

Learning-related Neuronal Responses in Prefrontal Cortex Studied with Functional Neuroimaging

Paul Fletcher¹, Christian Büchel¹, Oliver Josephs¹, Karl Friston¹ and Raymond Dolan^{1,2}

¹The Wellcome Department of Cognitive Neurology, Institute of Neurology, London WC1N 3BG and ²The Royal Free Hospital Medical School, Pond Street, London NW3, UK

We assessed time-dependent neuronal activity accompanying learning using functional magnetic resonance imaging (fMRI). An artificial grammar learning paradigm enabled us to dissociate activations associated with individual item learning from those involved in learning the underlying grammar system. We show that a localized region of right prefrontal cortex (PFC) is preferentially sensitive to individual item learning during the early stages of the experiment, while the left PFC region is sensitive to grammar learning which occurred across the entire course of the experiment. In addition to dissociating these two types of learning, we were able to characterize the effect of rule acquisition on neuronal responses associated with explicit learning of individual items. This effect was expressed as modulation of the time-dependent right PFC activations such that the early increase in activation associated with item learning was attenuated as the experiment progressed. In a further analysis we used structural equation modelling to explore time-dependent changes in inter-regional connectivity as a function of both item and grammar rule learning. Although there were no significant effects of item learning on the measured path strengths, rule learning was associated with a decrease in right fronto-parietal connectivity and an increase in connectivity between left and right PFC. Further fronto-parietal path strengths were observed to change, with an increase in left fronto-parietal and a decrease in right fronto-parietal connectivity. path strength from right PFC to left parietal cortex. We interpret our findings in terms of a left frontal system mediating the semantic analysis of study items and directly influencing a right fronto-parietal system associated with episodic memory retrieval.

Introduction

Because learning occurs across time, a description of the associated neuronal events requires repeated, temporally specified measurements. Functional neuroimaging techniques that describe regional brain activations associated with human memory function (Fletcher *et al.*, 1997) have focused on specific stages of memory such as encoding or retrieval (Squire *et al.*, 1992; Kapur *et al.*, 1994; Shallice *et al.*, 1994; Tulving *et al.*, 1994b; Schacter *et al.*, 1996a). Earlier attempts to address learning-related changes in brain activation (Friston *et al.*, 1992; Raichle *et al.*, 1994; Karni *et al.*, 1995) have contrasted naive and fluent task performance, interpreting the time-by-condition interaction in terms of learning processes. Such work has produced interesting findings with respect to prefrontal cortical function. Raichle *et al.* (1994) showed that a word generation task was associated with left prefrontal cortex (PFC) activation initially but not on subsequent performance, providing evidence that learning may be associated with prefrontal changes. Further work has examined the effect of skill acquisition (mirror reading) on repeated versus novel stimuli (Poldrack *et al.*, 1998). Such an approach, manipulating two types of learning as separate factors, enables a more detailed characterization of the regional cerebral changes associated with each since it becomes

possible to dissociate specific from more general learning-related changes in activation. Our study employed an analogous approach, measuring time-dependent changes in neuronal activity associated with learning the grammatical status of individual items and with learning the rules of the grammar system, enabling us to characterize responses to a behavioural task as these two distinct types of learning actually progressed. A crucial advantage of this approach is that neurophysiological changes which are common to both types of learning can be separated from those which are specific to one or the other. Thus, incidental changes (such as increasing success or diminishing effort) can be accounted for since they would be common to both types of learning.

An artificial grammar system embodies a set of arbitrary rules governing the concatenation of symbols. Subjects exposed to such a grammar system learn to categorize, as 'grammatical' (i.e. conforming to the hidden rules) or 'ungrammatical', subsequently presented items with an accuracy greater than chance (Reber, 1967). This accuracy can be disproportionate to their declarative knowledge of the underlying rules (Mathews *et al.*, 1989) and has been interpreted as indicating that rule-learning is implicit. However, it has also been suggested that the knowledge of the rule system may reflect more explicit remembering of surface features, i.e. bigrams and trigrams, from the studied set (Meulemans and Van der Linden, 1997). While this argument has yet to be resolved, recent evidence has suggested that a study task which predisposes subjects to learn the surface features of items will result in a tendency to characterize test items on the basis of surface similarity (Johnstone and Shanks, 1997). Conversely, a task requiring subjects to generate and test hypotheses about the underlying rules will result in subsequent rule-based classification. Thus, the instructions given to subjects, and the nature of the study phase, are crucial to the types of learning which occur.

The current experiment set out to distinguish the neurophysiological correlates of these two types of learning by emphasizing, within learning blocks, memorization of items on the basis of surface features, and, across learning blocks, classification of items on the basis of rules which subjects had been encouraged to infer and implement. The experimental design embodied repeated learning of individual items within blocks in the context of gradual rule acquisition across blocks (see Fig. 1). There is evidence that this approach can dissociate two types of learning in that amnesic patients, studied with an artificial grammar system, cannot learn specific instances of items but are, nevertheless, able to learn underlying grammatical rules (Knowlton *et al.*, 1992). In addition to exploring time-dependent regional changes in neuronal activity produced by item and rule learning, we were able to characterize neurophysiologically the nature and locus of the interaction between the two learning processes.

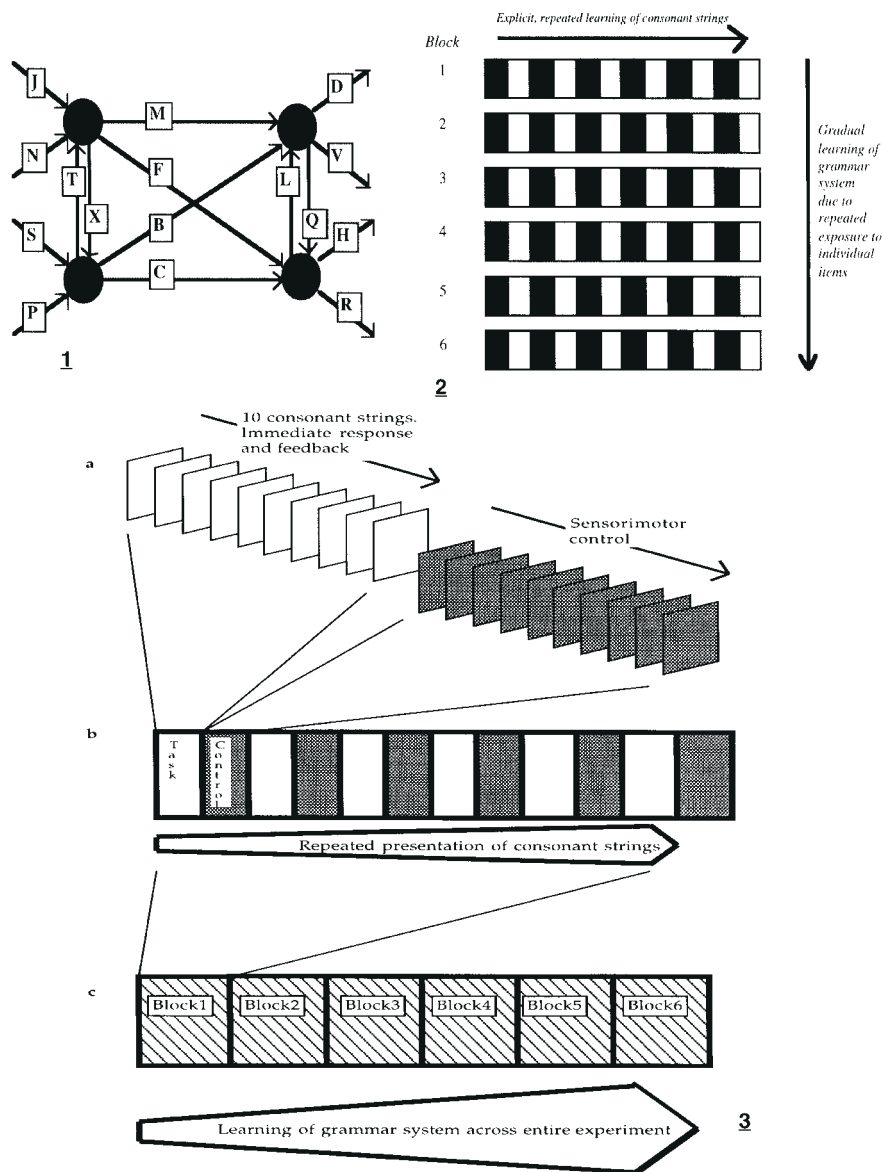


Figure 1. Experimental design. (1) The finite-state grammar system (hidden from subjects throughout) designed for use in the experiment. Grammatical items are produced by starting at the left of the system and moving to the right, following the direction of arrows moving from one node to a successive one. Every arrow traversed generates the attached letter. We added the constraint that all items consisted of four letters. Thus JMQH, STMV and JFLV would all be deemed grammatical, whereas JQLD, SCHR and JMFR would be deemed ungrammatical. (2) The overall design of the experiment. Dark boxes represent epochs of activation condition, white boxes epochs of baseline. Each row represents a single block within which item-specific learning occurs. Across the rows general rule learning occurs as shown. (3) This part of the diagram presents a more detailed description the way in which epochs and blocks were configured. (3a) represents a single epoch of the activation condition followed by a single epoch of baseline. During an activation epoch, 10 consonant strings were presented serially and subjects were required to make an immediate push-button response to indicate its grammatical status. Feedback ('right' or 'wrong') was given visually after each response. (3b) represents the six epochs of activation task (alternating with six epochs of baseline task) occurring within one block. During a given block, subjects saw the same 10 letter strings (the order of presentation was changed for each epoch). Thus, due to continuing feedback, explicit learning of grammatical status of individual items occurred across a block. (3c) represents the entire experiment which consisted of six identically structured blocks. During each successive block, 10 new consonant strings were presented and learnt, thus allowing the gradual acquisition of rule knowledge.

These analyses, which explore learning-related changes in regional brain activations, implicitly uphold a segregationist approach to the characterization of brain function. That is, separate cognitive processes are considered in terms of isolated brain regions. This approach, however, provides an incomplete picture of brain function. An alternative perspective is that captured by effective connectivity, i.e. intra-cerebral influences (Gerstein and Perkel, 1969). One approach to its measurement is through structural equation modelling (McIntosh and Gonzales-Lima, 1994) in which changes in activity in the sub-

components of a predefined anatomical model are explained in terms of directionally specified pathways connecting the regions. Task-related changes in the strengths of these pathways can be explored and interpreted in terms of cognitively mediated modulations of inter-regional influences. For example, Büchel and Friston (1997) have characterized cortical responses to a visual stimulus in terms of an integrated system in which path strengths are modulated as a function of whether or not subjects attend to stimuli. We therefore performed an analysis on our data in which this technique was used to evaluate how path strengths

in an anatomical model (incorporating the main task-sensitive regions) were modulated by learning. Thus, in the first section of the paper, we describe the regions showing learning-related changes in activation. In the second section, we explore how these changes may be characterized in terms of mutual influences or effective connectivity.

Materials and Methods

Informed consent was obtained from seven subjects (four male, three female; all right-handed; age range 24–36 years; mean age 28 years). A Siemens VISION system (Siemens, Erlangen) operating at 2 Tesla was used to acquire both T_1 anatomical and gradient-echo echo-planar T_2^* -weighted image volumes with blood oxygenation level dependent (BOLD) contrast. For all subjects, data were acquired in two scanning sessions separated by a 5 min rest period. Aside from six ‘dummy’ volumes, which were subsequently discarded to allow for T_1 equilibration effects, a total of 360 functional volumes per subject were acquired. Volumes were acquired continuously every 6400 ms. Each volume comprised 64 3 mm axial slices with in-plane resolution 3×3 mm positioned to cover the whole brain.

We used a finite state grammar system (see Fig. 1) to generate consonant strings under the constraint that all strings must consist of four letters. A total of 30 strings, mixed with 30 arbitrarily chosen non-grammatical lures, were presented over the course of the entire experiment. The experiment comprised six blocks (see Fig. 1) and within each block stimuli were presented on a screen, at a rate of one per 3.2 s, with a total of 10 consonant strings (50% grammatical; 50% lures). Subjects were required to respond to items by indicating with a key pad whether the item was grammatical or ungrammatical. Subjects received immediate visual feedback indicating whether their response was correct or incorrect. In a single block the same 10 items were presented a total of six times, each time followed by a sensorimotor control condition. The latter consisted of serial visual presentation of either rows of the letter P or N which signalled a right or left key-pad response. At the end of each block, a new block began and 10 new items (five grammatical; five ungrammatical) were presented. In the early stage of the experiment subjects had no knowledge of the underlying rules of the grammar and responded on the basis of guesses. Six identically structured blocks were presented in total with each successive block containing 10 new items.

Analysis 1

Functional Segregation

The imaging time series was realigned, spatially normalized and smoothed with a Gaussian kernel of 8 mm full width half-maximum (Talairach and Tournoux, 1988; Friston *et al.*, 1995a,b). Low-frequency cosine waves modelled and removed subject-specific low-frequency drifts in signal (Holmes *et al.*, 1997) and the global means were normalized by proportional scaling. To characterize activations we used a model embodying three effects. The first simply examined task versus baseline activations by contrasting all scans in which subjects were performing the grammaticality judgement with all those in which they were performing the sensorimotor, baseline task. The resulting SPM(t) was used to constrain the second and third analyses which were concerned with the learning effects. This constraint meant that we explored for learning-related changes in activation only in those regions that showed an activation relative to the baseline. The second analysis examined task versus baseline activations which were expressed linearly within blocks of repeated item learning. The third analysis evaluated linear changes expressed across blocks, i.e. over the entire duration of the experiment and, thus, reflected rule learning. We also assessed the product of the second and third models to reveal interactions between these two types of learning. We present results for linear changes only; however, a separate analysis embodying more flexible, second-order effects was also

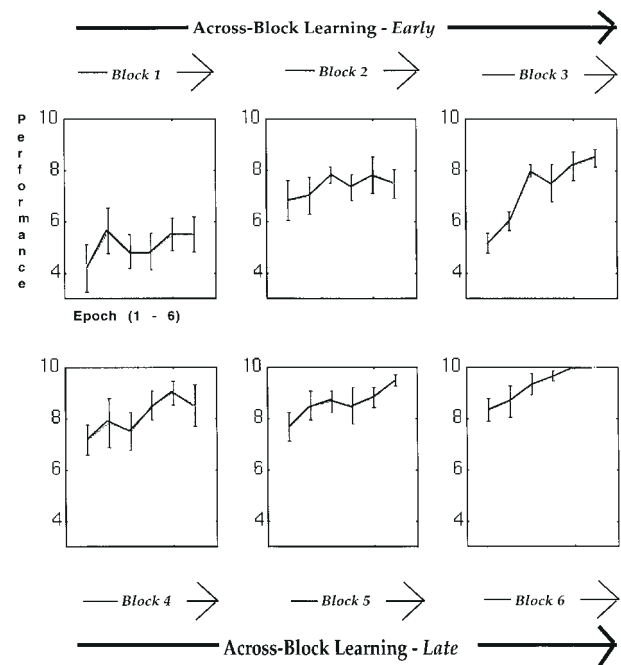


Figure 2. Mean of subjects' performance (with standard error bars) within and across blocks. Within each block (from epoch 1 to epoch 6) there was an improvement in performance due to the repeated presentation of items and feedback as to their status. The rule learning occurring from block to block is evidenced by the fact that there is a significant increase in performance on the first epoch of each successive block even though items seen in these cases were novel and the subject had received no prior knowledge of their grammatical status.

carried out but did not add to the characterization of the learning effects over and above that of the linear model.

For all effects, subject's data were modelled separately and group results are presented as the conjunction of activations across all seven subjects. In essence, this means that we explored the data only for those changes which were common to all subjects. We felt it important to do this since, in the standard SPM analysis in which variation across different subjects is treated as a fixed effect, there is a possibility that a strong effect in only one or two subjects may show up in the group analysis. The conjunction analysis indicates those effects which do not differ significantly between subjects in terms of magnitude and location. Thus, the only activations presented are those which occurred in all subjects. Results presented are those that survived an uncorrected threshold of $P < 0.001$.

Results

Behavioural Results

As predicted, performance levels showed a linear improvement (see Fig. 2) both within blocks ($df = 5$; $F = 14.6$; $P < 0.001$) and across blocks ($df = 5$; $F = 15$; $P = 0.001$). There was no significant block by epoch interaction. A separate test on the first epoch of each block showed a significant improvement in classifying the novel items as the experiment went on ($df = 5$; $F = 7.42$; $P < 0.001$). This latter observation provides further evidence of a progressive learning of the rule system since, in all cases, the first item of a block consisted entirely of novel items which could not be classified correctly on the basis of remembered instances.

Functional Neuroimaging

The overall main effect of performing the classification task,

compared to the baseline task ($P < 0.05$, corrected for multiple comparisons) is shown in Figure 3 and Table 1. It revealed wide areas of prefrontal, parietal, anterior cingulate and occipital cortex to be more active in association with the activation task. This analysis subsequently served as a 'mask' to constrain the subsequent analyses of the effects of learning. That is, only those voxels which were activated in association with this comparison were explored in the subsequent analyses of the effects of item and rule learning. Thus the effects described and discussed below can be more confidently concluded to be occurring within the brain system concerned with carrying out the task. Additionally, this analysis served to identify regions of the system explored in the assessments of effective connectivity carried out in the second analysis of the data.

Three brain regions showed significant activation changes in association with learning. Medial parietal and medial frontal cortex were sensitive to item learning occurring within blocks (Table 2), whereas a left PFC activation was sensitive to rule learning across blocks (Fig. 4 and Table 2). We found no evidence of a significant within-block effect in left PFC, suggesting that this region was relatively insensitive to learning the status of individual items. This issue was directly addressed by explicitly comparing within-block with across-block effects in left PFC. Exploration of the data for interactions between the two types of learning, i.e. to assess within-block changes which were modulated across blocks and vice versa, was carried out. This analysis showed a significant effect in right PFC and in occipital and parietal regions (see Fig. 5 and Table 2). In brief, the effect for right PFC was as follows: there was an increase in activity associated with item learning occurring only in early blocks and this effect was attenuated during later blocks (see Fig. 5).

Analysis 2

Functional Integration

The above approach uses multiple regression or ANOVA. Parameters of the model are based on minimization of the sums of squared differences between predicted and observed dependent variables. In structural equation modelling, the variables are considered in terms of the covariance structure with parameters (inter-regional connection strengths) being estimated by minimizing differences between observed covariances and those implied by a predefined model. This model comprises anatomical and functional features. The anatomical part of the model consists of the anatomically separable regions, the connections specified between those regions and their directions. The strengths of those connections constitutes the functional part of the model. We chose to specify the regions involved on the basis of a categorical comparison between the activation (grammaticality judgement) condition and the baseline condition (Fig. 3 and table 1). The system chosen comprised left and right PFC, left and right medial parietal cortex, and left and right occipital cortex, and is diagrammatically represented, with directionally specified connections between the regions, in Figure 3. As well as the use of the task versus baseline analysis to guide our model and to select the voxels from the regions of interest, the features of the model have their theoretical basis in a number of previous functional neuroimaging observations. Firstly, it is clear that memory function at the encoding and retrieval stages are associated with both parietal and prefrontal activations (Squire *et al.*, 1992; Shallice *et al.*, 1994; Tulving *et al.*, 1994b; Fletcher *et al.*, 1997). We thus included bilateral parietal and prefrontal regions in the model. The two occipital

foci were included due to the strong visual activations occurring in association with the classification task. Unidirectional connections from occipital to parietal regions were included. Further connections from right parietal to right PFC and from left parietal to left PFC were added and, in order to encompass the notion of a possibility of a frontal modulation of parietal activity, each frontal region was allowed connections to both left and right parietal regions and to each other. It has been noted that right prefrontal and parietal cortical regions are activated almost ubiquitously in association with explicit memory retrieval tasks (Squire *et al.*, 1992; Shallice *et al.*, 1994; Tulving *et al.*, 1994b; Buckner *et al.*, 1996; Fletcher *et al.*, 1997) and that left prefrontal activation is commonly associated with memory encoding tasks, most particularly those that emphasize processing of studied material in terms of its semantic attributes (Kapur *et al.*, 1994; Tulving *et al.*, 1994a; Dolan and Fletcher, 1997; Fletcher *et al.*, 1997, 1998). We suggest that our own experiment – which, we believe, involves a changing emphasis from a classification of consonant strings based upon episodic memory retrieval to one based on an increasing semantically based knowledge of the grammar system – would result in changing strength of these fronto-parietal connections. Since there was a gradual improvement in the subjects' ability to classify novel consonant strings in terms of their meaning (with respect to the rules of the grammar system) rather than merely on the basis of explicit recall of prior presentation of those items, we predicted that a left-sided system would strengthen whereas the right fronto-parietal pathways, reflecting explicit retrieval, would weaken as this process was relied upon less. We also predicted that this gradual change in the nature of the task, occurring as a function of the rule-learning, would be reflected in changing connectivity between the prefrontal regions.

The analysis was carried out in MATLAB (Mathworks, Sherborn, MA) and the operational procedures have been fully described previously (Büchel and Friston, 1997). The fMRI time series were normalized and therefore the estimated covariances are equal to the correlations. Regions were defined on the basis of the statistical parametric map of the contrast between all activation and all baseline epochs. Thus, single voxels locating the locus of maximal activation for six regions were identified, and around these peak activations regions were defined using a volume of interest of 8 mm radius. The first eigen time series of each volume was used for calculation of the correlation matrices. The residual variance was estimated, but its path coefficient was set to unity.

Statistical inference was made using the χ^2 statistic to test the difference between two models. That is two models are constructed. In the first of these, the 'Null Model', the estimates of some path coefficients are constrained to be equal in the pre-learning and post-learning measurements. The alternative model allows the path coefficient to change as a function of learning. The significance of the difference between the models is expressed by the difference in the χ^2 goodness of fit indicator [χ^2 difference test or likelihood ratio test (Bollen, 1989)]. Simply put, to test whether there is a significant effect of learning on a given connection, we explore the implied and observed covariance structures occurring pre- and post-learning when this connection is allowed to change and when it is constrained. If the first model (when the pathway is allowed to change) produces a significantly better fit, then this pathway may be considered to be affected by the learning process.

It is important to point out that the structural model used in the analysis was preceded by a simpler one encompassing fewer

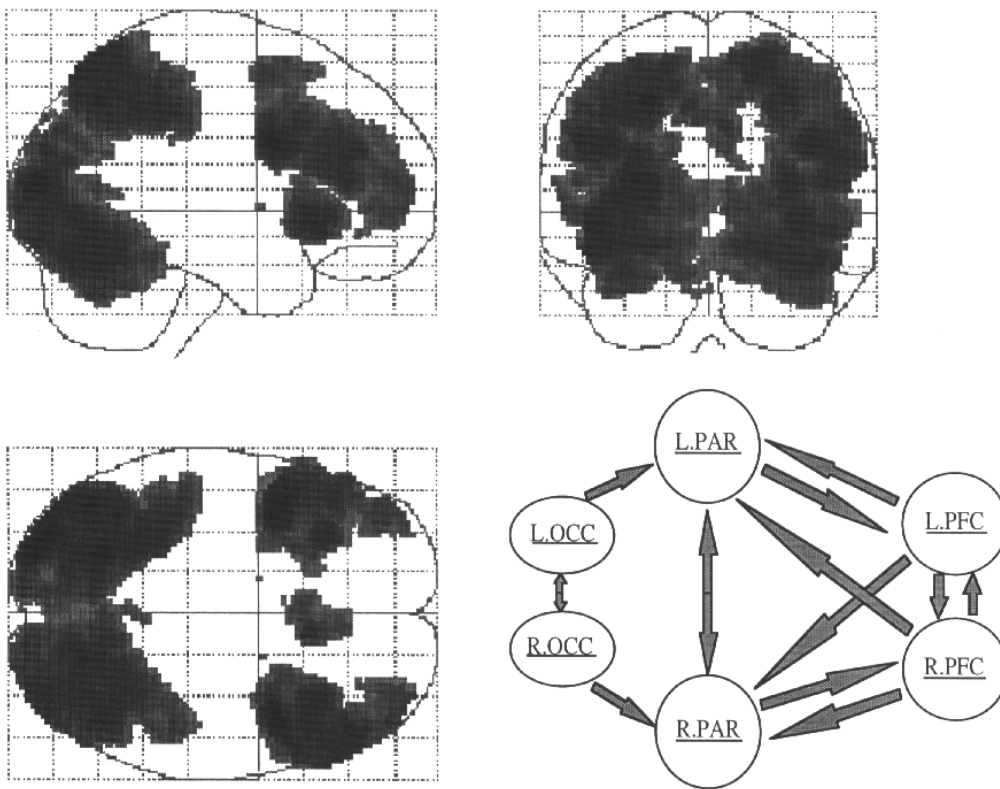


Figure 3. A statistical parametric map of the direct comparison ($P < 0.05$, corrected) of activation with baseline tasks (ignoring learning-related changes). ‘Glass brain’ views are shown from the right (top left panel), from behind (top right panel) and from above (bottom left panel). The regions activated are also detailed in table 2. This simple contrast was used to mask subsequent analyses and to identify regions for inclusion in the effective connectivity analysis shown in Figure 6. The bottom right panel shows a diagrammatic representation of the regions subjected to effective connectivity analysis, together with directionally specified connections. The regions were selected on the basis of the SPM and were as follows: right prefrontal cortex (R.PFC: $x, y, z = 42, 6, 30$), left prefrontal cortex (L.PFC: $x, y, z = -42, 8, 28$), right occipital cortex (R.OCC: $x, y, z = 32, -86, 14$), left occipital cortex (L.OCC: $x, y, z = -28, -88, 8$), right medial parietal cortex (R.PAR: $x, y, z = 18, -70, 58$), left medial parietal cortex (L.PAR: $x, y, z = -24, -68, 48$).

Table 1
Regional activations associated with the consonant-string classification task compared to baseline ($P < 0.05$, corrected for multiple comparisons)

Region	Coordinates {000186}	Z score
Right middle/inferior frontal gyrus	42 6 30	9.1
	48 36 14	8.8
	36 22 -2	9.0
Left middle/inferior frontal gyrus	-42 8 28	8.9
	-32 28 -2	8.7
Medial frontal/anterior cingulate gyrus	6 20 36	8.4
	6 32 28	7.8
	-2 18 40	7.4
Posterior cingulate gyrus	14 -60 10	6.0
Left medial occipital gyrus	-28 -88 8	9.3
Left medial parietal cortex (precuneus)	-24 -64 48	9.1
	-12 -84 38	6.8
Right medial occipital gyrus	32 -86 14	8.9
Right medial parietal cortex (precuneus)	18 -70 58	8.9
	10 -68 22	6.2
Cerebellum	2 -46 -16	6.0

areas and fewer connections. Our original model showed, in essence, the effects which are described below. While this is a heartening indication of the robustness of our findings, it also means that the results of the statistical testing should be treated with caution since the application of more than one model to

Figure 4. (a) Orthogonal sections of a T_1 -weighted anatomical image that conforms to a standard stereotactic space. Superimposed on these sections is a SPM ($P < 0.001$, prepared as detailed in the text) indicating regions showing increases in brain activation occurring across the whole experiment. Sections have been chosen at one of the voxels of maximal change (coordinates $x, y, z = -38, 34, 26$; see Table 2) to show the left PFC region. In the bottom right panel is a graph showing relative activation (compared to the repeating baseline task) at this voxel for blocks 1–6. Activations across epochs 1–6 have been averaged for each subject for each block. It can be seen that the average level of activation in this region increases across the course of the experiment. (b) Graphic representation of changes in activation at the same voxel (relative to the baseline condition). Here, activations averaged across subjects for each epoch in each block are shown. Thus, there is a significant increase in left PFC activation from early to late blocks but no significant pattern of changes within blocks.

these data constitutes multiple comparisons which have not been corrected for.

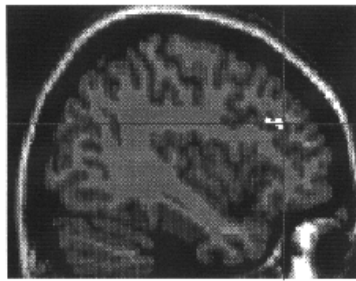
Results

Evaluation of changes in the system occurring in association with item learning, i.e. changes in path strengths occurring from the beginning to the end of each block did not reveal any significant effect. However, there were five effects ($P < 0.01$, $df 1$) occurring in association with learning to classify items on the basis of learning the rules of the grammar system (see Fig. 6).

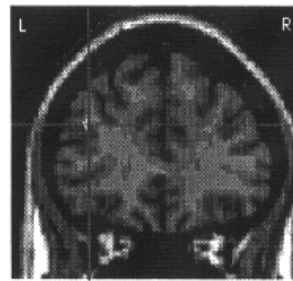
1. An increase in path strength from left to right PFC.

a

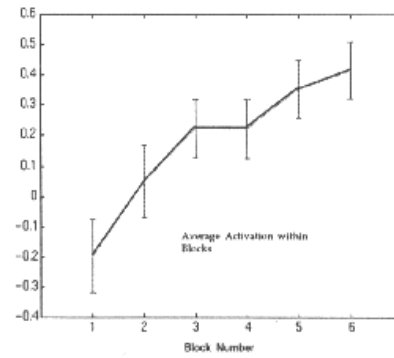
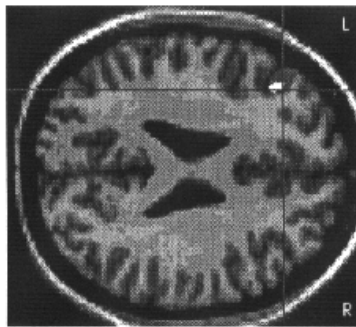
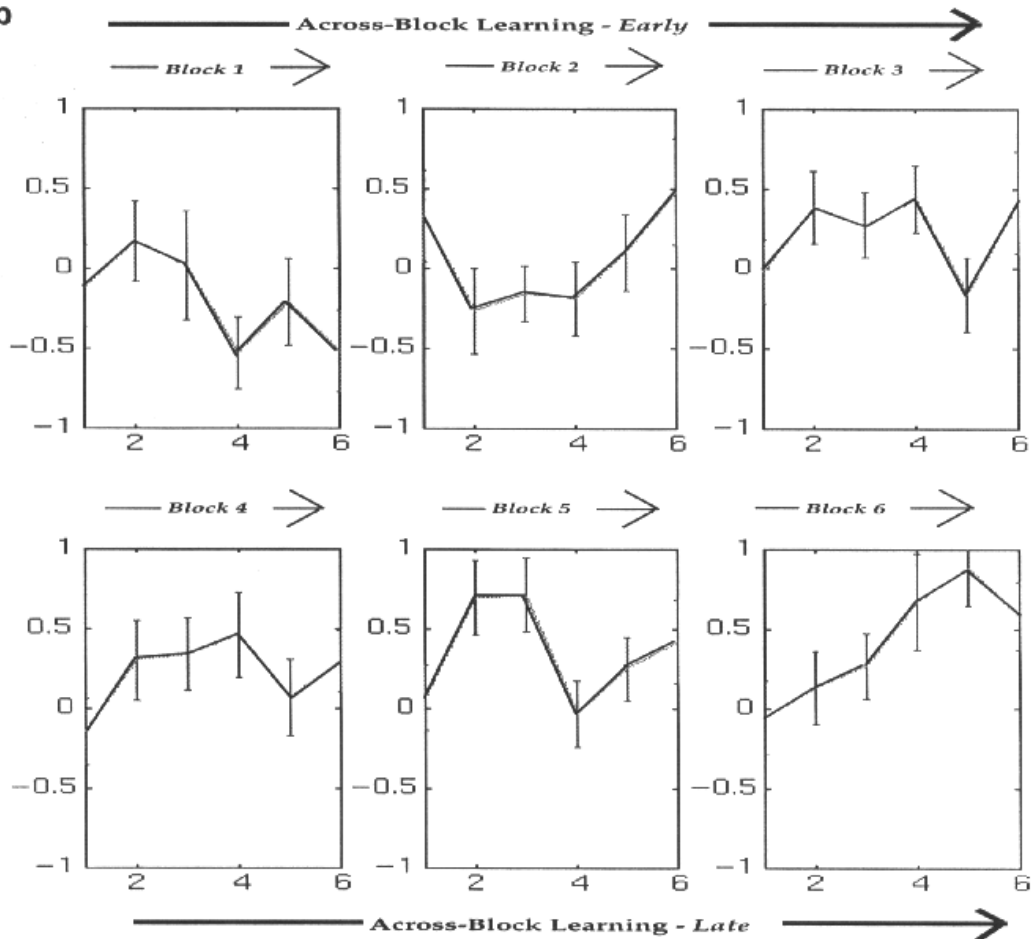
sagittal



coronal



transverse

**b**

2. A change in influence from right to left PFC
3. An increase in left parieto-prefrontal path strength.
4. A decrease in right parieto-prefrontal path strength.
5. A change in the right prefrontal to left parietal path.

These changes are shown diagrammatically in Figure 6.

Discussion

It is not possible to dissociate fully item from rule learning using a finite state grammar system since the grammatical rules are bound up in the arrangement of the consonants such that items which obey the rules have similar surface features. However, several aspects of our experimental design make us confident that the two types of learning have been dissociated. First, we used a modified finite state grammar system which had an increased number of entry and exit points (see Fig. 1). This allows for high variance in the surface features of presented consonant strings, reducing the extent to which judgements could be made on the basis of item similarity alone. Second, within-block item learning was maximized by repeatedly presenting a small set of items, encouraging subjects to explicitly learn each item in the light of previous presentations and feedback within the block. Third, across-block rule learning was maximized by informing subjects, prior to study, about the existence of a rule system and encouraging them to look for rules as the experiment progressed. Thus, from block to block, although item sets changed, the underlying rules remained constant and could be used in making decisions about each new set. That this behavioural interaction actually occurred can be seen from the performance data (see Fig. 2).

The right PFC activation, which reflected an interaction of rule learning with item learning, was located dorsolaterally in the middle frontal gyrus. This region is consistently implicated in memory retrieval (Squire *et al.*, 1992; Shallice *et al.*, 1994; Tulving *et al.*, 1994b; Buckner *et al.*, 1996; Fletcher *et al.*, 1997). Since subjects reported that their within-block performance improvements were based upon explicit recall of items' grammatical status on the basis of earlier presentations of those items, it can be construed that the finding of the early, within-block, increase in right PFC activation reflects explicit item learning. This pattern of within-block increase was also seen in the medial parietal and medial frontal cortex. These latter effects were seen across every block in the experiment and showed no interaction with rule learning. Of course, it might be argued that a change in activation could conceivably be related solely to performance improvements rather than learning [it has been suggested that both prefrontal and parietal cortex are sensitive to performance measures (Kapur *et al.*, 1995; Rugg *et al.*, 1996)]. However, there was no significant increase in overall activity in the right PFC across blocks despite clear improvements in performance. Indeed, during the latter part of the last blocks, when performance was at, or approaching, ceiling, activations in right PFC were attenuated (see Fig. 5). An across-block increase in activation would be expected if there was any simple relationship between activation and performance *per se*. In other words, increases in activation in right PFC, during the early blocks, are unlikely to reflect behavioural changes contingent upon, but not directly related to, the two learning processes.

We propose that regions showing an increase in activation across all blocks of the experiment reflect increasing knowledge of the underlying rule system rather than episodic memory of individual items since, from block to block, novel items were presented. The regions implicated by this analysis were the left

Table 2

Regional activations that show significant task-by-time interactions ($P < 0.001$, uncorrected)

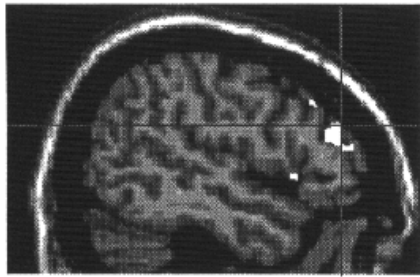
Region	Coordinates {000186}			Z score
<i>(a) Regions showing an increase in activation within blocks, i.e. with explicit learning of the grammatical status of individual items</i>				
Medial parietal/occipital cortex	4	-78	32	4.1
	18	-66	32	3.4
	10	-66	44	3.4
Medial frontal gyrus	2	50	24	3.7
<i>(b) Regions showing an increase in activation across blocks, i.e. with an increasing ability to use the grammatical rules</i>				
Left medial frontal gyrus	-38	34	26	3.5
Right cerebellum	32	-40	-20	3.3
<i>(c) Regions showing an increase in activation within blocks, i.e. attenuated over the whole course of the experiment</i>				
Right middle/inferior frontal gyrus	3	6	26	4.4
	52	38	24	3.6
Right superior parietal cortex	26	-68	50	4.9
Left medial occipital gyrus	-34	-78	8	4.0

Figure 5. (a) Orthogonal sections of a T_1 -weighted anatomical image that conforms to a standard stereotactic space. Superimposed on these sections is a SPM ($P < 0.001$, prepared as detailed in the text) indicating regions showing increases in brain activation occurring within, but not across, experimental blocks. Sections have been chosen at one of the voxels of maximal change (coordinates $x, y, z = 52, 38, 24$; see Table 2) to show the right PFC region. In the bottom right panel is a graph showing relative activation (compared to the repeating baseline task) at this voxel for blocks 1–6. Activations across epochs 1–6 have been averaged for each subject for each block. It can be seen that the average level of activation in this region shows no obvious pattern across the course of the experiment. (b) Graphic representation of changes in activation at this voxel relative to the baseline condition. Here, activations averaged across subjects for each epoch in each block are shown. Thus, there is a significant increase in right dorsolateral PFC activation from early to late epochs in the earlier blocks. In the late blocks, this pattern is lost and activations in this region show a tendency to fall away at the ends of blocks.

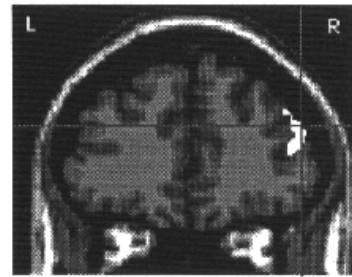
PFC (see Fig. 4) and the right cerebellum. This left PFC change in activation was located in the middle frontal gyrus. The region showed no evidence of any increase in activation within individual blocks that might reflect item learning (see Fig. 4). Thus, as with right PFC, changes in activation in this region cannot be accounted for in terms of non-specific effects. We propose that the left PFC activation changes reflect processes involved in categorizing items on the basis of knowledge of a rule system rather than episodic memory. Interestingly, this region has been implicated in memory encoding (Shallice *et al.*, 1994; Kapur *et al.*, 1994; Haxby *et al.*, 1996; Fletcher *et al.*, 1997) and our own data suggest a role in meaning-based elaboration of verbal material (Fletcher *et al.*, 1998). There is also evidence that activity in this region is particularly associated, in the context of encoding, with conditions in which prior associations interfere with formation of new meaningful associations (Dolan and Fletcher, 1997). The present study used a learning paradigm that emphasized an increasing knowledge of rules that make stimuli 'meaningful' (within the context of the grammar system). Thus, as the experiment progressed, each item was processed within a 'semantic' context that was unavailable at the beginning. This phenomenon, the acquisition and implementation of a meaning-based framework, is, we sug-

a

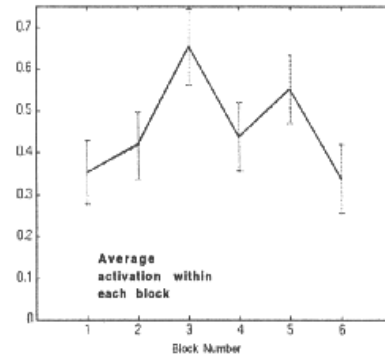
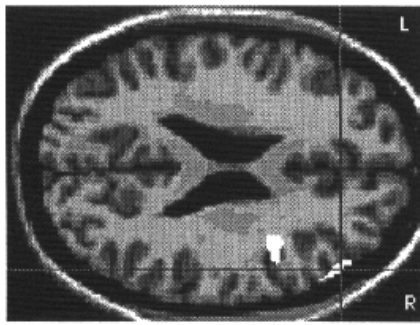
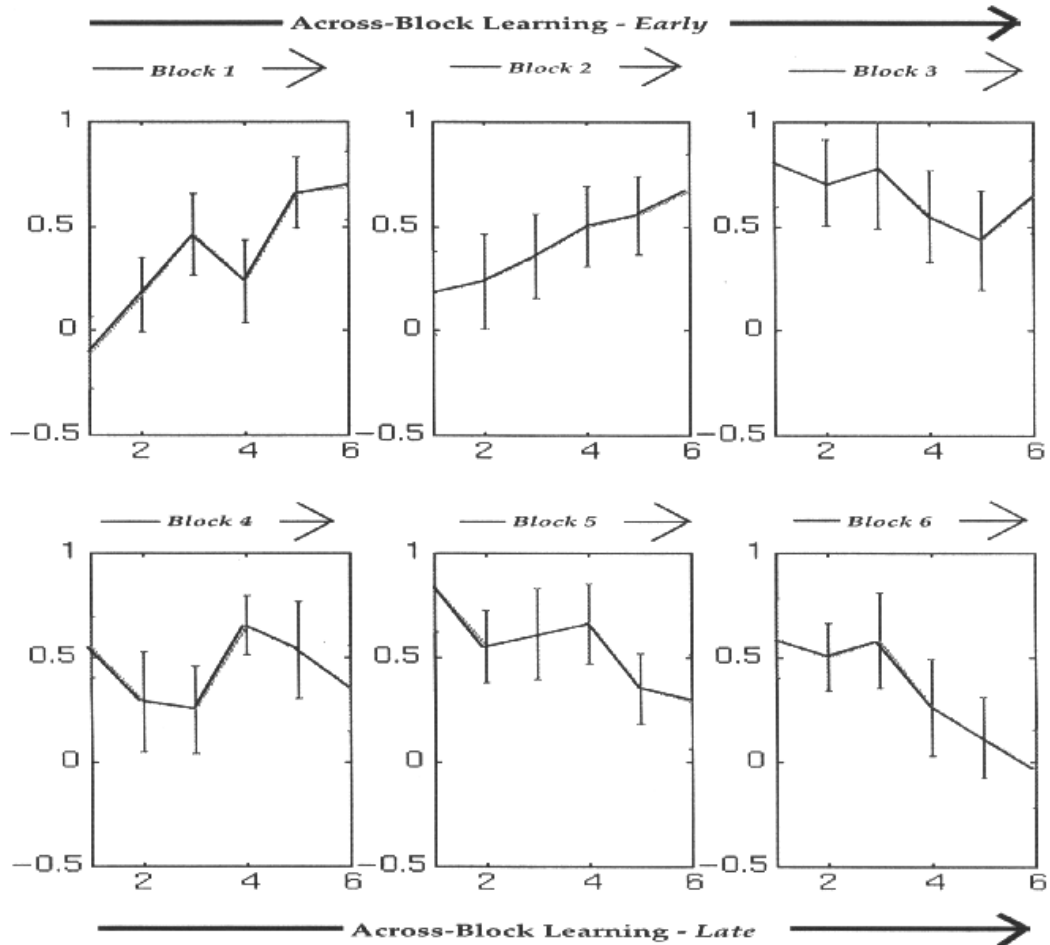
sagittal



coronal



transverse

**b**

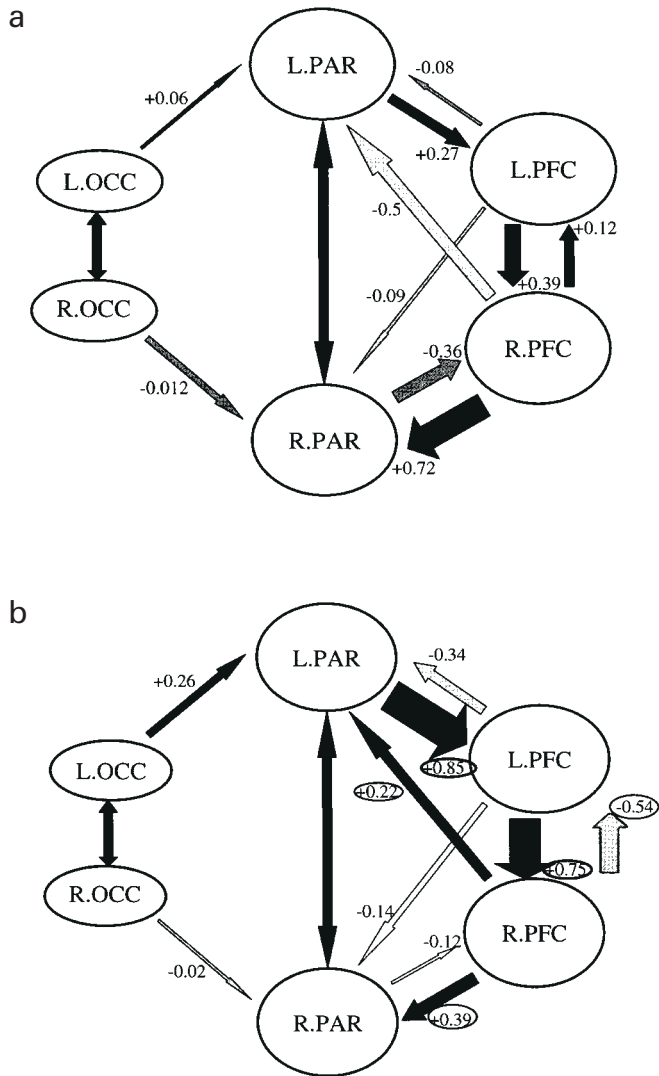


Figure 6. Diagrammatic representation of the functional model with path strengths estimated pre- and post rule-learning in (a) and (b) respectively. Pathways which changed significantly as result of this learning are circled. Pathways representing a negative influence are diagonally striped.

gest, reflected in increasing left PFC activation across learning blocks.

How are the two types of learning, engaged by this application of the artificial grammar system, best characterized? The repeated presentation of items within blocks produces an explicit, episodic memory of the grammatical status of individual items such that, with repeated presentation within each block, subjects employed episodic retrieval mechanisms. The component of learning occurring across blocks is more controversial. Classically, artificial grammar learning has been held to dissociate implicit from explicit learning. However, recent work using a biconditional grammar system has suggested that rules are not learnt implicitly (Johnstone and Shanks, 1997). In our experiment, we made both item and rule learning explicit and conceptualise our findings using a dichotomy of episodic versus semantic processes (Tulving, 1983). The former refers to memory about a specific encoding event, the latter to general knowledge concerning the meaning of items and the linkages between them. Interaction between these two forms of memory

is suggested by observations that episodic memories are retrieved with greater efficiency when subjects have contextual knowledge (Bransford and Johnson, 1972) or when items are studied according to their semantic attributes (Craik and Lockhart, 1972). In our study, subjects were required to learn the grammatical status of consonant strings against a background of increasing knowledge of the set of rules generating these strings.

The acquisition of semantic knowledge influences acquisition of item knowledge (Bransford and Johnson, 1972) and, clearly, improves performance on the grammatical decision task (see Fig. 2). Strikingly, the within-block profile of right PFC activations was sensitive to the subjects' knowledge and use of the rule system. The earliest blocks (blocks 1 and 2) showed a statistically significant linear increase in right dorsolateral PFC activation. During later blocks (i.e. as the subjects began to use the rule system) this effect was not seen. Indeed, by the end of blocks 5 and 6, activity in this region appeared to fall (see Fig. 5). Thus, the profile of right dorsolateral PFC activation can be framed as an early episodic memory retrieval response modulated, over time, by increasing fluency with the rule system. This rule learning enables the categorization of items on the basis of semantic knowledge, so decreasing subjects' reliance on explicit item recall of the status of items based on remembered instances of previous presentations. This observation is in keeping with suggestions that right PFC may have an important role in representing prior experience in the act of episodic recall (Wheeler *et al.*, 1997).

A number of suggestions have been made about the functional significance of the right PFC in memory retrieval, e.g. an unspecified retrieval mode (Tulving *et al.*, 1994a; Nyberg *et al.*, 1995), post-retrieval monitoring and verification of retrieval products (Schacter *et al.*, 1996b; Shallice *et al.*, 1994; Fletcher *et al.*, 1996). Our findings provide further evidence that this region is related to explicit retrieval processes in that the ability to make episodic judgements based upon prior occurrence during early stages of the experiment (when such processes are emphasized) is associated with significant increases in activation. When responses are based on rule knowledge these activations are attenuated. We suggest that this attenuation, occurring in the context of improving performance, reflects the implementation of semantic knowledge. It is in direct contrast to the activation profile in the left PFC, which shows an across-block but no within-block effect, indicating that left and right PFC have highly differentiated and context-specific roles in learning. Additionally, the findings are consistent with previous suggestions that memory-related left and right PFC activations reflect semantic and episodic retrieval respectively (Tulving *et al.*, 1994a).

Our findings suggest that attempts to conceptualize memory-related lateralization of prefrontal function purely in terms of encoding and retrieval are incomplete. The observation of learning-related changes in left and right PFC responses provides evidence for neurophysiological sensitivity to changing contexts in which encoding and retrieval occur. The findings further indicate that behavioural evidence of interactions between episodic and semantic memory systems is reflected in changing relationships between left and right PFC. During early learning, when subjects rely on specific instances, right PFC activation predominates. As learning acquires a semantic basis, left PFC activation increases with an associated modulation of the right PFC response. These complex interactions, also reflected behaviourally in performance improvements, provide a more dynamic view of memory-related prefrontal function than existing models based upon strict lateralization for encoding and

retrieval, and are in keeping with more recent suggestions that the lateralization of PFC activation found in studies of memory is strongly dependent upon the types of encoding and retrieval processes engaged (Nolde *et al.*, 1998).

A more complete view of this dynamic interaction between brain regions is afforded by the use of path analysis to evaluate connectivity within the selected group of brain regions. In the current study, this approach has enabled a more complete characterization of the changes related to rule learning. These changes can be summarized as follows: firstly, an increase in the left parietal influence on left PFC and a decrease in the right parietal influence on right PFC; secondly, a reduced influence of right PFC on right parietal cortex and an increased influence on left parietal cortex; thirdly, an increase in the influential relationships between left and right PFC. It has been suggested (Squire *et al.*, 1992; Shallice *et al.*, 1994; Tulving *et al.*, 1994b; Fletcher *et al.*, 1997) that a mainly right-sided system (comprising right PFC and medial parietal cortex) subserves episodic memory retrieval. Our findings suggest that this system becomes less important during the latter stages of learning since we observed decreasing effective connectivity between parietal and right PFC. Further, the suggestion that the left-sided activations reflect semantic processing of study material (Dolan and Fletcher, 1997; Fletcher *et al.*, 1998) is supported by the observation of increased effective connectivity within this system (an increase in path strength from left parietal to PFC). Furthermore, with the acquisition of rule knowledge there is less reliance on an episodic retrieval system. This is reflected in an increase in path strength from left to right PFC, reflecting direct interaction between semantic and episodic retrieval systems.

The use of structural equation modelling to characterize causal relationships between brain systems and regions is open to a number of criticisms. Firstly, the selected anatomical and functional models are almost certainly simplifications and heavily dependent upon preconceptions of the experimenter. We have tried to counter this criticism, as far as possible, by selecting components of the model on the basis of a broad analysis, indicating those regions which were sensitive to the activation task. Pathways were also selected on the basis of known neuroanatomical connections and further informed by our hypotheses about the brain systems likely to be involved in the two types of learning. It has been suggested that structural equation modelling, when used in the analysis of functional neuroimaging data, is less open to the possibility of introducing erroneous pathways into the model since these pathways can be derived from anatomical observations (McIntosh and Gonzales-Lima, 1994), a luxury not afforded exploring causality in, for example, the social domain.

Another criticism concerns the extent to which the technique enables inferences concerning causal or influential relationships between brain regions. Broadly, the contention is that causality may never be truly proved, an argument which dates back to Hume (1777). One can, at best, state that data accord with a given causal model. We accept this general view and treat the data with necessary caution with respect to inferring causal relationships in the brain. Notwithstanding these limitations, we suggest that effective connectivity provides a useful conceptual approach to characterizing interactions between brain regions. Characterizing learning-related changes in terms of evolving path strengths within brain systems is likely to provide a powerful analytical tool, and such techniques, which conceive brain function as the interaction of anatomically separate

regions, are likely to provide the basis for more comprehensive accounts of functional integration in the brain.

Notes

We are indebted to Richard Frackowiak and to the Physics group at the Wellcome Department of Cognitive Neurology, particularly Robert Turner and Chloe Hutton. We also thank radiography staff, particularly Helen Gallagher. P.C.F., O.J., C.B., K.F. and R.J.D. are supported by the Wellcome Trust.

Address correspondence to Paul Fletcher, C. and O. Vogt Institut für Hirnforschung, Heinrich-Hein Universität, Universitätsstrasse 1, D 40225, Düsseldorf, Germany. Email: fletcher@hirn.uni-duesseldorf.de.

References

- Bollen KA (1989) Structural equations with latent variables. New York: Wiley.
- Bransford JD, Johnson MK (1972) Contextual prerequisites for understanding: some investigations of comprehension and recall. *J Verb Learn Verb Behav* 11:717-726.
- Buckner RL, Raichle ME, Miezin FM, Petersen SE (1996) Functional anatomic studies of memory retrieval for auditory words and visual pictures. *J Neurosci* 16:6219-6235.
- Büchel C, Friston KJ (1997) Modulation of connectivity in visual pathways by attention: interactions evaluated with structural equation modelling and fMRI. *Cereb Cortex* 7:768-778.
- Craik FI, Lockhart RS (1972) Levels of processing: a framework for memory research. *J Verb Learn Verb Behav* 11:671-684.
- Dolan RJ, Fletcher PC (1997) Dissociating prefrontal and hippocampal function in episodic memory encoding. *Nature* 388:582-585.
- Fletcher PC, Shallice T, Frith CD, Frackowiak RSJ, Dolan RJ (1996) Brain activity during memory retrieval: the influence of imagery and semantic cueing. *Brain* 119:1587-1596.
- Fletcher PC, Frith CD, Rugg MD (1997) The functional neuroanatomy of episodic memory. *Trends Neurosci* 20:213-218.
- Fletcher PC, Shallice T, Dolan RJ (1998) The functional roles of the prefrontal cortex in episodic memory: I Encoding. *Brain* 121:1239-1248.
- Friston KJ, Frith CD, Passingham RE, Liddle PF, Frackowiak RSJ (1992) Motor practice and neurophysiological adaption in the cerebellum: a positron tomography study. *Proc R Soc Lond B* 248:223-228.
- Friston KJ, Ashburner J, Frith CD, Poline J-B, Heather JD, Frackowiak RSJ (1995a) Spatial registration and normalisation of images. *Hum Brain Map* 2:165-189.
- Friston KJ, Holmes AP, Worsley KJ, Poline JB, Frith CD, Frackowiak RSJ (1995b) Statistical parametric maps in functional imaging; a general linear approach. *Hum Brain Map* 2:189-210.
- Gerstein GL, Perkel DH (1969) Simultaneously recorded trains of action potentials: analysis and functional interpretation. *Science* 164:828-830.
- Haxby JV, Ungerleider LG, Horwitz B, Maisog JM, Rapoport SI, Grady CL (1996) Face encoding and recognition in the human brain. *Proc Natl Acad Sci USA* 93:922-927.
- Holmes AP, Josephs O, Büchel C, Friston KJ (1997) Statistical modelling of low frequency confounds in fMRI. *NeuroImage* 5:S480.
- Hume D (1777) Enquiries concerning human understanding and concerning the principles of morals. Oxford: Clarendon Press (10 edn).
- Johnstone T, Shanks DR (1997) Similarity effects in artificial grammar learning. In: Proceedings of SimCat 1997: an interdisciplinary workshop on similarity and categorisation (Ramsar M, Hahn U, Cambouropoulos E, Pain H, eds). Edinburgh: Department of Artificial Intelligence, Edinburgh University.
- Kapur S, Craik FIM, Tulving E, Wilson AA, Houle S, Brown GM (1994) Neuroanatomical correlates of encoding in episodic memory: levels of processing effect. *Proc Natl Acad Sci, USA* 91:2008-2011.
- Kapur S, Craik F, Brown GM, Houle S, Tulving E (1995) Functional role of the prefrontal cortex in memory retrieval: a PET study. *NeuroReport* 6:1880-1884.
- Karni A, Meyer G, Jezzard P, Adams MM, Turner R, Ungerleider L (1995) Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature* 377:155-158.
- Knowlton BJ, Ramus SJ and Squire LR (1992) Intact artificial grammar

- learning in amnesia: dissociation of classification learning and explicit memory for specific instances. *Psychol Sci* 3:172-179.
- Mathews RC, Buss RR, Stanley WB, Blanchard-Fields F, Cho JR, Druhan B (1989) Role of implicit and explicit processes in learning from examples: a synergistic effect. *J Exp Psychol Learn Mem Cognit* 15:1083-1100.
- McIntosh AR, Gonzales-Lima F (1994) Structural equation modelling and its application to network analysis in functional brain imaging. *Hum Brain Map* 2:2-22.
- Meulemans T, Van der Linden M (1997) Associative chunk strength in artificial grammar learning. *J Exp Psychol Learn Mem Cognit* 23:1007-1028.
- Nolde SF, Johnson MK, Raye CL (1998) The role of prefrontal cortex during tests of episodic memory. *TICS* 2:399-406.
- Nyberg L, Tulving E, Habib R, *et al.* (1995) Functional brain maps of retrieval mode and recovery of episodic information. *NeuroReport* 7:249-252.
- Poldrack RA, Desmond JE, Glover GH, Gabrieli JDE (1998) The neural basis of visual skill learning: an fMRI study of mirror reading. *Cereb Cortex* 8:1-10.
- Raichle MA, Fiez JA, Videen TO, *et al.* (1994) Practice-related changes in human brain functional anatomy during nonmotor learning. *Cereb Cortex* 4:8-26.
- Reber AS (1967) Implicit learning of artificial grammars. *J Verb Learn Verb Behav* 6:855-863.
- Rugg MD, Fletcher PC, Frith CD, Frackowiak RSJ, Dolan RJ (1996) Differential activation of the prefrontal cortex in successful and unsuccessful memory retrieval. *Brain* 119:2073-2083.
- Schacter DL, Alpert NM, Savage CR, Rauch SL, Albert MS (1996a) Conscious recollection and the human hippocampal formation: evidence from positron emission tomography. *Proc Natl Acad Sci USA* 93:321-325.
- Schacter DL, Curran T, Galluccio L, Milberg WP, Bates JF (1996b) False recognition and the right frontal lobe: a case study. *Neuropsychologia* 34:793-808.
- Shallice T, Fletcher P, Frith CD, Grasby P, Frackowiak RSJ, Dolan RJ (1994) Brain regions associated with acquisition and retrieval of verbal episodic memory. *Nature* 368:633-635.
- Squire LR, Ojemann JG, Miezin FM, Petersen SE, Videen TO, Raichle ME (1992) Activation of the hippocampus in normal humans: a functional anatomical study of memory. *Proc Natl Acad Sci USA* 89:1837-1841.
- Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. Stuttgart: Thieme-Verlag.
- Tulving E (1983) *Elements of episodic memory*. Oxford: Oxford University Press.
- Tulving E, Kapur S, Craik FIM., Moscovitch M, Houle S (1994a) Hemispheric encoding/retrieval asymmetry in episodic memory: positron emission tomography findings. *Proc Natl Acad Sci USA* 91:2016-2020.
- Tulving E, Kapur S, Markovitsch HJ, Craik FIM., Habib R, Houle S (1994b) Neuroanatomical correlates of retrieval in episodic memory: auditory sentence recognition. *Proc Natl Acad Sci USA* 91:2012-2015.
- Wheeler MA, Stuss DT and Tulving E (1997) Towards a theory of episodic memory: the frontal lobes and auto-noetic consciousness. *Psychol Bull* 121:331-354.