

Dynamics of Prefrontal and Cingulate Activity during a Reward-Based Logical Deduction Task

Claire Landmann¹, Stanislas Dehaene^{1,2}, Sabina Pappata³,
Antoinette Jobert¹, Michel Bottlaender⁴, Dimitri Roumenov⁴
and Denis Le Bihan⁵

¹Institut National de la Santé et de la Recherche Médicale 562, Cognitive Neuroimaging, Service Hospitalier Frédéric Joliot, Commissariat à l'Énergie Atomique/Département de Recherche Médicale/Direction des Sciences du Vivant, 91401 Orsay Cedex, France, ²Collège de France, 75231 Paris Cedex05, France, ³Istituto di Biostrutture e Bioimmagini, Consiglio Nazionale delle Ricerche, 80131 Napoli, Italy, ⁴Service Hospitalier Frédéric Joliot, Département de Recherche Médicale/Direction des Sciences du Vivant/Commissariat à l'Énergie Atomique, Orsay, France and ⁵Unité de Neuroimagerie Anatomofonctionnelle, Service Hospitalier Frédéric Joliot, Commissariat à l'Énergie Atomique/Direction des Sciences du Vivant, 91401 Orsay Cedex, France

We used behavioral and functional magnetic resonance imaging (fMRI) methods to probe the cerebral organization of a simple logical deduction process. Subjects were engaged in a motor trial-and-error learning task, in which they had to infer the identity of an unknown 4-key code. The design of the task allowed subjects to base their inferences not only on the feedback they received but also on the internal deductions that it afforded (autoevaluation). fMRI analysis revealed a large bilateral parietal, prefrontal, cingulate, and striatal network that activated suddenly during search periods and collapsed during ensuing periods of sequence repetition. Fine-grained analyses of the temporal dynamics of this search network indicated that it operates according to near-optimal rules that include 1) computation of the difference between expected and obtained rewards and 2) anticipatory deductions that predate the actual reception of positive reward. In summary, the dynamics of effortful mental deduction can be tracked with fMRI and relate to a distributed network engaging prefrontal cortex and its interconnected cortical and subcortical regions.

Keywords: autoevaluation, fMRI, learning, prefrontal cortex, reward prediction error

Introduction

All animal species need to adapt their behavior to the current environmental situation. Such learning is often based on external rewards. The phasic firing of midbrain dopamine neurons is thought to code for a “prediction error” signal that corresponds to the discrepancy between expected and actual outcomes and serves as a reinforcement signal altering subsequent response probabilities (Schultz and others 1997; Sutton and Barto 1998). Recent neuroimaging studies have shown that projection sites of dopamine neurons, such as parts of striatum (Berns and others 2001; Breiter and others 2001; McClure and others 2003; O'Doherty and others 2003) and prefrontal cortex (PFC) (Berns and others 2001; Breiter and others 2001; Knutson and others 2003; O'Doherty and others 2003; Ramnani and others 2004), exhibit activation profiles that are consistent with the rational use of a prediction error signal.

Particularly developed in the human species, however, is an additional capacity for reasoning that enables subjects to adapt their strategies even in the absence of external reward signals. Using autoevaluation, that is, internally driven deduction pro-

cesses, humans can evaluate the outcome of potential actions by mentally envisaging their likely consequences (Dehaene and Changeux 1991, 2000). According to the “global neuronal workspace” hypothesis (Dehaene and others 1998), such effortful, internally oriented mental activity engages a set of cortical neurons with long-distance axons, distributed mostly in prefrontal, cingulate, and other cortical association areas, and is capable of fast flexible modifications of their activity depending on task demands. Indeed, a variety of neuroimaging studies have examined autoevaluation and internal reasoning processes and have emphasized the participation of distributed prefrontal and parietal areas (Deglin and Kinsbourne 1996; Goel and Dolan 2003). These regions are known to be involved in high-level cognitive functions, such as action monitoring, working memory, flexibility of strategy, and planning processes, as assessed with the Wisconsin Card Sorting Test (WCST) and Tower of London test. In particular, rostral prefrontal, anterior cingulate, and dorsolateral prefrontal cortices appear to play a key role in strategy selection and planning processes (Morris and others 1993; Baker and others 1996; Dagher and others 1999; Newman and others 2003; van den Heuvel and others 2003).

In the present study, we introduce a novel task in order to study the dynamics of parietofrontal activity during reward-based and deduction-based learning and working memory. Our main goal is 2-fold: 1) to investigate the temporal dynamics of the parietofrontal system and, in particular, whether it responds quickly to transient variations in task demands; 2) to determine how these temporal variations are driven both by the reward prediction error, that is, the difference between expected and obtained rewards, and by the availability of anticipatory deductions that predate the actual reception of positive reward (autoevaluation). Detailed analyses of behavioral data allow us to understand, step-by-step, the algorithm used by subjects to solve the task. We can then correlate some of these steps with functional magnetic resonance imaging (fMRI) data obtained during the resolution of many successive problems and thus begin to dissect the broad cerebral network involved in feedback processing.

We adapted a trial-and-error learning task, similar in principle to the Mastermind game (though much simpler) and first used in an even simpler form in the macaque monkey by Procyk and others (2000). In this task (which we refer to as the “Master-brain” task), subjects search for a hidden sequence of 4-key

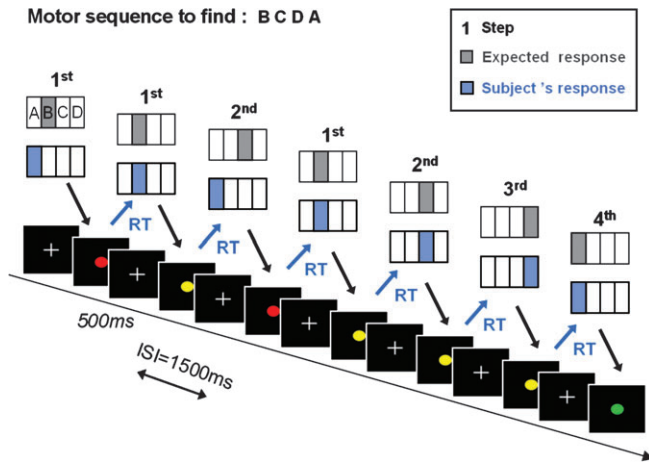


Figure 1. Temporal organization of the behavioral task. The figure illustrates a simple example of the unfolding of a search period, where the target sequence was B C D A. Subject had to find this hidden sequence of 4-key presses by trial and error. The subject first tried response A and was told that it was wrong via a negative feedback on the screen (red circle). Then he tried B, which was correct (yellow circle). At this point, the first element of the sequence (B) was known. As the second element, the subject tried A and was told that it was false. He had to restart the whole sequence from start (B), tried out C, then D, and then A which all turned out to be correct. At this point the problem was fully solved (green circle). RTs were measured from the onset of the feedback circle to the following button press (purple arrow).

presses by trial and error. For instance, during a given search period, they may have to discover that they have to press the keys in the order B, C, D, A (Fig. 1). After each key press, they receive a simple visual feedback (also referred to as “reward” because there is the prospect of a financial bonus; see Materials and Methods). When they press the correct key, they receive a positive feedback (yellow circle) and are allowed to search for the next element of the sequence. If, however, they press a wrong key, the feedback is negative (red circle), and they have to resume the whole sequence from the start before testing another hypothesis.

When subjects have found and executed the entire sequence correctly, they receive a green feedback signal. They then have to repeat the sequence a variable number of times, while still getting feedback on the screen. This period of routine execution ends abruptly upon receipt of negative feedback, which signals that the old sequence is no longer valid and that subjects have to search for a new one. Thus, the task alternates between periods of search and periods of routine execution. In that respect, it is reminiscent of the WCST, in which subjects also have to discover a criterion for sorting cards and then apply it for a few trials before switching to a new criterion. However, relative to the WCST, the present task presents the specific interest of allowing both reward-based and deduction-based learning. When searching for the correct sequence, subjects can occasionally fall upon it by chance (chance discovery). Conversely, when subjects have tested all but one of the possibilities for a given step, they can deduce that the last remaining possibility is necessarily correct (logical discovery). Note that on such trials, they receive negative feedback and have to resume the execution of the entire sequence from the beginning. Nevertheless, after the third step, logic allows them to be certain that they have now found the correct sequence. By contrasting the brain events associated with chance versus logical discovery, we can dissociate the moment when subjects know the correct sequence from the moment when they actually perform it.

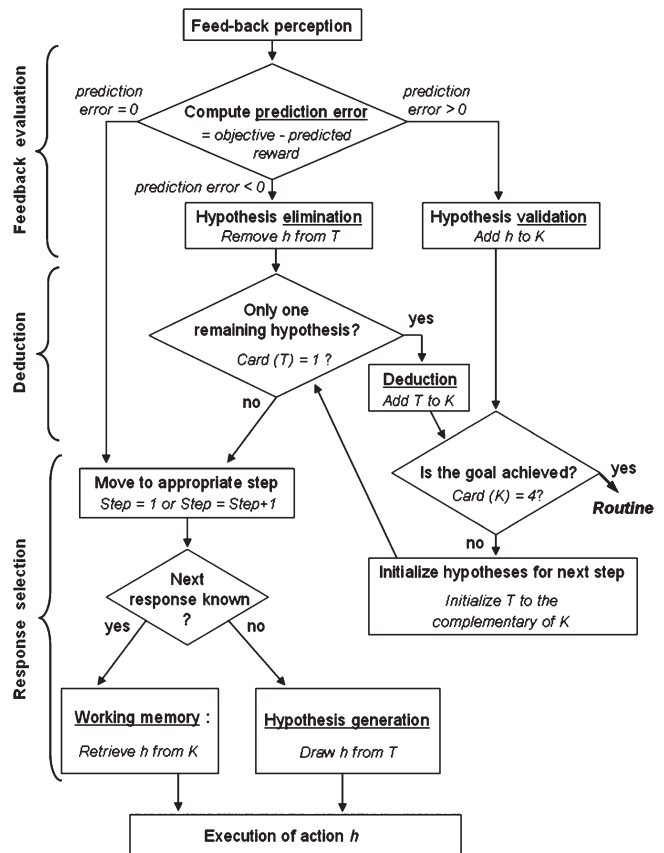


Figure 2. Functional model of the search process. We designed a simple hypothetical search algorithm that might be followed by our subjects. The algorithm assumes that subjects can maintain a memory of the hypotheses that remain to be tested (set T) as well as of the elements of the sequence that they have already validated (set K). It also supposes that subjects evaluate the discrepancy between the feedback that they received and what they expected (reward prediction error, top diamond). Subjects use this information to either validate or reject their ongoing hypothesis (h). Finally, the algorithm also assumes that subjects can deduce the solution once they realize that only a single hypothesis remains to be tested (central diamond). $Card(T)$ = cardinal of set T .

To facilitate fMRI analysis, a precise model of the cognitive operations involved was developed (see Fig. 2). This is one member of a large class of algorithms that are able to discover the unknown sequence in an optimal time. Assuming that human subjects adopt such an optimal search strategy, the following predictions were made:

1. Sharp distinction between routine and search periods. Because deduction and working memory updating processes occur only when the subject is in a “search mode,” there should be a sudden collapse of brain activity, particularly in prefrontal, cingulate, and parietal cortices, as soon as subjects switch from search to routine, and a restoration of this activation as soon as subjects have to search for a new sequence.
2. Time locking to mental discovery. This collapse of the search mode should be time locked to the time when subjects first mentally discover the solution, which differs from the time when they first correctly execute it.
3. Dependency on the reward prediction error. Consistent with the model of Schultz and others (1997), we expect fully predictable rewards to have little or no impact on brain activation. Rather, the reward prediction error and the

information carried by feedback should be good predictors of switches in parietofrontal brain activation.

Materials and Methods

Subjects

Sixteen healthy male right-handed volunteers participated in the study (mean age \pm standard deviation [SD]: 23 ± 2.2). Informed consent was obtained from all subjects. Ethics approval was given by the regional ethical committee (Comité Consultatif de Protection des Personnes dans la Recherche Biomédicale, Hôpital Bicêtre, Paris, France).

Experimental Paradigm

Subjects were trained before the scan and were informed that their performance during the experiment would be evaluated based on the total number of positive feedback received, with the best volunteer among all participants receiving a financial bonus (€75). According to their subjective reports, the challenge was highly motivating.

The participants' hand (right hand for 8 subjects, left hand for the other 8) rested on a 4-button device (for all fingers except the thumb). Each trial lasted 1500 ms, during which subjects pressed one key and received feedback, consisting in colored circles presented during 500 ms (red if the tested button was wrong, yellow if it was correct, and green after the entire correct sequence execution). These stimuli were presented even when subjects did not respond (in which case it was considered as a wrong response). When the feedback disappeared, it was replaced by a fixation cross.

During the scan, the paradigm was divided into three 10-min blocks (totaling 1029 ± 12 trials per subject). Subjects began each block by a routine period (i.e., repeating a known sequence), and then search and routine periods alternated at a semirandom rate. During the search, subjects attempted to discover a hidden 4-key sequence by trial and error, namely, to discover in which order they should press these 4 keys (see Fig. 1). Once they had found the sequence, they repeated it 3–6 times (routine phase), still getting feedback. The switch from routine to a new search period was indicated by an unexpected negative feedback.

A resting period was inserted in about half of the routine periods to define an fMRI baseline. It was announced by the word "REPOS" (rest) and ended with the word "ATTENTION," both presented on screen during 1000 ms. During this period subjects had to keep on looking at the screen (a fixation cross), and afterward they returned to the routine mode, executing the same known sequence as before the resting period. Presentation of stimuli and behavioral data collection were done using Expe (Pallier and others 1997).

Scanning Procedures

A 3 T Bruker scanner was used to acquire both T_1 -weighted anatomical images (voxel size 1.2 mm) and gradient-echo T_2^* -weighted echo-planar images with blood oxygenation level-dependent (BOLD) contrast. Twenty-six volumes of 4.5-mm-thick axial slices were acquired sequentially every 2.4 s, the first 5 volumes being discarded to allow for T_1 equilibration effects.

Data Analysis

Data were analyzed using statistical parametric mapping software (SPM99; Wellcome Department of Cognitive Neurology, London, UK). The image time series were realigned to correct for interscan movement and normalized to the standard anatomical space of the Montreal Neurological Institute. The data were then smoothed with a Gaussian kernel of 5 mm full-width half-maximum. Following pre-processing of data, statistical analysis was conducted for each subject using the general linear model, which included regressors drawn from behavioral results.

To evaluate our hypotheses about the microstructure of the task, we designed a hierarchy of statistical fMRI models similar to the use of stepwise regression in cognitive psychology, first capturing the block structure of the task (search vs. routine), then attempting to capture more variance due to local trial-to-trial variations in the type of cognitive activity involved (see Discussion).

The first model included 6 regressors: search trials, routine trials, errors during search and errors during routine as effects of interest, and

word presentation before and after resting periods as effects of no interest. Error trials correspond to nonoptimal behavior, such as testing an already tested button during a search period or making an execution error during routine. Each regressor was convolved with a canonical hemodynamic response function and, concerning word presentation, with its temporal derivative. Voxelwise t -statistics were carried out on linear contrasts of the regressors in order to produce individual statistical maps. Finally, a random-effect analysis was applied across the 16 subjects in order to determine group activation for the conditions of interest (voxelwise threshold $P < 0.001$, cluster extent threshold $P < 0.05$ corrected). This analysis also included a between-subject factor of right-hand versus left-hand groups, but no effect or interaction involving the effector hand was observed outside of the contralateral motor regions, and all results were therefore pooled across both groups.

In a second model, 2 additional regressors modeled separately the first correct execution of the searched sequence as a function of whether the sequence was discovered by logic or by chance (see text). Finally, the last model included, in addition to the regressors of the first model, regressors quantifying, trial by trial, the objective reward value of the feedback signal (1 for yellow or green circles, -1 for red circles), the reward prediction error (objective reward minus mean expected value of the reward), and the quantity of information carried by feedback ($-\log_2(P)$), with P the probability of occurrence of the obtained objective reward given the subject's current knowledge).

Model Algorithm

The algorithm used to model task performance assumes that subjects can maintain a memory of the hypotheses that remain to be tested (set T) as well as of the elements of the sequence that they already have validated (set K). It also supposes that subjects evaluate the discrepancy between the feedback that they receive and what they expected (reward prediction error; see Schultz and others 1997). Subjects use this information to either validate or reject their ongoing hypothesis. Finally, the algorithm also assumes that subjects can deduce the solution once they realize that only a single hypothesis remains to be tested.

In detail, the sequence of events that unfold upon reception of reward information is the following (see Fig. 2):

1. The visual feedback stimulus (yellow, red, or green) is evaluated and transformed into an internal reward signal (referred to as the "objective reward"). This signal is then compared with an expected reward, and the reward prediction error is generated by subtraction.
2. If the reward prediction error is positive (and therefore the reward is better than expected), then the hypothesis for the current step is valid and can be registered in set K and eliminated from set T . If, conversely, the reward prediction error is negative (and therefore the reward is worse than expected), then the hypothesis for the current step is invalid and can be eliminated from set T . Finally, if the reward is just as expected, this means that the subject is merely executing a known part of the sequence and hence must simply advance one step on the next trial.
3. Whenever set T has been modified, logical deduction can occur: if set T contains only one remaining item, it is necessarily the correct one and can be added to set K .
4. If set K is complete, meaning that all 4 steps have been discovered, the subject switches to routine mode. Otherwise, a third stage of response preparation occurs in which subject selects an action for the next step, drawing it either from the set K of known steps (if having to resume execution of the entire sequence) or from the set T of remaining hypotheses (if having to generate a new hypothesis).

The complete algorithm was implemented as a Matlab program, and we verified that this program identified the solution to all problems in an optimal time.

Results

Behavioral Data

We analyzed the behavioral data from 16 right-handed, healthy subjects who performed a total of 513 search periods during

fMRI scanning—half of them with the right hand, the other half with the left hand. Figure 3 shows the distribution of the duration (i.e., the number of trials until the first execution of the entire sequence) of those search periods. The mean number of trials before reaching the correct solution was 8.94 trials (SD across subjects = 0.48), which was only slightly higher than the mean of 8.5 trials predicted if subjects used an optimal search algorithm. Despite a small departure between model and data, the observed distribution was tightly fitted by the ideal error-free model ($\chi^2_{14} = 2.22$, not significant; see Fig. 3), suggesting that subjects solve the search task in a near-optimal way.

We therefore used the above-described algorithm to infer, from the actual data, the subject's progression through the search task. For each trial, the model was used to label the cognitive processes used by the subject. Consistently with the above distribution analysis, only 4% of trials did not fit with the predictions of the algorithm (plausibly corresponding to errors in the execution of this algorithm). We conducted a 2-factor repeated-measures analysis of variance (ANOVA) on error rates with task (search, routine) and hand (right, left) as factors. Only marginal effects were found. Subjects were essentially equally accurate in search periods and in routine periods (4.7% vs. 3.4% of errors, $F_{1,14} = 3.91$, $P = 0.07$) and with either hand (5.1% with the right hand, 2.9% with the left hand, $F_{1,14} = 4.24$, $P = 0.06$). These errors were eliminated from further analyses.

On each trial, response times (RTs) were recorded, reflecting the time elapsing between reception of visual feedback about the previous response and emission of the next response. We therefore expected these RTs to reflect cognitive processes involved in feedback evaluation, deduction, updating of internal representations, and response selection.

RTs were analyzed with a 2-factor repeated-measures ANOVA with task (search, routine) and hand (right, left) as factors. As expected, RTs showed a main effect of task: subjects were faster in routine periods than in search periods (333 vs. 490 ms)

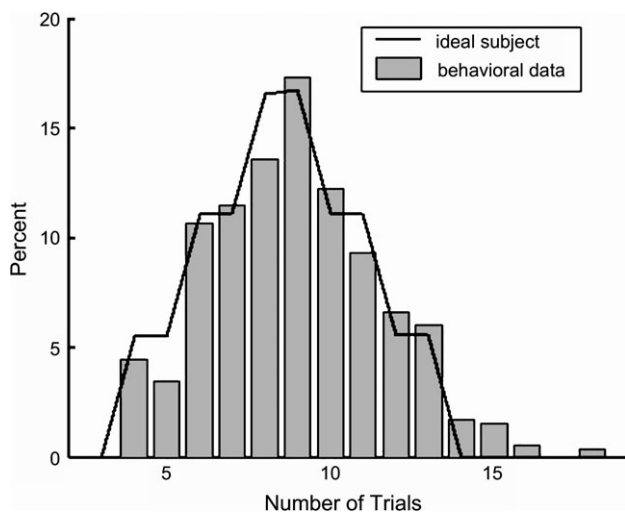


Figure 3. Distribution of search period durations. The graph represents the percentage of problems that subjects solved using a given number of trials. We confronted this empirical distribution (bars) with the predictions of an “ideal subject” performing according to our hypothetical algorithm (solid line). The model predicted accurately the distribution of search durations ($\chi^2_{14} = 2.22$, not significant)—note that introducing an error rate of as little as 3% into the model allowed an even better fit of the empirical distribution (not shown). This indicated that subjects behaved close to optimally and, in particular, took advantage of the possibility of making logical deductions.

($F_{1,14} = 183$, $P < 0.001$). The main effect of hand on RTs was not significant (430 ms with the right hand vs. 393 ms with the left one) ($F_{1,14} < 1$). Finally, the interaction was slightly significant ($F_{1,14} = 5.66$, $P < 0.05$), corresponding to paradoxically faster RTs during routine blocks with the left hand, perhaps due to great attention and/or training in the left-hand group.

We then analyzed in more detail the transition from search to routine. As noted above, the first correct sequence execution can be achieved in 2 qualitatively different manners: chance versus logical discovery. We expected RTs to reflect this distinction. On logical discovery trials, there might be an initial delay due to deduction and the ensuing representation of the correct sequence in working memory, but subjects should then quickly enter routine mode. On chance discovery trials, however, RTs should exhibit the slowness typical of search trials.

As shown in top panel of Figure 4, the results conformed to this prediction. A 2-factor repeated-measures ANOVA was conducted on RTs with type of sequence (chance discovery

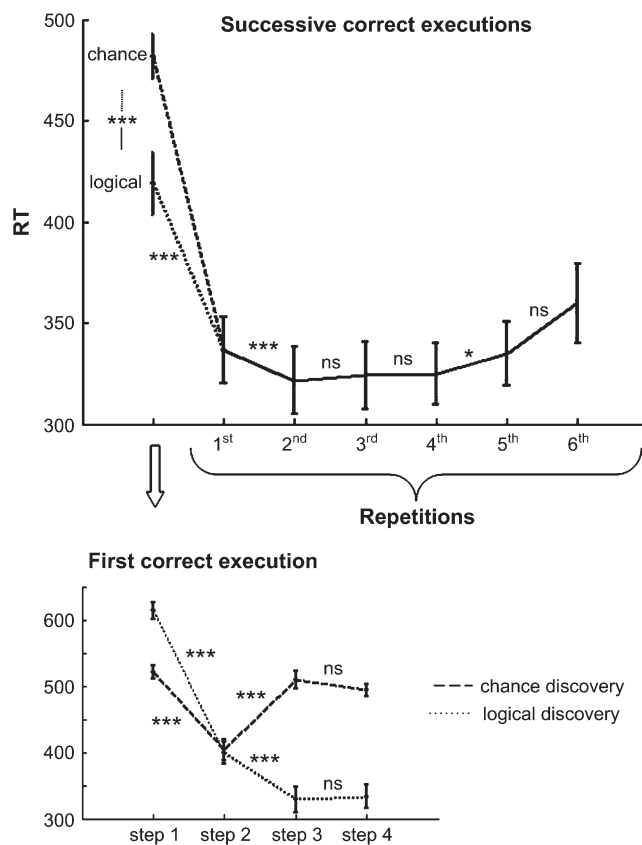


Figure 4. Evolution of RTs as subjects discover and execute the hidden sequence. Top panel: RTs averaged across the 4 steps of the correct sequence as a function of successive correct executions. Statistics refer to *t*-tests for pairwise comparisons between 2 conditions (statistics: * $P < 0.05$; *** $P < 0.001$; ns: not significant). Error bars = 1 standard error across subjects. During the first correct execution (left 2 points), subjects were faster in executing the sequence discovered by logic than the sequence discovered by chance. After this first correct execution, RT showed a progressive reduction toward a plateau (routinization) as well as a late increase probably reflecting an anticipation of the switch to a new search. Bottom panel: detailed evolution of mean RTs during the first correct execution as a function of execution step and whether the hidden sequence was discovered by chance or by logic. Subjects were slow in executing the sequence discovered by chance, presumably reflecting their ongoing search for the sequence. When the sequence was discovered by logic, subjects were initially slow, presumably reflecting the deduction process and the need to reload the whole sequence from start but afterward they executed it very fast.

vs. logical discovery) and steps (1–4) as factors. As predicted, subjects were faster overall in logical discovery than in chance discovery (415 vs. 481 ms, $F_{1,14} = 23.1$, $P < 0.001$). Furthermore, a type by step interaction ($F_{1,14} = 88.4$, $P < 0.001$) showed that subjects kept responding slowly until the end of chance discovery sequences, while they were slow in the first step of logical discovery sequence but fast in the following steps (see Fig. 4, bottom panel). This acceleration is consistent with an early switch from search to routine after logical discovery occurred.

Top panel of Figure 4 also shows RTs during the repeated execution of the discovered sequence. Following the first correct sequence, subjects became progressively faster across repetitions, and RTs converged toward a “plateau.” Finally, RTs increased slightly during the last 2 repetitions of the routine period, presumably reflecting the subject’s anticipation of the switch to a new sequence’s search.

Overall, these behavioral observations support our theoretical analysis of the task and, in particular, the postulate that it

involves cognitive processes of search and deduction. We next turned to the neural bases of these processes.

fMRI Differences between Search and Routine

We first modeled the data using separate regressors for the routine and search blocks (excluding error trials; see Materials and Methods). A random-effect analysis was conducted using a 2-way ANOVA with trial type (routine or search) as a within-subject factor and hand as a between-subject factor.

The main effect of routine relative to rest elicited a bilateral activation in frontoparietal regions (principally in the right hemisphere), including the precentral gyri and the supplementary motor area (SMA), in occipital cortex (bilateral inferior occipital gyri), and in cerebellum (Fig. 5). We also observed an activation of the left striatum (mainly the putamen, extending into the caudate).

Contrasted with resting periods, search periods yielded an extensive activation not only in the same regions but also in bilateral frontal cortices, particularly the lateral orbitofrontal

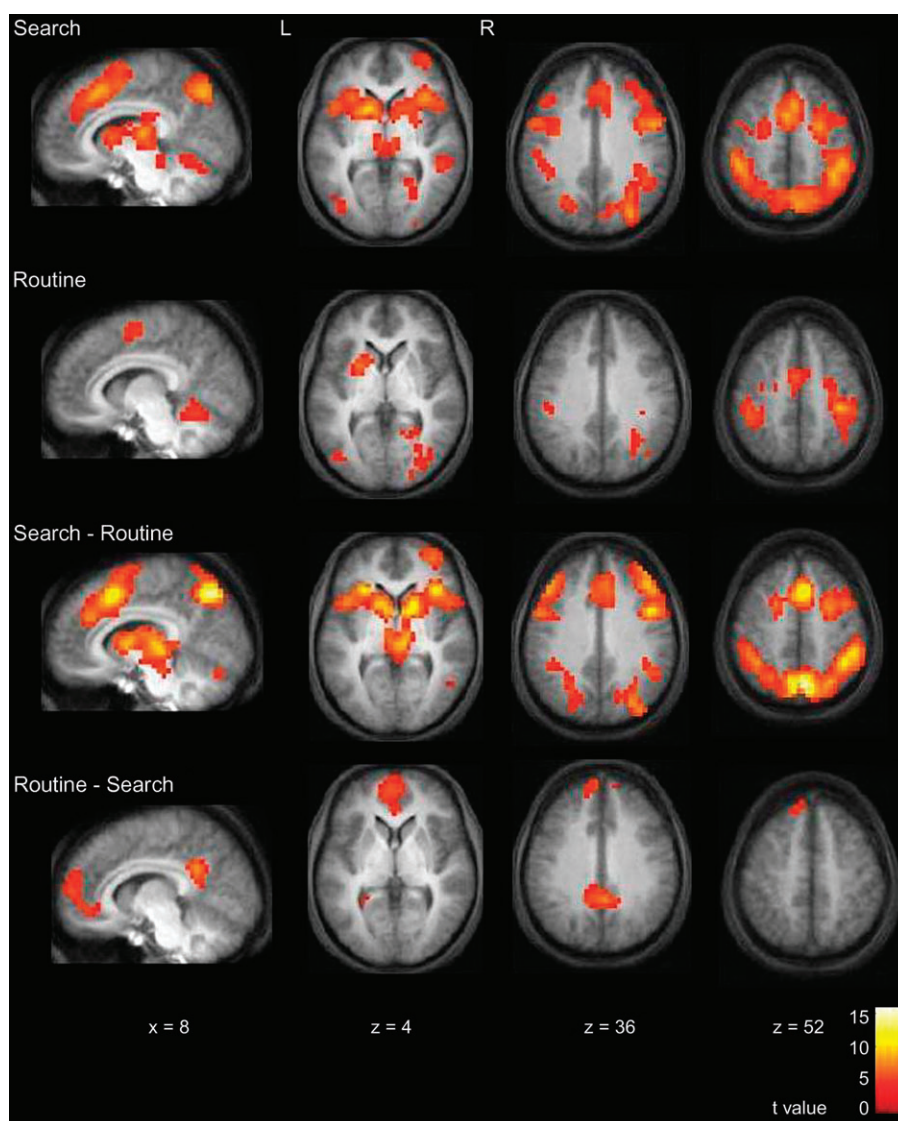


Figure 5. Neural correlates of search and routine stages. Statistical t -maps for search minus rest, routine minus rest, search minus routine, and routine minus search are displayed on the average anatomical image from all subjects, shown in sagittal (left) and transverse (right) sections. Voxelwise threshold was set to $P < 0.001$ and cluster extent threshold to $P < 0.05$ corrected.

and dorsolateral PFCs, and the anterior cingulate cortex (ACC), extending into the SMA. The striatum was bilaterally activated (caudate nucleus and putamen), as was the thalamus. Finally, the bilateral cerebellum and midbrain, probably including the medial portion of the substantia nigra and the ventral tegmental area, were activated in this condition.

The direct contrast between search and routine periods revealed virtually the same pattern of bilateral parietofrontal and subcortical regions (Table 1). Thus, most regions activated in the routine task showed enhanced activation during search. Few regions showed up in the converse contrast (routine > search): right posterior insula, bilateral anterior medial PFC (including anterior part of ACC), posterior cingulate cortex, and left middle temporal gyrus. Most of these regions fit within a previously described “resting state” network (Raichle and others 2001) that deactivates during cognitive tasks. Indeed, they were not activated in routine relative to rest and showed even greater deactivation during search than during routine.

In summary, a large network consisting mostly of prefrontal, cingulate, parietal, and striatal regions was specifically engaged during search, whereas some focal regions of this network also participated in routine periods. Figure 6 shows the temporal evolution of the BOLD signal in 3 representative regions of the search network (ACC, right lateral PFC, and right caudate nucleus). In each of them, we observed that activation rose suddenly as soon as subjects began searching for a new sequence and collapsed a few seconds later once they had found it. Indeed, the duration of activation was directly related to the duration of the search period, as revealed by a significant time by search duration interaction ($F_{2,45} > 10$, $P < 0.001$ in ACC).

Brain Activity during Logical and Chance Discovery

Figure 7 shows the evolution of the BOLD signal in these 3 brain regions at the end of the search period when subjects could either execute the first correct sequence by chance or discover

it by logical deduction. Search periods are matched for length: the figure displays separately short search periods (7 or 8 trials to find and execute the entire sequence) and longer search periods (9 or 10 trials). We observed the same pattern in the 3 regions. Regardless of search length, the activation curve corresponding to logical discovery collapsed about 3 s earlier than the one corresponding to chance discovery. In other

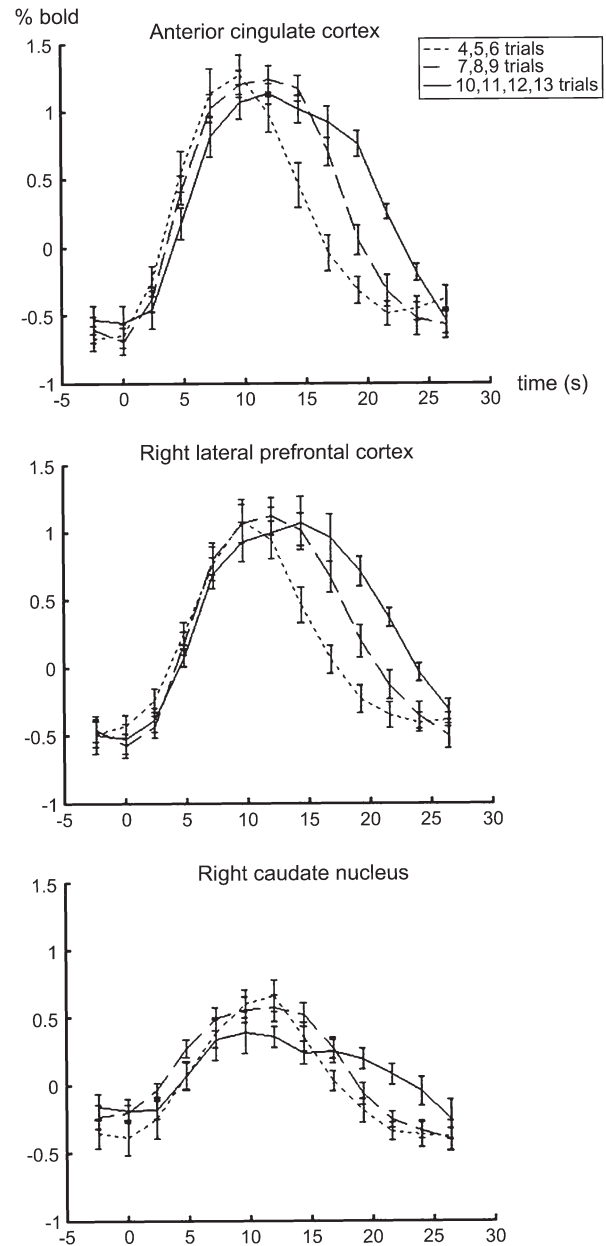


Figure 6. Cingulate, prefrontal, and caudate fMRI activation indexes the duration of the search process. The graphs show activation curves (% BOLD signal) averaged across all subjects, time-locked to the onset of the search periods, as a function of time. Bars indicate one standard error of the mean across subjects. Search periods were grouped into 3 categories based on their duration—dotted line: search periods lasting 4, 5, or 6 trials (mean = 7.5 s); dashed line: search periods lasting 7, 8, or 9 trials (mean = 12 s); solid line: search periods lasting 10, 11, 12, or 13 trials (mean = 17.25 s). In all cases, activation collapsed suddenly as soon as subjects enter the routine phase. Note that for each subject, data were obtained from the local maxima closest to the highest peaks observed in the random-effect analysis of search versus routine. The mean and standard deviation of the coordinates were: ACC: $x = 3$ mm, $SD = 2$, $y = 19$, $SD = 7$, $z = 48$, $SD = 5$; lateral prefrontal cortex: $x = 43$, $SD = 5$, $y = 29$, $SD = 5$, $z = 32$, $SD = 7$; caudate: $x = 12$, $SD = 3$, $y = 9$, $SD = 4$, $z = 10$, $SD = 4$.

Table 1
Significant peaks observed in whole-brain analyses of the search minus routine contrast

Area	Coordinates			Z-score
	x	y	z	
R precuneus	4	-64	52	6.36
L precuneus	-4	-64	52	5.83
R inferior/superior parietal lobule (interparietal sulcus)	40	-48	52	5.58
L inferior/superior parietal lobule (interparietal sulcus)	-36	-48	52	4.83
R anterior cingulate cortex	12	32	28	4.79
L superior frontal gyrus/precentral sulcus	-20	0	68	5.69
R superior frontal gyrus/precentral sulcus	24	0	60	5.52
R insula (anterior part)	36	24	0	5.60
L insula (anterior part)	-28	24	4	5.51
R lateral PFC (middle frontal gyrus)	44	8	36	5.53
L lateral PFC (middle frontal gyrus)	-48	4	36	4.65
R DL PFC (middle frontal gyrus)	44	28	28	5.38
L DL PFC (middle frontal gyrus)	-44	28	32	5.27
R VL PFC (middle/inferior frontal gyrus)	36	52	0	5.05
R putamen	16	8	4	5.49
L putamen	-16	12	4	5.19
R thalamus	8	-24	4	5.42
L thalamus	-8	-16	8	4.44
Midbrain	0	-24	-16	3.85
L cerebellum	-28	-60	-24	4.43
R cerebellum	32	-64	-24	4.21
L fusiform gyrus	-36	-72	-12	3.83
L inferior temporal gyrus	-44	-60	-8	3.60
R inferior temporal gyrus	56	-64	-8	4.00

Note: This table of coordinates summarizes the main peaks of activation belonging to the large network represented on Figure 5 (voxelwise threshold $P < 0.001$, cluster extent threshold $P < 0.05$ corrected). R, right; L, left; PFC, prefrontal cortex; DL, dorsolateral.

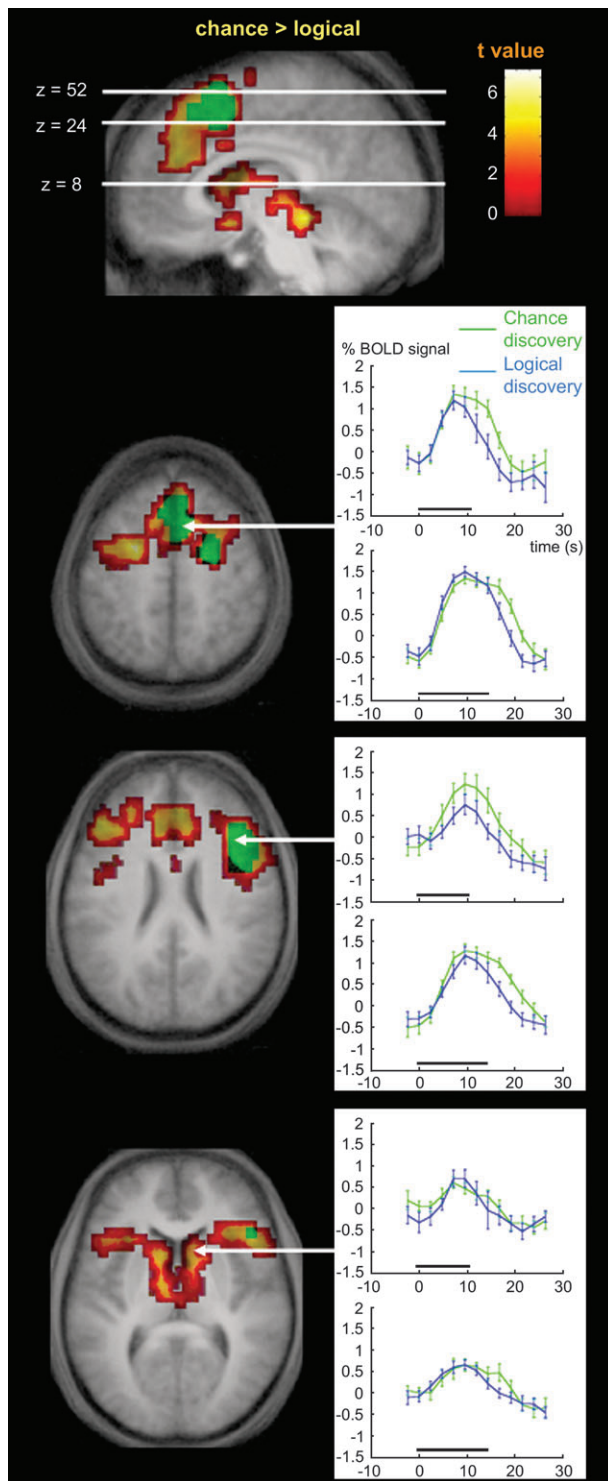


Figure 7. Neural correlates of the shortening of effortful search afforded by logical deduction. When subjects could deduce the correct sequence by logic, activation collapsed earlier in a broad anterior network including anterior cingulate, prefrontal, and striatal regions. Slices represent the t -maps for chance discovery minus logical discovery at 2 statistical thresholds: green map, voxelwise $P < 0.001$ and cluster extent $P < 0.05$ corrected; yellow-red map, voxelwise $P < 0.05$ and cluster extent $P < 0.05$ corrected. Curves show the temporal evolution of activation (% BOLD signal) in anterior cingulate cortex, right lateral PFC, and right caudate nucleus averaged across all subjects. Two different search durations are shown: top graph, 7/8 trials (mean = 11.25 s); bottom graph, 9/10 trials (mean = 14.25 s). In both cases, logical deduction led to shorter activation.

words, the activation collapsed earlier when the correct sequence could be deduced and thus known before it was executed than when it was discovered at the same time as it was executed.

This difference between logical and chance discovery was confirmed by a dedicated SPM model that concentrated on the first correct execution of the searched sequence (see Materials and Methods). During this period, greater activation was observed on chance trials in anterior cingulate/mesial frontal cortex, right lateral PFC, precentral gyrus, and inferior frontal cortices (see Fig. 7, green map, and Table 2). All these regions were parts of the search network described above. At a lower threshold, essentially all of the anterior part of the search network was observed, including the left prefrontal and bilateral caudate regions (Fig. 7, yellow-red map). This observation confirms that anterior brain regions remain active, and the subject remains in search mode, until the correct sequence is discovered whether by chance or by logic. In other words, the sudden drop of activity in these regions indexes internal knowledge of the correct sequence, independently of whether or not it has already been executed.

Feedback Processing during Search Periods

To further clarify the respective contributions of the observed brain regions to different components of our task, we distinguished 3 types of content carried by the visual feedback signal. First, we called “objective reward” the positive or negative value assigned to the feedback signal. Second, we computed, for each trial, the reward prediction error obtained by subtracting from the objective reward the average value that subjects could expect. Note that the 2 parameters were largely uncorrelated in our experiment. For instance, there were trials in which the objective reward was positive but entirely expected by the subjects (for instance, when received following an already tested or deduced hypothesis). Finally, we also computed the absolute information carried by the feedback signal regardless of its positive or negative rewarding value. Note that negative feedback can be highly informative (when the prior probability of its occurrence was small), although it is not rewarding.

These 3 variables were therefore entered as regressors in a new SPM analysis of the search period. As predicted, the

Table 2

Significant peaks observed in whole-brain analyses of the chance discovery minus logical discovery contrast

Area	Coordinates			Z-score
	x	y	z	
Mesial PFC (R frontal superior gyrus)	4	20	56	4.73
<i>R cingulate gyrus</i>	8	36	24	3.31
<i>L cingulate gyrus</i>	-8	28	28	3.29
R lateral PFC (middle frontal gyrus)	44	20	28	4.72
<i>L DL PFC (middle frontal gyrus)</i>	-44	28	28	3.71
<i>L frontal superior gyrus</i>	-20	-4	56	4.00
R frontal superior gyrus	24	-4	52	3.99
R insula (inferior frontal gyrus)	36	24	4	4.30
<i>L insula (inferior frontal gyrus)</i>	-36	20	-8	2.94
<i>R caudate</i>	12	8	16	2.97
<i>L caudate</i>	-8	8	8	2.92
<i>R putamen</i>	16	8	4	2.63
<i>Midbrain</i>	-4	-12	-4	2.78
<i>L thalamus</i>	-4	-8	12	2.85

Note: voxelwise $P < 0.001$, cluster extent $P < 0.05$ corrected, unless areas in italic font: voxelwise $P < 0.05$, cluster extent $P < 0.05$ corrected. R, right; L, left; PFC, prefrontal cortex; DL, dorsolateral.

objective reward did not correlate with any brain region activity even at a very low threshold (voxelwise $P < 0.05$). By contrast, a strong positive correlation between reward prediction error and brain activation was found in parts of the search network, including bilateral dorsolateral PFCs, frontal eye field (FEF), SMA, inferior PFC/insula (mainly in the left hemisphere), bilateral putamen and a part of right caudate body, left thalamus, bilateral parietal regions, and bilateral cerebellum (Fig. 8 and Table 3). No negative correlation was found. Finally, information carried by feedback correlated with brain activation in a separate set of brain areas at voxelwise $P < 0.005$, mainly in the right hemisphere: right inferior PFC/insula, rostral ACC, right dorso-lateral PFC and essentially the head of right caudate (Fig. 8 and Table 3).

Discussion

Our behavioral results demonstrate that, without explicit training of a resolution strategy, human subjects perform the Masterbrain task almost optimally. Analysis of their performance gives evidence that they rely on internal deduction processes and show a fast transition from an effortful search mode to a routinized mode of execution. Assuming that they follow a quasi-optimal algorithm, we can infer, from the observed sequence of actions, the exact stage where subjects are in their mental algorithm (including the hypotheses that they currently maintain in working memory, those they have already rejected, etc). Cross-correlation with fMRI then allows us to verify the accuracy of this model of the subject's mental state and to identify some of its neural correlates.

The Masterbrain task calls for the activation of a widespread brain network, extending from frontoparietal cortices to striatal and midbrain regions. These brain regions have been repeatedly associated with higher-order executive functions (Fuster 1989) including problem solving (Owen and others 1996) and reward coding (Hollerman and Schultz 1998). PFC, and in particular its dorsolateral part, is viewed by many theorists as a rule-coding device that is capable of very fast switching between different states, corresponding to different cognitive strategies, as a function of the reward or the information received by the organism. This view receives support from modeling (Dehaene and

Changeux 1991; Cohen and others 1992; Dehaene and others 1998; Rougier and others 2005), electrophysiological (Nakahara and others 2002; Wallis and Miller 2003), and neuroimaging studies (Botvinick and others 2001; for a review, see Miller and Cohen 2001). Our results regarding the dynamics of activation support this hypothesis and also suggest that switches in PFC activation are accompanied by large-scale changes in other interconnected distant cortical and subcortical regions forming a distributed network.

At the beginning of a new search period, the entire network immediately switches on and stays active for a duration directly related to that of the search period (Fig. 6). Conversely, activation drops suddenly as soon as the subject knows the entire sequence (Fig. 7), even when this knowledge was acquired thanks to a deduction process, before actually receiving a positive reward and achieving the task. This profile of activation is congruent with results using a similar paradigm in the macaque monkey (Procyk and others 2000). Procyk and others recorded task-related neurons in the anterior cingulate cortex during alternation between search and routine periods of the task. Some neurons showed elevated firing rates during search; crucially, their firing ended abruptly as soon as the animal had accumulated enough information to infer the solution of the problem. In particular, this collapse of firing occurred even on trials described here as involving logical discovery, where the animal could deduce the correct sequence before having actually tested it. These findings indicate that a self-evaluation process (Dehaene and Changeux 1991, 2000), capable of drawing conclusions about future rewards through purely internal reflection processes, is available to the macaque monkey.

Although fMRI is notoriously poor in temporal resolution, it could easily track fast global changes in deduction-related activation occurring across trials separated by only 1.5 s. In fact, electrophysiological studies indicate that the state of activity of prefrontal neural populations can switch extremely rapidly (10–100 ms), either as an adaptive response to external stimuli (Rainer and others 1998; Wallis and Miller 2003) or as a reflection of endogenous sequential processes (Seidemann and others 1996; Procyk and others 2000). How are such distributed dynamical changes generated? One possibility is that there exist source regions that broadcast signals requesting a change in the

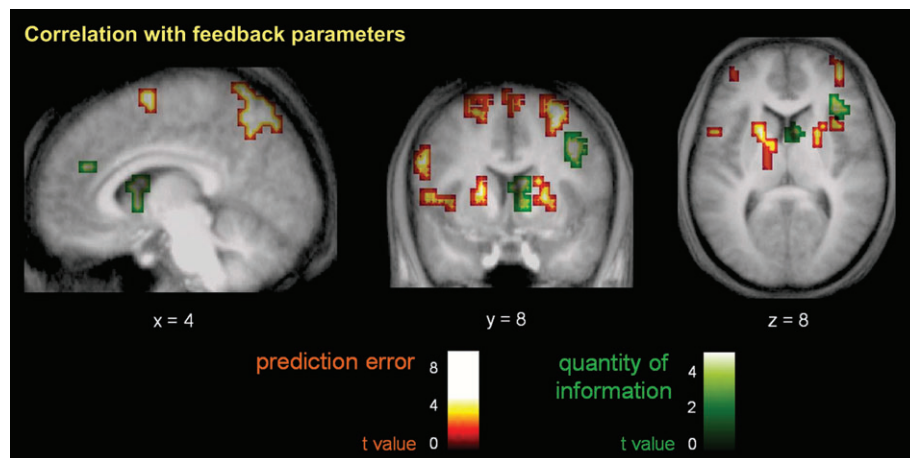


Figure 8. Neural correlates of feedback processing during the search period. Statistical t -maps for signal correlations with feedback parameters, masked by the search minus routine network. The figure shows regions whose activation correlates with prediction error (red–yellow scale: voxelwise threshold $P < 0.001$ and cluster extent threshold $P < 0.05$ corrected) and with the information content of the feedback signal (green scale: voxelwise $P < 0.005$ and cluster extent $P < 0.05$ uncorrected (higher than 20 voxels). No correlation was found with objective reward even at a low threshold (voxelwise $P < 0.05$ and cluster extent $P < 0.05$ corrected).

Table 3

Significant peaks observed in the correlation analyses of brain activity with feedback parameters

Area	Coordinates			Z-score
	x	y	z	
Prediction error				
R parietal superior lobule/precuneus	12	-60	52	4.54
L parietal superior lobule/precuneus	-8	-56	64	5.33
R inferior parietal lobule	44	-40	56	5.09
L inferior parietal lobule	-40	-40	52	5.14
R DL PFC (middle frontal gyrus)	36	44	28	5.05
L DL PFC (middle frontal gyrus)	-44	28	28	4.00
R superior frontal gyrus	28	0	64	4.99
L superior frontal gyrus	-20	0	68	4.42
R insula	40	12	8	4.66
L insula	-44	12	4	3.97
R caudate	20	4	20	4.35
R putamen	24	4	8	3.47
L putamen	-20	-4	8	3.97
L thalamus	-12	-12	12	3.45
L precentral gyrus	-52	4	28	4.28
L cerebellum	-24	-68	-20	3.92
R cerebellum	32	-64	-24	3.86
Information in feedback				
<i>R insula (inferior frontal gyrus)</i>	40	28	8	3.80
R anterior cingulate cortex	8	32	8	3.67
R anterior cingulate cortex	4	32	24	2.89
L anterior cingulate cortex	-8	24	32	2.97
<i>R DL PFC (middle frontal gyrus)</i>	44	4	40	3.57
R caudate	8	8	12	3.44

Note: Prediction error: voxelwise threshold $P < 0.001$, cluster extent threshold $P < 0.05$ corrected. Information in feedback: voxelwise $P < 0.005$, cluster extent $P < 0.05$ corrected, unless areas in italic font: cluster extent $P < 0.05$ uncorrected (higher than 20 voxels). No brain region was found to correlate with the objective reward value even at a low threshold (voxelwise $P < 0.05$, cluster extent $P < 0.05$ corrected). R, right; L, left; DLPFC, dorsolateral prefrontal cortex.

current plan. One possible candidate is an ascending neuro-modulator signal arising from the nucleus basalis or locus coeruleus, which are thought to convey “uncertainty” signals (Yu and Dayan 2005). Norepinephrine, in particular, has been recently proposed to encode “unexpected uncertainty” associated with large uncued changes in the environment, as found in the present task. Another candidate region is the frontopolar cortex, which has been associated with planning (Baker and others 1996; Owen and others 1996) and with the temporary switching between tasks contingent on unpredictable events (Koechlin and others 2000). Finally, a third candidate is the anterior cingulate, whose involvement in error detection is well known and is thought to be capable of causing major subsequent changes in prefrontal executive control processes (Botvinick and others 2001). All 3 of these regions were activated in the present study and could play a key role in fast adaptation of a behavioral plan when a major change is required. To resolve their respective roles, gaining access to their temporal order of activation is essential and could be obtained using an adaptation of the present task for electroencephalography or magnetoencephalography recording.

Based on the temporal difference model (Schultz and others 1997; Sutton and Barto 1998), we expected brain activation changes to correlate with the reward prediction error, that is, the subtraction of the actual and predicted rewards, rather than with the objective value of the reward. The analysis of feedback parameters confirmed this prediction. First, no activation was found to correlate with the objective value of the reward. Second, the reward prediction error correlated with activation in an extended corticosubcortical network including bilateral putamen and part of the right caudate as well as the bilateral

prefrontal, parietal, and cerebellar regions. Third, reward information, rather than reward prediction error, correlated with activation in the head of the right caudate, the right lateral PFC/insula, the right dorsolateral PFC, and the rostral ACC.

Although the detailed functions of these 2 circuits are beyond reach of the present experiment, our results do point to an interesting dissociation within the striatum, with the bilateral putamen and a part of the right caudate body indexing the sign of the reward prediction error and the head of the right caudate indexing the available information that can be extracted from it. These observations are consistent with previous neuroimaging results relating striatal activity to executive functions that implement behavioral changes following feedback, for instance, when shifting between objects (Rogers and others 2000; Cools and others 2004) or between contingency rules (Haruno and others 2004; Shohamy and others 2004). In particular, recent studies seeking to define the computation site of the reward prediction error have shown the involvement of striatal structures such as the putamen (McClure and others 2003; O’Doherty and others 2003) and the nucleus accumbens (Berns and others 2001; Breiter and others 2001). Differences among paradigms, such as the requirement of a motor response and the type of reward (primary vs. secondary), may account for these discrepancies in activation topography. Furthermore, the distinction between rewarding and purely informational aspects of feedback can be considered as an exemplification of the difference between the ventral striatum, which would be dedicated in particular to emotional and reward processing (Burgdorf and Panksepp 2005), and the dorsal striatum, which is considered as being involved, in particular, in rule learning and memory for motor behaviors (Alexander and others 1986; White 1997). In this respect, our results are also consistent with the notion of a ventral “critic” guiding a dorsal “actor” system (Schultz and others 1997; Sutton and Barto 1998; O’Doherty and others 2004). In this model, the ventral striatum (critic) computes and updates predictions of future reward, whereas the dorsal striatum (actor), through its involvement in motor and cognitive control, behaves as an actor that uses the error prediction signal to optimize the selection of actions (via the modification of stimulus-response or stimulus-response-reward associations) (O’Doherty and others 2004). In fMRI, the BOLD signal in bilateral ventral putamen and right nucleus accumbens was shown to correlate with the prediction error during both instrumental and Pavlovian conditioning, whereas activity in the left anterior caudate correlated with prediction error only in instrumental conditioning (O’Doherty and others 2004). Our observations provide another dissociation consistent with this actor-critic architecture.

Moreover, our results suggest that the dissociation between the processing of rewarding and informational values of feedback extends to cortical regions. Reward prediction error correlated with a large bilateral frontoparietal network including dorsolateral PFCs, FEF, SMA, lateral PFC/insula, and parietal regions. These regions may be involved in operations such as working memory stabilization and reorientation of attention toward the next element to be tested, which are engaged whenever a positive reward validates the subject’s current hypothesis. On the other hand, brain activation correlated with reward information was found in anterior parts of the search network, especially in rostral ACC and right PFC. These regions have been associated with performance monitoring through response inhibition (Carter and others 1998; Braver and others 2001); the ACC is also associated with error

detection and response conflict (Carter and others 1998). The present results, showing activation increases whenever the feedback (positive or negative) differed from expectations, tentatively suggest a broader role in alerting, as postulated in models of executive attention (Bush and others 2000).

As noted above, our experimental design aimed at characterizing the temporal dynamics of a global frontoparietal network involved in an effortful deduction task and investigating how these variations are driven by reward processing. Although it would be desirable to further decompose the brain mechanisms of the various cognitive components of the model presented in Figure 2, they are currently too intercorrelated in time to be distinguished by the present fMRI study. Further research should therefore attempt to temporally decorrelate the different cognitive operations, for example, by using variable interstimulus intervals (see, e.g., Dale 1999). In this manner, the cognitive functions associated with each active region, as well as their interactions, might be delineated. An alternative approach would be to study separately the different cognitive components of interest with a hierarchy of simpler tasks involving a subset of the current cognitive processes. Examples are given by studies of working memory (e.g., Haxby and others 2000) and error processing (e.g., Braver and others 2001). However, we note that it is often necessary to use a cognitively complicated task in order to reliably activate the most anterior regions of PFC (see studies by Koechlin and others 2000, 2003). Thus, certain cognitive control processes might remain beyond reach when using simpler tasks.

We close with a methodological note. To identify task-related areas, we used a hierarchy of increasingly refined linear models for the same fMRI data. One potential problem with this approach is that some of the regressors may be intercorrelated. To address this criticism, we verified that a single global model, where all the variables were gathered, yielded essentially identical results (C. Landmann, unpublished data). Nevertheless, our hierarchical approach, similar to stepwise regression in behavioral experiments, may be more insightful in resolving the mechanisms of activation. For instance, part of the frontocingulate network that collapsed following logical discovery periods compared with chance discovery periods, also correlated with the reward information and reward prediction error. We speculate that the latter may provide an explicit mechanism for the former. Following logical discovery, the positive reward no longer provides information because it is fully anticipated and thus the reward prediction error is null (the subject anticipates being rewarded). What appears, in the first model, as an unexplained difference between logical and chance discovery is later seen, in the second model, as a putative consequence of the reward prediction mechanism.

Conclusion

We designed a dynamic, cognitive search task where the different cognitive processes involved could be inferred based on behavioral data. This task allows exploration of fundamental processes of reward-based behavioral adjustment and internal deduction processes. Taken together, our results suggest that effortful mental search engages a distributed set of prefrontal, cingulate, parietal, and striatal brain regions whose dynamics closely track the sequence of reasoning processes. These regions collectively allow the subject to adapt in a flexible way to the rapidly changing contingencies

between stimulus and reward. None of these activations can be explained properly based solely on the minimal stimuli and responses involved (colored circles and motor key presses). Rather, understanding them requires assuming highly structured internal processes of anticipatory reward prediction and deduction whose detailed architecture remains to be fully resolved.

Notes

We thank Mariano Sigman and Andreas Kleinschmidt for helpful comments. We also thank Christophe Pallier for his assistance regarding statistical procedures. CL was supported by a Fondation pour la Recherche Médicale fellowship and SD by a centennial fellowship of the McDonnell Foundation. *Conflict of Interest*: None declared.

Address correspondence to email: landmannclaire@yahoo.fr.

References

- Alexander GE, Delong MR, Strick PL. 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Ann Rev Neurosci* 9:357-381.
- Baker SC, Rogers RD, Owen AM, Frith CD, Dolan RJ, Frackowiak RSJ, Robbins TW. 1996. Neural systems engaged by planning: a PET study of the Tower of London task. *Neuropsychologia* 34: 515-526.
- Berns GS, McClure SM, Pagnoni G, Montague PR. 2001. Predictability modulates human brain response to reward. *J Neurosci* 21: 2793-2798.
- Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD. 2001. Conflict monitoring and cognitive control. *Psychol Rev* 108:624-652.
- Braver TS, Barch DM, Gray JR, Molfese DL, Snyder A. 2001. Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cereb Cortex* 11:825-836.
- Breiter HC, Aharon I, Kahneman D, Dale A, Shizgal P. 2001. Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30:619-639.
- Burgdorf J, Panksepp J. 2006. The neurobiology of positive emotions. *Neurosci Biobehav Rev* 30:173-187.
- Bush G, Luu P, Posner MI. 2000. Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn Sci* 4:215-222.
- Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD. 1998. Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* 280:747-749.
- Cohen JD, Servan-Schreiber D, McClelland JL. 1992. A parallel distributed processing approach to automaticity. *Am J Psychol* 105:239-269.
- Cools R, Clark L, Robbins TW. 2004. Differential responses in human striatum and prefrontal cortex to changes in object and rule relevance. *J Neurosci* 24:1129-1135.
- Dagher A, Owen AM, Boecker H, Brooks DJ. 1999. Mapping the network for planning: a correlational PET activation study with the Tower of London task. *Brain* 122(Pt 10):1973-1987.
- Dale AM. 1999. Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8:109-114.
- Deglin VL, Kinsbourne M. 1996. Divergent thinking styles of the hemispheres: how syllogisms are solved during transitory hemisphere suppression. *Brain Cogn* 31:285-307.
- Dehaene S, Changeux JP. 1991. The Wisconsin Card Sorting Test: theoretical analysis and modeling in a neuronal network. *Cereb Cortex* 1:62-79.
- Dehaene S, Changeux JP. 2000. Reward-dependent learning in neuronal networks for planning and decision making. *Prog Brain Res* 126:217-229.
- Dehaene S, Kerszberg M, Changeux JP. 1998. A neuronal model of a global workspace in effortful cognitive tasks. *Proc Natl Acad Sci USA* 95:14529-14534.
- Fuster JM. 1989. The prefrontal cortex: anatomy, physiology, and neuropsychology of the frontal lobe. New York: Raven.

- Goel V, Dolan RJ. 2003. Reciprocal neural response within lateral and ventral medial prefrontal cortex during hot and cold reasoning. *Neuroimage* 20:2314-2321.
- Haruno M, Kuroda T, Doya K, Toyama K, Kimura M, Samejima K, Imamizu H, Kawato M. 2004. A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. *J Neurosci* 24:1660-1665.
- Haxby JV, Petit L, Ungerleider LG, Courtney SM. 2000. Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. *Neuroimage* 11:145-156.
- Hollerman JR, Schultz W. 1998. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* 1:304-309.
- Knutson B, Fong GW, Bennett SM, Adams CM, Hommer D. 2003. A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *Neuroimage* 18:263-272.
- Koechlin E, Corrado G, Pietrini P, Grafman J. 2000. Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. *Proc Natl Acad Sci USA* 97:7651-7656.
- Koechlin E, Ody C, Kouneiher F. 2003. The architecture of cognitive control in the human prefrontal cortex. *Science* 302:1181-1185.
- McClure SM, Berns GS, Montague PR. 2003. Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339-346.
- Miller EK, Cohen JD. 2001. An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167-202.
- Morris RG, Ahmed S, Syed GM, Toone BK. 1993. Neural correlates of planning ability: frontal lobe activation during the Tower of London test. *Neuropsychologia* 31:1367-1378.
- Nakahara K, Hayashi T, Konishi S, Miyashita Y. 2002. Functional MRI of macaque monkeys performing a cognitive set-shifting task. *Science* 295:1532-1536.
- Newman SD, Carpenter PA, Varma S, Just MA. 2003. Frontal and parietal participation in problem solving in the Tower of London: fMRI and computational modeling of planning and high-level perception. *Neuropsychologia* 41:1668-1682.
- O'Doherty J, Dayan P, Friston K, Critchley H, Dolan RJ. 2003. Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329-337.
- O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304:452-454.
- Owen AM, Doyon J, Petrides M, Evans AC. 1996. Planning and spatial working memory: a positron emission tomography study in humans. *Eur J Neurosci* 8:353-364.
- Pallier C, Dupoux E, Jeannin X. 1997. EXPE: an expandable programming language for psychological experiments. *Behav Res Methods Instrum Comput* 29:322-327.
- Procyk E, Tanaka YL, Joseph JP. 2000. Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. *Nat Neurosci* 3:502-508.
- Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. 2001. A default mode of brain function. *Proc Natl Acad Sci USA* 98:676-682.
- Rainer G, Asaad WF, Miller EK. 1998. Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature* 393:577-579.
- Ramnani N, Elliott R, Athwal BS, Passingham RE. 2004. Prediction error for free monetary reward in the human prefrontal cortex. *Neuroimage* 23:777-786.
- Rogers RD, Andrews TC, Grasby PM, Brooks DJ, Robbins TW. 2000. Contrasting cortical and subcortical activations produced by attentional-set shifting and reversal learning in humans. *J Cogn Neurosci* 12:142-162.
- Rougier NP, Noelle DC, Braver TS, Cohen JD, O'Reilly RC. 2005. Prefrontal cortex and flexible cognitive control: rules without symbols. *Proc Natl Acad Sci USA* 102:7338-7343.
- Schultz W, Dayan P, Montague PR. 1997. A neural substrate of prediction and reward. *Science* 275:1593-1599.
- Seidemann E, Meilijson I, Abeles M, Bergman H, Vaadia E. 1996. Simultaneously recorded single units in the frontal cortex go through sequences of discrete and stable states in monkeys performing a delayed localization task. *J Neurosci* 16:752-768.
- Shohamy D, Myers CE, Grossman S, Sage J, Gluck MA, Poldrack RA. 2004. Cortico-striatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. *Brain* 127:851-859.
- Sutton RS, Barto AG. 1998. Reinforcement learning: an introduction. Cambridge, MA: MIT Press.
- van den Heuvel OA, Groenewegen HJ, Barkhof F, Lazeron RH, van Dyck R, Veltman DJ. 2003. Frontostriatal system in planning complexity: a parametric functional magnetic resonance version of Tower of London task. *Neuroimage* 18:367-374.
- Wallis JD, Miller EK. 2003. From rule to response: neuronal processes in the premotor and prefrontal cortex. *J Neurophysiol* 90:1790-1806.
- White NM. 1997. Mnemonic functions of the basal ganglia. *Curr Opin Neurobiol* 7:164-169.
- Yu AJ, Dayan P. 2005. Uncertainty, neuromodulation, and attention. *Neuron* 46:681-692.