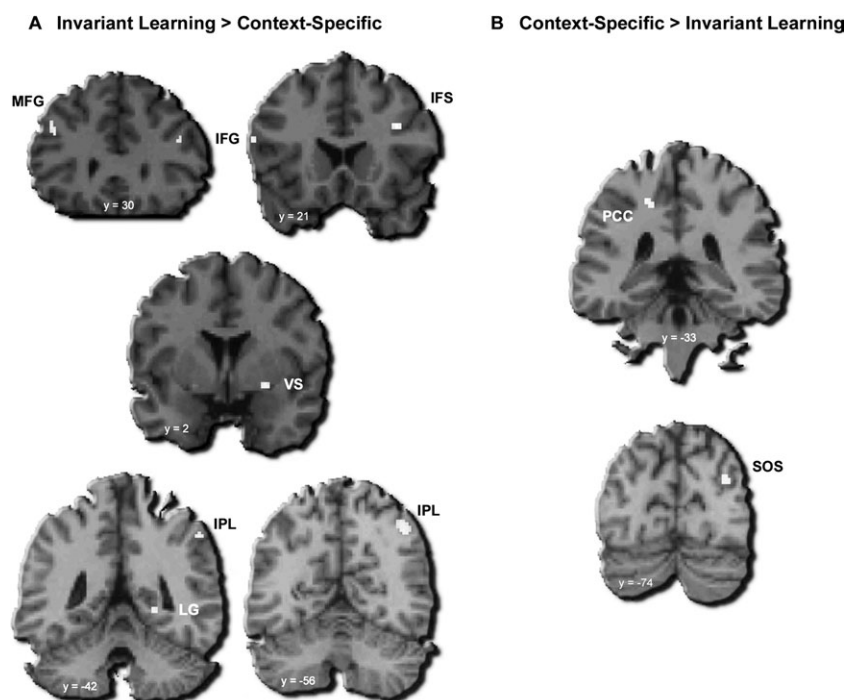

Figure 3. Direct comparisons between conditions. SPMs superimposed on coronal sections of the MNI T1-weighted MRI template showing activated regions for the contrasts comparing the invariant learning and the context-specific condition. Talairach y-coordinate is given below each image, respectively. (A) Trials in the invariant learning condition revealed greater hemodynamic activity than trials in the context-specific condition in the left middle frontal gyrus (MFG), the opercular part of the right inferior frontal gyrus (IFG), the triangular part of the left IFG, the right inferior frontal sulcus (IFS), the right ventral striatum (VS), the right inferior parietal lobule (IPL) and the right lingual gyrus (LG). (B) The left posterior cingulate cortex (PCC) and the right superior occipital sulcus (SOS) showed stronger BOLD responses in the context-specific condition as compared with the invariant learning condition (for further details, see Table 1).

Table 1

fMRI activation foci: direct contrasts (all trials)

Region	BA	Hemisphere	Talairach coordinates			Z-score
			x	y	z	
Invariant learning > context-specific						
Inferior middle frontal gyrus	46	L	−40	30	24	3.66
Inferior frontal gyrus (opercular part)	45	R	40	28	19	3.50
Inferior frontal gyrus (triangular part)	45	L	−57	20	17	3.86
Inferior frontal sulcus	9	R	36	21	25	4.30
Ventral striatum	–	R	26	2	−7	3.67
Inferior parietal lobule	40	R	51	−42	44	3.58
	40	R	42	−56	45	4.02
Lingual gyrus	30	R	22	−47	2	3.81
Context-specific > invariant learning						
Posterior cingulate cortex	31	L	−20	−33	48	4.11
Superior occipital sulcus	19	R	34	−74	28	3.76

Regions activated in the direct contrasts between the context-specific and the invariant learning condition (from anterior to posterior), described in terms of Brodmann area (BA), hemisphere (L, left; R, right), Talairach coordinates (mm; transformed from the MNI-space) and peak Z-score. All regions pass the statistical threshold of $P < 0.0005$ (uncorrected). All trials were included in this analysis.

Table 2

fMRI activation foci: direct contrasts (correct trials)

Region	BA	Hemisphere	Talairach coordinates			Z-score
			x	y	z	
Invariant learning > context-specific						
Inferior middle frontal gyrus	46	L	−40	28	24	3.75
Inferior frontal gyrus (opercular part)	45	R	40	28	19	3.66
Inferior frontal gyrus (triangular part)	45	L	−55	26	19	3.88
Inferior frontal sulcus	9	R	36	21	25	4.22
Ventral striatum	–	R	28	4	−5	3.83
Inferior parietal lobule	40	R	53	−41	44	3.45
	40	L	−42	−45	41	3.77
	40	R	40	−56	45	4.10
Lingual gyrus	30	R	22	−47	2	3.82
Context-specific > invariant learning						
Posterior cingulate cortex	31	L	−18	−35	46	4.06
Superior occipital sulcus	19	R	34	−74	28	3.70

Regions activated in the direct contrasts between the context-specific and invariant learning condition. Only correct trials were included in this analysis (for further details, see Table 1 legend).

z_{max} = 2.53) and left (−18, −10, −13; z_{max} = 3.41 / − 26, −18, −16; z_{max} = 4.46) hippocampus.

On the basis of the interaction between condition and time, suggesting a differential involvement of several brain regions in both conditions as a function of time, we further examined learning-related activation pattern in more detail by means of parametric fMRI analyses. For this purpose we weighted the fMRI time series separately for both conditions with individually fitted logarithmic learning functions. Learning-related changes in hemodynamic activation could thereby be modeled explicitly (see Materials and Methods). All individual learning functions were well approximated by a logarithmic function [mean functions and goodness of fit: $y = 0.10 \times \ln(t_i) + 0.67$; R^2 = 0.72 (invariant learning condition); $y = −0.07 \times \ln(t_i) + 0.72$; R^2 = 0.71 (context-specific condition)]. In a first parametric analysis, we tested our a-priori hypotheses about a differential learning-related hippocampal activation pattern for both conditions. For this purpose, we conducted the following conjunction analysis (Friston *et al.*, 1999), restricted to both hippocampi ($P < 0.05$, small volume corrected): learning-related decrease in the invariant learning condition and constant activity in the context-

specific condition. This conjunction analysis revealed an activation of the right hippocampus (34, −12, −16; Fig. 4A; Table 3). As apparent from Figure 4A (middle panel), the hippocampus showed continuously enduring activation within blocks of the context-specific condition. To further control whether the hippocampus was responsive across all trials in the context-specific condition, we conducted a *t*-test against the null-hypothesis of no activation for the peak BOLD response of the hippocampus. For each subject, the peak BOLD response was determined for each trial of the blocks and subsequently averaged across all trials. This analysis revealed a significant activation of the hippocampus [$t(8) = 3.84$, $P < 0.01$].

A second parametric analysis tested whether brain regions showing a main effect of condition in the direct comparisons also exhibit a learning-related activation pattern in the invariant learning condition. Thus, this analysis was restricted to the activation foci found in the direct comparisons between conditions ($P < 0.05$, small volume corrected; see Methods). The triangular part of the left inferior frontal gyrus (BA 45; −53, 24, 19) and the right ventral striatum (28, 2, −7) showed learning-related increases of activity within the experimental blocks (Fig. 4B). By contrast, right inferior parietal lobule (BA 40; 38, −56, 47) activity decreased as a function of learning within blocks (Fig. 4C). Moreover, a learning-related decrease of activation could be observed in the superior occipital sulcus (BA 19; 34, −76, 26). None of the additional regions found in the direct comparison between both conditions exhibited a learning-related activation pattern in this analysis.

Discussion

Our behavioral data show that subjects learned spatial regularities across trials and did benefit from this knowledge in the invariant learning condition. This was reflected in increased *Pr* values, due to increasing correct detection and decreasing false alarm behavior relative to the context-specific condition during the time-course of experimental blocks. The data also indicate that learning is based on reducing false alarm responses to those probes that include new positions. Here, subjects benefit most of all from learning spatial regularities, probably based on an enhanced selectivity of the spatial representations (Yeshurun and Carrasco, 1998).

The fMRI results point to dissociable neural correlates for context-specific and invariant memories. In the parametric analysis the right hippocampus showed continuously enduring activity in the context-specific condition, in which subjects encode variable objects at variable positions. Importantly, when subjects benefit from learning invariant spatial features across trials, right hippocampal activity decreased in the invariant learning condition. Furthermore, we could dissociate a prefrontal-striatal-parietal network supporting learning regularities. The lateral prefrontal cortex and the ventral striatum showed a learning related increase of activity, whereas activity in the inferior parietal lobule decreased as a function of learning.

Hippocampus and Pattern-separated, Relational Representation

It has been proposed that the hippocampus uses sparse, pattern-separated representations to encode arbitrary conjunctions or bindings of features defining an episode (O'Reilly and Norman, 2002). This mechanism results in highly distinct,
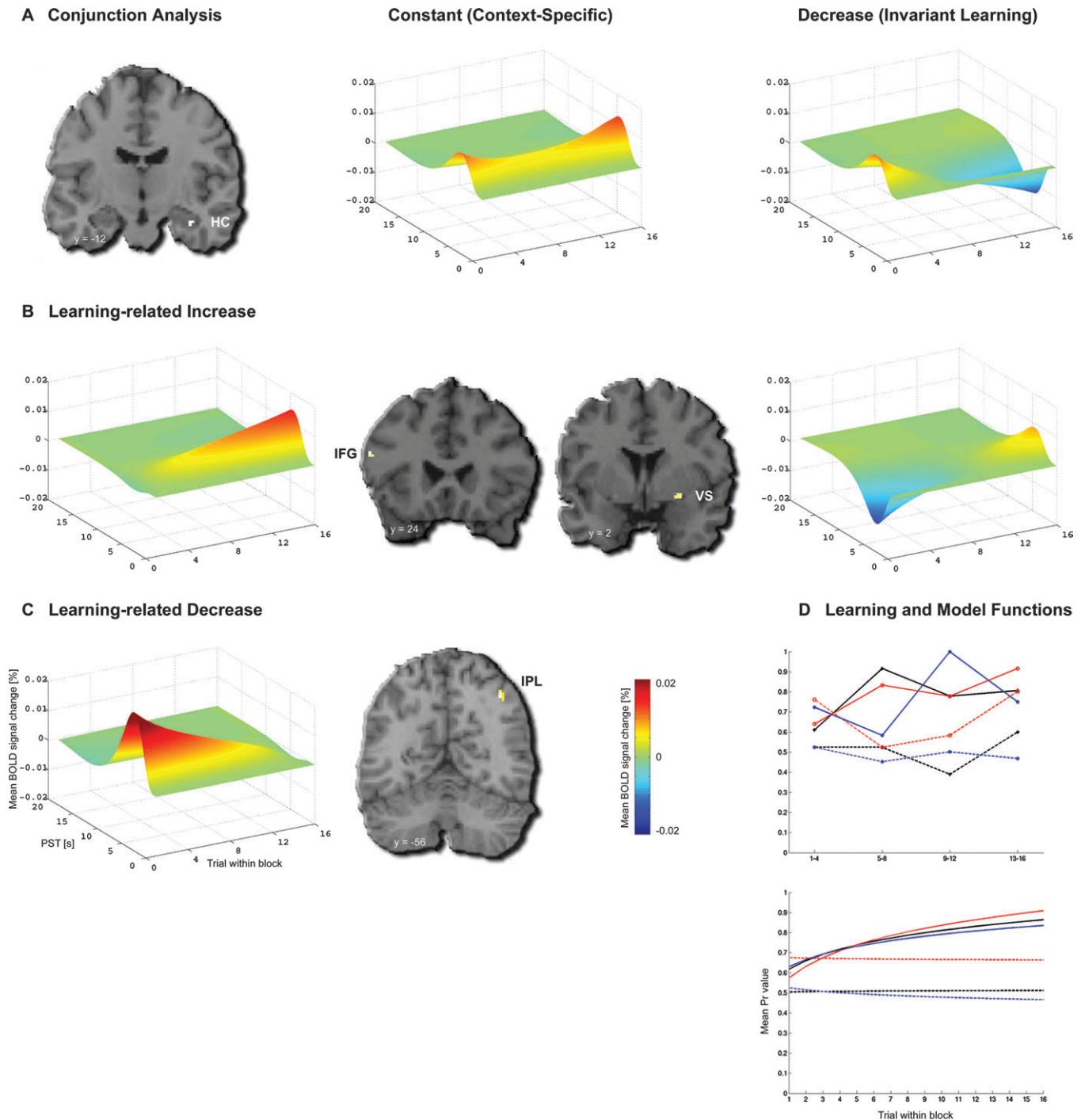
Figure 4. Brain regions showing a learning-related activation pattern. (*A*) Results of the conjunction analysis. The right hippocampus (HC) showed constant activity in the context-specific condition and decreased activity during the time course of learning in the invariant learning condition. (*B*) Regions demonstrating increased learning-related activity in the invariant learning condition included the opercular part of the left IFG and the right VS. (*C*) Furthermore, the parametric analysis revealed a learning-related activation decrease in the right IPL. Time courses of the best-fitting parametric BOLD response relative to grand mean over voxels, time-locked to the sample phase onset against post-stimulus time (PST), are plotted across trials of the experimental blocks. The response is collapsed across experimental blocks and averaged across participants. For the conjunction analysis (*A*), parametric BOLD responses are plotted for the hippocampus separately for the context-specific (middle panel) and the invariant learning condition (right panel). For (*B*) and (*C*), parametric responses are depicted on the left and right side of the structural scan, respectively (for further details, see Table 3). (*D*) Examples of learning functions (upper panel) and model functions (lower panel) from three selected subjects (black, red, blue), separately for the invariant learning (solid) and the context-specific condition (dashed). Learning functions were derived by averaging mean *Pr* values for four consecutive trials across all blocks separately for both conditions. The model functions were derived from the individual learning functions by fitting a logarithmic function $y = a \times \ln(t_i) + b$.

non-overlapping representations of episodes and minimizes interference between episodes. In a similar vein, Eichenbaum (2000) argues that the hippocampus is responsible for binding multiple inputs, and by this optimally represents the relations

between temporally and spatially disparate features comprising a complex episode (relational memory framework) (Cohen and Eichenbaum, 1993; Cohen *et al.*, 1999). Accordingly, in our task four different objects have to be bound to four different

Table 3

fMRI activation foci: learning-related activity

Region	BA	Hemisphere	Talairach coordinates			Z-score
			x	y	z	
(1) Conjunction analysis						
Hippocampus	–	R	34	−12	−16	1.73
(2) Learning-related increase						
Inferior frontal gyrus (triangular part)	45	L	−53	24	19	2.06
Ventral striatum	–	R	28	2	−7	1.88
(3) Learning-related decrease						
Inferior parietal lobule	40	R	38	−56	47	2.37
Superior occipital sulcus	19	R	34	−76	26	2.34

Regions showing a learning-related activation pattern. (1) Conjunction analysis: learning-related decrease (invariant learning condition) and constant activity (context-specific condition). (2–3) Regions with a learning-related (2) increase and (3) decrease in the invariant learning condition. SPMs were thresholded at $P < 0.05$ (corrected), using small volume corrections (for further details, see Table 1 legend).

positions and these four object–position bindings have further to be bound together in each trial to provide a complete, relational representation of each episode. In accordance with both frameworks (Eichenbaum, 2000; O'Reilly and Norman, 2002), in the present experiment the hippocampus was activated continuously across experimental blocks in the context-specific condition. Here, each trial comprised a unique episode, i.e. the probability of each possible object–position conjunction was constant across trials, resulting in highly variable feature conjunctions in each trial. Consequently, constant hippocampal relational binding operations were required to encode unique episodes into separated memory representations.

In the invariant learning condition, by contrast, there were four invariant positions in each block. Here, the relational binding requirements of each episode decreased during the time-course of each block, since the probability for specific object–position conjunctions was substantially increased (e.g. a specific object presented at one of the four positions held constant in a block), whereas the probability for other object–position conjunctions was reduced to zero (e.g. an object presented at one of the remaining 12 positions never occurring in one block). Thus, invariant positions in object–position conjunctions might cause a reduced relational representation of the current episode. As the objects can be bound to the same positions within a block, learning presumably results from facilitated object–position binding. The decreasing hippocampal activation in the invariant learning condition may thus reflect these lower relational binding requirements. It is important to note that in the invariant learning condition a singular object–position conjunction is presented only once in each block, i.e. is not repeated within one block. By this, our experimental manipulation did not entail the repetition of a specific object–position conjunction. Thus, the decrease of hippocampal activation cannot be attributed to a higher amount of repetition of object–position conjunctions in the invariant learning condition.

This line of arguments is confirmed by previously observed decreased hippocampal activity during the time-course of probabilistic cue–outcome learning (Poldrack et al., 2001), associative learning (Zeineh et al., 2003), artificial grammar learning (Opitz and Friederici, 2003) and sequence learning (Grafton et al., 1995; Schendan et al., 2003). For instance, Schendan et al. (2003) showed a learning-related modulation of hippocampal activity using a serial reaction time task. Subjects learned complex sequential structures of stimuli and responses.

The authors observed activation of the hippocampus that was more pronounced during the initial phases of learning than in the final phase. In line with the relational memory account (Cohen and Eichenbaum, 1993; Eichenbaum, 2000), Schendan et al. (2003) argue that the hippocampus is involved in the acquisition of higher-order associations, i.e. relations among temporally discontiguous events, most pronounced at the beginning of learning. It is conceivable that in the Schendan et al. Study, as well as in our study, relational representations of episodes get reduced due to overlapping episodic features and by this less hippocampal relational processing is required.

Moreover, the hippocampal activation was right-lateralized in the present study. Numerous studies have demonstrated that the hippocampus mediates the processing of spatial relationships (Eichenbaum et al., 1999; Burgess et al., 2002). Neuroimaging and neuropsychological studies suggest a preferential role of the right hippocampus in spatial memory (Smith and Milner, 1981; Maguire et al., 2003) and spatial navigation (Grön et al., 2000; Maguire et al., 2000; Burgess et al., 2001). Consistent with these findings, the present right hippocampal activation decrease in the invariant learning condition may reflect the lower spatial processing requirements as compared with the context-specific condition, i.e. the processing of invariant spatial features from trial to trial. Given the small set of 16 objects and 16 positions used in the present study, it could be argued that learning was not restricted to the invariant learning condition but also took place in the context-specific condition. To test this hypothesis, we conducted a post-hoc analysis, contrasting mean performance (Pr values) in blocks 1–3 with mean performance in blocks 5–7 for the context-specific condition. This analysis revealed no differences between both blocks [$t(8) < 1$; mean Pr: 0.65 ± 0.05 (block 1–3); 0.66 ± 0.06 (block 5–7)]. Furthermore, in an additional fMRI analysis we contrasted hemodynamic responses in block 1–3 with block 5–7 (and vice versa) in the context-specific condition. Hippocampal activation did not differ significantly between the first and the last blocks ($P < 0.05$, small-volume corrected). These supplementary analyses suggest that learning was absent or negligible and hippocampal activation did not change across blocks in the context-specific condition.

Brain regions exhibiting greater activity for context-specific as compared with invariant learning trials include the posterior cingulate cortex (PCC) and the superior occipital sulcus (SOS). The PCC is part of the extended hippocampal diencephalic system (Aggleton and Brown, 1999). The integrity of this system is a prerequisite for successful memory operations. However, the precise function of the PCC is an issue of current debate. Furthermore, the SOS as part of the dorsal visual processing stream (Ungerleider and Mishkin, 1982) might be recruited due to processing variable spatial features from trial to trial in the context-specific relative to the invariant learning condition. In accordance with this view, SOS activity decreased during the time course of experimental blocks in the invariant learning condition, since spatial processing requirements get reduced.

Lateral Prefrontal Cortex and Learning Regularities

As predicted, lateral prefrontal regions were activated in the invariant learning relative to the context-specific condition. Moreover, the triangular part of the left inferior frontal gyrus showed an activation increase as a function of learning in the invariant learning condition. Based on these results, we propose

that the increase of lateral PFC activity in the invariant learning condition reflects (i) the extraction and the maintenance of the invariant positions across trials and (ii) the adjustment of object–position conjunctions in each trial on the basis of the extracted information, resulting in an efficient encoding strategy.

First, we argue that the lateral PFC is involved in extracting task-relevant invariant spatial features during the initial trials. After regularities are successfully extracted, these regularities have to be actively maintained during the whole experimental block. This view receives support from numerous studies demonstrating that the lateral PFC plays a crucial role in active maintenance of information against distraction (Miller *et al.*, 1996; Sakai *et al.*, 2002) and in suppressing interference from varying sources (Bunge *et al.*, 2001; Mecklinger *et al.*, 2003). These processes are important prerequisites for rule extraction and formation, and might therefore mediate rule-based learning (Ashby and Ell, 2001).

Secondly, during the time-course of the experimental blocks, object–position conjunctions could be reorganized in light of the representations of the extracted regularities. Supporting evidence is provided by a recent fMRI study (Bor *et al.*, 2003). When subjects could integrate items into higher-level chunks, performance as well as lateral prefrontal activity increased. In the same vein, lateral prefrontal regions showed increased activity, when information is stored in a bound, rather than in a separated representation (Prabhakaran *et al.*, 2000). In both studies, subjects benefit from a reorganization of items during encoding. Lateral PFC involvement has been consistently found during episodic memory tasks (Wagner, 2002). The lateral PFC has been posited to subserve specific control operations during memory encoding, like the reorganization, evaluation and manipulation of the items to be memorized (Fletcher and Henson, 2001; Simons and Spiers, 2003). In a recent computational model of prefrontal memory control, Becker and Lim (2003) could show that the PFC represents internal mnemonic codes, rapidly adjusted to current task requirements. Furthermore, Becker and Lim's simulations reveal that these mnemonic codes evolve via reinforcement mechanisms during the time-course of the experiment. In accordance with these accounts, we propose that the PFC is involved in the adjustment of representations of incoming object–position conjunctions on the basis of the extracted regularities, i.e. the knowledge of four invariant positions in one block. It could be argued that the extracted regularities are represented as an internal mnemonic code (Becker and Lim, 2003). In light of this mnemonic code, object–position conjunctions are encoded more efficiently, as reflected in increased PFC activity in the invariant learning as compared with the context-specific condition. Moreover, the learning-related increase of lateral PFC activity indicates that this process evolves across trials and by this entails increased task performance in the invariant learning condition. Taken together, the dynamic PFC activation pattern seems to reflect the implementation of an efficient encoding strategy in the face of task-relevant regularities. These data underscore the special importance of the lateral prefrontal cortex during regularity learning.

Beyond the lateral prefrontal cortex, right ventral striatum activity increased as a function of learning. The ventral striatum as part of the basal ganglia has been associated with habit learning, the gradual acquisition of stimulus–stimulus and stimulus–response associations (Jog *et al.*, 1999; Packard and Knowlton,

2002), e.g. during probabilistic classification learning (Knowlton *et al.*, 1996; Poldrack *et al.*, 2001) and sequence learning (Willingham *et al.*, 2002; Schendan *et al.*, 2003). It has been proposed that the striatum and the hippocampus comprise interactive memory systems, specialized for distinct memory processes, namely relational binding of features defining episodes (hippocampus) and the acquisition of stimulus–stimulus associations (striatum), respectively (Poldrack *et al.*, 2001; Packard and Knowlton, 2002; Poldrack and Packard, 2003). This is in accordance with the complementary learning-related activation pattern observed in the hippocampus and the ventral striatum, respectively. Hippocampal activity decreased, whereas ventral striatal activity increased as a function of learning.

Another line of neurophysiological (Waelti *et al.*, 2001; Lauwereyns *et al.*, 2002) and neuroimaging research (McClure *et al.*, 2003; O'Doherty *et al.*, 2003) suggests that the meso-limbic dopamine system and its main target areas, especially the striatum, play a pivotal role during classical conditioning. Based on assumptions of formal reinforcement learning theories (Rescorla and Wagner, 1972; Sutton and Barto, 1990), it has been proposed that these brain regions represent predictions of future reward delivery and by this drive learning of associations between stimuli (Schultz, 2002). In accordance with these views, it is conceivable that the observed learning-related activation of the right ventral striatum might reflect the processing of increased predictability of spatial features during the time-course of learning. It could further be argued that these modified expectations of positions serve as internal reinforcement signals supporting regularity learning (Koechlin *et al.*, 2002; Schultz, 2002). Moreover, it has been proposed that mesolimbic reinforcement signals provided to the PFC play a crucial role in maintaining and updating of prefrontal memory representations (Miller, 2000; Cohen *et al.*, 2002). Thus, it could be speculated that bottom-up reinforcement signals from mesolimbic and striatal dopamine neurons, sensitive to the predictability of spatial features might modulate processing in higher-level brain structures like the PFC (Schultz, 2002). On the basis of these reinforcement signals, the PFC might implement an efficient encoding strategy (Becker and Lim, 2003).

The right inferior parietal lobule (IPL) showed stronger BOLD responses in the invariant learning as compared with the context-specific condition. However, in contrast to the PFC, the IPL activation decreased as a function of learning. The parietal cortex is assumed to be a core structure for attention-based selection and representation of spatial features (Gottlieb, 2002). In addition, the right lingual gyrus was activated in the invariant learning condition. Similar to the IPL, this region has been associated with attentional feature processing (Hopfinger *et al.*, 2000). In light of these findings, we assume that the IPL is involved in the allocation of attention to upcoming positions and in maintaining an attentional set (Corbetta and Shulman, 2002) across trials and by this supports the extraction of spatial regularities. During the time-course of experimental blocks, attentional maintenance requirements get reduced as reflected in a decrease of IPL activity.

To summarize, the present results point to a dynamic interplay between medial temporal, striatal and lateral prefrontal brain regions during the formation of invariant memories. During initial trials, the right hippocampus is involved, since unique object–position conjunctions require hippocampal

relational binding processes. After regularities are extracted, hippocampal activation decreases as a function of learning due to the reduced relational binding requirements. In contrast, both the lateral PFC and the ventral striatum showed a learning-related increase of activity. However, further studies will be required in order to examine the differential involvement of the hippocampus, the striatum and the lateral PFC in other forms of regularity learning.

In conclusion, these data indicate a transition of the relative roles of distinct neural systems during the time-course of learning, i.e. learning is accompanied by a shift from a hippocampal to a prefrontal-striatal brain system (Poldrack *et al.*, 2001; Opitz and Friederici, 2003).

Notes

Address correspondence to Christian F. Doeller, Saarland University, Department of Psychology, Experimental Neuropsychology Unit, PO Box 15 11 50, Saarbrücken D-66041, Germany. Email: c.doeller@mx.uni-saarland.de.

References

Aggleton JP, Brown MW (1999) Episodic memory, amnesia, and the hippocampal–anterior thalamic axis. Behav Brain Sci 22:425-489.

Aizenstein HJ, Stenger VA, Cochran J, Clark K, Johnson M, Nebes RD, Carter CS (2004) Regional brain activation during concurrent implicit and explicit sequence learning. Cereb Cortex 14:199-208.

Ashburner J, Friston KJ (1999) Nonlinear spatial normalization using basis functions. Hum Brain Mapp 7:254-266.

Ashby FG, Ell SW (2001) The neurobiology of human category learning. Trends Cogn Sci 5:204-210.

Becker S, Lim J (2003) A computational model of prefrontal control in free recall: strategic memory use in the California Verbal Memory Task. J Cogn Neurosci 15:821-832.

Birn RM, Cox RW, Bandettini PA (2002) Detection versus estimation in event-related fMRI: choosing the optimal stimulus timing. Neuroimage 15:252-264.

Bor D, Duncan J, Wiseman RJ, Owen AM (2003) Encoding strategies dissociate prefrontal activity from working memory demand. Neuron 37:361-367.

Büchel C, Holmes AP, Rees G, Friston KJ (1998) Characterizing stimulus–response functions using nonlinear regressors in parametric fMRI experiments. Neuroimage 8:140-148.

Bunge SA, Ochsner KN, Desmond JE, Glover GH, Gabrieli JDE (2001) Prefrontal regions involved in keeping information in and out of mind. Brain 124:2074-2086.

Burgess N, Maguire EA, Spiers HJ, O'Keefe J (2001) A temporoparietal and prefrontal network for retrieving the spatial context of lifelike events. Neuroimage 14:439-453.

Burgess N, Maguire EA, O'Keefe J (2002) The human hippocampus and spatial and episodic memory. Neuron 35:625-641.

Cleeremans A, Destrebecqz A, Boyer M (1998) Implicit learning: news from the front. Trends Cogn Sci 2:406-416.

Cocosco CA, Kollokian V, Kwan RKS, Evans AC (1997) Brainweb: online interface to a 3D MRI simulated brain database. Neuroimage 5:425.

Cohen JD, Braver TS, Brown JW (2002) Computational perspectives on dopamine function in prefrontal cortex. Curr Opin Neurobiol 12:223-229.

Cohen NJ, Eichenbaum H (1993) Memory, amnesia, an the hippocampal system. Cambridge, MA: MIT Press.

Cohen NJ, Ryan J, Hunt C, Romine L, Wszalek T, Nash C (1999) Hippocampal system and declarative (relational) memory: summarizing the data from functional neuroimaging studies. Hippocampus 9:83-98.

Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3:201-215.

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 USA 100:2157-2162.

Eichenbaum H (2000) A cortical-hippocampal system for declarative memory. Nat Rev Neurosci 1:41-50.

Eichenbaum H, Dudchenko P, Wood E, Shapiro M, Tanila H (1999) The hippocampus, memory, and place cells: is it spatial memory or a memory space? Neuron 23:209-226.

Feenan K, Snodgrass JG (1990) The effect of context on discrimination and bias in recognition memory for pictures and words. Mem Cogn 18:517-527.

Fernández G, Effern A, Grunwald T, Pezer N, Lehnertz K, Dümpelmann M, Van Roost D, Elger CE (1999) Real-time tracking of memory formation in the human rhinal cortex and hippocampus. Science 285:1582-1585.

Fletcher PC, Henson RNA (2001) Frontal lobes and human memory: insights from functional neuroimaging. Brain 124:849-881.

Fletcher PC, Anderson JM, Shanks DR, Honey R, Carpenter TA, Donovan T, Papadakis N, Bullmore ET (2001) Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. Nat Neurosci 4:1043-1048.

Fletcher PC, Büchel C, Josephs O, Friston K, Dolan RJ (1999) Learning-related neuronal responses in prefrontal cortex studied with functional neuroimaging. Cereb Cortex 9:168-178.

Fortin NJ, Agster KL, Eichenbaum HB (2002) Critical role of the hippocampus im memory for sequences of events. Nat Neurosci 5:458-462.

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2001) Categorial representation of visual stimuli in the primate prefrontal cortex. Science 291:312-316.

Friston KJ, Fletcher P, Josephs O, Holmes AP, Rugg MD, Turner R (1998) Event-related fMRI: characterizing differential responses. Neuroimage 7:30-40.

Friston, KJ Holmes AP, Price CJ, Büchel C, Worsley KJ (1999) Multi-subject fMRI studies and conjunction analyses. Neuroimage 10: 385-396.

Friston KJ, Penny W, Kiebel S, Hinton G, Ashburner J (2002) Classical and Bayesian inference in neuroimaging: theory. Neuroimage 16:465-483.

Gottlieb J (2002) Parietal mechanisms of target representation. Curr Opin Neurobiol 12:134-140.

Grafton ST, Hazeltine E, Ivry R (1995) Functional mapping of sequence learning in normal humans. J Cogn Neurosci 7:497-510.

Grön G, Wunderlich AP, Spitzer M, Tomczak R, Riepe MW (2000) Brain activation during human navigation: gender-different neural networks as substrate of performance. Nat Neurosci 3:404-408.

Hagberg GE, Zito G, Patria F, Sanes JN (2001) Improved detection of event-related functional MRI signals using probability functions. Neuroimage 14:1193-1205.

Holmes AP, Friston KJ (1998) Generalizability, random effects and population inference. Neuroimage 7:S754.

Hopfinger JB, Buonocore MH, Mangun GR (2000) The neural mechanisms of top-down attentional control. Nat Neurosci 3:284-291.

Huettel SC, Mack PB, McCarthy G (2002) Perceiving patterns in random series: dynamic processing of sequence in prefrontal cortex. Nat Neurosci 5:485-490.

Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM (1999) Building neural representations of habits. Science 286:1745-1749.

Knowlton BJ, Mangels JA, Squire LR (1996) A neostriatal habit learning system in humans. Science 273:1399-1402.

Koechlin E, Danek A, Burnod Y, Grafman J (2002) Medial prefrontal and subcortical mechanisms underlying the acquisition of motor and cognitive action sequences in humans. Neuron 35:371-381.

Lauwereyns J, Takikawa Y, Kawagoe R, Kobayashi S, Koizumi M, Coe B, Sakagami M, Hikosaka O (2002) Feature-based anticipation of cues that predict reward in monkey caudate nucleus. Neuron 33:463-473.

Maguire EA, Gadian DG, Johnsrude IS, Good CD, Ashburner J, Frackowiak RSJ, Frith CD (2000) Navigation-related structural

change in the hippocampi of taxi drivers. Proc Natl Acad Sci USA 97:4398-4403.

Maguire EA, Valentine ER, Wilding JM, Kapur JN (2003) Routes to remembering: the brains behind superior memory. Nat Neurosci 6:90-95.

McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. Neuron 38: 339-346.

Mecklinger A, Weber K, Gunter TC, Engle RW (2003) Dissociable brain mechanisms for inhibitory control: effects of interference content and working memory capacity. Cogn Brain Res 18:26-38.

Miller EK (2000) The prefrontal cortex and cognitive control. Nat Rev Neurosci 1:59-65.

Miller EK, Erickson CA, Desimone R (1996) Neural mechanism of visual working memory in prefrontal cortex of the macaque. J Neurosci 16: 5154-5167.

Norman KA, O'Reilly RC (2003) Modeling hippocampal and neocortical contributions to recognition memory: a complementary learning systems approach. Psychol Rev 110:611-646.

O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. Neuron 38:329-337.

Opitz B, Friederici AD (2003) Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. Neuroimage 19:1730-1737.

O'Reilly RC, Norman KA (2002) Hippocampal and neocortical contributions to memory: advances in the complementary learning systems approach. Trends Cogn Sci 6:505-510.

Packard MG, Knowlton B (2002) Learning and memory functions of the basal ganglia. Annu Rev Neurosci 25:563-593.

Poldrack RA, Packard MG (2003) Competition among multiple memory systems: converging evidence from animal and human brain studies. Neuropsychologia 41:245-251.

Poldrack RA, Clark J, Paré-Blagoev EJ, Shohamy D, Creso Moyano J, Myers C, Gluck MA (2001) Interactive memory systems in the human brain. Nature 414:546-550.

Prabhakaran V, Narayanan K, Gabrieli JDE (2000) Integration of diverse information in working memory within the frontal lobe. Nat Neurosci 3:85-90.

Reber PJ, Stark CEL, Squire LR (1998) Cortical areas supporting category learning identified using functional MRI. Proc Natl Acad Sci USA 95:747-750.

Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In: Classical conditioning II. Current research and theory (Black AH, Prokasy WF, eds), pp. 64-99. New York: Appleton-Century-Crofts.

Sakai K, Rowe JB, Passingham RE (2002) Active maintenance in prefrontal area 46 creates distractor-resistant memory. Nat Neurosci 5:479-484.

Schendan HE, Searl MM, Melrose RJ, Stern CE (2003) An fMRI study of the medial temporal lobe in implicit and explicit sequence learning. Neuron 37:1013-1025.

Schultz W (2002) Getting formal with dopamine and reward. Neuron 36:241-263.

Shanks DR, St John MF (1994) Characteristics of dissociable human learning systems. Behav Brain Sci 17:367-447.

Shastri L (2002) Episodic memory and cortico-hippocampal interactions. Trends Cogn Sci 6:162-168.

Simons JS, Spiers HJ (2003) Prefrontal and medial temporal lobe interactions in long-term memory. Nat Rev Neurosci 4:637-648.

Smith ML, Milner B (1981) The role of the right hippocampus in the recall of spatial location. Neuropsychologia 19:781-793.

Strange BA, Henson RNA, Friston KJ, Dolan RJ (2001) Anterior prefrontal cortex mediates rule learning in humans. Cereb Cortex 11: 1040-1046.

Sutton RS, Barto AG (1990) Time-derivative models of Pavlovian reinforcement. In: Learning and computational neuroscience: foundations of adaptive networks (Gabriel M, Moore J, eds), pp. 497-537. Cambridge, MA: MIT Press.

Talairach J, Tournoux P (1988) Co-planar stereotaxis atlas of the human brain. New York: Thieme.

Toni I, Ramnani N, Josephs O, Ashburner J, Passingham RE (2001) Learning arbitrary visuomotor associations: temporal dynamic of brain activity. Neuroimage 14:1048-1057.

Tulving E (1983) Elements of episodic memory. Oxford: Clarendon Press.

Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In: Analysis of visual behavior (Ingle DJ, Goodale MA, Mansfield RJW, eds), pp. 549-586. Cambridge, MA: MIT Press.

Waelti P, Dickinson A, Schultz W (2001) Dopamine responses comply with basic assumptions of formal learning theory. Nature 412:43-48.

Wagner AD (2002) Cognitive control and episodic memory: contributions from prefrontal cortex. In: Neuropsychology of memory (Squire LR, Schacter DL, eds), pp. 174-192. New York: Guilford Press.

Wallis JD, Anderson KC, Miller EK (2001) Single neurons in prefrontal cortex encode abstract rules. Nature 411:953-956.

Warner JJ (2001) Atlas of neuroanatomy. Boston, MA: Butterworth-Heinemann.

Willingham DB, Salidis J, Gabrieli JDE (2002) Direct comparison of neural systems mediating conscious and unconscious skill learning. J Neurophysiol 88:1451-1460.

Wood ER, Dudchenko PA, Eichenbaum H (1999) The global record of memory in hippocampal neuronal activity. Nature 397:613-616.

Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC (1996) A unified statistical approach for determining significant signals in images of cerebral activation. Hum Brain Mapp 4:58-73.

Yeshurun Y, Carrasco M (1998) Attention improves or impairs visual performance by enhancing spatial resolution. Nature 396:72-75.

Yonelinas AP, Kroll NEA, Quamme JR, Lazzara MM, Sauvé MJ, Widaman KF, Knight RT (2002) Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. Nat Neurosci 5:1236-1241.

Zeineh MM, Engel SA, Thompson PM, Bookheimer SY (2003) Dynamics of the hippocampus during encoding and retrieval of face-name pairs. Science 299:577-580.

Zola SM, Squire LR, Teng E, Stefanacci L, Buffalo EA, Clark RE (2000) Impaired recognition memory in monkeys after damage limited to the hippocampal region. J Neurosci 20:451-463.