On the Benefits of not Trying: Brain Activity and Connectivity Reflecting the Interactions of Explicit and Implicit Sequence Learning

Under certain circumstances, implicit, automatic learning may be attenuated by explicit memory processes. We explored the brain basis of this phenomenon in a functional magnetic resonance imaging (fMRI) study of motor sequence learning. Using a factorial design that crossed subjective intention to learn (explicit versus implicit) with sequence difficulty (a standard versus a more complex alternating sequence), we show that explicit attempts to learn the difficult sequence produce a failure of implicit learning and, in a follow-up behavioural experiment, that this failure represents a suppression of learning itself rather than of the expression of learning. This suppression is associated with sustained right frontal activation and attenuation of learning-related changes in the medial temporal lobe and the thalamus. Furthermore, this condition is characterized by a reversal of the fronto-thalamic connectivity observed with unimpaired implicit learning. The findings demonstrate a neural basis for a well-known behavioural effect: the deleterious impact of an explicit search upon implicit learning.

Keywords: explicit, fMRI, frontal, implicit, learning, medial temporal, thalamic

Introduction

Brain systems associated with motor sequence learning have been studied widely using functional neuroimaging. Variations of this paradigm have proven valuable in the exploration of brain systems concerned with both declarative (explicit) and procedural (implicit) memory (Jenkins et al., 1994; Grafton et al., 1995, 1998; Hazeltine et al., 1997; Rauch et al., 1997, 1998; Toni et al., 1998). One frequently used learning paradigm is the serial reaction time (SRT) task (Nissen and Bullemer, 1987). In this task, subjects are usually required to respond rapidly to shifting spatial positions, with or without the awareness that there is a sequence to these shifts. Reductions in reaction time, over and above those seen when there is no repeating sequence, are indicative of learning and the accompanying changes in task-dependent brain systems have been a focus for PET (Grafton et al., 1995, 1998; Rauch et al., 1995; Doyon et al., 1996; Hazeltine et al., 1997; Honda et al., 1998) and functional magnetic resonance imaging (fMRI) (Rauch et al., 1997, 1998; Ungerleider et al., 2002; Willingham et al., 2002; Schendan et al., 2003) studies. More recently, attempts have been made to evaluate the impact of modifications in parameters such as sequence length and complexity upon accompanying brain activity (Catalan et al., 1998; Harrington et al., 2000).

The patterns of learning-related brain changes in these studies of the SRT task depend upon a number of task and performance-related features. Striatal activation for example correlates with the degree of learning across individuals (Rauch et al., 1997). Cortical changes appear to be sensitive to subjects' awareness of the sequence. For example, the motor cortex is active when awareness is diminished by dual task performance, that is, when learning is likely to be implicit. Conversely, the prefrontal, premotor and temporal cortex are sensitive to explicit sequence awareness that occurs during single task performance (Hazeltine et al., 1997). This distinction between implicit and explicit sequence learning provides an important clue to the origins of inconsistency among the SRT task imaging studies (Honda et al., 1998). A number of studies have manipulated sequence awareness to allow a direct comparison of implicit and explicit learning (Grafton et al., 1995; Rauch et al., 1995; Hazeltine et al., 1997; Honda et al., 1998; Willingham et al., 2002; Schendan et al., 2003). Some of these studies show little evidence of implicit-explicit overlap (Grafton et al., 1995; Rauch et al., 1995; Hazeltine et al., 1997; Honda et al., 1998) and thus appear to contradict an important behavioural observation: that implicit SRT sequence learning occurs even when the task is made explicit (Willingham and Goedert-Eschmann, 1999; Willingham et al., 2002). Willingham et al. (2002) suggest that this absence of overlap may arise from a number of experimental confounds, including insensitivity, order effects and the use of dual task designs. Using a design that obviated these confounds, they demonstrated that a sub-group of areas (ventrolateral prefrontal cortex (PFC), parietal cortex and putamen) show common activation during learning. This overlap is important: it suggests that there are a set of brain regions signifying implicit learning and that, even when explicit learning occurs, these show the same patterns of activity, suggesting that the presence of explicit learning does not impair implicit learning. Subsequently, Schendan et al. (2003) have replicated these findings with respect to PFC and putamen and provided evidence too that medial temporal cortex changes reflect learning both for implicit and explicit conditions.

Thus, behavioural (Curran and Keele, 1993; Willingham and Goedert-Eschmann, 1999) and imaging (Willingham et al., 2002; Schendan et al., 2003) studies provide evidence that implicit learning during SRT task performance proceeds unmolested by explicit learning. Nevertheless, there is also evidence that, under some circumstances, explicit processing does produce impairment. This has been shown in other forms of implicit learning, for example of artificial grammars (Reber, 1976). More recently, Howard and Howard (2001) have demonstrated that, under certain circumstances, explicit attempts to learn a sequence attenuates the implicit learning of that sequence. They suggest that, since sequence learning demands the detection of covariance across time, then it demands cognitive resources.
Usually, this demand is meagre but, if it is increased (e.g. if there is noise in the sequence) or if cognitive capacity is reduced (e.g. in the elderly) then implicit learning becomes vulnerable. They demonstrated this using a more demanding SRT task (an ‘alternating SRT task’ in which sequence elements alternate with non-sequence or noise elements). They showed that, in an elderly population, the predicted hindrance effects of explicitly trying to learn the sequence did, indeed, occur (Howard and Howard, 2001).

To summarize, the interaction between explicit and implicit learning systems appears to be condition-dependent. The evidence that explicit learning of sequences does not impair implicit learning (Curran and Keele, 1993; Willingham and Goedert-Eschmann, 1999; Willingham et al., 2002; Schendan et al., 2003) is restricted to conditions in which there is sufficient cognitive capacity available. When this is no longer the case, there is a deleterious effect of explicit processing (Howard and Howard, 2001). While a possible brain basis for the implicit–explicit interaction has been shown under categorical learning conditions (Poldrack et al., 2001), it remains unclear with respect to sequence learning tasks. We used fMRI to address this question. Subjects were exposed to two types of sequences: a standard SRT task sequence (Nissen and Bullemer, 1987) and a more complex alternating sequence comparable to that used by Howard and Howard (1997, 2001). In addition, we used a control task with random occurrence of elements. For each of these sequence types, learning was either explicit (i.e. subjects were instructed to find and learn the sequence) or implicit (i.e. subjects were instructed just to carry out the button-push response). We predicted that explicit instructions in association with the more difficult alternating SRT task would produce an attenuation of implicit learning. Thus, this condition allowed us to identify directly the brain correlates of the explicit–implicit interaction. We did this through a series of analyses determining the main differences in activation levels between the explicit and implicit alternating SRT tasks, through an exploration of the differences in time-dependent changes accompanying repeated exposure to the sequence and, finally, through identification of changes in connectivity produced by the requirement to process sequences explicitly.

We report two studies: experiment 1 explored the interaction using fMRI; experiment 2 was a follow-up behavioural study aiming to establish that the explicit search did indeed impair sequence learning, rather than simply preventing its expression. In keeping with recent functional neuroimaging work on sequence learning (Schendan et al., 2003), we chose a within-subjects design. Of course, a potential drawback with such a design is that explicit learning during implicit conditions may be more likely since subjects will be aware that sequences may be present in some cases. This does not, however, produce a confound with respect to interpreting differences between the implicit and explicit learning conditions, which is the focus of interest in our analysis.

**Experiment 1 — fMRI**

**Materials and Methods**

Eleven healthy volunteers (average age = 29 years, range = 24–42 years, seven males) were recruited by advertisement and each attended the Institute of Medicine at the Forschungszentrum Jülich on one occasion. All were free from neurological or psychiatric illness and showed normal structural MRI scans. The local ethics committee approved the study and written informed consent was obtained from each volunteer. Due to technical problems, data from one of the subjects was unusable and 10 datasets were therefore used in the final imaging analysis. Furthermore, because of technical difficulties, data from one of the conditions for one of the ten subjects in whom scanning data were successfully acquired were lost. Data from nine subjects therefore contributed to the behavioural analysis.

**Behavioural Task**

During a single visit, each subject underwent six 6 min scanning sessions. Each session encompassed a different experimental condition (alternated with a baseline task). After pre-scan instructions and practice at pushing the appropriate buttons in response to stimuli on the computer screen, subjects entered the scanner and, prior to each set of image acquisition, were given a single instruction that informed them as to whether they should, during the forthcoming task, look for a repeating sequence or simply concentrate on carrying out the task irrespective of whether a sequence was or was not present. They were informed that sometimes a sequence would indeed be present and at other times it would not.

The presentation program used was MEL (MEL Professional, Version 2.0, Psychology Software Tools Inc., Pittsburgh PA) and a Polaroid Polarview 222 projector was used to deliver stimuli to a screen positioned comfortably within the subject’s field of view. Prior to scanning, we ensured that each subject was able to see stimuli without difficulty.

For each activation block (in each of the six conditions), stimulus presentation was the same: four empty boxes were present throughout and, when the SRT task started, one of them would be filled by a black square (see Fig. 1). The subject was required to press the button on a four-button keypad corresponding to the position of the filled box. They received prior instructions to make the button push as rapidly as they could while striving to avoid errors. We suggested to them that the most important goal was accuracy. The response and reaction time at each trial were recorded. Boxes were filled for 800 ms, there was then a 200 ms period during which all four boxes were empty before the next trial began (a different box was filled and the subject was required to respond accordingly). This design produced an inter-trial interval of 1 s. These were held constant across all trials and conditions in order to avoid different numbers of trials occurring within scanning sessions. Each scanning session comprised twelve 30 s blocks (six blocks of activation/SRT task alternating with six blocks of visual fixation). Each activation block consisted of 30 button-push trials. Since each sequence, if present, was spread over 10 trials, subjects were exposed to three
repetitions of a sequence per block. Thus, each subject was exposed to a total of 18 repetitions of each sequence.

**Experimental Factors**

Scanning data were acquired in six conditions. There were four conditions of primary interest and these embodied the manipulation of two factors:

**Sequence Difficulty.** In two of the scanning sessions, subjects were exposed to a standard SRT task (Nissen and Bullemer, 1987) in which a repeating sequence of 10 positions occurred (see Fig. 1). In the other two activation conditions, subjects were exposed to an alternating SRT task (Howard and Howard, 1997, 2001), in which only alternative elements were part of a sequence, with intervening elements differing across repetitions of the sequence. Thus, for the standard SRT, a typical sequence might be 1 3 4 2 3 1 3 4 2 3, where 1-4 are the four positions on the screen. For the alternating SRT, a sequence might be 1 x 3 x 4 x 2 x 1, where x = positions 1-4 which changed on subsequent repetitions. As we outline below, there were certain constraints placed upon these non-sequence elements, meaning that they were not truly random but probabilistic in nature.

**Prior Instructions.** For each of the sequence types, subjects were either given explicit instructions to search for a sequence or were told merely to concentrate on responding to successive stimuli as accurately and rapidly as possible. In keeping with convention, we shall refer to the former conditions as explicit learning conditions and the latter as implicit learning. We do so with the caution that, this manipulation in itself is unlikely to produce a reliable distinction between implicit and explicit memory systems. Rather, as will be described below, the effects that we were most interested in were produced by the interaction between the type of sequence and the task instructions.

Thus, this factorial manipulation generated four conditions of interest:

1. Explicit learning in standard SRTT (EsSRT)
2. Explicit learning in alternating SRTT (EaSRT)
3. Implicit learning in standard SRTT (IsSRT)
4. Implicit learning in alternating SRTT (IaSRT)

In addition to these tasks of interest, images were obtained for each subject during two further conditions. In both of these, successive positions were random. In one case (implicit random), subjects were given no search instructions (as for the implicit sequence conditions; see above). In the other, despite the absence of the sequence (unknown to subjects) they were instructed to search for a sequence. These were designed to serve as control conditions, matched for visual input and response patterns. In fact, unsurprisingly, subjects found the explicit search for a sequence in the random condition a rather frustrating exercise, one which they quickly gave up on. In view of this we decided not to use the imaging data from this condition in our full analysis but used solely the non-search random condition as our baseline task. The condition did, however, serve another important purpose — indicating to subjects that there were indeed conditions in which no sequence was present.

All conditions (sequence and non-sequence) were subject to a number of constraints [no immediate repetitions, no runs (e.g. 4, 3, 2, 1), no trills (e.g. 1, 2, 1, 2, 1, 2)]. These constraints mean that, for the alternating SRT conditions, the non-sequence elements could not be randomly selected from the four positions, but had to be selected in order to avoid repetition. We do not believe that this is a critical problem for two reasons. First, the contrasts of interest compared sequences of the same type (with only instructions differing). Secondly, the constraint on the inter-sequence noise might predispose us, if anything, to failing to find reaction time (RT) differences between sequence and non-sequence elements.

The order of the six conditions was randomized for each of the subjects.

**Behavioural Data**

On-line behavioural data were collected in the form of reaction-times and accuracy. For every repetition of a sequence the median RT was calculated and the medians were averaged within blocks. Thus for each subject, the median RT values for each condition entered into a repeated measures ANOVA. Since we were primarily concerned with identifying differential learning effects across the four activation conditions, the block-by-block RTs for these conditions were entered into the same 2 x 3 x 6 (instruction x sequence type x block) ANOVA in order to allow the relevant comparisons. Learning associated with each of the four activation conditions (EsSRT, IsSRT, EaSRT and IaSRT) was tested against non-specific time-dependent change in the random (no-sequence) conditions. Note that, for the IsSRT and EaSRT median RTs only for the elements composing the sequence were taken (i.e. not the non-sequence elements in alternate positions alternated).

In addition to these measures, after each condition, subjects were debriefed over the scanner intercom. Subjects were asked if they had noticed any repeating sequence. If they answered to the contrary, they were further probed as to whether they got the feeling at any stage that they could predict which position would be highlighted next and, if possible to give an example. If they answered that they had noticed a sequence, or fragments of sequence, they were asked to reproduce these by reciting the positions (in terms of the numbers 1-4) constituting those parts of the sequence that they had noted.

**Scanning**

Images were acquired using a Siemens Vision scanner with echo-planar imaging (EPI) capability operating at 1.5 T. A gradient echo EPI sequence was used \((T_R = 66 \text{ ms}, T_E = 5 \text{ ms}, \text{ flip angle } = 90^\circ)\). Thirty slices, each of 4mm thickness (inter-slice gap = 0.4 mm, field of view = 200 mm, in-plane resolution = 3.125 x 3.125 mm) were imaged. Using a mid-sagittal scout image, slices were orientated in the plane of the anterior-posterior commissure line. An additional high resolution anatomical image was acquired for all subjects, using the 3-D magnetization-prepared, rapid-acquisition gradient-echo sequence with the following parameters: \(T_E = 4.4 \text{ ms}, T_R = 11.4 \text{ ms}, \text{ flip angle } = 15^\circ, \text{ inversion time } = 300 \text{ ms}, \text{ matrix } = 200 \times 256, \text{ field of view } = 200 \text{ mm}, 128 \text{ sagittal slices, slice thickness } = 1.33 \text{ mm}\).

For each subject, a total of six functional images per block were acquired, giving a total of 72 per session and 432 images per subject.

**fMRI Analysis**

Analysis was carried out using statistical parametric mapping software: SPM2 (Wellcome Department of Imaging Neuroscience, London, UK, www.fil.ion.ucl.ac.uk), comprising the following stages:

**Image Pre-processing**

Each volume was realigned to the first volume in the session and then realigned across sessions for each subject. Following estimation of movement parameters, this realignment was carried out with resampling of voxels using a sinc interpolation in space. A mean image was then normalized to standard stereotactic space [based on the Montreal Neurological Institute reference brain (Cocosco et al., 1997)] using a 12 parameter affine transformation and non-linear warping using basis functions (Ashburner and Friston, 1999). The parameters derived from this transformation were used to warp all volumes to the standard brain space. Finally, images were spatially smoothed using a Gaussian kernel (8 mm full-width at half-maximum).

**Modelling BOLD Responses**

The time series in each session was high-pass filtered to 1/120 Hz to remove low-frequency noise and corrected for temporal autocorrelation using an AR(1)+ white noise model. Activation blocks within each session were modelled using a box car function convolved with a haemodynamic response function. As well as a static function modelling a consistent effect across the six blocks comprising a session, we also introduced into the model a linear time-dependent modulation of the block effect in order to identify regional activations whose magnitude changed as a result of learning.

In each session, for each subject, we computed the magnitude of effect (change in BOLD signal relative to the recurring baseline task). This was expressed as a parameter estimate, beta, and these were compared across conditions within subjects to produce contrast images. Each of the contrast images for each subject was taken to a second level analysis in which group effects were explored using a t test, treating inter-subject variability as a random effect.
Group effects were explored and expressed using the following comparisons.

(i) Time-independent Effects. Here, we identified regions in which there was a sustained increase (i.e. independent of time-dependent learning effects) in task-related activation (relative to recurring baseline fixation task) that differed according to the two experimental factors. We thus carried out an analysis of main effects of task instructions: (EsSRT plus EaSRT) versus (IsSRT plus IaSRT) and analysis of the interaction between task instructions and the nature of the sequence.

(ii) Time-dependent/Learning Effects. Main effects of learning in each of the four activation tasks were identified by direct comparison with time dependent effects in association with the random (no-sequence) condition in which subjects were not issued instructions to learn. In brief, we wished to identify brain regions showing a time-dependent change in the sequence learning conditions and to be confident that these changes were not non-specific time effects. These non-specific time-effects were estimated using the random condition (i.e. time-dependent changes in activation during this task relative to its recurring baseline) and subtracted away from those in the sequence conditions. This left us with changes in activation magnitude for each of the four conditions of interest that we would be able to interpret more specifically in terms of learning.

Learning was modelled in terms of time-dependent linear increases and decreases in task-related activity. We were primarily interested in the effects of sequence type and task instructions upon learning effects. We therefore identified the interactions in time-dependent effects produced by manipulations in these factors.

Analysis of Task-dependent Brain Interactions. Having identified the brain regions in which there was a time-independent difference in task-related activation resulting from sequence type and task instructions and regions in which these experimental variables were reflected in modulations of learning-dependent activations, we finally sought to specify the effects of the experimental manipulation of primary interest (EsSRT versus IaSRT) in terms of within-system correlations. We did this through implementing an analysis of ‘psychophysiological interactions’ (Friston et al., 1997). In this approach, we identify condition-specific correlations between activity in a chosen region and activity elsewhere. The correlation was formulated as an index of the influence of regionally-mediated processes engendered by explicit instructions upon regions associated with implicit learning. Since our goal was to investigate the deleterious effects of explicit processing during EaSRT our psychophysiological interaction determined regions in which there was a significantly different correlation with activity in right PFC as a function of whether subjects were performing the aSRT task or were visually fixating and, critically, whether these differences in correlation differed between the EsSRT and IaSRT conditions. For completeness, we also report key interactions for the sSRT tasks.

Thresholding Strategy

In order to maximize sensitivity while at the same time reducing the risk of type II error, the strategy for analysis was to use a set of regions of interest based upon an a priori mask (Maldjian et al., 2003) drawn from regions that have been shown to be involved in sequence learning: prefrontal and anterior cingulate cortices, thalamus, striatum, and medial temporal lobes. In this case, only regions surviving the false discovery rate (FDR) correction for this entire volume are reported. This formed the initial basis for thresholding and the strategy was as follows. In analysing time-independent block effects, we used an approach controlling for false discovery rate (Benjamini and Hochberg, 1995) implemented with SPM2 (Genovese et al., 2002) within the regions of interest. Previous approaches to the multiple comparisons problem in fMRI have used a Bonferroni-type correction but it has been noted that the level of type II error may be unacceptably high with such an approach (Benjamini and Hochberg, 1995) (Genovese et al., 2002). Controlling for false discovery rate (FDR) obviates the multiple comparisons problem. Its goal is to control the proportion of false positives results among the activations (i.e. it controls the proportion of falsely rejected null hypotheses among all rejected null hypotheses). Thus for each region identified using FDR control, we can specify the likelihood that this was a type I error. We chose a FDR threshold of \( P < 0.05 \). In addition, for completeness, we additionally report uncorrected data (\( P < 0.001 \)) lying within this system of interest.

In modelling time-dependent (i.e. learning) effects, an initial FDR threshold of \( P < 0.05 \) for the entire volume of interest was selected. We then explored condition-specific learning effects using a small volume correction for the regions identified by this comparison [medial temporal lobes (MTL) and left head of caudate (see below)].

In order to take into account condition-specific learning effects in regions lying outside the system showing a main effect of learning for all conditions, we also report regions outside of this system that survived FDR thresholding \( P < 0.05 \). Subsequent connectivity analyses employed small volume corrections based upon regions identified by the above analyses.

Results

Behavioural Results

Plots of RTs across the learning period are presented in Figure 2 for each of the activation tasks (EsSRT, IsSRT, EaSRT and IaSRT). Within the ANOVA, we restricted ourselves to the identification of linear time effects, ignoring all non-linear effects. We tested for learning-related effects by comparing, in each case, the profile of time-dependent change relative to that found in the relevant random condition (i.e. the condition in which any time-dependent change may be considered a non-specific effect of practice). For the two sSRT tasks compared with this random condition, there were clear time-dependent effects on RT \( [\text{IsSRT}: F(1,8) = 10.1; \ P < 0.05; \text{EsSRT}: F(1,8) = 8; \ P < 0.05] \). For the IaSRT, there was a strong trend towards a main behavioural effect of learning \([F(1,8) = 4.2; \ P = 0.07]\) but not for EaSRT \([F(1,8) = 0.01; \ P = 0.9]\).

The critical finding is that there was a significant attenuation of the time-dependent reduction in RT for the aSRT when subjects were issued explicit instructions to look for this sequence (Fig. 2, lower panel). This attenuation was significant as a task by time interaction \([F(1,8) = 0.1, \ P = 0.05]\). That is, there was a difference in the shapes of the learning curve depending upon whether or not subjects were issued explicit instructions to look for the sequence (Fig. 2, lower panel).

Figure 2. Reaction time data. The median reaction time was calculated for each sequence repetition and the mean of these was calculated to produce the average reaction time in each of the six blocks. These are presented above for the explicit standard SRT task (EsSRT) and the implicit standard SRT task (IsSRT) in the upper panel, and for the explicit alternating SRT task (EaSRT) and the implicit alternating SRT task (IaSRT) in the lower panel.
not explicit instructions had been given. This provides confirmation that a high-demand alternating SRT task is vulnerable to interference when subjects are actively looking for the sequence. Comparison of EsSRT with IsSRT showed a strong trend towards a main effect of the task \( F(1.8) = 5.1; P = 0.054 \). It thus appears that the instructions to explicitly learn the sequence lead to a general RT penalty but not to a failure of learning-related RT decreases \( F(1.8) = 0.002; P = 0.97 \).

With regard to post-task debriefing, in both sSRT tasks, subjects noticed a sequence and were able to reproduce parts of it. For the IsSRT task, all subjects were conscious that there had been a sequence and were able to reproduce a mean \( \pm 3.6 \pm 1.8 \) elements, with \( 0.3 \pm 0.5 \) errors. Thus, in this case, explicit learning had taken place to a degree even though no instructions had been issued to look for a sequence. This is not surprising, given that subjects were exposed to the sequence as part of a within-subjects design and were therefore aware that one might be present. For the EsSRT task, all subjects were conscious that there had been a sequence and were able to reproduce \( 7.1 \pm 1.4 \) elements, with \( 0.5 \pm 0.7 \) errors. The difference in post-learning explicit retrieval (in terms of number of sequence elements produced) was significant (two-tailed, paired \( t \)-test, \( df = 9, P < 0.01 \)).

For the alternating SRT task, explicit recall was substantially less. For the IsSRT task, no subject was able to reproduce any of the sequence. Two subjects conceded that, perhaps there may have been some repeating pattern but were unable to be more specific than this. In the EaSRT task, eight subjects were conscious that there had been a sequence and the mean number of elements reproduced was \( 2.5 \pm 1.8 \) elements, with \( 0.2 \pm 0.4 \) errors. The difference in post-learning explicit retrieval (in terms of number of sequence elements produced) was significant (two-tailed, paired \( t \)-test, \( df = 9, P < 0.001 \)).

**Brain Activations**

**Time-independent Effects.** We performed comparisons of activations (relative to a recurring visual fixation baseline) for the following contrasts:

1. **EsSRT versus IsSRT** — This comparison was set up to identify the effects of explicit versus implicit sequence processing in a standard SRT task. No significant time-independent brain activations differentiated these two tasks.
2. **EaSRT versus IsSRT** — The explicit attempt to learn the alternating sequence was associated with increased levels of activity in prefrontal cortex, predominantly on the right — see Table 1 and Figure 3.

A direct test of the task by sequence type interaction [(EaSRT versus IsSRT) versus (EsSRT versus IsSRT)] showed evidence for a trend \( (P < 0.005, \) uncorrected) towards greater frontal activation when explicitly trying to learn the alternating sequence than when trying to learn the standard sequence. This is unsurprising given the differential levels of difficulty of the two sequences.

**Time-dependent (Learning) Effects.** (i) Main Effects. In the first analysis, we identified regions showing a time-dependent change across all learning conditions combined. Time effects in the learning condition were expressed relative to time-effects in the implicit random condition, the latter being considered a reflection of non-specific time-dependent changes. We examined the data both for increases and decreases in activation across the six blocks of each of the conditions. Significant decreases were found in left medial temporal cortex in the region of the hippocampus and the parahippocampal gyrus. A trend towards a learning-related decrease in activation was also found in caudate nucleus (FDR threshold, \( P = 0.08 \)). These results are summarized in Table 2 and Figure 4. The images in Figure 4 are thresholded at \( P < 0.01 \) (uncorrected for multiple comparisons) in order to show a trend towards bilaterality of time-dependent decreases in MTL and caudate.

No significant increases in activation were found for the combined-conditions learning analysis. In view of the fact that a number of previous studies have identified and interpreted learning-dependent increases in brain activation in association with comparable tasks (Grafton et al., 1995, 1998; Hazeltine et al., 1997; Honda et al., 1998), we extended this analysis to identify increases in an unmasked comparison at a lower threshold \( (P < 0.005, \) uncorrected for multiple comparisons). The results of this analysis are presented in Table 2. Consistent with previous work, we identified increases in activation accompanying learning in right lateral frontal, medial frontal, motor and parietal cortex.

(ii) Condition-by-learning Interactions. This analysis was carried out to identify the effects of sequence type and of instructions upon time-dependent changes in activation. We made the following comparisons:

1. **EaSRT versus IsSRT** — Two regional differences were noted in this comparison. The MTL learning-related decrease was significantly greater for IsSRT than for EaSRT. A region within the left thalamus also showed significantly different learning effects between implicit and explicit learning in the asSRT task. A significantly greater linear decrease in activation was seen in the EaSRT task. These results are summarized in Table 2 and Figure 5.

Note, in Figure 5, an apparent non-linearity in the plots, wherein it appears that the early stages of learning (the first three blocks, corresponding to the first nine presentations of the sequence) were associated with sustained (and possibly even increasing) MTL activation, with reductions appearing from the fourth block onwards. While wishing to draw attention to this, we are cautious about interpreting these apparent non-linearities since our statistical model was set up to look for linear increases and decreases in activation. (To remodel our data on the basis of our observations would be circular since it would involve attempting to test a hypothesis that was indicated by the initial analysis of the data.)

2. **EsSRT versus IsSRT** — No significant differences in learning-related modulations of activation were seen in either direction.

---

Table 1

<table>
<thead>
<tr>
<th>Region</th>
<th>x*</th>
<th>y*</th>
<th>z*</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>EsSRT vs. IsSRT</td>
<td>-48</td>
<td>20</td>
<td>23</td>
<td>3.7</td>
</tr>
<tr>
<td>Right lateral PFC</td>
<td>44</td>
<td>20</td>
<td>12</td>
<td>3.9</td>
</tr>
<tr>
<td>Left lateral PFC</td>
<td>46</td>
<td>36</td>
<td>13</td>
<td>3.9</td>
</tr>
<tr>
<td>EaSRT vs. IsSRT</td>
<td>-48</td>
<td>20</td>
<td>23</td>
<td>3.7</td>
</tr>
<tr>
<td>Right lateral PFC</td>
<td>48</td>
<td>20</td>
<td>23</td>
<td>3.1</td>
</tr>
</tbody>
</table>

**Figure 4** are thresholded at \( P < 0.01 \) (uncorrected for multiple comparisons).
We did not compare directly the EaSRT with the EsSRT or the IaSRT with the IsSRT. This was because there were necessarily different lengths of sequence associated with the aSRT and the sSRT conditions, making direct contrasts difficult to interpret.

**Analysis of Task-dependent Brain Interactions.** As described above, this final analysis — a psychophysiological correlation — aimed to specify the effects of the experimental manipulation of primary interest (EaSRT versus IaSRT) in terms of within-system correlations (Friston et al., 1997). In this approach, we identify condition-specific correlations between activity in a chosen region and activity elsewhere. We wished to explore correlations between right prefrontal cortex and regions that showed time-dependent changes in activation. This correlation is suggested to index the influence of frontally mediated processes engendered by explicit instructions upon regions associated with implicit learning.

We observed a striking impact of task instructions upon fronto-thalamic connectivity. For EaSRT compared with baseline, right PFC is a strong positive predictor of bilateral thalamic activity (see Fig. 5a and Table 3). For IaSRT compared with baseline, it was a strong negative predictor (see Fig. 5b and Table 3). A direct comparison of EaSRT and IaSRT confirmed a significant task-dependent effect upon fronto-thalamic connectivity (see Table 3 and Fig. 6). We also show in Table 3 the fronto-thalamic connectivity for the sSRT tasks. Note that, with the standard sequences, negative fronto-thalamic interactions were observed irrespective of task instructions.

**Summary**
We have shown that the explicit attempt to identify and learn a sequence has a number of effects both upon behaviour and the underlying brain activity. These effects were seen for the more difficult alternating sequence.

Reaction times showed significant time-dependent reductions (compared with a random sequence) for the standard SRT task. This was also the case for the alternating SRT task, but only...
if subjects were not actively seeking a sequence. In cases where subjects were looking for the alternating sequence, there was little improvement as measured by reaction time, although there was greater evidence of explicit post scan sequence recall.

With regard to brain activations, searching for the alternating sequence (EaSRT condition) was associated with the highest level of right PFC activity. Furthermore, it was differentiated from the other three conditions by an attenuation in learning-dependent reductions in left MTL activation and by a pattern of learning-dependent change in thalamic activity that differed significantly from that in the other conditions (a reduction in EaSRT as compared with trends towards increase in IaSRT, EsSRT and IsSRT).

Finally, expressing the brain responses to the alternating SRT tasks in terms of connectivity, fronto-thalamic patterns of interaction were strikingly different across EaSRT and IaSRT. Specifically, EaSRT was marked by positive and IaSRT by negative fronto-thalamic correlations.

**Experiment 2 — A Behavioural Study Exploring the Extent to Which the Behavioural Findings from Experiment 1 were Attributable to Suppressed Learning or to Prevention of its Expression**

It is possible that explicit instructions in the EaSRT task might suppress not the occurrence but the expression of implicit learning. That is, under explicit conditions, implicit learning might be continuing in the normal way but the decrease in RTs that would normally signify this could be masked by the cognitive demands of the explicit search. We therefore conducted a further behavioural study focusing upon the alternating SRT task under both implicit and explicit conditions (again using a within-subjects design to ensure comparability with the fMRI study). In this study, subjects were exposed to alternating sequences in precisely the same way as during the fMRI study, but, in this case, following the interposition of 20 random stimuli and a short break, they were re-exposed to the sequence following a briefing in which they were told not to search but simply to push the relevant buttons as quickly as possible. If the
suggestion is correct, that the search suppresses the expression of learning without actually preventing learning itself, then this re-exposure, when the search no longer occurs, should be associated with an advantage for sequence over non-sequence elements following both EaSRT and IaSRT. If the search does indeed suppress learning, then there should be a relative disadvantage following EaSRT compared with IaSRT. Twelve healthy volunteers (age range = 23–40 years, six males) were recruited for this study. Each volunteer was required to carry out two tasks over an ~15 min period.

**Behavioural Task**

As for the fMRI study, subjects were given basic instruction and practice in pushing the appropriate buttons in response to stimuli on the computer screen. They carried out modified versions of two of the SRT tasks used in the fMRI study: IaSRT and EaSRT. This was a within-subjects design and order of performance was counter-balanced across subjects. Prior to the EaSRT task, subjects were instructed to push the buttons quickly, avoiding errors and simultaneously to look for a repeating sequence. Prior to IaSRT, they were told simply to concentrate on pushing the buttons.

DMDX ([www.u.arizona.edu/~jforster/dmdx/official.htm](http://www.u.arizona.edu/~jforster/dmdx/official.htm)) was used as the presentation programme. It was run on a Dell Latitude laptop computer. The programme was set up to run with the same time parameters as in the fMRI study above. The two conditions were run as follows: the subject received instructions either to look for a sequence (EaSRT) or simply to

---

**Table 3**

Regions showing task-dependent correlation with right PFC

<table>
<thead>
<tr>
<th>Regions showing task-dependent functional connectivity right PFC (44, 20, 12)</th>
<th>$x^*$</th>
<th>$y^*$</th>
<th>$z^*$</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EaSRT versus visual fixation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increases</td>
<td>Right thalamus</td>
<td>12</td>
<td>-14</td>
<td>18</td>
</tr>
<tr>
<td>Left thalamus</td>
<td>-14</td>
<td>-8</td>
<td>20</td>
<td>2.7</td>
</tr>
<tr>
<td>Decreases</td>
<td>None significant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IaSRT versus visual fixation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increases</td>
<td>None significant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decreases</td>
<td>Right thalamus</td>
<td>8</td>
<td>-14</td>
<td>8</td>
</tr>
<tr>
<td>Left thalamus</td>
<td>-10</td>
<td>-4</td>
<td>10</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td>-18</td>
<td>-32</td>
<td>2</td>
<td>2.4</td>
</tr>
<tr>
<td><strong>EaSRT versus IaSRT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greater positive connectivity in EaSRT</td>
<td>Right thalamus</td>
<td>16</td>
<td>-14</td>
<td>10</td>
</tr>
<tr>
<td>Left thalamus</td>
<td>-16</td>
<td>-12</td>
<td>18</td>
<td>3.1</td>
</tr>
<tr>
<td>Greater positive connectivity in IaSRT</td>
<td>None significant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EaSRT versus visual fixation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decreases</td>
<td>Right thalamus</td>
<td>16</td>
<td>-15</td>
<td>10</td>
</tr>
<tr>
<td>Left thalamus</td>
<td>-10</td>
<td>-7</td>
<td>6</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>-22</td>
<td>-23</td>
<td>1</td>
<td>2.8</td>
</tr>
<tr>
<td><strong>IaSRT versus visual fixation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decreases</td>
<td>Right thalamus</td>
<td>6</td>
<td>-15</td>
<td>3</td>
</tr>
<tr>
<td>Left thalamus</td>
<td>-4</td>
<td>-13</td>
<td>4</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>-20</td>
<td>-21</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

---

**Figure 5.** Condition-specific differences in learning-related activation. Upper panel: comparison of EaSRT with IaSRT conditions. Left MTL shows an attenuation of learning-related reductions in activation when subjects are explicitly trying to learn the sequence. The region showing this interaction is superimposed upon a representative MRI ($x, y, z = -30, -30, -16$). The inlaid plot shows averaged levels of activity in this region (relative to the fixation condition) for the standard SRT tasks (upper plot) and the alternating SRT (lower plot). The lower plot reflects the significant attenuation of learning-related decrease when explicit instructions are given. Lower panel: posterior thalamic differences in learning-related activation. Activation differences are shown for the direct comparison between EaSRT and IaSRT and are superimposed upon a standard MRI ($x, y, z = -22, -24, 4$). The accompanying plots shows a time-dependent decrease in activation for these regions during EaSRT. This effect was not seen in any of the other conditions, which showed trends in the opposite direction.

**Figure 6.** Regions showing condition-dependent differences in functional connectivity with right PFC (EaSRT versus IaSRT). The upper panel shows the maximum intensity projection (thresholded at $P < 0.01$, uncorrected for multiple comparisons) produced by the comparison between frontal correlations (relative to visual fixation) during EaSRT and IaSRT. The thalamus shows bilaterally a significant difference in task-dependent frontal connectivity, being positively correlated with right PFC during EaSRT and negatively correlated during IaSRT. The lower panel shows the results of the same comparison rendered onto a MRI in standard space ($x, y, z = 16, -14, 10$).
push the buttons (IaSRT) and was then exposed to an alternating sequence of the same length (five sequence elements alternating with five non-sequence elements) as in the fMRI study. Eighteen repetitions of the sequence were seen. Following this, and continuous with it, 20 random elements were presented. Then, there was a pause during which subjects were asked if they had detected a sequence. If they had, they were asked to try to reproduce as much of it as they could using the appropriate buttons on the keyboard. If they had not, they were further asked if they had noticed any sort of patterns or if sometimes they felt that they could predict what button they would have to push next. If they had, they were asked to try to give examples. Following this break, they were told that they could now relax and they would be asked to continue pressing the buttons in response to stimuli. They were asked to refrain from looking for a sequence but simply to press the buttons as rapidly as they could. They were told that continuing to search for the sequence would impair performance so they should try to avoid doing so. Reaction time and accuracy data were recorded for each stimulus.

Data Analysis
We were primarily concerned with addressing the question of whether the previously observed absence of RT reductions in association with sequence elements in the EaSRT reflects an attenuation of learning or whether it simply occurs because searching for the sequence temporarily suppresses the expression of learning. Thus, our primary interest was in whether there was a RT reduction for sequence elements following the EaSRT condition when the search demands were removed. We focused therefore upon the reaction times associated with sequence and non-sequence elements during the first re-exposure to the sequence. For completeness, we also report here data from the learning stages. In order to present the data in a way that is comparable to that in association with experiment 1, we calculated the median RTs for each third repetition of the sequence in order to produce six values across the 18 presentations of the sequence for each of the two conditions (IaSRT and EaSRT). These values were entered into a 2 x 6 repeated measures ANOVA with the two factors being task (IaSRT versus EaSRT) and learning stage (1–6). (Note that this analysis was for the purpose of comparability with the fMRI study and perhaps does not constitute the optimal treatment of the behavioural data across the learning period.)

With respect to the primary goal of experiment 2, this was addressed by a paired t test comparing sequence versus non-sequence advantages following IaSRT with that following EaSRT.

Results
Findings are summarized in Figure 7. As for experiment 1, there was a reduction in reaction times for sequence compared with non-sequence elements during IaSRT \[F(1,11) = 6.19; P < 0.05\] but not EaSRT \[F(1,11) = 0.99; P = 0.34\]. The condition-by-learning interaction failed to reach significance \[F(1,11) = 2.0; P = 0.19\] although a post-hoc t test comparing the degree of learning achieved during IaSRT (as estimated by the overall increase in advantage at the end of learning compared with the beginning) was significantly greater than that achieved during EaSRT (one-tailed t test, df = 11, P < 0.05).

For the IaSRT task, no subject was able to reproduce any of the sequence. One subject felt that there may have been some

Figure 7. Data from behavioural study (experiment 2). Learning phase: the data are presented as for RTs in Figure 2. Median RTs for the (alternate) sequence elements across each set of three sequence repetitions was averaged across the twelve subjects. Data for the IaSRT and EaSRT tasks are shown with standard error bars. Re-exposure to sequence post-learning: again, for each subject, we subtracted RTs to each non-sequence element from RT to the neighbouring sequence element to produce a measure of the amount that had been learned. Median values for the first re-exposure were calculated for each subject and are presented for re-exposure following IaSRT and EaSRT together with standard error bars. The negative value following IaSRT indicates the impact of prior learning. No such effect was significant following EaSRT.

Summary
Thus, this follow-up study has replicated the observation that exposure to the alternating sequence is associated with reaction time decreases (for the sequence elements compared with the non-sequence elements) only when subjects are not searching. Critically, in this follow-up study we allowed subjects a break and then re-exposed them to the sequence that they had just seen. This re-exposure was preceded by instructions to ignore the possibility that there might be a sequence and simply to push the buttons as quickly as possible. If the absence of RT decreases reflects the impact of the explicit search upon the expression of implicit learning, then one would predict that, when subjects were exposed to the same sequence under non-search conditions, they should show evidence for a RT advantage for sequence over non-sequence elements. This would not
be the case if the explicit search had prevented the sequence learning. The fact that we found that the RT advantage following the implicit aSRT task was significantly greater than following the explicit aSRT task is strongly supportive of our contention that the search suppresses implicit learning itself, over and above any effect that it may have upon its expression.

Discussion
We provide further evidence for the involvement of a number of key regions in SRT task sequence learning — right PFC, medial temporal lobe, caudate and thalamus. We extend previous findings, however, by showing task-dependent interactions in the specific functions of these regions: the condition in which sustained right PFC activity was greatest (EaSRT) was one in which implicit learning, as estimated by RT reductions, was significantly reduced. This behavioural impairment was accompanied by an attenuation of time-dependent activation changes in medial temporal cortex and in thalamus. Additionally, a more direct exploration of the influence of frontal cortical activation on other brain regions, using connectivity analysis, shows that explicit learning of the more difficult aSRT task is associated with a profound change in relationship between right PFC and thalamus.

Before discussing these results in more detail, we consider the possibility that explicit learning instructions act not to prevent implicit learning but rather to suppress its expression. Seidler et al. (2002), for example, showed that a concurrent distracting task prevents a decrease in RTs during standard sequence learning but that, with cessation of the distracting task, there is a reduction in RTs suggesting that learning had indeed occurred during the distracting phase. It is possible that, in our study, the explicit search has a comparable effect: it prevents the effects of learning from being expressed without actually hindering learning itself. However, the results of our subsequent behavioural experiment (experiment 2) suggest that this is not the case. Following each aSRT learning session, and following a short period of exposure to a non-sequence condition, subjects were re-exposed to the sequence. We observed that, following lSRT, the relative advantage for sequence over neighbouring non-sequence elements, during initial re-exposure, was preserved and was significantly greater than following EaSRT (in which there was no measurable advantage). This difference occurred when subjects were no longer trying explicitly to learn the sequence. It contends strongly in favour of the suggestion that searching for the sequence disrupts learning itself rather than its expression.

Frontal Activation and Explicit Sequence Learning
Time-independent effects differentiating explicit from implicit processing were found in right PFC and, to a lesser extent, in left PFC (only the former survived the FDR thresholding). These effects were specific to the comparison between explicit and implicit processing of the more difficult, aSRT sequence. Previous studies have shown pronounced frontal activation for explicit learning in standard SRT tasks (Rauch et al., 1995; Hazeltine et al., 1997; Willingham et al., 2002; Schendan et al., 2003). However, a direct comparison between explicit and implicit learning of the standard sequence did not elicit frontal differences. This is perhaps unsurprising given our within-subjects counter-balanced design. Subjects were aware that sequences might sometimes be present. Thus, there was an increased likelihood that they would become aware of the standard sequence even when they were not searching for it (a suggestion supported by the post-scan debriefing). This incidental awareness of the presence and nature of the sequence may account for the lack of a significant frontal difference between the explicit and implicit standard sequence learning conditions. This is compatible with previous work showing that a condition in which incidental sequence awareness might occur is associated with right ventrolateral PFC activation (Hazeltine et al., 1997). Doyon et al. (1996), too, showed evidence that VLPFC activation accompanies sequence awareness, although, in contrast to the current study, they have shown evidence for a non-linear increase in frontal activity as subjects become well-practised on an explicit sequence learning task (Doyon et al., 2002). In a more recent fMRI study, Aizenstein et al. (2004) demonstrated that, when subjects were engaged in concurrent learning of two features of a stimulus sequence (implicit learning of colour sequences and explicit learning of shape sequences), PFC activation was more extensive in association with explicit learning. Uniquely, they produced their dissociation in the setting of a probabilistic sequence task: one in which 30% of trials were violations of expectation. This noise within the sequence presumably engenders difficulty that is comparable to that produced by our EaSRT task.

We do not argue that PFC activation reflects solely sequence awareness. Our critical observation with respect to time-independent task effects is an increased level of right PFC activation for the more difficult condition EaSRT (see graph in Fig. 3). The aSRT task involves sequences that are very difficult to detect and would require that subjects generate and test hypotheses, as well as making and monitoring the outcome of the predictions that these hypotheses generate. The engagement of such processes in the EaSRT condition provides the most likely explanation of maximal PFC activation in this condition. Lateral PFC has been repeatedly implicated in such processing in learning and memory tasks (Fletcher and Henson, 2001), and in those in which the requirement for monitoring processes is likely to be maximized (Fink et al., 1999).

Dynamic Changes in Implicit Sequence Learning
All conditions in which there was evidence of learning, in the form of decreasing reaction times, were associated with a time-dependent decrease in left posterior hippocampus/parahippocampal gyrus activation. This is consistent with recent animal work indicating a role for the hippocampus in sequence learning (Agster et al., 2002; Fortin et al., 2002), with a more general formulation of its role in the binding of temporally or spatially contiguous events (Wallenstein et al., 1998), and with recent findings in humans that reductions in medial temporal activity characterize both implicit and explicit learning of sequences (Schendan et al., 2003). Since it was associated with a condition in which there was no evidence of sequence awareness, Schendan and colleagues interpreted their observation of a decrease in MTL activation in terms of implicit learning. Of course, this is a dangerous assumption given that some of the subjects did indeed show some sequence awareness and, even though this study employed sensitive behavioural measures, there is always the possibility that subtle indices of awareness may be missed. Our study, however, supports their contention and their interpretation that MTL is critical to learning higher order associations between temporally distinct stimuli, irrespective of whether those associations are available to awareness or not. Our experimental design extends theirs in a key
way. We have been able to show that, when there is behavioural evidence of a reduction in this implicit associative learning, the pattern of modulation in MTL is disrupted even when post-scan debriefing indicates that a degree of awareness occurs. Thus, while they defined their implicit learning condition in terms of absence of awareness, we have shown that the MTL change is significantly greater in the condition in which no awareness was reported compared with one (EaSRT) in which it was. This is more strongly suggestive of the role for this region in implicit sequence learning.

The mechanism for hippocampal involvement in learning the associations embedded in sequences is likely to be a complex one. Simple Hebbian learning provides an insufficient explanation since, in common with previous studies, we have used higher order sequences wherein a future position is predicted not merely by the current position but also by the prior positions. Furthermore, synaptic plasticity arising from purely Hebbian learning demands that associated cells are coactivated within a tight time window (~100 ms). The alternating SRT sequence extends the time-interval well beyond this. Wallenstein and colleagues have suggested a complementary mechanism by which the hippocampus could drive detection and learning of sequences (Wallenstein et al., 1998), invoking the notion of context-sensitive cells in the CA3 region, which fire as more temporally extended patterns emerge. In doing so, they have a binding effect, allowing associations to form when stimuli are spatially or temporally remote. This is critical during the aSRT task and it may be that the absence of learning-related change in hippocampus during EaSRT reflects a failure to form these associations. We consider possible mechanisms for this below.

We also found learning-related changes in caudate activation. We must, however, be cautious about interpreting this caudate effect since it was seen only as a trend towards significance (FDR, \( P = 0.08 \)). Nevertheless, it is worth considering our findings in light of previous comparable studies. Although a number of sequence learning tasks have implicated the caudate nucleus, the pattern of activation has not yet been consistently related to whether learning is implicit (Rauch et al., 1995; Peigneux et al., 2000; Poldrack et al., 2001), explicit (Willingham et al., 2002) or both. It is notable that our observations are partly inconsistent with time-dependent increases shown by Schendan et al. (2003). Interestingly, they found apparent functional dissociations within the caudate nucleus with activity in the left heads of caudate increasing with implicit learning and activity in the body and right head decreasing with explicit learning. It is noteworthy that the coordinates of the learning-dependent region showing activity decreases (6, 24, 6) are comparable to those found in the current study (8, 20, 0). Clearly, the differential contributions of caudate nucleus to sequence learning merit further study.

We should consider more generally the absence, at the preset threshold, of the learning-dependent increases in activation in our study when comparing all learning conditions with the random condition. This is an important consideration in view of the fact that a large number of studies have identified, and focused primarily upon, learning-related increases in activation (Grafton et al., 1995, 1998; Hazeltine et al., 1997; Honda et al., 1998), while only a handful have been concerned with non-linear changes (Toni et al., 1998) or with linear decreases (Jueptner et al., 1997; Schendan et al., 2003). The main difference between our study and those that have focused upon learning-related increases in activation lies in the fact that we expressed learning changes as a time by task interaction. That is, in identifying the main effects of learning sequences (across all four task conditions) we have subtracted out time-dependent changes occurring in the non-sequence (random) condition. In previous studies, experimenters have either made this dissociation in different ways, or have not attempted to make it. In one study which suggested learning-related increases in activation in the supplementary motor area (SMA) and the parietal, premotor and prefrontal cortex (Honda et al., 1998) the authors acknowledge that similar increases were seen in a random condition. Alternatively, in three previous PET studies that have shown learning-related increases in activation in the PFC, motor and premotor cortex, putamen, SMA, parietal cortex and insular cortex (Grafton et al., 1995, 1998; Hazeltine et al., 1997), the method of ruling out ‘non-specific’ activation changes was critically different from ours. They identified regions that did not show such changes in a random sequence condition and excluded them from further analysis, focusing only upon those regions that did not. Such an approach is likely to be more sensitive to learning-changes since it does not involve a direct test of the interaction alluded to above. In view of this consideration, we carried out the analysis of learning-related increase at a reduced threshold, as described. The results, presented in Table 2, in identifying learning-related increases in motor, prefrontal and parietal cortices, are consistent with previous studies.

**Interactions between Explicit and Implicit Processing: the Deleterious Effects of Trying to Learn**

Our findings are readily interpretable within the framework of a hybrid model of human learning comprising an associative system, concerned with the detection of statistical regularities in the environment, and a cognitive system concerned with the search for sequential structure (Mclaren et al., 1984). While McLaren and colleagues’ formulation of this model emphasized the complementarity within this hybrid system, our neuro-imaging findings are relevant to evidence that this complementarity is conditional. Howard and Howard (2001) propose that, while associative learning is possible in the absence of intention and awareness, it nevertheless makes subtle demands upon cognitive capacity. In the case of elderly subjects, a reduced cognitive capacity is more quickly absorbed by explicit search processes and it is this that produces the deficit in implicit learning. In our experimental design, we emphasized the conscious search and monitoring processes necessary for explicit learning through a reduced emphasis on response speed. In doing so, we have produced a learning deficit in EaSRT that is similar to the one reported by Howard and Howard. We note also that RTs were generally slower for EsSRT than IsSRT but there was no difference in slopes of RT reduction, showing that learning was preserved for explicit attempts to learn the standard sequence (Fig. 2). In addition to replicating these behavioural observations, we have shown that trying to learn produces sustained right lateral PFC activation with a corresponding attenuation of the medial temporal and thalamic changes that accompany all tasks in which there is behavioural evidence of learning. Furthermore, this condition was characterized by positive fronto-thalamic correlation, while the IsSRT task was associated with the opposite pattern. This negative correlation underlying unimpaired learning may be consistent with a previous observation made by Rauch et al. (1998), who explored implicit learning of a sequence. They showed early thalamic
deactivations and suggested that this may reflect a suppression of thalamic activation, mediated by the indirect cortico-striato-thalamic pathway. We suggest that the negative fronto-thalamic connectivity observed here during implicit learning would be predicted by the findings of Rauch and colleagues.

We have considered our medial temporal findings in terms of an associative or covariance-detection mechanism. However, while Howard and Howard (2001) suggest that the impact of an explicit search is to absorb the cognitive capacity that this process demands, an alternative view suggests that the behavioural observation might arise not from limitations of cognitive capacity but rather as a consequence of attentional allocation (Keele et al., 2003). This model suggests that two sequence-learning systems, a unidimensional (implicit) and a multidimensional (implicit or explicit) system, may be in operation at any given time. Critically, interference occurs when correlations between elements are disrupted by the occurrence of other elements. That is, interference is produced not by overburdening of cognitive capacity but rather by disruptions in correlations produced by the interposition of attended elements. This may be the case in dual task SRT learning, when attention to another dimension (e.g. tones) produces a disruption. Schmidtke and Heuer (1997) have provided compelling evidence that the source of dual task interference lies in attending to non-sequence elements rather than to exhaustion of cognitive capacity. While the interference in our aSRT task is intra-dimensional, our results should, nevertheless, be considered in terms of the model of Keele and colleagues. If we posit that the explicit instructions engender attention towards successive sequence elements, this should not pose a problem during the sSRT task since successive elements are correlated. However, there would be a clear disadvantage in the aSRT task since attention to successive elements actually disrupts existing correlations. Thus, the model would predict our observation that they will exert a maximal interference effect in the explicit condition.

Thus, we might consider the explicit attempt to learn the sequences in terms of a top-down, attentional enhancement of some contingencies and a suppression of others, these effects occurring on the basis of guesses as to what the sequence might be. If the top-down processes favour trial-to-trial contingencies, they will prove a relative hindrance and will interfere with learning. Put another way, it is possible to get ‘stuck’ in the search for superficial contingencies at the expense of the more abstract ones (Tecumseh Fitch and Hauser, 2004). In terms of brain activity and connectivity, the top-down processes are associated with higher levels of PFC activity and with positive prefronto-thalamic connectivity, while the bottom up processes are reflected in a negative prefronto-thalamic connectivity. The supposition that the PFC activation, in this setting, reflects a top-down biasing is consistent with an fMRI analysis of attempts to extract a remembered melody interleaved with distractor tones (Bey and Zatorre, 2003).

Our data do not distinguish between the competing perspectives of the cognitive capacity and the attentional allocation models. However, in suggesting a neuronal signature for the interaction between searching for the sequence and learning it, our observations may prove useful in further neuroimaging studies addressing this question. For example, the cognitive capacity model would predict that, if we used a secondary task (rather than a search/no-search manipulation) performance on the aSRT task would be impaired in a way that is comparable, both behaviourally and neuronally, to that seen in the current study. The attention-based model of Keele et al. (2003) suggests that the aSRT becomes relatively vulnerable only when attention to non-sequence elements is augmented and this model, therefore, would predict that the impact of a secondary task on the aSRT should be relatively small.

Leaving speculations aside, our key findings are clear: explicit and implicit sequence learning tasks are dissociable in terms of sustained and time-dependent changes in frontal, medial temporal and thalamic regions. However, there is an interaction between processes associated with these two types of learning. This interaction is reflected in impaired learning under certain conditions and, as we have shown, in changes in the patterns of task-dependent brain activity in these regions. Furthermore, we suggest that an important feature of this interaction may be observed in a profound change in fronto-thalamic integration.

Notes

P.C.F., C.D.F., R.A.E.H. and P.R.C. are supported by the Wellcome Trust. G.R.E. and K.Z. are supported by the Deutsche Forschungsgemeinschaft (DFG-KFO 112). We are grateful to N.J. Shah and the MR-team for expert technical help with MR-scanning. The work was completed within the MRC Center for Behavioural and Clinical Neuroscience.

Address correspondence to Paul Fletcher, Box 189, Department of Psychiatry, University of Cambridge, Addenbrooke’s Hospital, Hills Road, Cambridge CB2 2QQ, UK. Email: pcf22@cam.ac.uk.

References


