Striatum Forever, Despite Sequence Learning Variability: A Random Effect Analysis of PET Data

P. Peigneux,[1, 2,3] P. Maquet,[1, 3] T. Meulemans,[2] A. Destrebecqz,[4] S. Laureys,[1,3] C. Degueldre,[1] G. Delfiore,[1] J. Aerts,[1] A. Luxen,[1] G. Franck,[1,3] M. Van dei- Linden,[2] A. Cleeremans[4]

[1]*Centre de Recherches du Cyclotron, Université de Liège, Belgium*

[2]*Service de Neuropsychologie, Université de Liège, Belgium*

[3]*Service de Neurologie, Université de Liège, Belgium*

[4]*Cognitive Science Research Unit, Université Libre de Bruxelles, Belgium*

Abstract

This PET study is concerned with the what, where, and how of implicit sequence learning. In contrast with previous studies imaging the serial reaction time (SRT) task, the sequence of successive locations was determined by a probabilistic finite-state grammar. The implicit acquisition of statistical relationships between serially ordered elements (i.e., what) was studied scan by scan, aiming to evidence the brain areas (i.e., where) specifically involved in the implicit processing of this core component of sequential higher-order knowledge. As behavioural results demonstrate between- and within-subjects variability in the implicit acquisition of sequential knowledge through practice, functional PET data were modelled using a random-effect model analysis (i.e., how) to account for both sources of behavioural variability. First, two mean condition images were created per subject depending on the presence or not of implicit sequential knowledge at the time of each of the 12 scans. Next, direct comparison of these mean condition images provided the brain areas involved in sequential knowledge processing. Using this approach, we have shown that the striatum is involved in more than simple pairwise associations and that it has the capacity to process higher-order knowledge. We suggest that the striatum is not only involved in the implicit automatization of serial information through prefrontal cortex-caudate nucleus networks, but also that it plays a significant role for the selection of the most appropriate responses in the context created by both the current and previous stimuli, thus contributing to better efficiency and faster response preparation in the SRT task.

Key words

Implicit learning; serial reaction time; individual variability; PET; statistical parametric mapping; striatum; caudate nucleus

INTRODUCTION

Implicit learning is often described as the non-episodic learning of complex information in an incidental manner, without awareness of what has been learned [Seger, 1994; see also Berry and Dienes, 1993; Reber, 1993; Shanks and St John, 1994]. Inferences about the mechanisms involved in implicit learning situations rely on various experimental paradigms, among which the most widely used is the serial reaction time (SRT) task. Numerous studies have now explored sequence learning performance in this task from cognitive, neuropsychological, neurophysiologi-cal, and functional brain imaging perspectives.

In a typical SRT task, as initially developed by Knopman and Nissen [1987], participants face a computer and are simply instructed to respond as quickly as possible to the appearance of a stimulus at one of four locations on the screen by pressing on one of four corresponding keys arranged in a spatially compatible layout. Unbeknownst to participants, the sequence of locations at which successive stimuli appear is manipulated: a fixed (e.g., ten-element) sequence is repeated throughout successive blocks of trials, except for the penultimate block, during which the sequence is random. Typical results indicate that while reaction time (RT) decreases with practice, transfer to the random block elicits slower RTs. This interference suggests that participants have acquired sequence-specific knowledge over training and use it to improve their performance beyond what could be expected from simple visuo-motor adaptation. Further, participants are generally not aware that the material contained sequential structure, albeit this issue still remains controversial [e.g., see Shanks and Johnstone, 1998]. Sequence learning in the SRT task is a robust phenomenon, as indicated by the variety of paradigms in which it has been observed. Nevertheless, the nature and the mechanisms of sequence learning in the SRT task remain unclear. Many questions are still debated, such as the explicit or implicit nature of the acquired information, the form(s) of the representations being learned (e.g., the serial order of the elements, simple pairs of stimuli, higher probabilities of

occurrence for particular groups of stimuli, statistical relationships between sequential elements), and various methodological concerns influencing performance and its interpretation [for reviews, see Cleeremans et al., 1998a; Curran, 1995; Meulemans, 1998a; Stadler and Frensch, 1998].

From a neuropsychological perspective, sequence learning in SRT tasks is altered in patients suffering from degenerative disorders involving lesions of the basal ganglia, primarily Huntington's disease (HD) [Knopman and Nissen, 1991; Willingham and Korosh-etz, 1993] and Parkinson's disease (PD) [Doyon et al., 1997, 1998; Ferraro et al, 1993; Jackson et al., 1995; Pascual-Leone et al., 1993; Westwater et al., 1998]. In accordance with the proposal that skill learning may depend on the integrity of these nuclei [e.g., Squire, 1992], these findings suggest that the basal ganglia play an important role in the implicit acquisition of sequential information. However, their role is not exclusive because sequence learning may also be impaired in patients with circumscribed lesions of cerebellum [Doyon et al., 1997,1998; Gomez-Beldarrain et al., 1998; Molinari et al., 1997] and supplementary motor area [Ackermann et al., 1996].

Functional neuroimaging and neurophysiological studies have further demonstrated that implicit sequence learning may involve a more extended set of cerebral regions. With positron emission tomography (PET), Grafton et al. [1995] showed, using dual-task conditions, that regional cerebral blood flow (rCBF) increases were related to the implicit processing of structured sequences in contralateral motor cortex, supplementary motor area and putamen, a result replicated with colour-coded rather than spatially-coded stimuli [Hazeltine et al., 1997]. Under single-task conditions, implicit sequence processing was associated with rCBF activation in the right hemisphere in ventral premotor cortex, caudate nucleus and thalamus, and bilaterally in visual association areas [Rauch et al., 1995]. Another study pointed to a specific involvement of the cerebellum and the striatum in an advanced—and more automated—late stage of the implicit learning process [Doyon et al., 1996]. At variance, using a probabilistic rather than a fixed (i.e., deterministic) sequence of numeral-coded stimuli, Berns et al. [1997] found a ventral striatum activation more important in the early phase of sequence learning, whereas rCBF progressively increased in the dorsolateral prefrontal cortex (DLPFC). Functional magnetic resonance imaging (fMRI) studies further showed that the thalami were deactivated while the striatum was recruited in the early phase of learning [Rauch et al., 1997], suggesting a thalamic gating mediated by the indirect striato-pallido-thalamic pathway [Rauch et al., 1998]. Finally, a recent PET study showed a correlation between performance improvement and increased activity in the contralateral primary sensorimotor cortex when the RT for repeated sequence of numeral-coded stimuli was correlated with rCBF in the implicit learning phase [Honda et al., 1998], but failed to find any basal ganglia or cerebellum activation.

The diversity of the results reported by these neu-roimaging studies may be partly explained by several factors: the experimental design, the stimulus type (e.g., spatial locations, colours, arabic numerals, letters), the structure of the material (i.e., fixed vs. probabilistic sequences), or the relative contribution of explicit knowledge to performance as well as the way in which this contribution was assessed (verbal reports, structured questionnaires, generation or recognition tasks). Specific limitations of each neuroimaging technique and of analysis methods (whole-brain or region-of-interest investigation, a priori or post-hoc hypotheses, categorical or parametric analyses) also influence the results of such experiments. As a consequence, the diversity of brain areas activated in the SRT task may reflect various components of functional networks responsible for implicit sequential learning. In the present PET study, our goal was to identify the core cerebral correlates of sequence learning in the SRT task. To correlate rCBF with behavioural evidence of the acquisition of complex higher-order sequential knowledge, one must ensure that the task could not be completed on the basis of lower-order knowledge acquisition with practice. In this respect, it has been claimed that implicit sequential knowledge is essentially statistical in nature, regardless of whether the task involves deterministic or probabilistic material [Cleeremans and Jiménez, 1998; Shanks and Johnstone, 1998]. Probabilistic sequences present the advantage of controlling for the influence of the simple memorization strategies that can be effective when the material is deterministic. Indeed, probabilistic material allows for many more combinations of sequence elements to be presented since any stimulus can occur as the successor of any other with some definite conditional probability. As a result, no element can be predicted with certainty on the only basis of lower-order serial knowledge of pairwise associations, or based on memory for entire sequences. Hence, this study explored sequence learning using probabilistic sequences of visual stimuli in the SRT task. Sequences were generated on the basis of a finite-state grammar. It was assumed that sequence-related RT improvement depends on the acquisition of higher-order statistical relationships between disjoint sequence elements [Cleeremans and McClelland, 1991; Jiménez et al., 1996]. To our knowledge, few studies are based on this type of task. Baldwin and Kutas [1997] have recorded event-related brain responses (ERPs) over cen-troparietal scalp locations, and documented the development of expectancies about predictable events with practice. Berns et al. [1997], using PET, showed a striatum reactivation in response to the presentation of a novel probabilistic sequence after training on another such sequence. These results were interpreted as suggestive of a metabolic response to novelty in the absence of awareness. However, a significant response to a new sequence does not inform us about which aspects of the sequential structure were responsible for this response.

Another difficulty to cope with was to take into account the typically high within- and between-sub-jects performance variability in the analysis of PET data. This second point deserves further explanation, as we believe that important sources of variation in individual behavior should be heeded when looking for the haemodynamic correlates of subjects' performance. Up until recently, most brain imaging group studies tacitly assumed that each subject makes approximately the same fixed contribution to the observed patterns of activation (i.e., a statistical "fixed-effect" model). In other words, it is assumed that there is no Subject by Condition interaction, and that individual differences between subjects are constant across conditions during the experiment. If this were effectively the case, then fixed-effect analyses would indeed be statistically valid. However, in the present experiment, consistent variability in the evolution of individual performance in the SRT task was observed. It was thus necessary to choose more appropriate statistical models that would account for both between- and within-subject variability. Accordingly, we used the so-called "random effect" model analysis [Holmes and Friston, 1998], based on the approach of mean summary statistics on repeated measures [Frison and Pocock, 1992]. It is important to note that this type of analysis is usually used in neuroimaging studies to accommodate for the interindividual variability of PET or fMRI data. In the present study, we used it to capture the variability of behavioural performance in the analysis of functional data. As an advantage, the inferences could be validly generalised to the population from which our subjects were drawn because both sources of variance are accounted for.

Hence, and to paraphrase a recent review [Curran, 1998], this experiment is concerned with the what, where, and how of the functional neuroimaging of implicit sequence learning. That is, we tried to identify the neural correlates (i.e., where) of the implicit processing of statistical higher-order sequential information (i.e., what), and took advantage of recent advances in the statistical analysis of functional neuroimaging data (i.e., how) to capture both between- and within-subjects variability in the development of the acquisition and use of such knowledge.

MATERIALS AND METHODS

Subjects

Fourteen right-handed healthy volunteers (four men and ten women, range 20-29 years) gave their informed consent to participate to this PET experiment, approved by the Ethical Committee of the University of Liège. All subjects were recruited through advertisement and participated as paid volunteers. Based on a preliminary interview, none of them evidenced any history of any significant psychiatric, neurologic, or medical illness, nor had they ever participated before in a PET or in a SRT experiment.

Apparatus

Participants lay in the supine position on the bed of the PET scanner, with a Macintosh PowerPC computer 17-inch screen located in front and upside to them at a mean distance of 80 cm from the eyes (visual angle: 42.1° X 31.6°). The screen itself was 25 cm above a reclining table with a six-button response keyboard. Keys were spaced 4 cm apart, to elicit whole-arm movements, and discourage finger-specific responses. Head-holder and keyboard position were individually adjusted for optimal comfort during task performance, ensuring that all responses may be given with minimal forearm displacements. A thermoplastic face-mask secured to the scanner head-holder maintained the position of the subject's head, but did not prevent comfortable vision of the screen and of the keyboard. During the PET session, room lights were dimmed and ambient noise kept constant.

SRT Task

Because the probabilistic SRT task has been previously described elsewhere [Cleeremans and McClelland, 1991; Jiménez et al., 1996], we only present information relevant to this specific study. Before the PET session and SRT practice, participants were simply told that the goal of the experiment was to study with PET the cerebral effects of the sustained practice of a simple motor task, and that they could earn additional money depending on how well they performed. The session consisted of 12 blocks of 410 successive trials each. On each trial, a black circle appeared 2 cm below one of six markers arranged horizontally on the screen, and the task consisted of pressing as fast and as accurately as possible with the right hand on the spatially corresponding key. The computer emitted a short beep on incorrect responses. The next stimulus was then displayed after a fixed 200 msec response-stimulus interval (RSI). Following completion of each block, the computer displayed information about the mean RT, the accuracy of the performance, and the financial incentives estimation based on response accuracy and speed.
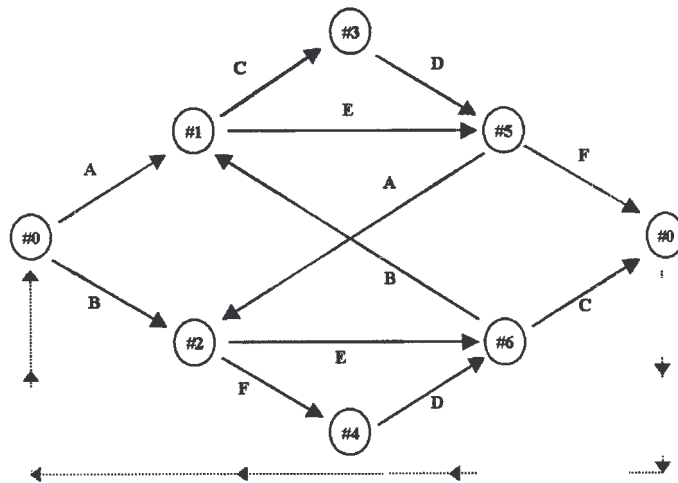
Figure I.

The finite-state grammar used to generate the stimulus material, from Jiménez et al. [1996]. Each letter corresponds to a stimulus location on the screen. At each node (#), one of the possible arrows (→) is randomly selected, and the corresponding label is collected and added to the sequence of so-called grammatical (G) stimuli.

Unknown to participants, the sequential structure of the material was manipulated by generating sequences of stimuli based on the finite-state grammar illustrated in Figure 1. Finite-state grammars instantiate a set of rules that describes permissible transitions between successive stimuli. As Figure 1 shows, each label (i.e., the screen location, corresponding to label A, B, C, D, E, or F) can be legally followed by several so-called grammatical (G) stimuli. Therefore, G stimuli each have a specific probability of appearing in a given context (i.e., a given set of previous stimuli). To best prepare their responses to successive stimuli during the SRT task, subjects therefore need to encode the temporal context (i.e., previous elements of the sequence) in which stimuli occur so as to reduce the uncertainty associated with the next element. For instance, the label A appears twice in the grammar, and may be followed in the first case by labels C or E, and in the second case by labels E or F. Thus, labels C and F each have a 25% chance each to appear after A, whereas E occurs with a 50% probability and B, D, and A never occur after an A. The higher probability of occurrence for C, E, and F at this point in the sequence is thus contingent on the temporal context set by one previous element (i.e., L1 context)—in this example, the label A. It is important that uncertainty associated with the legal successors of each instance of A can be further significantly reduced by encoding two elements of temporal context (i.e., L2 context). Indeed, when D precedes A, only E and F may occur as successors, each with a probability of 50%, whereas all the remaining stimuli now have a null probability of occurring. In some further cases, the context created by pairs of previous stimulus remains ambiguous as well because the two-elements sequences themselves appear twice in the grammar (e.g., A-E). In such cases, optimal predictions require additional previous elements to be encoded. With the grammar used in this experiment, up to three previous labels (i.e., L3 contexts) may be necessary to optimise prediction of the next stimulus. Because sequence-specific RTs improvement depends on better preparation of the motor response for the next event, and because our material controlled for the influence of lower-order regularities, observing improved RTs can only be interpreted as reflecting encoding of the context set by previous elements of the sequence.

To assess whether participants were sensitive to the sequential structure, there was a 15% chance on each trial of replacing the grammatical (G) stimulus by a nongrammatical (NG) stimulus that violated the rules of the grammar. Assuming that motor response preparation is proportional to stimulus predictability, predictable G stimuli should thus elicit faster responses than nonpredictable NG stimuli, albeit only if the context in which stimuli may occur has been encoded by participants. Returning to our example, C, E, and F would elicit faster RTs after A, but only if the LI context has been encoded. Likewise, E and F would elicit faster RTs than C after a D-A context, but only if this L2 context has been encoded by participants. Insertion of random elements in the grammatical sequence therefore makes it possible to conduct detailed comparisons between RTs for G and NG stimuli in specific matched contexts set by the number of previous trials to be considered (e.g., in the L1 context set by the element A, RTs elicited by G stimuli C, E, and F will be compared to RTs elicited by NG stimuli B and D). Previous studies [Cleeremans and McClelland, 1991; Jiménez et al., 1996] have shown that participants can learn about the constraints set by two previous trials at most, and that this sensitivity emerges through practice as a gradually increasing difference between the RTs elicited by G and NG stimuli (NG^{RT}-G^{RT}) occurring in specific sets of controlled contexts of length 1 to 3. Given the limited amount of practice that the constraints of PET studies

impose (4,920 trials), as well as the results of previous validations of this task in our lab (with practice up to 19,680 trials), we did not expect our participants to learn more than the L1 context.

Immediate repetitions (i.e., A-A) leading to short-term priming effects [Cleeremans and McClelland, 1991] were discarded regardless of stimulus type. Likewise, interference effects due to response-related differences between G and NG stimuli were controlled. Each location occurred proportionally at the same rate in the two conditions (i.e., based on the total number of G stimuli generated for one 410-elements block, we attempted to make it so that each of the six positions occurs in 16.6% of the cases). We likewise controlled the frequency of NG stimuli in the same blocks. A two-way analysis of variance failed to show any significant effect of the interaction between the grammatical (G vs. NG) and the location (A vs. B vs. C vs. D vs. E vs. F) types on this proportion of occurrence [$F(5,115) = 2.17$; $p > .05$], indeed suggesting that each specific location was proportionally equally likely to occur regardless of the grammatical status of the associated stimulus. To further control motor, serial order, and spatial effects between locations, the mapping between labels in the grammar and the screen positions was randomly chosen and then systematically modified by shifting the screen locations one step to the right for each participant. Hence, differences between RTs elicited by G and NG stimuli cannot be simply attributed to facilitatory effects unrelated to the learning of the conditional probabilities of the grammar.

Generation Task

Immediately following completion of the SRT task and outside the PET scanner, subjects were told that a set of complicated rules had been used to determine the sequence of locations where successive stimuli could appear, and that they would now be exposed to a generation task to find out whether they had any explicit knowledge about the material. In the generation task, participants faced the same apparatus as used for the SRT task, but instead of reacting to the current stimulus by pressing on the corresponding key, each trial now required that participants predict the location of the next stimulus, again by pressing on the corresponding key, and without any time constraint. If participants felt that they could not predict the location of the next stimulus, they were encouraged to guess and to "follow their instinct." The sequence of stimuli consisted of one block (410 trials) of stimuli that had been previously presented during the SRT task, but that did not incorporate the nongram-matical substitutions. No explicit verbal feedback was given until completion of the task, but the correct stimulus was displayed on the screen following each response, and regardless of whether the response was correct or not. Participants could therefore assess the quality of their prediction responses just as during the SRT task (during which a sound was emitted to indicate erroneous responses).

The benefits of using closely related direct and indirect measures of sequence learning were discussed by Jiménez et al. [1996]. First, the similarity between the two tasks makes it possible to analyze performance in a detailed way and on exactly the same set of temporal contexts (L1 to L3). Second, comparison between compatible direct and indirect tests of the same knowledge provides us with a strong method to assess the degree with which the acquired knowledge is implicit. Indeed, knowledge can be assumed to be implicit whenever participants exhibit greater sensitivity to the relevant information in the context of an indirect test rather than in the corresponding direct test [Merikle and Reingold, 1991; Reingold and Mer-ikle, 1988]. This is simply because if the relevant knowledge were conscious, then participants would certainly make better use of it when explicitly instructed to do so, that is, in the direct, but not in the indirect test.

The analyses were conducted in a way exactly analogous to the way SRT data were analysed: the accuracy of each prediction was assessed depending on both contexts L1 and L2 set by previously displayed stimuli, considering any generated sequence element that is grammatical after a given context as a hit. To assess whether accuracy differed from a random prediction strategy, 14 computerised simulations of continuous random prediction were generated with the simple constraints to avoid immediate repetitions and to begin a block with a G stimulus. These simulations were then assessed referring to given contexts set by previously displayed stimuli, and compared to participants' performance.

Imaging Protocol

PET acquisitions were obtained with a Siemens CTI 951 R 16/31 scanner (CTI, Knoxville, TN, USA; in plane resolution: 8.7 mm) in 3D mode. A transmission scan was acquired for attenuation correction using three rotating sources filled with ^{68}Ge. Emission scans consisted of a 30-sec background frame followed by a 90-sec frame. The rCBF was measured using 12 iterative infusions of oxygen-15 labelled water (6 mCi/222 MBq in 5 cc saline each), automatically infused during 20 sec through a cannula into an antebrachial vein of the left arm. Tasks were begun 15 sec before the second frame, and experimental SRT blocks presented during each of the 12 scans. Data were reconstructed using a Harming filter (cutoff frequency: 0.5 cycle/ pixel), and corrected for attenuation and background activity.

PET Data Analysis

PET data were analysed using the statistical parametric mapping software SPM96 (Wellcome Department of

Cognitive Neurology, London; http://www-.fil.ion.ucl.ac.uk/spm) implemented in MATLAB (Mathworks Inc., Sherborn, MA, USA). Data from each subject were realigned using a least square approach and the first scan as a reference. Following realignment, all images were normalised in a stereotaxic standard anatomical space [Friston et al., 1995; Talairach and Tournoux, 1988] and smoothed using a 16 mm full-width half-maximum (FWHM) isotropic kernel.

Regression Analysis

First, a regression analysis was conducted to highlight systematic relationships between rCBF and performance measures thought to reflect cognitive and sensorimotor processing. Two variables for each scan were entered as separate covariates of interest in the design matrix: (1) the mean overall RT, and (2) the mean RT advantage for G stimuli in L1 context [L1(NG^{RT}- G^{RT})]. The specific rCBF responses to these covariates of interest were estimated at each and every voxel according to the general linear model, using a design matrix which included the global activity as covariate of no interest [Friston et al., 1995b]. The resulting set of voxel values for each contrast constituted a map of the t statistic (SPM{t}), transformed to the unit normal distribution (SPM{Z}) and thresh-olded at $p < .001$ ($z = 3.09$). Statistical inferences were then obtained at the voxel level (in terms of peak height at $p < .05$) corrected for multiple comparisons, using the conjoint test provided in SPM96, by reference to the theory of Gaussian random fields [Poline et al., 1997].

Random Analysis

A two-step procedure [Holmes and Friston, 1998], implemented in SPM96, was applied to appropriately model within and between-subject variances in scans related to the behavioural conditions. First, all images were averaged into two adjusted-mean conditions per subject on the basis of the successfulness (S) or unsuc-cessfulness (U) of L1 context-related RT performance during this particular scan, determined as follows. We assumed that the very first scan was not likely to be associated with any relevant knowledge, and it was therefore systematically associated to the U condition regardless of the differences between RTs elicited by G and NG stimuli. Second, if the L1 (NG^{RT}- G^{RT}) value associated to the remaining scans was a null or negative result (no advantage for G stimuli processing), these scans were also included in the U condition. Otherwise, the L1 (NG^{RT}- G^{RT}) positive measure at each scan was expressed as a percentage of the maximal L1 (NG^{RT}- G^{RT}) value across the 11 remaining scans for each subject. To protect the analysis against spurious positive results, only those scans associated with a percentage higher than 50% of the maximal L1 (NG^{RT}- G^{RT}) value were considered successful and included in the S condition, the others belonging to condition U. All subjects had a minimum of two consecutive S scans, ensuring minimal consistency in the acquired knowledge, and the possibility to create adjusted mean images for each condition. Hence, adjusted mean condition images were created with different numbers of scans in each condition, regarding to the subject performance for context L1.

Because a separable model is required to allow subsequent modelling of the between-subject variance, we ensured that the adjustment for one subject was independent from other subjects by using proportional scaling adjustment of the global activity. The resulting 28 estimates (14 subjects X 2 conditions) fitted the within-subject component of the variance and could then be used for a subsequent second-level analysis, that took between-subject variance into account. In this case, a simple one-sample *t*-test was considered. The design matrix used in this subtraction included the two mean images per subject as conditions. Grand mean scaling and global normalisation were not considered as they had already been taken into account during the within-subject level of this analysis. Statistical inferences were obtained at the voxel level (in terms of peak height at $p < .05$), corrected for multiple comparisons. We also report regional effects at cluster-level ($p < .05$) as well as voxel values above the uncorrected threshold *(p < .001)*. This enables comparisons with other research, and addresses the criticism that reporting only the "hottest voxel" in a peak misses out on subtler activations that can be biologically meaningful in terms of identifying neural circuitry [Gold et al., 1997].
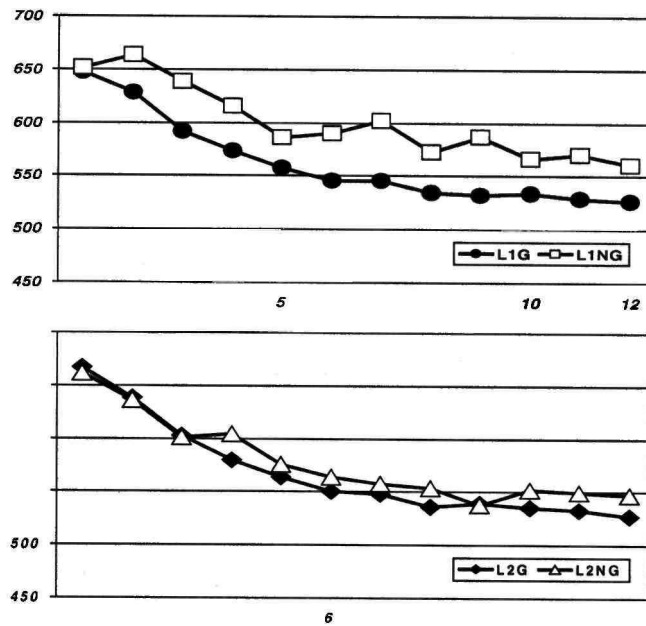
Figure 2.

Group behavioural results. Mean RTs for grammatical (G) and nongrammatical (NG) stimuli in specific temporal contexts defined by one (LI) or two (L2) previous elements, at each of the 12 scans.

BEHAVIOURAL RESULTS

SRT Task

Mean RTs for G and NG stimuli in contexts L1 and L2 were computed for each of the 12 SRT blocks practised during the 12 scans (12 X 410 trials). Incorrect responses and extreme values falling outside of 2 standard deviations from the mean were discarded, as well as the first ten trials of each block. Figure 2 shows that practice exerts a strong effect on performance: participants improved their global RT as training progressed. In addition, a comparison between RTs for G and NG stimuli shows that grammaticality influences RTs for L1 contexts, but not yet for L2 contexts. Two-way analyses of variance with grammaticality (G vs. NG, two levels) and practice (RTs, 12 levels) as within-subjects factors confirmed these observations. A significant main effect of grammaticality was observed for the L1 context, with mean RT elicited by G stimuli faster than mean RT elicited by NG stimuli, 562 vs. 600 msec [F(l, 13) = 13.73, MSE = 8986.7, $p < .005$], confirming that participants were sensitive to the sequential constraints set by a single previous stimulus.

In contrast, the grammaticality effect did not reach significance for L2 contexts, although RTs elicited by G stimuli tended to be slightly faster than RTs elicited by NG stimuli, 568 vs. 578 msec *(p = .06)*. A significant main effect of practice was also found for L1 [F(ll, 143) = 28.06, MSE = 1373.1, $p < .0001$], and L2 context [F(ll, 143) = 21.80, MSE = 2291.6, $p < .0001$], showing improvement of global RT performance with practice irrespective of the grammatical status of the stimuli. Lastly, the Grammaticality X Practice interaction was significant for L1 contexts [F(ll, 143) = 2.8, MSE = 581.6, $p < .015$], but not L2 contexts ($p > .7$), suggesting that with such a restricted practice, participants were able only to learn about the probabilistic constraints set by the preceding trial (i.e., the L1 context).

Beyond these results, close inspection of the individual data reveals different sources of variability clearly illustrated by the graphical representation of three individuals' RT performance for G and NG stimuli in L1 contexts (see Fig. 3, top row). First, whereas overall RT performance may vary widely from subject to subject, all of them improve their performance over practice in the same way. There is therefore no Subject X Condition (block) interaction. Hence, even if initial levels of global RT performance differ between subjects, all of them demonstrate a similar Practice effect, and between-subject differences are not critical in this case insofar as analysis of the functional PET data is concerned.

However, this is far from being the case for the differentiation between G and NG stimuli over practice, illustrated for the three example participants by plotting at each scan the difference between RTs for G and NG stimuli in L1 contexts [(NG^{RT}- G^{RT}); see Fig. 3, middle row]. Although this differentiation is markedly apparent very early on for the first subject (S12), it occurs later for another (S11), and not at all for a third one (S5). Moreover, it is

apparent that fluctuations are present in the magnitude of the differentiation between RTs for G and NG stimuli during successive scans of the same subject, and that this differentiation neither increases monotoni-cally throughout practice nor reaches a plateau at some point.
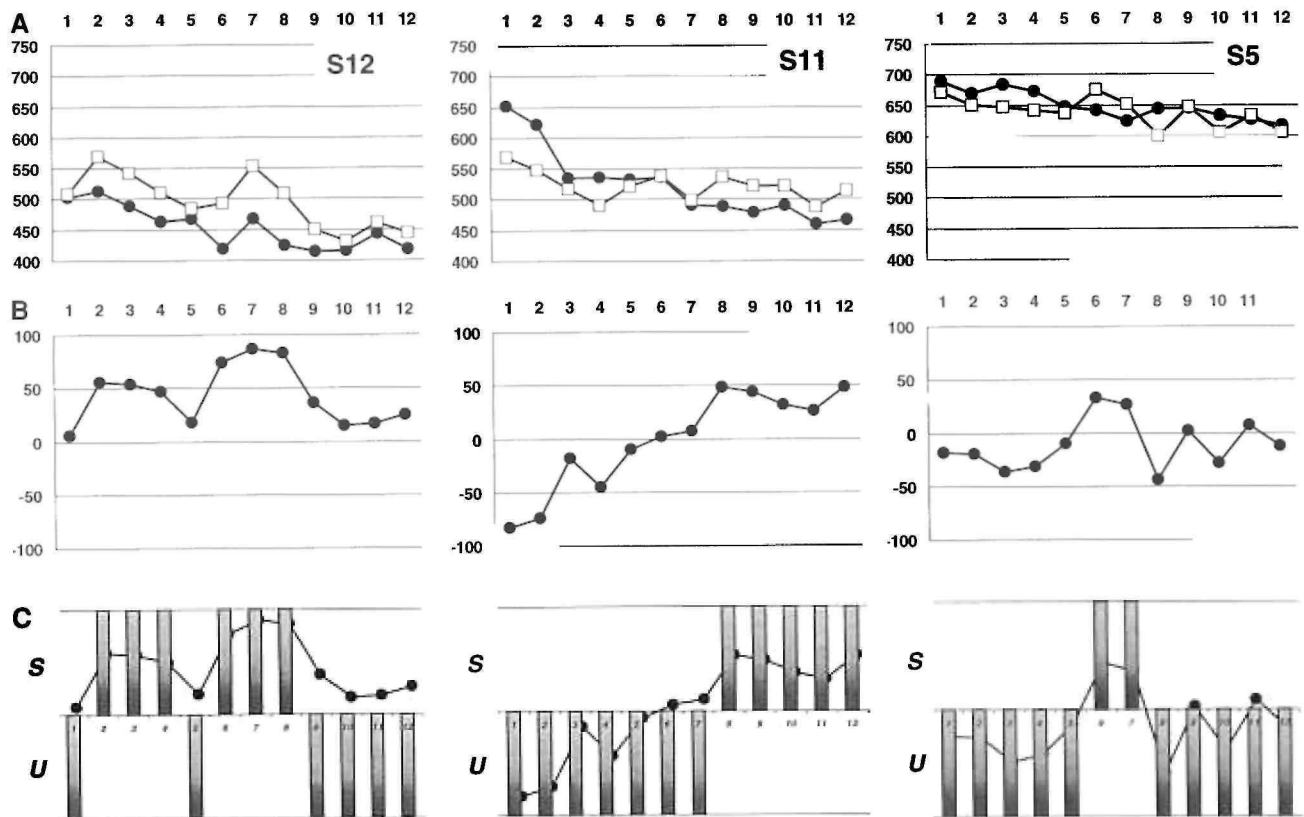

Figure 3.

Subject by condition interaction. Illustration of the within- and between-individual variability of the LI context learning for three sample participants, (a) In the first row, mean RTs (msec) for grammatical (G) and nongrammatical (NG) stimuli in LI context at each of the 12 scans are displayed. Global RT monotonically decreases in all subjects, (b) In the middle row, LI context learning for these three subjects is expressed through RT (msec) differences between G and NG stimuli at each scan [LI(NG^{RT}-G^{RT})]. LI context learning varies both between subjects and within subjects, in a nonlinear way, as shown by improvements and regressions of the performance (e.g., S12). (c) In the bottom row, LI context performance at each scan was modelled under two possible behavioural conditions: [S]uccesful or [U]nsucessful. Adjusted-mean condition images are created with different numbers of scans in each condition, regarding to the subject performance for context LI.

Generation Task

Independent analyses of variance with group (Subjects vs. Simulation) and grammaticality (G vs. NG) variables conducted on the data corresponding to each context fail to show any significant Group main effect on prediction accuracy both for L1 contexts [$F(1,26) = 0.24, p < .63$], and for L2 contexts [$F(1,26) = 0.01, p < .93$]. A main effect of Grammaticality was observed, with more stimuli generated accurately (G) than inaccurately (NG) both for L1 contexts (means = 218 vs. 115) [$F(1,26) = 171.46, p < .001$], and for L2 contexts (means = 59 vs. 45) [$F(1,26) = 78.00, p < .001$]. These differences between the number of G and NG predictions were expected, however, because there are more possible G successors than NG successors in the L1 and L2 contexts [see Jiménez et al., 1996, p. 955, for a detailed presentation]. Finally, the interaction between Group and Grammaticality failed to reach significance, both for L1 contexts [$F(1,26) = 0.78, p < .39$], and L2 contexts [$F(1,26) = 2.47, p < .13$], thus suggesting that subjects' performance on the generation task did not differ from chance. Therefore, it seems that subjects show greater sensitivity to the sequential structure during the indirect (SRT) task than during the direct (generation) task. If interpreted according to Reingold and Merikle's conceptualisation [Merikle and Reingold, 1991; Reingold and Merikle, 1988], this pattern of results is highly suggestive that most of the knowledge acquired during practice with the SRT task was unconscious.

FUNCTIONAL IMAGING RESULTS

Regression analyses

The SPM{Z} had a volume of 196744 voxels, with 11.8 X 12.8 X 13.3 mm FWHM smoothness (786.0 RESELS) and 150 residual degrees of freedom. Several cerebral sites were identified in which rCBF variations across the 12 scans were monotonically negatively correlated to the mean global RT. These representative sites (Table I) demonstrating an increasing rCBF activity with improvement in RT performance include, mainly in the left hemisphere contralateral to the response hand, the inferior occipital (BA 18), lingual (BA 17, 18), transverse temporal (BA 41), posterior cingu-late (BA 23, 30), postcentral (BA 3) and precentral (BA 4, 6), middle frontal (BA 6, 8, 9) and inferior frontal (BA 45) gyri, the inferior parietal lobule (BA 40), the cerebellum, the superior longitudinal bundle, and the insula; in the right hemisphere, the cuneus (BA 17), the lingual (BA 17, 18), fusiform (BA 18) and medial frontal (BA 6) gyri, the superior longitudinal bundle and the cingulum.

On the other hand, the regression analysis failed to show any significant systematic relationship between rCBF variations across scans and the evolution of the L1 (NG^{RT}- G^{RT}) measure, hence we turn toward the random effect analysis.

Random Effect Analysis

The SPM{Z} had a volume of 201479 voxels, with 9.0 X 9.4 X 9.5 mm FWHM smoothness (2014.1 RESELS) and 13 residual degrees of freedom. The final subtraction between mean-conditions estimates across subjects reveals that condition S is significantly associated with rCBF increase (see Fig. 4) in the striatum at the voxel-level (coordinates in the standard Talairach and Tournoux stereotaxic space: -16 8 10 mm, $p < .05$, corrected for multiple comparisons). Regional effects at the cluster-level *(p < .05)* and uncorrected voxel-level ($p < .001$, see Table II) were observed in the left hemisphere in the caudate nucleus, the putamen, and the middle (BA 10) and inferior frontal gyri (BA 44, 45, 46, 47).

DISCUSSION

Following restricted practice, participants exposed to a probabilistic SRT task demonstrated limited sensitivity to the statistical features of a structured sequential material, that is, to the temporal context set by one previous element of the sequence (L1). Given that this sensitivity to sequential structure failed to be expressed by participants in the context of a closely matched direct test (the generation task), this knowledge may be considered to be mostly implicit or unconscious. Our results further suggest that this sensitivity cannot be accounted for by acquisition of lower-order knowledge, but rather reflects the unconscious use of higher-order L1 contextual information. In addition, detailed analysis of the data has not only revealed between-subject variability in this higher-order learning process, but also that within-subject variability may be observed across time. Presumably, this variability is a common observation in implicit learning studies, but its importance has not been explicitly considered in previous neuroimaging experiments. Our analysis based on the random effect model (which embodied this between- and within-subject variability in behavioural results) showed that the striatum was significantly more activated during scans where subjects responded significantly faster for predictable than for unpredictable stimuli. It is also worth noting that in addition to the striatal involvement in the successful processing of the L1 context, we found (at a lower statistical threshold) a related regional activation in the ipsilateral prefrontal areas (see Fig. 5), which suggests that there is a functional interrelationship between these regions and the basal ganglia. The finding of a striatal activation in implicit sequence learning in the SRT task confirms the results of most functional brain imaging studies [Berns et al., 1997; Doyon et al., 1996; Grafton et al., 1995; Hazeltine et al., 1997; Rauch et al, 1995, 1997, 1998], and extends this inference to the population level by using the random effect model.

At a lower-order level, we have shown that all participants tend to improve their overall RT performance with time in a similar way, which reflects simple effects related to the processing of a sequential multiple-choice reaction time task, independently of the acquisition of statistical higher-order information. On the one hand, this pattern is consistent with the results of the regression analysis showing, with overall RT improvement, the increasing involvement of a large perceptive-motor network, including motor areas (primary and premotor cortex, SMA) recruited for response preparation and selection on the basis of external or internal information, occipito-parietal structures contributing to visuospatial perception and attentional selection, and frontal cortex contribution in anticipation and inhibition of responses [Curran, 1995, 1998], wheeas the cerebellum might be concerned with monitoring and optimising movements using sensory feedback [Jueptner et al., 1998]. On the other hand, several prior studies have also associated rCBF increases in such motor, supplementary motor, premo-tor, or visual cortices with implicit sequence learning in humans [Grafton et al, 1995; Hazeltine et al, 1997; Honda et al., 1998; Rauch et al., 1995], which may suggest that these regions are also involved in learning-related, even if lower-order, levels of processing in the SRT task.

TABLE I. Regression analysis results. Brain areas where rCBF significantly correlated with global RT improvement*

Area of activation	Hemisphere L(eft) R(ight)	BA	Coordinates x	y	z	Z-score	*p*-corr. value
Inferior occipital gyrus	L	18	-24	-102	-6	7.09	0.000
Cuneus	R	17	6	-104	6	4.87	0.015
Cerebellum		Cb	0	-92	-24	5.07	0.006
Lingual gyrus	L	18	-20	-98	-12	6.93	0.000
Lingual gyrus	R	18	8	-104	-4	5.39	0.001
Fusiform gyrus	R	18	24	-102	-10	6.12	0.000
Transverse temporal gyrus	L	41	-34	-26	20	4.87	0.015
Inferior parietal lobe	L	40	-36	-30	34	4.65	0.040
Cingulate posterior gyrus	L	30	-8	-46	16	4.98	0.008
Fasciculus longitudinalis superior	R	FLS	34	-42	22	5.97	0.000
Fasciculus longitudinalis superior	L	FLS	-34	-16	24	5.09	0.005
Postcentral gyrus	L	3	-38	-20	34	4.95	0.011
Precentral gyrus	L	6	-26	-18	62	5.72	0.000
Precentral gyrus	L	4	-38	-12	30	5.48	0.001
Cingulum	R	Ci	12	36	6	5.39	0.001
Insula	L	Ins	-26	-18	18	4.95	0.011
Middle frontal gyrus	L	6	-36	2	38	5.08	0.006
Middle frontal gyrus	L	9	-34	20	24	4.80	0.021
Middle frontal gyrus	L	8	-32	12	28	4.59	0.050
Inferior frontal gyrus		45	-40	24	16	5.54	0.001
Medial frontal gyrus (SMA)	R	6	8	-18	54	4.85	0.016

BA: Brodmann area; p^{corr}: *p*-value corrected for multiple comparisons. Coordinates *x, y, z* refer to the standard Talairach and Tournou: [1988] stereotaxic space. Only the most representative voxel for each area is displayed here.

Peculiarities of the present study may explain the fact that we found striatal and motor/premotor areas separately associated to distinct levels of processing in implicit sequence learning. First, unlike what happens with probabilistic material, the deterministic nature of the sequences often used in prior studies makes it possible that sequence elements can be *univocally* predicted on the basis of the serial knowledge of two previous elements, or even that the entire sequence can be memorised. Such fixed-associations of serial elements may result in encapsulated motor programs during the implicit learning phase, stored in motor or premotor cortices, leading to an improvement in RT performance. It was suggested that such a motor sequence process corresponds to a cortico-subcortical loop between putamen and premotor-motor cortex [e.g., Hikosaka et al., 1999], which may explain the fact that striatal and motor regions were both activated in some studies [Grafton et al., 1995; Hazeltine et al., 1997; Rauch et al., 1995]. Second, our measure of higher-order knowledge is actually independent of the global improvement in RT across scans (which was the dependent measure in prior studies). Indeed, what was computed here is the difference between RTs elicited by G and NG items for L1 or L2 contexts in the same scan. Consequently, the measure of the implicit acquisition of the sequential contingencies in a probabilistic task reflects a higher-order cognitive process that is not as dependent on motor performance as the implicit acquisition of deterministic sequences. This difference explains the absence of motor activation in the results of the random-effect analysis, but its presence in the results of the regression analysis with RT improvement.

As we will now discuss, our results suggest that the striatum plays a critical role in the processing of already acquired higher-order sequential knowledge. This hypothesis is further supported by neuropsychological

observations as well as by basal ganglia function modelling. From a neuropsychological perspective, sequence learning in the SRT task has already been shown to be impaired in HD and PD patients [Ferraro et al., 1993; Knopman and Nissen, 1991; Pas-cual-Leone et al., 1993; Willingham and Koroshetz, 1993]. For some authors, these deficits may reflect an interruption of the processing loop involving the prefrontal cortex and the caudate nucleus of the striatum [Ferraro et al., 1993; Jackson et al., 1995; Knopman and Nissen, 1991; Pascual-Leone et al., 1993; Willingham and Koroshetz, 1993]. Furthermore, it was suggested that the striatum might be more particularly involved in the advanced stages of the automation phase of learning a motor sequence [Doyon et al., 1997], and even in the very long-term retention (up to 1 year) of an overlearned visuomotor sequence [Doyon et al., 1998]. This hypothesis may be consistent with our observation that the striatum is more activated when subjects demonstrate successful performance, a finding that implies that some knowledge might have been already learned. The striatum would be particularly active when a previously acquired information has to be used for optimising performance. In contrast, the preserved performance in patients with frontal lobe lesions (in contrast with altered performance in patients with striatal damage) in the same Doyon et al. studies [1997, 1998] are intriguing and partly contradict our results. Indeed, such findings suggest that frontal regions are not as critical as the striatum in the implicit processing of a sequence of movements. However, the normal performance of frontal patients might equally reflect the wide functional heterogeneity of frontal cortex or a lack of statistical power.

Alternative, but not exclusive, hypotheses have also suggested that the striatum might be involved in the attentional mechanisms upon which sequence learning may depend [Cohen et al., 1990; Curran and Keele, 1993; Knopman and Nissen, 1991; Stadler, 1995; Will-ingham and Koroshetz, 1993], or that it is not really be involved in learning per se, but rather plays a major role for the correct execution of the motor programs that are required for sequential knowledge to enhance motor performance [Jackson et al., 1995; Knopman and Nissen, 1991]. However, the latter interpretation seems less likely because we show that the caudate is involved in the processing of higher-order conditional probabilities.

In a different perspective, modelling studies have also highlighted the role and the functioning of basal ganglia and related cortical structures in the processing of sequences. Based on the modular anatomic organisation of "parallel loops" linking frontal cortex, basal ganglia, and thalamus [Alexander et al., 1986], it was proposed that the loop through area 46 in the prefrontal cortex, caudate nucleus, internal segment of the globus pallidus, thalamus, and back to the prefrontal cortex, has an inherent capacity for encoding the serial order of events [Beiser and Houk, 1998]. In a serial task, the caudate nucleus would be able to detect contextual intrinsic signals such as working memory representations of previous stimuli, and forward this information in cortico-thalamic loops for sustaining the representation of these contextual events in working memory. Applying this hypothesis to our observation would suggest that striatal activity associated with the successful processing of L1 contexts is subsequent to the detection of behaviourally significant signals in preceding trials. The encoding of the L& contexts would take place in the working memory cortical-basal ganglia network.
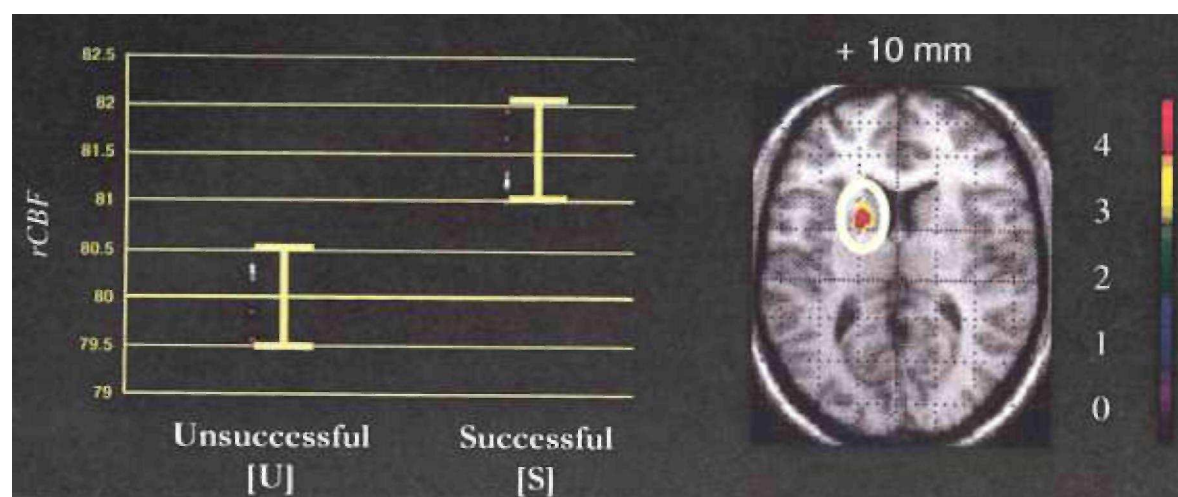

Figure 4.

Significant activation at the voxel-level ($p^{corr} < .05$) in the anterior striatum; coordinates -16 8 10 mm in the standard stereotaxic atlas of Talairach and Tournoux [1988]. Left panel show rCBF response pattern at this voxel in [S]uccesful or [U]nsuccessful conditions. Horizontal bars indicate minimal and maximal observed rCBF values across all participants.

TABLE II. *Random analysis results. Local and regional rCBF activations in [S]uccessful-related adjusted-mean conditions [S] compared to [Unsuccessful adjusted-mean conditions [U]**

Area of activation	Hemisphere L(eft) R(ight)	BA	Coordinates x	y	z	Voxel-level {Z}	Cluster-level {k, Z}	Set-level {c}
Caudate nucleus	L	NC	-16	8	10	4.86 (p^{corr}<.05)	446,4.86 ($p < .001$)	2 ($p < .005$)
Caudate nucleus	L	NC	-14	6	18	4.14		
Caudate nucleus	L	NC	-18	18	6	3.22		
Caudate nucleus	L	NC	-20	2	24	3.20		
Putamen		Pu	-18	8	2	4.33		
Putamen	L	Pu	-24	4	16	3.71		
Middle frontal gyrus	L	10	-34	42	4	4.30	287,4.30 ($p < .005$)	
Middle frontal gyrus	L	10	-28	36	4	3.36		
Inferior frontal gyrus	L	47	36	32	4	3.74		
Inferior frontal gyrus	L	45	56	20	18	3.72		
Inferior frontal gyrus		44	-50	12	12	3.62		
Inferior frontal gyrus		44/45	-40	18	12	3.38		
Inferior frontal gyrus		46	-42	26	12	3.30		

BA: Brodmann area; p^{corr} ; p-value corrected for multiple comparisons. Coordinates *x, y, z* refer to the standard Talairach and Tournoux [1988] stereotaxic space.
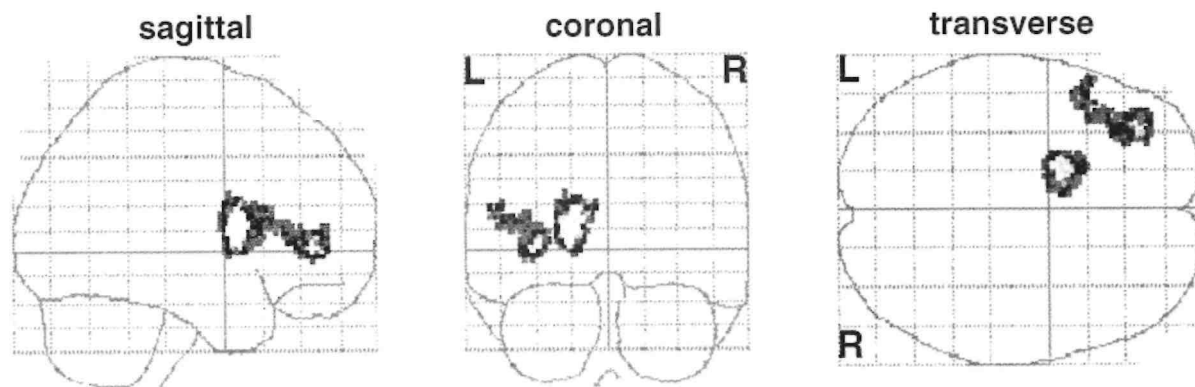

Figure 5.

Stereotaxic projection of the regional metabolic activity associated to [S]uccesful condition (cluster-level, p < .05) in the striatum (caudate nucleus and putamen) and prefrontal cortex (BA44/45/46).

Still another hypothesis to further explain how a complex sequence of actions could be implemented by the basal ganglia was proposed by Berns and Se-jnowski [1998]. They proposed that the basal ganglia

(1) project to cortical areas that implement actions, and

(2) filter multimodal information by selecting previously learned optimal actions based on the instantaneous cortical state. Similarly, a frontostriatal-based model of sensory-motor sequence learning was described, in which the prefrontal cortex is modelled as a recurrent network that encodes a sequence of internal states (e.g., information about previous events), and the caudate nucleus in the striatum is modelled as an associative memory structure that binds internal states to their corresponding motor outputs [Dominey, 1995]. Alternatively, Hikosaka et al. [1999] proposed that two main components of a sequential procedure could be learned throughout complementary cortico-subcortical loops. The spatial components of the sequence are learned via a loop circuit comprising the associations (prefrontal and parietal) cortices and the anterior basal ganglia, whereas the motor components depends on the motor cortices and the middle basal ganglia.

It was also proposed that the basal ganglia may chunk the representations of motor and cognitive action sequences so that they can be implemented as performance units [Graybiel, 1998], in line with Miller's notion [1956] of information chunking. Sequence encoding may occur both at neocortical and basal ganglia levels, but the basal ganglia will reorder cor-tically derived information to produce a more efficient action-oriented representation, which might reduce the number of possible, distinctively represented, units. In Graybiel's [1998] perspective, this view centred on habit or stimulus-response (S-R) learning is complementary to the models described above addressing other functions, such as response selection, and is interestingly compatible with a slow kinetic of learning without awareness. Slow learning must be necessary to avoid collapsing all temporally ordered acts into chunks: only relevant sequences must be selected, because once formed, one of the characteristics of a chunk is that it will be difficult to break apart. Absence of awareness is also an advantageous property for a chunking mechanism in that chunks can be treated as encapsulated units; intervention of consciousness might disrupt this implementation and force the system to consider elements of the sequence as a response chain of separate elements.

Summarising the common properties of various computational, methodological, and neuropsychological approaches, Cleeremans and co-workers convincingly argued that implicit learning may be construed as a complex form of priming whereby experience continuously shapes memory, and through which stored traces in turn continuously influence further processing, in the absence of awareness that this knowledge was acquired or that it is currently influencing processing [Cleeremans et al., 1998]. One of the central features of such processes is that learning involves elementary association or recoding processes that are highly sensitive to the statistical features of the training set. It seems that the neural mechanisms underlying those processes may take place, at least in part, in the basal ganglia, the role of which might be twofold in the present experiment. First, the basal ganglia may act so as to implicitly automate the serial information processed at the prefrontal (and possibly also at parietal or motor cortical sites [see Hikosaka, 1999]) level, slowly and gradually creating chunks of increasing length for the most frequent sequences or associations of stimuli, forwarded back in the cortical-subcortical loops. Therefore, the acquisition of knowledge of the L1 and L2 context will correspond partly to this gradual chunk formation (i.e., the elaboration of encapsulated representations of the particular contexts set by the previous element(s) at the current location). Second, these chunks will be helpful for response selection in that they reduce the number of plausible serial patterns at a given point, that is, they enhance their associated probability of occurrence. The basal ganglia may be active at this level in selecting the most appropriate response (i.e., the response with the highest probability of occurrence) in the context created by both the current stimulus introduced in the system and the chunks already processed and available for this computation, leading to faster response preparation for these particular chunks in this context. Therefore, the basal ganglia will become increasingly activated with the gradual acquisition of contextual knowledge because the material to process becomes increasingly complex. Accordingly, Boecker et al. [1998] report a complexity-correlated rCBF increase in the basal ganglia, and propose that it may play a specific role in the process of sequence facilitation and control, possibly by acting to filter and focus inputs from motor cortical areas as patterns of action becomes increasingly complex.

Nevertheless, it must be noted that implicit learning was demonstrated in Parkinson's disease patients using an artificial grammar learning task [Peigneux et al., 1999; Reber and Squire, 1999] that involves processing of the same kind of probabilistic relationships between elements than our SRT task, but for which there is no main motor component. This may suggest that the striatum might be more specifically involved in motor-linked cognitive learning, in line with the suggestion of Seger [1997, 1998].

CONCLUSIONS

In this study, we highlighted the importance of accounting for both between- and within-subject variability in the implicit processing of sequential knowledge during practice of an SRT task. We used a random-effect model analysis, where functional PET data may be summarised according to these sources of behavioural variability. Using this approach, we demonstrated that the striatum may be involved in more than simple pairwise associations, and has the capacity to process higher-order knowledge. Given the use of a probabilistic SRT task, the core component of this higher-order knowledge is thought to be concerned with statistical relationships between serially ordered elements. We suggest that the role of the striatum in sequence learning may be twofold. First, it is involved in the implicit automation of serial information through its participation in the cortical-subcortical motor loop linking prefrontal and caudate areas. Second, the striatum is particularly active for the selection of the most appropriate response in the context created by both the current stimulus and sequences of encapsulated previous stimuli, leading to higher efficiency and faster response preparation in the SRT task. However, further dedicated studies will be necessary to assess these hypotheses for striatum functions in implicit sequence learning.

ACKNOWLEDGMENTS

REFERENCES

Ackermann H, Daum I, Schugens MM, Grodd W (1996): Impaired procedural learning after damage to the left supplementary motor area (SMA). J Neurol Neurosurg Psychiat 60:94-97.

Alexander DN, DeLong MR, Strick PL (1986): Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci 9:357-381.

Baldwin KB, Kutas M (1997): An ERP analysis of implicit structured sequence learning. Psychophysiology 34:74-86.

Beiser DG, Houk JC (1998): Model of cortical-basal ganglionic processing: encoding the serial order of sensory events. J Neuro-physiol 79:3168-3188.

Berns GS, Sejnowski TJ (1998): A computational model of how the basal ganglia produce sequences. J Cognitive Neurosci 10:108-121.

Berns GS, Cohen JD, Mintun MA (1997): Brain regions responsive to novelty in the absence of awareness. Science 276:1272-1275.

Berry DC, Dienes Z (1993): Implicit learning: theoretical and empirical issues. Hove, England: Lawrence Erlbaum. 195 p.

Boecker H, Dagher A, CeballosBaumann AO, Passingham RE, Samuel M, Friston KJ, Poline JB, Dettmers C, Conrad B, Brooks DJ (1998): Role of the human rostral supplementary motor area and the basal ganglia in motor sequence control: Investigations with H-2 0-15 PET. J Neurophysiol 79:1070-1080.

Cleeremans A, Jiménez L (1998): Implicit sequence learning: the truth is in the details. In: Stadler MA, Frensch PA, editors. Handbook of implicit learning. Newbury Park, CA: Sage, p 323-364.

Cleeremans A, McClelland JL (1991): Learning the structure of event sequences. J Exp Psychol Gen 120:235-253.

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

Cohen A, Ivry R, Keele SW (1990): Attention and structure in sequence learning. J Exp Psychol Learn Mem Cog 16:17-30.

Curran T (1995): On the neural mechanisms of sequence learning. PSYCHE: an interdisciplinary journal of research on consciousness (on-line) 2:12.

Curran T (1998): Implicit sequence learning from a cognitive neuroscience perspective. What, how, and where? In: Stadler MA, Frensch PA, editors. Handbook of implicit learning. Newbury Park, CA: Sage, p 365-400.

Curran T, Keele SW (1993): Attentional and nonattentional forms of sequence learning. J Exp Psychol Learn Mem Cog 19:189-202.

Dominey PF (1995): Complex sensory-motor sequence learning based on recurrnt state representation and reinforcement learning. Biol Cybern 73:265-274.

Doyon J, Owen AM, Petrides M, Sziklas V, Evans AC (1996): Functional anatomy of visuomotor skill learning in human subjects examined with positron emission tomography. Eur J Neurosci 8:637-648.

Doyon J, Gaudreau D, Laforce R, Castonguay M, Bedard PJ, Bedard F, Bouchard JP (1997): Role of the striatum, cerebellum, and frontal lobes in the learning of a visuomotor sequence. Brain Cogn 34:218-245

Doyon J, Laforce R, Bouchard JP, Gaudreau D, Roy J, Poirier M, Bedard PJ, Bedard F (1998): Role of the striatum, cerebellum and frontal lobes in the automatization of a repeated visuomotor sequence of movements. Neuropsychologia 36:625-641.

Ferraro FR, Balota DA, Connor LT (1993): Implicit memory and the formation of new associations in nondemented Parkinson's disease individuals and individuals with senile dementia of the Alzheimer type: A serial reaction time (SRT) investigation. Brain Cogn 21:163-180.

Frison L, Pocock SJ (1992): Repeated measures in clinical trials: an analysis using mean summary statistics and its implications for design. Stat Med 11:1685-1704.

Friston K, Ashburner J, Frith C, Poline JB, Heather J, Frackowiak R (1995): Spatial realignment and normalization of images. Hum Brain Mapp 2:165-189.

Gold S, Arndt S, Johnson D, O'Leary DS, Andreasen NC (1997): Factors that influence effect size in 150 PET studies: a meta-analytic review. Neuroimage 5:280-291.

Gomez-Beldarrain M, Garcia-Monco JC, Rubio B, Pascual-Leone A (1998) Effect of focal cerebellar lesions on procedural learning inthe serial reaction time task. Exp Brain Res 120:25-30.

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

Graybiel AM (1998): The basal ganglia and chunking of action repertoires. Neurobiol Learn Mem 70:119-136.

Hazeltine E, Grafton ST, Ivry R (1997): Attention and stimulus characteristics determine the locus of motor-sequence encoding—A PET study. Brain 120:123-140.

Hikosaka O, Nakahara H, Rand MK, Sakai K, Lu X, Nakamura K, Miyachi S, Doya K (1999): Parallel neural networks for learning sequential procedures. Trends Neurosci 22:464-471.

Holmes A, Friston K (1998): Generalisability, random effects and population inference. Neuroimage 7:754.

Honda M, Deiber MP, Ibanez V, Pascual-Leone A, Zhuang P, Hallett M (1998): Dynamic cortical involvement in implicit and explicit motor sequence learning—A PET study. Brain 121:2159-2173.

Jackson GM, Jackson SR, Harrison J, Henderson L, Kennard C (1995): Serial reaction time learning and Parkinson's disease: evidence for a procedural learning deficit. Neuropsychologia 33:577-593.

Jiménez L, Mendez C, Cleeremans A (1996): Comparing direct and indirect measures of sequence learning. J Exp Psychol Learn Mem Cog 22:948-969.

Jueptner M, Weiller C (1998): A review of differences between basal ganglia and cerebellar control of movements as revealed by functional imaging studies. Brain 121:1437-1449.

Knopman D, Nissen MJ (1987): Implicit learning in patients with probable Alzheimer's disease. Neurology 37:784-788.

Knopman D, Nissen MJ (1991): Procedural learning is impaired in Huntington's disease: evidence from the serial reaction time task. Neuropsychologia 29:245-254.

Merikle PM, Reingold EM (1991): Comparing direct (explicit) and indirect (implicit) measures to study unconscious memory. J Exp Psychol Learn Mem Cog 17:224-233.

Meulemans T (1998a): L'apprentissage implicite. Marseille: Solal. 227 p.

Miller GA (1956): The magic number seven, plus or minus two: some limits on our capacity for processing information. Psychol Rev 63:81-97.

Molinari M, Leggio MG, Solida A, Ciorra R, Misciagna S, Silveri MC, Petrosini L (1997): Cerebellum and procedural learning: evidence from focal cerebellar lesions. Brain 120:1753-1762.

Pascual-Leone A, Grafman J, Clark K, Stewart M (1993): Procedural learning in Parkinson's disease and cerebellar degeneration. Ann Neurol 34:594-602.

Peigneux P, Meulemans T, Van der Linden M, Salmon E, Petit H (1999): Exploration of implicit artificial grammar learning in Parkinson's disease. Acta Neurol Belg 99:107-117.

Poline JB, Worsley KJ, Evans AC, Friston KJ (1997): Combining spatial extent and peak intensity to test for activations in functional imaging. Neuroimage 5:83-96.

Rauch SL, Savage CR, Brown HD, Curran T, Alpert NM, Kendrick A, Fischman AJ, Kosslyn SM (1995): A PET investigation of implicit and explicit sequence learning. Hum Brain Mapp 3:271-286.

Rauch SL, Whalen PJ, Savage CR, Curran T, Kendrick A, Brown HD, Bush G, Breiter HC, Rosen BR (1997): Striatal recruitment during an implicit sequence learning task as measured by functional magnetic resonance imaging. Hum Brain Mapp 5:124-132.

Rauch SL, Whalen PJ, Curran T, McInerney S, Heckers S, Savage CR (1998): Thalamic deactivation during early implicit sequence learning: a functional MRI study. Neuroreport 9:865-870.

Reber AS (1993): Implicit learning and tacit knowledge: an essay on the cognitive unconscious. New York: Oxford University Press. 188 p.

Reber PJ, Squire LR (1999): Intact learning of artificial grammars and intact category learning by patients with Parkinson's disease. Behav Neurosci 113:235-242.

Reingold EM, Merikle PM (1988): Using direct and indirect mesures to study perception without awareness. Percept Psychoph 44: 563-575.

Seger CA (1994): Implicit learning. Psychol Bull 115:163-196.

Seger CA (1997): Two forms of sequential implicit learning. Conscious Cogn 6:108-131.

Seger CA (1998): Independent judgment-linked and motor-linked forms of artificial grammar learning. Conscious Cogn 7:259-284.

Shanks DR, Johnstone T (1998): Implicit knowledge in sequential learning tasks. In: Stadler MA, Frensch PA, editors. Handbook of implicit learning. London: Sage Publications, p 533-572.

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

Squire LR (1992): Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. Special Issue: Memory systems. J Cognitive Neurosci 4:232-243.

Stadler MA (1995): Role of attention in implicit learning. J Exp Psychol Learn Mem Cog 21:674-685.

Stadler MA, Frensch PA (1998): Handbook of implicit learning. London: Sage Publications. 648 p.

Talairach J, Tournoux P (1988): Co-planar stereotaxic atlas of the human brain. New York: Thieme. 122 p.

Westwater H, McDowall J, Siegert R, Mossman S, Abernethy D (1998): Implicit learning in Parkinson's disease: evidence from a verbal version of the serial reaction time task. J Clin Exp Neu-ropsychol 20:413-418.

Willingham DB, Koroshetz WJ (1993): Evidence for dissociable motor skills in Huntington's disease patients. Psychobiology 21:173-182.