doi:10.1093/brain/awr147 Brain 2011: 134; 2287–2301 | 2287

BRAIN
A JOURNAL OF NEUROLOGY

Dopamine-dependent reinforcement of motor skill learning: evidence from Gilles de la Tourette syndrome

Stefano Palminteri,[1,2] Maël Lebreton,[1,2] Yulia Worbe,[2,3] Andreas Hartmann,[2,3,4] Stéphane Lehéricy,[2] Marie Vidailhet,[2,3] David Grabli[2,3] and Mathias Pessiglione[1,2]

1 Team Motivation, Brain and Behaviour (MBB), Institut du Cerveau et de la Moelle épinière (ICM), F-75013 Paris, France
2 Université Pierre et Marie Curie (UPMC- Paris 6), CNRS UMR 7225, INSERM UMRS 975, F-75013 Paris, France
3 Centre d'Investigation Clinique, INSERM CIC 9503, F-75013 Paris, France
4 Centre de Référence 'Syndrome Gilles de la Tourette', Pôle des Maladies du Système Nerveux, Assistance Publique - Hôpitaux de Paris (AP-HP), F-75013 Paris, France

Correspondence to: Mathias Pessiglione,
Institut du Cerveau et de la Moelle épinière,
Groupe Hospitalier Pitié-Salpêtrière,
47 Boulevard de l'Hôpital,
75013 Paris,
France
E-mail: mathias.pessiglione@gmail.com

Reinforcement learning theory has been extensively used to understand the neural underpinnings of instrumental behaviour. A central assumption surrounds dopamine signalling reward prediction errors, so as to update action values and ensure better choices in the future. However, educators may share the intuitive idea that reinforcements not only affect choices but also motor skills such as typing. Here, we employed a novel paradigm to demonstrate that monetary rewards can improve motor skill learning in humans. Indeed, healthy participants progressively got faster in executing sequences of key presses that were repeatedly rewarded with 10 euro compared with 1 cent. Control tests revealed that the effect of reinforcement on motor skill learning was independent of subjects being aware of sequence-reward associations. To account for this implicit effect, we developed an actor-critic model, in which reward prediction errors are used by the critic to update state values and by the actor to facilitate action execution. To assess the role of dopamine in such computations, we applied the same paradigm in patients with Gilles de la Tourette syndrome, who were either unmedicated or treated with neuroleptics. We also included patients with focal dystonia, as an example of hyperkinetic motor disorder unrelated to dopamine. Model fit showed the following dissociation: while motor skills were affected in all patient groups, reinforcement learning was selectively enhanced in unmedicated patients with Gilles de la Tourette syndrome and impaired by neuroleptics. These results support the hypothesis that overactive dopamine transmission leads to excessive reinforcement of motor sequences, which might explain the formation of tics in Gilles de la Tourette syndrome.

Keywords: motor skill learning; reinforcement learning; dopamine; computational modelling; Gilles de la Tourette syndrome; dystonia

Received December 16, 2010. Revised April 26, 2011. Accepted May 1, 2011. Advance Access publication July 3, 2011

Introduction

Motor skills refer to the effortless execution of movement sequences, such as lacing a shoe or playing the piano, with optimal speed and accuracy. They are mostly acquired with practice, the main factor for motor skill improvement being the number of repetitions. However, parents and educators may share the intuition that positive reinforcements, such as congratulations, can also contribute to motor skill acquisition. To our knowledge, the influence of reinforcements on motor skill learning has never been specifically investigated in humans. This notion must be distinguished from two close phenomena, namely instrumental learning and incentive motivation. In instrumental learning, several actions are possible and rewards increase the frequency of choosing a particular action. Here, we wish to examine whether rewards improve the execution of a particular action that is imposed on subjects (with no choice). In incentive motivation, potential rewards are announced before motor performance to augment the desire to do well and adjust the effort accordingly. Here, we wish to study the impact of rewards that are incidentally obtained after performance completion. Both instrumental learning and incentive motivation processes have been extensively studied in a variety of species. In particular, incentive motivation has recently been found to improve procedural learning in humans (Wachter *et al*., 2009). The present study primarily aimed to extend this finding by investigating reinforcement effects on motor skill performance.

To this aim, we developed a motor learning task based on reaction time paradigms (Rhodes *et al*., 2004) with the addition of monetary reinforcements. The task involves subjects learning through repetition to quickly execute a series of key presses that are visually instructed on a computer screen. To assess reinforcement effects, the key manipulation was to associate one-half of the sequences with a negligible reward (1 cent) and the other half with a consequent reward (10 euros). In order to dissociate reinforcement from motivation effects, two groups of healthy subjects were assessed on different versions of the motor learning task. In the implicit version, rewards were displayed after movement execution, so as to reinforce visuo-motor representations. In the explicit version, the reward at stake was disclosed before motor execution, in order to modulate subjects' motivation to do well. Sequence-reinforcement associations were never mentioned to subjects in either group. To assess their knowledge of these associations, we asked them to make forced choices after completion of the motor learning task.

Beyond these psychological considerations, we wished to examine the underlying neural basis, more specifically the implication of neuromodulatory processes that are relevant to neurological diseases and pharmacological treatments. Patient and neuroimaging studies in humans suggest the implication of several cortical and subcortical brain structures in motor skill learning, such as the supplementary motor area, the pre-supplementary motor area, the basal ganglia and the cerebellum (Willingham, 1998; Ashe *et al*., 2006; Doyon *et al*., 2009). More precisely, the implication of the basal ganglia in motor skill learning has been supported by neuropsychological studies in patients with Parkinson's disease, Huntington's disease and dystonia (Agostino *et al*., 1992; Doyon, 2008). Within the basal ganglia, neuronal mechanisms underlying motor skill learning appear to involve dopamine modulating synaptic plasticity in the striatum (Kumari *et al*., 1997; Badgaiyan *et al*., 2007; Garraux *et al*., 2007; Karabanov *et al*., 2010). However, dopamine implication in motor learning is far from established, as several studies reported no effect or even detrimental effect of dopamine replacement therapy in patients with Parkinson's disease (Argyelan *et al*., 2008; Carbon *et al*., 2006; Kwak *et al*., 2010). Indeed, some authors have implicated dopamine in reinforcement rather than motor learning *per se* (Berns and Sejnowski, 1998; Suri and Schultz, 1998; Doya, 2000; Hikosaka *et al*., 2002).

In parallel to the motor learning literature, a growing body of evidence from electrophysiological studies in animals suggests that midbrain dopamine activity reflects reward-related (and not movement-related) signals. More precisely, phasic dopamine activity seems to encode reward prediction errors, i.e. discrepancies between obtained and expected rewards (Schultz *et al*., 1997; Bayer and Glimcher, 2005; Zaghloul *et al*., 2009). Reward prediction errors have been used as teaching signals in various formal accounts of reinforcement learning (Sutton and Barto, 1998). Dopamine-mediated prediction errors have also been implemented in formal models of frontal cortex–basal ganglia circuitry (Houk *et al*., 1995; Bar-Gad and Bergman, 2001; Frank *et al*., 2006). Generally, the idea is that dopamine release strengthens frontostriatal synaptic connections involved in generating actions that lead to more reward than expected, so that these actions are more easily selected in the future. Although most of these models aim at explaining improvement of action selection in instrumental learning, a side implication is that motor execution itself should be improved by dopamine-mediated reinforcement signals. This is because reinforcement processes should make frontal cortex-basal ganglia circuits underlying learned actions more strongly and more specifically activated when the same context repeats itself. A subclass of frontal cortex-basal ganglia models, termed 'actor-critic', dissociate dopamine functions in reinforcement versus motor learning (Joel *et al*., 2002). Dopamine-mediated reward prediction errors would serve to improve both reward prediction in the critic and movement execution in the actor. A further aim of this study was to use the actor-critic framework to allow insight into pathological conditions involving dopamine dysfunction.

To test the implication of dopamine in the reinforcement of motor skills, we chose to include patients with Gilles de la Tourette syndrome. This is a childhood-onset hyperkinetic disorder, characterized by motor and vocal tics, that has been related to anatomical and functional peculiarities in the basal ganglia and to constitutive hyperactivity of dopaminergic transmission (Graybiel and Canales, 2001; Albin and Mink, 2006). Symptoms of Gilles de la Tourette syndrome are markedly improved by antipsychotics (or neuroleptics, i.e. dopamine receptor antagonists), which remain the drugs of choice for treatment (Leckman, 2002; Singer, 2005). Dopaminergic hyperactivity in both the frontal cortex and basal ganglia has been confirmed by some, but not all, PET studies using ligand binding to dopamine receptors and transporters (Malison *et al*., 1995; Gilbert *et al*., 2006; Wong

et al., 2008). Post-mortem studies have also found increased concentrations of dopamine receptors and transporters in the frontal cortex and basal ganglia of patients with Gilles de la Tourette syndrome (Minzer *et al*., 2004; Yoon *et al*., 2007). We therefore hypothesized that unmedicated patients with Gilles de la Tourette syndrome exhibit stronger reinforcement effects on motor skill learning compared with healthy controls and that neuroleptics abolish or even reverse this difference. To test these predictions, we compared the performance of two matched groups of patients with Gilles de la Tourette syndrome (unmedicated and medicated) in the same implicit motor learning task as healthy controls. For further control, we included a group of patients with levodopa non-responsive focal dystonia (torticollis), which can be considered as an example of motor impairment without dopamine dysfunction (Mink, 2003; Jankovic, 2006; Vidailhet *et al*., 2009).

Subjects and Methods

Subjects

This study was approved by the Ethical Committee of the Pitié-Salpêtrière Hospital. All subjects gave informed consent prior to participation. A total of 123 participants were recruited, including patients and healthy subjects. General exclusion criteria for this study were left handedness and, because the task involved hand motor skills (see below), being a professional pianist or typist. For healthy subjects, we also excluded any history of neurological or psychiatric illness and use of psychotropic drug or medication. A first subgroup of 53 healthy subjects constituted the control group for the 52 patients and performed the main (implicit) motor learning task. A second subgroup of 18 healthy subjects performed a control (explicit) variant of the motor learning task. These two groups of healthy subjects ($n = 53$ and $n = 18$) were not different in age (mean age ± SEM: 23.8 ± 0.5 versus 23.2 ± 0.7 years, $P > 0.5$, *t*-test) or gender (per cent males: 45.3 versus 50.0%, $P > 0.1$, χ^2 test).

We also recruited 35 patients with Gilles de la Tourette syndrome, who were split into two groups according to their current treatment at the time of testing. The first group (unmedicated Gilles de la Tourette syndrome, $n = 20$) included patients who were not taking antipsychotics (i.e. neuroleptics) that would interfere with dopaminergic transmission. However, some patients in the unmedicated Gilles de la Tourette syndrome group were treated with non-dopaminergic medications such as venlafaxine ($n = 2$), alprazolam ($n = 2$), clonazepam ($n = 1$), propanolol ($n = 1$) and insulin ($n = 1$). The second group (medicated Gilles de la Tourette syndrome, $n = 15$) included patients who were taking neuroleptics, either in monotherapy or in association with other medications. More specifically, these patients were treated with aripiprazole ($n = 7$), risperidone ($n = 4$), pimozide ($n = 2$), haloperidol ($n = 1$), cyamemazine ($n = 1$), tetrabenazine ($n = 1$), tiapride ($n = 1$) and olanzapine ($n = 1$). Some of the medicated patients with Gilles de la Tourette syndrome were also taking medications other than neuroleptics, such as venlafaxine ($n = 2$), valium ($n = 1$), lithium carbonate ($n = 1$), topiramate ($n = 1$) and duloxetine ($n = 1$). The unmedicated and medicated Gilles de la Tourette syndrome groups were matched for age (33.0 ± 2.5 versus 29.9 ± 2.9 years, $P > 0.1$, *t*-test), gender (75.0 versus 73.3% males, $P > 0.5$, χ^2 test) and tic severity as measured with the Yale Global Tic Severity Scale (YGTSS/100 score: 24.0 ± 2.9 versus 30.3 ± 3.3, $P > 0.1$, *t*-test). Because of the preponderant co-morbidity between Gilles de la Tourette syndrome and obsessive compulsive disorder, patients were also assessed with the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS). We found eight patients (four in each Gilles de la Tourette syndrome group) presenting obsessive compulsive symptoms according to the Y-BOCS, but no significant between-group difference in Y-BOCS scores (2.4 ± 1.4 versus 4.5 ± 2.0, $P > 0.1$, *t*-test). Patients with Gilles de la Tourette syndrome did not exhibit any other concomitant psychiatric disease, assessed with the Mini International Neuropsychiatric Interview (MINI, French version) (Sheehan *et al*., 1998).

Finally, we recruited 17 patients with idiopathic dystonia, precisely spasmodic torticollis (dystonia group: $n = 17$; age = 52.8 ± 3.2 years; gender: 23.5% male). One patient also exhibited a blepharospasm and another presented a shaking form of dystonia. All dystonic patients were checked for non-response to levodopa and were treated with botulinum toxin at the time of the study. More precisely, they were all tested 3 months after the last toxin injection and immediately before a scheduled new injection in the next hour. Some dystonic patients were treated with other medications than botulinum toxin at the time of testing, but no patient was taking any neuroleptic medication. Other medications included: trihexyphenidyl ($n = 2$), lorazepam ($n = 2$), losartan ($n = 1$), levothyroxine ($n = 1$), tramadol ($n = 1$), mianserin ($n = 1$), ropinirole ($n = 1$), clonazepam ($n = 1$), escitalopram ($n = 1$), levocetirizine ($n = 1$), pentoxifylline ($n = 1$). Table 1 provides a summary of the subjects' demographic data.

Behavioural tasks

All subjects performed a motor learning task followed by a forced choice task. During the motor learning task (Fig. 1A), subjects had to maintain the five fingers of their right hand on five keys forming an inverse V shape at the centre of the computer azerty keyboard (C-F-T-H-N). In every trial they had to press three keys out of the five possible as fast as they could. The three keys were to be pressed one after the other, from the leftmost to the rightmost. Each trial started with a 1000 ms fixation screen in which five crosses were displayed, corresponding to the five fingers. The fixation screen was followed by the instruction and response screen, in which the triplet of keys to press was circled in red. This step was self-paced; its duration depended on the time subjects took to achieve the sequence of key presses. Once the subject had pressed the three keys, the

Table 1 Demographic data

Demographic features	CON-1 ($n = 53$)	CON-2 ($n = 18$)	U-GTS ($n = 20$)	M-GTS ($n = 15$)	DYS ($n = 17$)
Age (years)	23.8 ± 0.5	23.2 ± 0.7	33.0 ± 2.5*	29.9 ± 2.9*	52.8 ± 3.2*
Sex (% of males)	45.3	50.0	75.0*	73.3*	23.5*

*$P < 0.05$, compared to group CON-1, using two-tailed *t*-test for age and χ^2 test for sex.
CON-1 and CON-2: healthy controls who performed the implicit and explicit motor learning task, respectively. U-GTS and M-GTS = unmedicated and medicated Gilles de la Tourette syndrome patients, respectively; DYS = dystonic patients.

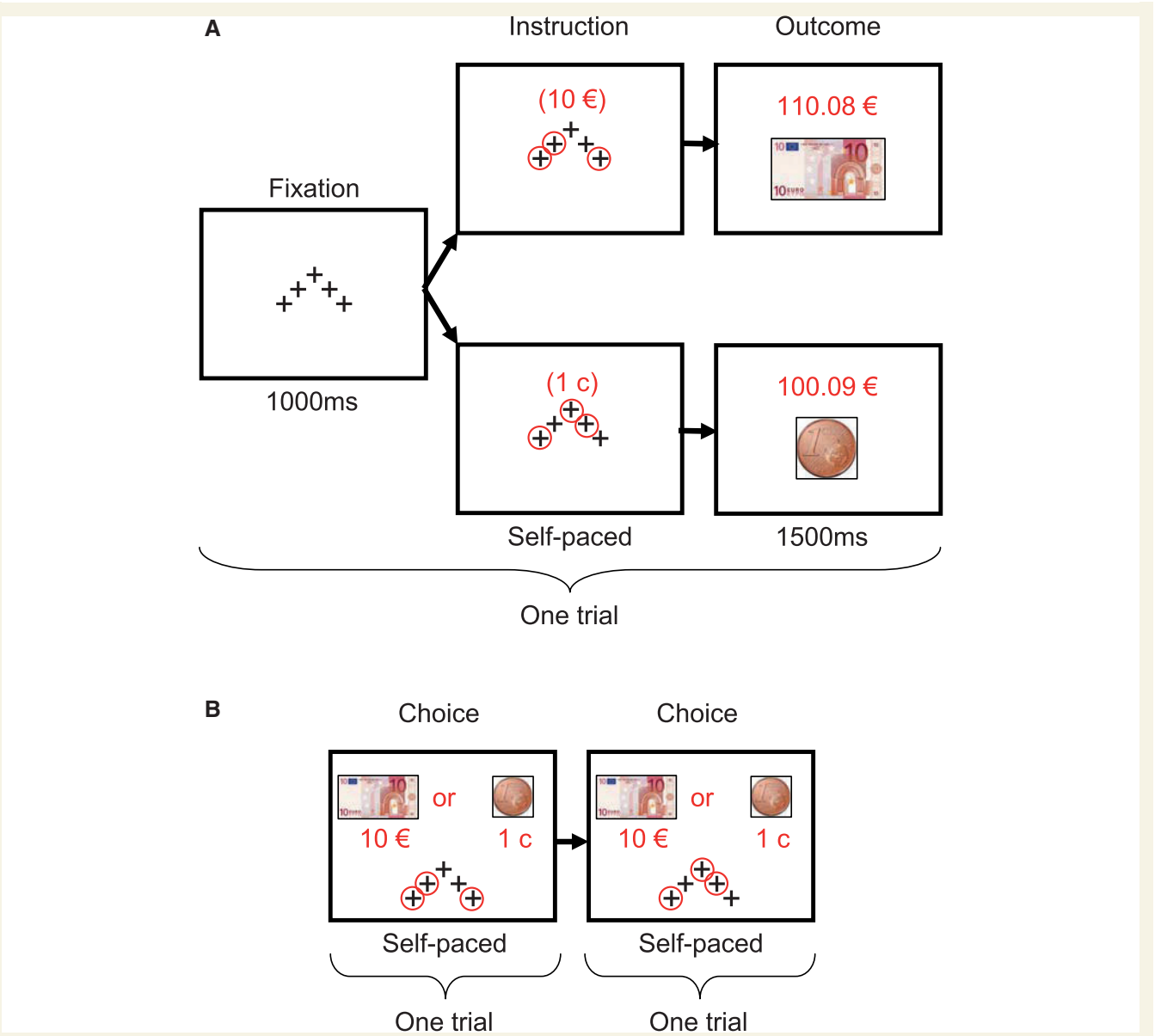

Figure 1 Behavioural tasks. (**A**) Successive screens displayed during a trial of the motor learning task. Subjects saw a triplet of crosses circled in red, pressed the corresponding keys and subsequently observed the outcome (10 euro or 1 cent). The potential outcome was announced on the instruction screen in the explicit, but not in the implicit, version of the task. (**B**) Two trials of the forced choice task. Subjects saw a triplet and guessed with which outcome (10 euro or 1 cent) it was associated during the motor learning task.

outcome was displayed on the last screen for 1500 ms. The outcome could either be a significant (10 euro note) or a negligible reinforcement (1 cent coin). It was displayed together with the cumulative amount of money gained in euros. If subjects pressed a wrong key the outcome was nil and the word 'error' appeared on the screen. In addition to the visual illustration, the three outcomes (10 euros, 1 cent or error) were also accompanied by evocative sounds: white noise for errors, coin jingling for 1 cent gains and cash registering for 10 euro gains.

Among the 10 possible motor sequences (triplets of key presses) to be executed, five were associated with the 10 euro outcome, and the other five with the 1 cent outcome. Associations between triplets and outcomes were randomly assigned for each subject and then fixed for the entire experiment. The task comprised 15 blocks of 10 trials, within which all 10 triplets were presented once, such that each triplet appeared 15 times in total. The presentation order of the 10 triplets was fully randomized within each block. Task performance lasted between 10 and 15 min, depending on subjects' speed. In the main version of the task, referred to as the 'implicit task', subjects were not instructed about the outcome associated with the current triplet. The first subgroup of healthy subjects ($n = 53$) and all patients ($n = 52$) performed this version of the task. In a control version of the task, referred to as the 'explicit task', the outcome was indicated on the instruction screen, together with the keys to press. However, even in this explicit task, subjects were not informed that the associations between triplets and outcomes were deterministic. This version of the task was only administered to the second subgroup of healthy subjects ($n = 18$). In both task versions, subjects were instructed to maximize their final outcome. As a motivation to speed up, they were not told that the number of trials was predetermined but instead

that the task duration would be fixed (to 15 min). Therefore, they were induced to believe that going faster would allow them to perform more trials and consequently to win more money.

After the motor learning task (implicit or explicit), subjects performed a forced choice task (Fig. 1B) in which they had to report, for each triplet, the associated outcome. In each trial, subjects were presented with one triplet and asked to choose between the two possible monetary outcomes (1 cent and 10 euros). The 10 triplets were presented only once, in a randomized order. Instructions for all tasks were first provided in a written form (Supplementary material) and then reformulated orally, if necessary. The tasks were programmed on a PC using the Cogent 2000 (Wellcome Trust Centre for Neuroimaging) utility running on MatLab® R2007b (MathWorks®).

Learning curves

For both motor learning tasks (implicit and explicit), the main dependent measure of interest was response time, defined as the interval between the instruction onset and the third (last) key press. We considered the three key presses as a single atomic action (or motor skill) because preliminary analyses showed similar results for the times of each key press taken separately. Response time learning curves were generated separately for the 1 cent and 10 euro conditions. For each 10-trial block, we extracted the median response time over the five triplets of each condition (1 cent and 10 euros). Thus, learning curves contained 15 data points for each condition, corresponding to the 15 blocks of trials. Response times were discarded when an error occurred as well as when more than three standard deviations above the mean for the current block. Error rates were also analysed as an index of accuracy. Error rate learning curves were generated in a similar manner to response time learning curves, by averaging data for each 10-trial block, separately for the 1 cent and 10 euro conditions. The reinforcement effect was defined as the cumulative difference between the 10 euro and 1 cent conditions (noted $\Delta T_{1c\text{-}10€}$ in figures). We considered the reinforcement magnitude as negligible in the 1 cent condition, being a thousand times smaller than in the 10 euro condition. Thus, we approximated the 1 cent condition as representing pure motor skill learning (i.e. with no reinforcement effect).

Computational model

We aimed to formalize reinforcement effects on motor performance using the framework of reinforcement learning theory. Here, we proposed a two-stage model to explain the behavioural data obtained in the implicit motor learning task (Fig. 2A). The first stage of our computational model was implemented to account for motor skill learning itself as observed in the 1 cent condition, which we approximated as a no reinforcement condition. This first model had to account for the specific dynamics of response time reduction across trials, which was characterized by a rapid decrease within the first five blocks of trials, followed by a slighter decrease for the remaining blocks that eventually reached a plateau. Historically, response time tasks have been modelled with the so-called power law of practice function. However, it has recently been suggested that an exponential function may provide a better fit of the individual learning curves (Heathcote and Brown, 2000). Here, we confirmed this observation when comparing power and exponential functions (Supplementary material). Therefore, we opted for an exponential decay function, $T_{1c}(n) = b \times \exp^{-(a \times n)} + c$, according to which response time in trial n of the 1 cent condition depends on three parameters: a, b and c. Parameter a can be seen as the motor learning rate, since it adjusts the extent to which response times are reduced from one trial to the next. Parameter c represents the plateau to which learning curves would converge after an infinite number of trials ($n = \infty$). Parameter b represents the difference between the initial performance and the plateau, since $T_{1c}(0) = b + c$, i.e. the magnitude of response time reduction across learning. We note that this model is purely phenomenological, meaning that it has been conceived to reproduce the observed motor learning dynamics, but not to formalize any actual neuronal mechanism. The three free parameters were optimized to fit the data on an individual basis using least square distance minimization. Precisely, we minimized for each subject the square distance between the model estimate and the learning curve observed for the 1 cent triplets. The ranges and steps used for exploring parameter spaces in the optimization process were the following: a = [0:0.1:1], b = [−10000:100:10000] ms and c = [500:100:5000] ms.

The second stage of our computational model has been implemented to account for reinforcement effects on motor skill learning, as observed in the 10 euro condition. We first modelled how subjects learned the predictive value of 10 euro triplets ($V_{10€}$). Here, we opted for a learning rule adapted from Rescorla and Wagner (1972), according to which the value of 10 euro triplets is updated as follows: $V_{10€}(n + 1) = V_{10€}(n) + \alpha \times PE(n)$, where n is the trial number, α the reinforcement learning rate and PE the prediction error, defined as $PE(n) = R - V_{10€}(n) + \gamma$, where R is the reinforcement magnitude (fixed at 10 euros) and γ a free parameter meant to adjust the size of the prediction error. To explain reinforcement effects on response times, we elaborated a second learning rule, according to which prediction errors dictate response time reduction due to reinforcement: $\Delta T_{1c\text{-}10€}(n + 1) = \Delta T_{1c\text{-}10€}(n) + \beta \times PE(n)$, where β is a reinforcement-to-time scaling parameter. Contrary to the first-stage model used to fit the 1 cent condition, this second-stage mechanism might be implemented in neuronal architectures. In particular, prediction error has been found to be reflected by the phasic bursting of midbrain dopaminergic neurons, in both human and non-human primates (Schultz *et al*., 1997; Zaghloul *et al*., 2009). Our second delta rule may thus approximate phasic dopamine release strengthening corticostriatal synapses and subsequently facilitating the activation of rewarded motor responses, as shown with neural network simulation (Frank *et al*., 2006). One possible architecture is an actor-critic model (Fig. 2B), in which the critic implements the first learning rule to predict reinforcement value, while the actor implements the second learning rule to reduce response time (Maia *et al*., 2010). Growing evidence suggests that actor and critic modules may be represented in the dorsal and ventral striatum, respectively (Doya, 2008). Each module would use the same teaching signal (prediction error), with different learning rates: α for the critic, β for the actor. Given the hypothesis that dopamine encodes prediction error, our prediction was that variations in dopaminergic activity would be captured by parameter γ.

For both learning rules, initial values, i.e. $V_{10€}(1)$ and $\Delta T_{1c\text{-}10€}(1)$, were set to 0. Response times observed in the 10 euro condition were obtained by subtracting reinforcement effects to motor learning: $T_{10€}(n) = T_{1c}(n) - \Delta T_{1c\text{-}10€}(n)$. Again, we optimized for each individual the free parameters (α, β and γ) using the least square method. The only difference with the optimization procedure used for the first-stage model is that for the second-stage model we generated five different learning curves, corresponding to the five different 10 euro triplets, which we then averaged and compared with the observed learning curve. This has been done to take into account the personal history of success and error trials (for which the prediction error term was zero). The ranges and steps of parameter values explored in the optimization process were the following: α = [0:0.1:1], γ = [−20:1:0]€ and β = [0:1:10] ms€$^{-1}$.

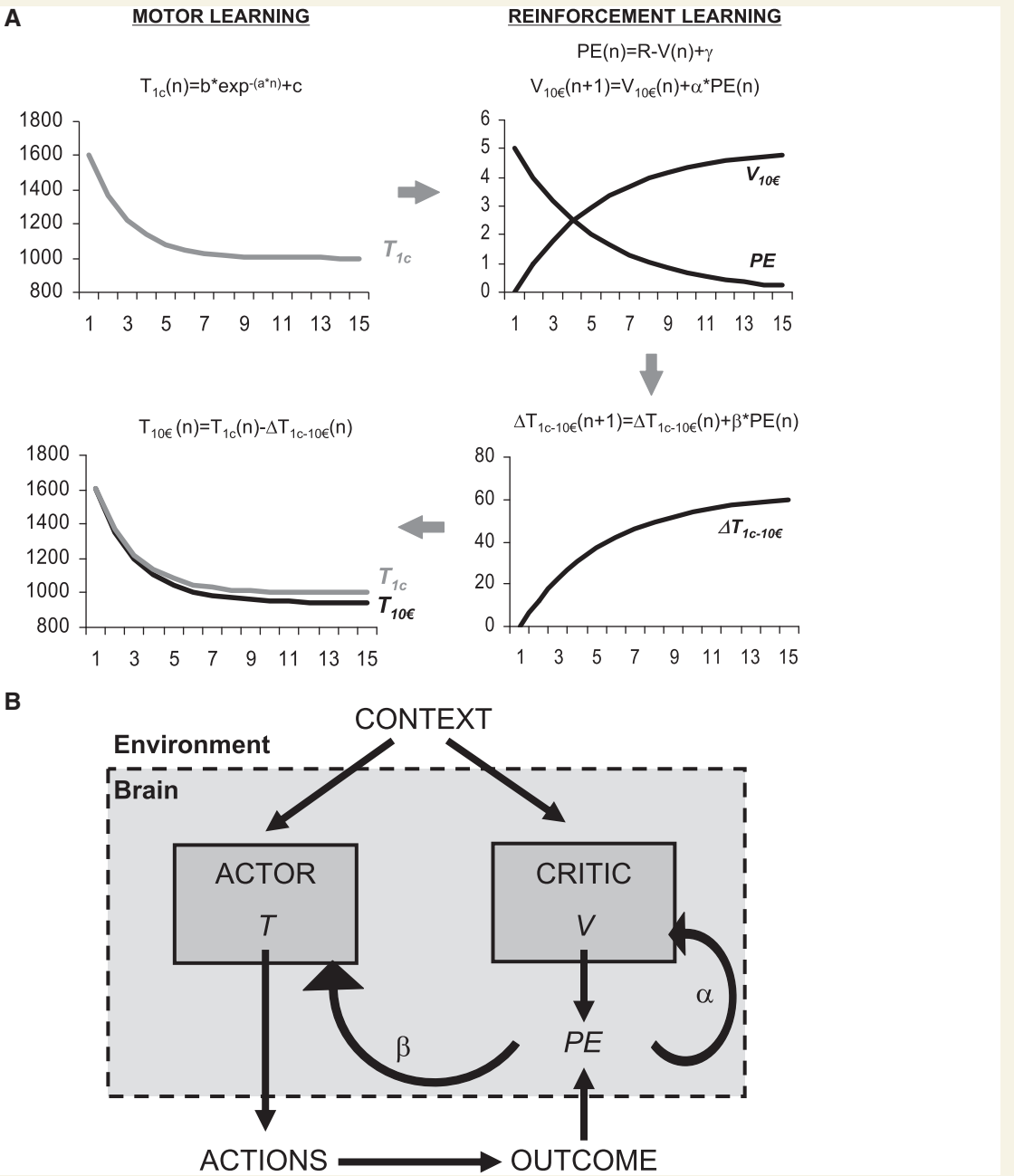

Figure 2 Computational model. (**A**) Decomposition of computational stages. The motor learning model accounted for performance in the 1 cent condition (*left*), while the reinforcement learning model accounted for the addition of reinforcement effects in the 10 euro condition (*right*). The arrows indicates the transitions between the different stages of the computational modelling. (**B**) Schematic illustration of the computational model. The motor learning model is implemented in the actor component and the reinforcement learning model in the critic component. Arrows represent causal influence. See explanations in main text.

Statistical analyses

First (model-free) analyses focused on comparing the behavioural performance of healthy subjects between the implicit and explicit tasks, in order to assess the importance of knowing the reinforcement at stake prior to pressing keys. An ANOVA with task as a between-subject factor and block as a within-subject factor was assessed on our four main dependent variables: response time (1 cent condition), error rate (1 cent condition), reinforcement effect (10 euro-1 cent condition) on response time and reinforcement effect on error rate. *Post hoc* comparisons between tasks (explicit versus implicit) were performed on the average dependent measures using two-tailed two-sample *t*-tests. We used one-tailed paired *t*-tests to assess for each group whether the reinforcement effect was significantly different from zero, on average and every three blocks across learning. Behavioural performance in the forced choice task,

expressed as the percentage of correct responses, was also compared with chance level guessing (50% of correct responses) using one-tailed paired *t*-test and between tasks (implicit versus explicit) using two-tailed *t*-test. A linear regression fit was used to assess the correlation between correct guessing and response time reinforcement effect, in order to test whether or not reinforcement effects were linked to explicit knowledge of triplet-outcome contingencies. We then compared behavioural performance between groups of subjects who performed the implicit motor learning task. An ANOVA with group (healthy controls, medicated and unmedicated Gilles de la Tourette syndrome and dystonic patients) as a between-subject factor and block as a within-subject factor was assessed on the same main dependent measures as above. Two-tailed *t*-tests were also employed for *post hoc* between-group comparisons on the average measures.

A second series of (model-based) analyses was aimed to characterize the effects of disorders and treatments by comparing the behavioural performance in the implicit learning task between the different groups (control, medicated Gilles de la Tourette syndrome, unmedicated Gilles de la Tourette syndrome and dystonia). Statistical analyses concentrated on model free parameters obtained for both motor learning (*a*, *b* and *c*) and reinforcement learning (α, β and γ) in addition to response time, error rate and reinforcement effect. Significance was assessed using an ANOVA with group as the main factor in interest. Since we noted certain gender differences between the groups (Table 1) in a preliminary analysis, we addressed this potential confound using two-way ANOVA with gender as an additional factor. Because gender did not significantly affect any behavioural or computational variable, we will not discuss this factor further. We also noted age differences between groups (Table 1). To eliminate this potential confound, we performed an ANCOVA with age as a covariate. Statistical results of between group comparisons reported hereafter refer to group effects of the ANCOVA tests; the detailed statistics of the variance explained by the age covariate is reported only when reaching significance ($P < 0.05$). ANCOVAs were followed by between-group two-tailed *t*-tests for *post hoc* comparisons. To assess the adequacy of the computational model, we tested correlations, using linear regression fit, between modelled and observed average response time in both the 1 cent and 10 euro condition, across subjects ($n = 105$, interindividual variability) included in the computational analysis and across blocks ($n = 15$, intertrial variability). For all statistical tests, the threshold for significance was set at $P < 0.05$.

Results

Model-free analysis: comparison between tasks

We first compared motor learning performance of healthy subjects between the implicit and explicit tasks (Fig. 3A). The block × task ANOVA for the response time in the 1 cent condition showed a significant effect of block [$F(14,966) = 13.605$, $P < 0.001$], with no significant task effect [$F(1,69) = 0.040$, $P > 0.5$] or block × task interaction [$F(14,966) = 0.385$, $P > 0.5$]. For the 1 cent error rate, the ANOVA showed a significant task effect [$F(1,69) = 4.134$, $P < 0.05$] with no block effect [$F(14,966) = 0.830$, $P > 0.5$] or block × task interaction [$F(14,966) = 0.636$, $P > 0.5$]. These results suggest that only response times (not error rates) were influenced by learning and that only error rates (not response times) differed between tasks, irrespective of learning effects. These conclusions were confirmed by *post hoc* *t*-tests, which showed no significant difference for cumulative response time (implicit task: 1060 ± 31 versus explicit task: 1048 ± 46 ms; $t_{69} = 0.200$, $P > 0.5$) but a significant task effect on average error rate (implicit task: 8.1 ± 0.9 versus explicit task 12.4 ± 2.4 % of errors; $t_{69} = -2.003$, $P < 0.05$).

The block × task ANOVA on response time reinforcement effect showed a significant block effect [$F(14,966) = 7.302$, $P < 0.001$] but no significant task effect [$F(1,69) = 0.046$, $P > 0.1$] or block × task interaction [$F(14,966) = 0.111$, $P > 0.5$]. The reinforcement effect on error rates was not significantly modulated by any factor of the ANOVA [block: $F(14,966) = 1.537$, $P > 0.05$; task: $F(1,69) = 0.102$, $P > 0.1$; interaction: $F(14,966) = 0.341$, $P > 0.5$]. These results suggest that rewards only impacted on response time reduction (not error rates) and that this effect was similar across tasks. These conclusions were confirmed by *post hoc* *t*-tests showing that reinforcement effect on cumulative response time was not different between the two tasks ($t_{69} = -0.095$, $P > 0.5$). More precisely, it was significantly different from zero for the implicit task (297.4 ± 105.7 ms; $t_{52} = 2.812$, $P < 0.01$), but not for the explicit task, even if numerically similar (318.8 ± 234.4 ms; $t_{17} = 1.360$, $P < 0.1$). The average reinforcement effect on error rates was not different between the two tasks ($t_{69} = -0.482$, $P > 0.5$) and not significantly different from zero for both the implicit (8.3 ± 10.4 % of errors; $t_{52} = 0.799$, $P > 0.1$) and explicit task (17.8 ± 13.99 % of errors; $t_{17} = 1.270$, $P > 0.1$). When analysing blocks three by three, we found that the response time reinforcement effect was not different from zero at the beginning (implicit task: 9.2 ± 12.3 ms; $t_{52} = 0.746$, $P > 0.1$; explicit task: 2.7 ± 33.2 ms; $t_{17} = 0.081$, $P > 0.1$) but became progressively significant towards the end of the learning session (implicit task: 32.1 ± 11.4 ms; $t_{52} = 0.281$, $P < 0.01$; explicit task: 45.5 ± 24.9 ms; $t_{17} = 1.827$, $P < 0.05$).

Next, we analysed the performance in the forced choice task, which was significantly above chance level (implicit task: 57.0 ± 2.4% of correct guessing; $t_{52} = 3.415$, $P < 0.001$; explicit task: 58.3 ± 4.4 % of correct guessing; $t_{17} = 1.905$, $P < 0.05$) and not different between the two tasks ($t_{69} = -0.312$, $P > 0.05$). To rule out the possibility that the response time reinforcement effect was driven by an acquired explicit knowledge of the deterministic triplet-outcome contingencies, we tested the regression against correct guessing in the forced choice task (Fig. 3B). We found no positive correlation between these variables, but a significant negative correlation for the implicit task (regression coefficient = -15.6, intercept at chance level correct guessing = 406.4 ms, $R^2 = 0.090$, $P < 0.05$) and a similar trend for the explicit task (regression coefficient = -21.5, intercept at chance level correct guessing = 498.1 ms, $R^2 = 0.161$, $P < 0.1$).

Model-free analysis: comparison between groups

We analysed learning performance in the implicit task to test for differences between patients and controls, as well as between

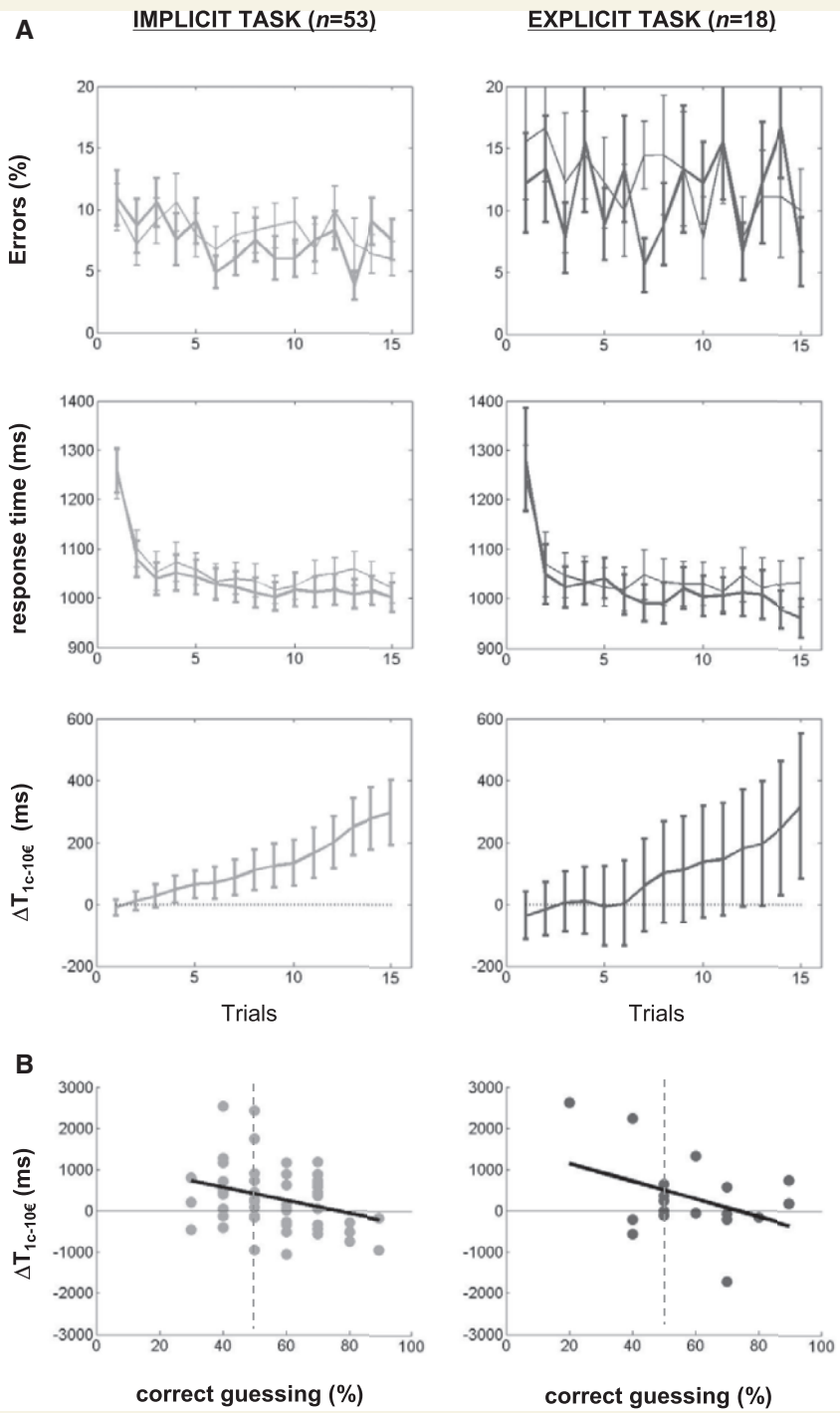

Figure 3 Comparison between implicit and explicit motor learning tasks. (**A**) Learning curves observed for error rate, response time and cumulative reinforcement effect across experimental blocks. Thin lines represent the 1 cent condition, bold lines the 10 euro condition. (**B**) Response time final reinforcement effect plotted against the percentage of correct responses obtained in the forced choice task. Each point represents one subject; dark lines represent linear regression fits.

unmedicated and medicated patients with Gilles de la Tourette syndrome. Regarding response time in the 1 cent condition (motor learning), the block × group ANOVA showed significant effect of block [$F(14,1414) = 31.956$, $P < 0.001$], group [$F(3,101) = 22.955$, $P < 0.001$] and block × group interaction [$F(14,1414) = 4.939$, $P < 0.001$]. The ANCOVA showed a significant effect of both group and age on the average response time [group: $F(3,100) = 3.762$, $P < 0.05$; age: $F(1,100) = 22.340$, $P < 0.05$)]. *Post hoc* comparisons confirmed significant differences between healthy controls and all patient groups, who exhibited longer response times (unmedicated Gilles de la Tourette syndrome: 1550 ± 60 ms, $t_{71} = -6.279$, $P < 0.001$; medicated Gilles

de la Tourette syndrome: 1491 ± 87 ms, $t_{66} = -4.135$, $P < 0.001$; dystonia: 2064 ± 102 ms, $t_{68} = -8.768$, $P < 0.001$), but no significant difference between Gilles de la Tourette syndrome groups ($t_{33} = 0.326$, $P > 0.5$).

The reinforcement effect was significantly modulated by block [$F(14,1414) = 16.848$, $P < 0.001$], group [$F(3,101) = 3.088$, $P < 0.05$] and block × group interaction [$F(14,1414) = 4.027$, $P < 0.001$]. The ANCOVA showed that the reinforcement effect on cumulative response time was significantly modulated by the group factor [$F(3,100) = 4.362$, $P < 0.01$], but not by the age covariate. More precisely, it was significantly positive in unmedicated Gilles de la Tourette syndrome (1152.9 ± 378.6 ms, $t_{19} = 3.044$, $P < 0.01$) and dystonia (786.2 ± 248.7, $t_{16} = 3.165$, $P < 0.01$), but not in medicated patients with Gilles de la Tourette syndrome (59.1 ± 262.9, $t_{14} = 0.225$, $P > 0.5$). *Post hoc* comparisons showed higher reinforcement effects in unmedicated Gilles de la Tourette syndrome compared with healthy controls ($t_{71} = 2.974$, $P < 0.01$) and medicated Gilles de la Tourette syndrome ($t_{33} = 2.214$, $P < 0.05$). The comparison between healthy controls and medicated patients with Gilles de la Tourette syndrome did not reach significance ($t_{66} = 0.983$, $P > 0.1$). In the forced choice task, percentages of correct responses were numerically higher than chance level in all patient groups, and significantly so in the unmedicated Gilles de la Tourette syndrome ($t_{19} = 4.187$, $P < 0.001$) and medicated Gilles de la Tourette syndrome ($t_{14} = 3.108$, $P < 0.01$) groups, but not the dystonia group ($t_{16} = 1.144$, $P > 0.1$). The ANCOVA showed no significant group effect on correct guessing [$F(3,100) = 1.291$, $P > 0.1$). Table 2 provides a summary of the behavioural results discussed in this section.

Model-based analysis: motor learning

We first analysed the parameter estimates obtained from the motor learning model fit of the 1 cent condition for each group (Figs 4A and 5, *left*). Regarding motor learning rate (parameter *a*), we found no significant group effect [$F(3,100) = 0.565$, $P > 0.5$] and no difference between Gilles de la Tourette syndrome groups ($t_{33} = -1.144$, $P > 0.1$). Regarding the magnitude of response time reduction (parameter *b*), we found a significant modulation by age covariate [$F(1,100) = 21.118$, $P < 0.01$] but not by group factor [$F(3,100) = 1.223$, $P > 0.1$]. Regarding response time plateau (parameter *c*), we found a significant modulation by both group factor and age covariate [group: $F(3,100) = 3.636$, $P < 0.05$; age: $F(1,100) = 15.540$, $P < 0.05$]. *Post hoc* comparisons confirmed significant differences between healthy controls and all patient groups (unmedicated Gilles de la Tourette syndrome: $t_{71} = -5.444$, $P < 0.001$; medicated Gilles de la Tourette syndrome: $t_{66} = -4.012$, $P < 0.001$; dystonia: $t_{68} = -7.743$, $P < 0.001$), but no significant difference between Gilles de la Tourette syndrome groups ($t_{33} = 0.655$, $P > 0.5$).

Model-based analysis: reinforcement learning

We then analysed the parameter estimates obtained from the reinforcement learning model fit of the 10 euro condition (Figs 4B and 5, *right*). ANCOVA indicated that the reinforcement learning rate (α) and the reinforcement-to-time scaling parameter (β) did not differ between groups [α: $F(3,100) = 1.348$, $P > 0.1$; β: $F(3,100) = 1.747$, $P > 0.1$]. Importantly, direct comparisons between Gilles de la Tourette syndrome groups for parameters α and β showed no significant difference (α: $t_{33} = 0.940$, $P > 0.1$; β: $t_{33} = 1.193$, $P > 0.1$). In contrast, γ showed a significant group effect [$F(3,100) = 3.238$, $P < 0.05$]; *post hoc* comparisons indicated that γ was significantly greater in unmedicated patients with Gilles de la Tourette syndrome compared with healthy controls ($t_{71} = 2.001$, $P < 0.05$) and to medicated patients with Gilles

Table 2 Experimental results

Variables	CON ($n = 53$)	U-GTS ($n = 20$)	M-GTS ($n = 15$)	DYS ($n = 17$)
Experimental measures				
Error rate (%)	8.1 ± 0.9	16.5 ± 2.2*	17.2 ± 1.4*	6.7 ± 0.8
Response time (ms)	1060 ± 31	1550 ± 60*	1491 ± 87*	2064 ± 102*
$\Delta T_{1c\text{-}10€}$ (ms)	296 ± 105	**1153 ± 378****	**58 ± 263**	786 ± 249
$\Delta E_{1c\text{-}10€}$ (%)	0.6 ± 0.7	3.1 ± 1.4	0.5 ± 1.5	0.2 ± 0.7
Correct guessing (%)	57.0 ± 2.0	62.0 ± 2.9	59.3 ± 3.0	53.5 ± 3.1
Model parameters				
Motor learning rate (*a*)	0.59 ± 0.06	0.55 ± 0.09	0.70 ± 0.10	0.50 ± 0.08
Response time reduction (*b*)	470 ± 80	600 ± 230	1120 ± 470	1780 ± 350
Response time plateau (*c*)	1030 ± 30	1510 ± 120*	1390 ± 140*	1790 ± 150*
Reinforcement learning rate (α)	0.54 ± 0.05	0.64 ± 0.09	0.51 ± 0.11	0.49 ± 0.09
Effective reinforcer ($R + \gamma$)	3.1 ± 1.0	**6.7 ± 1.3****	**− 0.7 ± 2.2**	3.7 ± 1.7
Reinforcement-to-time parameter (β)	3.6 ± 0.4	5.8 ± 0.8	4.3 ± 0.8	5.1 ± 0.7

Error rates and response times correspond to the 1 cent condition. ΔT and ΔE indicate the difference with the 10 euro condition (reinforcement effects) on response times and error rates, respectively.
For variables with a significant group effect in the ANCOVA: *$P < 0.05$, two-tailed *t*-test, compared with the control group; ** = $P < 0.05$, two-tailed *t*-test, comparing unmedicated Gilles de la Tourette syndrome versus medicated Gilles de la Tourette syndrome groups.
CON = healthy controls who performed the implicit motor learning task; U-GTS and M-GTS = unmedicated and medicated patients with Gilles de la Tourette syndrome, respectively; DYS = dystonic patients.

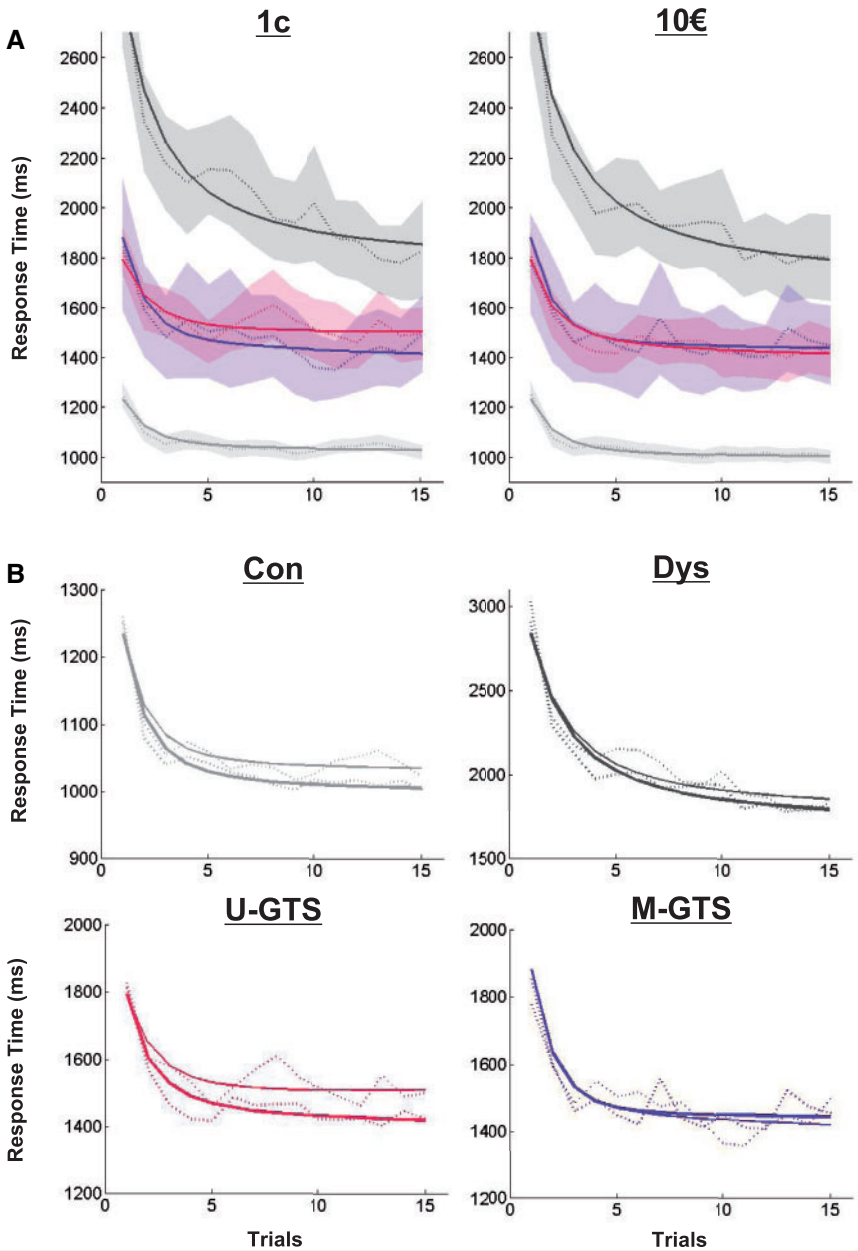

Figure 4 Model fits. (**A**) Response time learning curves are shown separately for the 1 cent (*left*) and 10 euro (*right*) conditions. Thin dotted lines surrounded by filled areas represent mean observed data ± standard errors; bold lines represent estimated data from the relevant computational model. (**B**) Response time learning curves for both 1 cent and 10 euro conditions are shown separately for the different groups. Thin lines represent the 1 cent condition, bold lines the 10 euro condition. Dashed lines represent mean observed data, solid lines represent model estimates. Healthy controls (Con) = light grey; unmedicated Gilles de la Tourette syndrome (U-GTS) = red; medicated Gilles de la Tourette syndrome (M-GTS) = blue; dystonic patients (Dys) = dark grey.

de la Tourette syndrome ($t_{33} = 3.048$, $P < 0.01$), but not when compared with patients with dystonia ($t_{68} = -0.301$, $P > 0.5$). The direct comparison between healthy controls and medicated Gilles de la Tourette syndrome was only marginally significant ($t_{66} = -1.789$, $P < 0.08$). The estimate of the effective reinforcer term ($R + \gamma$) was found to be positive and significantly different from zero for healthy controls ($t_{52} = 3.228$, $P < 0.001$), unmedicated Gilles de la Tourette syndrome ($t_{19} = 5.059$, $P < 0.001$) and patients with dystonia ($t_{16} = 2.480$, $P < 0.001$), but not for medicated patients with Gilles de la Tourette syndrome ($t_{14} = -0.336$, $P > 0.5$).

To rule out the possibility that between-group differences in parameter γ were driven by differences in explicit knowledge of the deterministic triplet-outcome contingencies, we also compared forced choice task performance. For groups showing a positive and significant effective reinforcement ($R + \gamma$) parameter, we computed a linear regression against percentage of correct guessing. We found a significant negative correlation for healthy

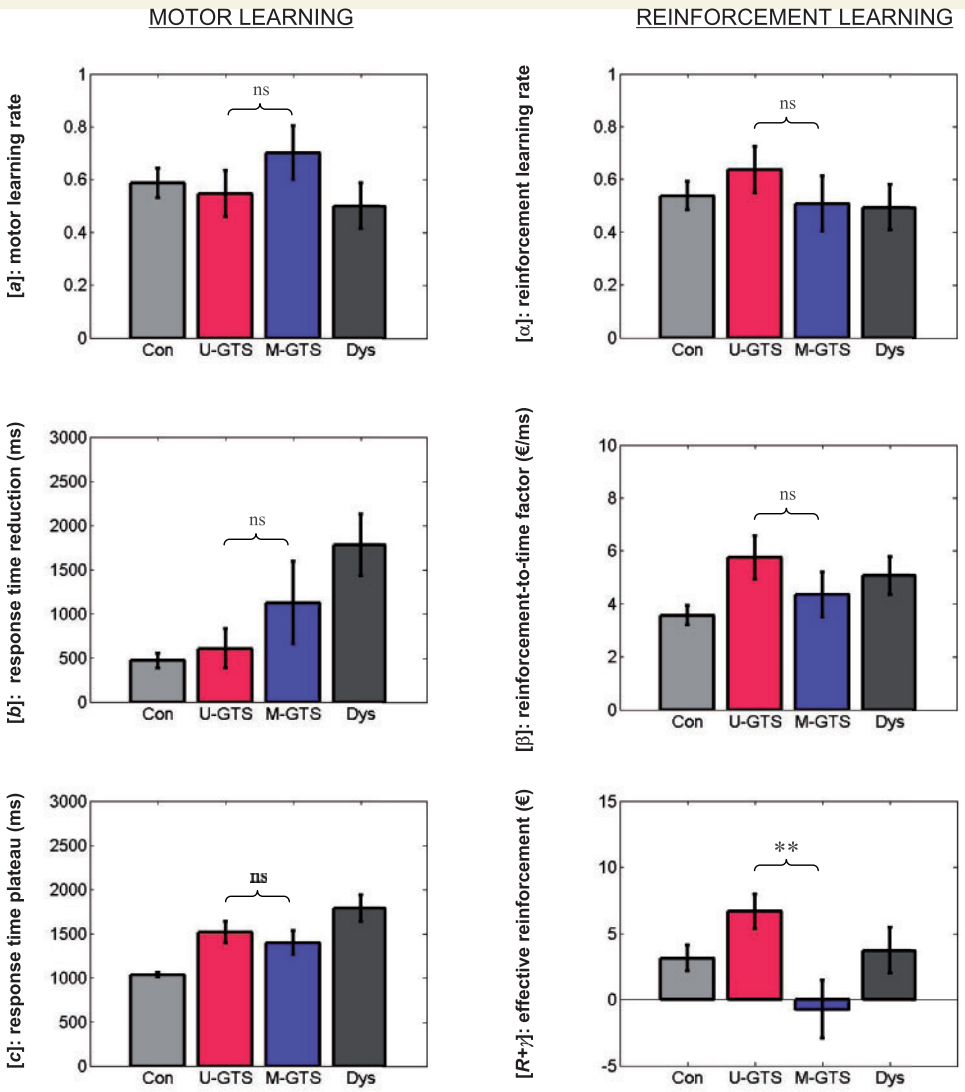

Figure 5 Model parameters. Parameter estimates obtained in the different groups are shown separately for the motor learning (*left*) and the reinforcement learning (*right*) models. Error bars are intersubject standard errors. Con = healthy controls who performed the implicit motor learning task; U-GTS and M-GTS = unmedicated and medicated patients with Gilles de la Tourette syndrome patients, repectively; Dys = dystonic patients. Comparisons between unmedicated Gilles de la Tourette syndrome and medicated Gilles de la Tourette syndrome groups: $^{**}P < 0.01$; ns = $P > 0.1$, two-tailed *t*-test; ns = not significant.

subjects (coefficient = −0.15, intercept = 4.1 euro, $R^2 = 0.094$, $P < 0.05$) and also a negative, but non-significant, correlation in patients groups (unmedicated Gilles de la Tourette syndrome: coefficient = −0.18, intercept = 8.8 euro, $R^2 = 0.151$, $P < 0.1$; dystonia: coefficient = −0.13, intercept = 4.2 euro, $R^2 = 0.054$, $P > 0.1$). Table 2 is a summary of the computational results discussed above.

Adequacy between model and data

To test the adequacy of our model relative to the real data, we regressed estimated response time against observed response time for the 1 cent and 10 euro conditions separately, including all subjects ($n = 105$) who performed the implicit task. First, to assess whether the model could capture interindividual variability, we tested the correlation of average response times across subjects (Fig. 6A). Both regressions were extremely strong (1 cent: coefficient = 1.00 ± 0.01, $R^2 = 0.999$, $P < 0.001$; 10 euro: coefficient = 1.01 ± 0.01, $R^2 = 0.996$, $P < 0.001$). Crucially, the 95% confidence intervals for the regression coefficient was ∼1, ruling out any systematic bias towards under or overestimation. Then we assessed whether the model could capture intertrial variability (i.e. learning effects) using correlations of average estimated and observed response times across blocks ($n = 15$). Again we found for both conditions extremely strong linear regression fits (1 cent: coefficient = 0.98 ± 0.13, $R^2 = 0.956$, $P < 0.001$; 10 euro: coefficient = 0.97 ± 0.13, $R^2 = 0.953$, $P < 0.001$), with 95% confidence intervals ∼1 (Fig. 6B). We also verified (Fig. 6C) the similarity of intergroup differences between the observed reinforcement effect ($\Delta T_{1c\text{-}10€}$) and the effective reinforcer term ($R + \gamma$) of the model.

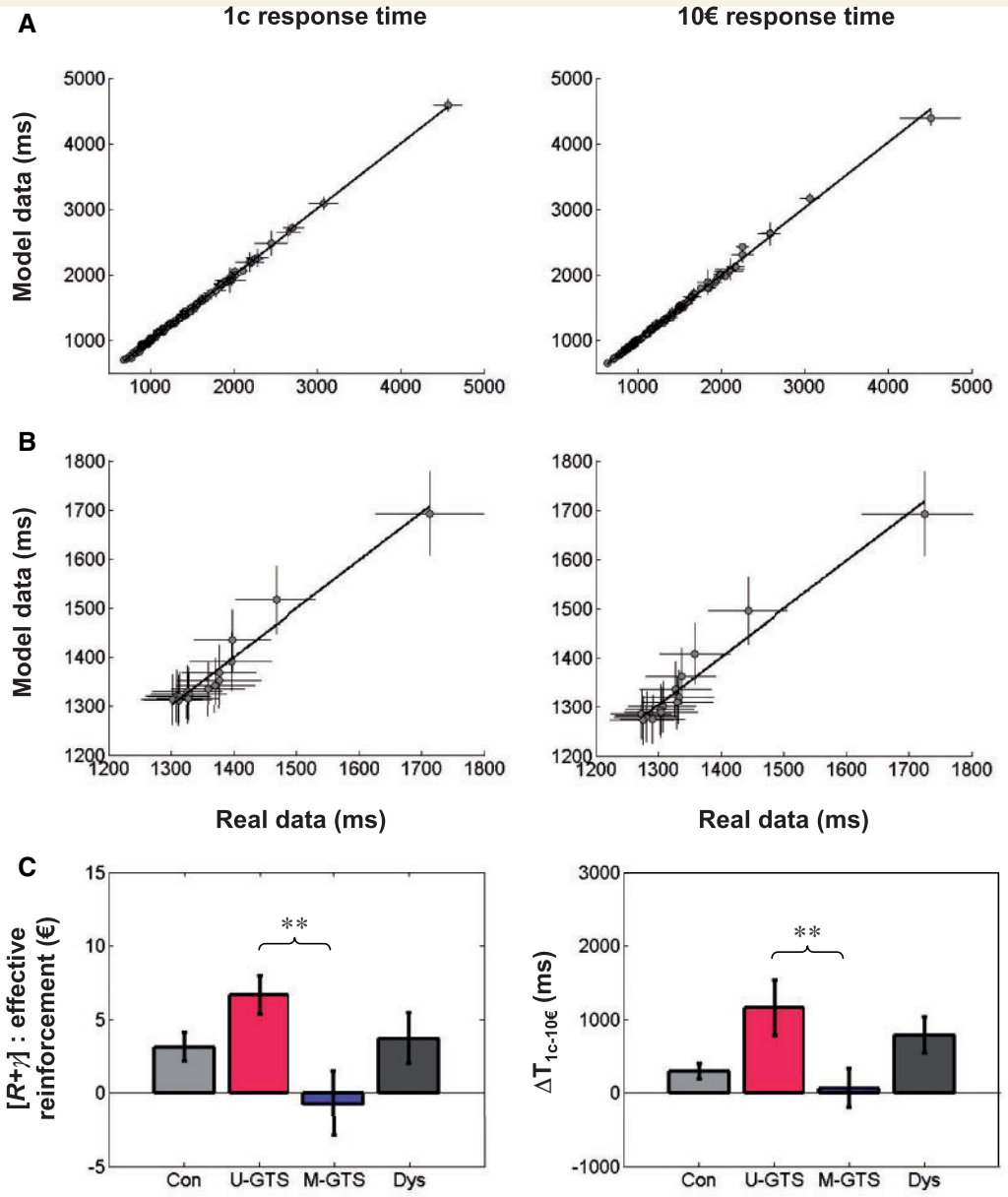

Figure 6 Model/data adequacy. (**A**) Intersubject ($n = 105$) correlation between the observed and modelled response times. (**B**) Intertrial ($n = 15$) correlation between the observed and modelled response times. *Left*, 1 cent condition; *right*, 10 euro condition. Bold black lines represent linear regression fits, thin black lines represent intertrial (in **A**) or intersubject (in **B**) standard errors. (**C**) Average estimated reinforcement magnitude parameter (*left*) and observed cumulative response time reinforcement effect (*right*) in the different groups. Error bars are intersubject standard errors. Con = healthy controls who performed the implicit motor learning task; U-GTS and M-GTS = unmedicated and medicated patients with Gilles de la Tourette syndrome patients, respectively; Dys = dystonic patients. Comparisons between unmedicated Gilles de la Tourette syndrome and medicated Gilles de la Tourette syndrome groups: $^{**}P < 0.01$, two-tailed *t*-test.

Discussion

Results show that reinforcements can improve motor skill learning, since execution of sequential key presses was shortened by monetary rewards. This effect was observed in healthy controls, dystonic patients and unmedicated patients with Gilles de la Tourette syndrome but not in patients with Gilles de la Tourette syndrome treated with neuroleptics, suggesting that it is mediated via dopaminergic transmission. The effect was well accounted for by reinforcement learning models, in accordance with the theory that dopamine encodes a reward prediction error signal that can drive the formation of visuomotor representations.

Reinforcement learning theory was principally meant to explain instrumental learning, i.e. how choices are improved by the experience of success or failure (Sutton and Barto, 1998). Here, reinforcement effects applied not to choice but to motor skill,

precisely the ability to sequentially press three keys in response to visual instructions. It has been shown that sequence learning can be improved by incentive motivation, meaning when subjects are told that shorter response times will be rewarded (Wachter *et al*., 2009). Here, reinforcements were purely incidental: subjects knew that reinforcement magnitude (1 cent or 10 euro) was independent of response time. Moreover, we ruled out motivation effects using an explicit version of the task where the reward at stake is announced prior to motor performance. If response times were sensitive to incentive motivation, they should be reduced by the 10 euro incentive from the first trial onwards. This was not observed: reward effects on response times built up progressively throughout learning sessions, similarly for the explicit and implicit versions of the task. This suggests that motivation does not induce subjects to go faster in our paradigm. The only effect of explicit incentives was to increase errors, as if they distracted subjects. In fact, slowing down may be rational here, in order to ensure correct execution of the 10 euro-associated sequences. Furthermore, with a forced-choice task we controlled to what extent subjects acquired clear explicit knowledge of sequence-outcome contingencies. Results showed poor discrimination between 1 cent and 10 euro triplets, albeit significantly above chance level. Between-subject regression indicated that reinforcement effects on response time were positive even for chance level correct guessing in the forced-choice task. Moreover, the correlation was negative, suggesting that reinforcement effects were blunted in subjects who developed explicit knowledge of sequence-outcome contingencies. This may be due to subjects strategically slowing down when aware that the outcome at stake is 10 euro, to avoid making costly errors. Thus, we conclude that the beneficial effect of incidental reinforcements on motor skill acquisition is largely implicit.

To capture this effect, we proposed a two-stage model, with reinforcement learning on top of pure motor learning, which was simply fitted with an exponential decay function. As in classical reinforcement learning algorithms (Sutton and Barto, 1998), the expected value of each sequence was learned using a delta rule (Rescorla and Wagner, 1972). This rule specifies that in every trial, the expected value of the current sequence is to be updated in proportion to the reward prediction error. We added a second delta rule in which reward prediction errors serve to reduce response times. The two delta rules might be implemented in two learning systems that respond to visual instructions: one system would generate a reward prediction, and the other would facilitate the execution of the associated action. This matches the actor-critic distinction in formal reinforcement learning models (Houk *et al*., 1995; Doya, 2002; Joel *et al*., 2002), in which both the actor and the critic improve their response using reward prediction errors as teaching signals. Previous functional neuroimaging studies suggested that the actor and critic components are implemented in the motor and limbic parts of the striatum, respectively (O'Doherty *et al*., 2004). Furthermore, both the motor and limbic striatum were shown to express reward prediction errors, whose magnitude was sensitive to modulation of dopaminergic transmission (Pessiglione *et al*., 2006). Thus, delta rules may be underpinned by dopaminergic signals modulating corticostriatal synapse efficacy in different parts of the striatum, improving reward prediction in the ventral striatum and movement execution in the dorsal striatum. This would predict perturbation of dopaminergic transmission to affect reinforcement effects on motor learning and not motor learning *per se*.

We indeed found that reinforcement effects were higher in unmedicated patients with Gilles de la Tourette syndrome compared with healthy controls and that this superiority was reversed by neuroleptics. This is consistent with the alleged hyperdopaminergy in Gilles de la Tourette syndrome (Malison *et al*., 1995; Minzer *et al*., 2004; Gilbert *et al*., 2006; Yoon *et al*., 2007; Wong *et al*., 2008) that could lead to excessive reinforcement of corticostriatal synapses, which is a candidate mechanism for explaining the progressive formation of motor tics (Leckman, 2002; Groenewegen *et al*., 2003; Saka and Graybiel, 2003; Maia and Frank, 2011). Here, we dissociate reinforcement from motor learning, which occurred when reward was negligible (in the 1 cent condition). We show that motor skill performance was not enhanced but impaired in unmedicated patients compared with controls. Importantly, medication had no effect on motor learning *per se*, indicating that dopamine transmission is specifically involved in reinforcement learning. This also suggests that differences in reinforcement effects are not related to patients being older and slower than healthy controls, as age and response times were similar between unmedicated and medicated patients with Gilles de la Tourette syndrome. We also controlled these factors by including dystonic patients, who were older and slower than patients with Gilles de la Tourette syndrome. In dystonic patients, we found a deficit in motor performance (overall slowing) but no difference in reinforcement learning compared with controls. Note that age was included as a covariate in statistical analyses and was found to affect response times but not reinforcement effects. Thus, while all patient groups showed motor skill impairment increasing with age, reinforcement effects specifically followed variations in dopamine transmission: enhanced in unmedicated patients with Gilles de la Tourette syndrome and impaired in patients with Gilles de la Tourette syndrome treated with neuroleptics.

Motor performance impairment was best captured by one parameter of the exponential decay function: response time plateau. This points out that patients ended with longer response times relative to controls, which denotes impaired motor skills. Variations in reinforcement effects were well captured by adjusting the size of prediction errors, not the reinforcement learning rate or reinforcement-to-time scaling factor. This agrees well with previous studies showing that dopaminergic drugs modulate the amplitude of reward-related signals (Pessiglione *et al*., 2006; Abler *et al*., 2007; Menon *et al*., 2007). It suggests that variations in dopamine transmission affect the prediction error signal itself, not the capacity to use it for adjusting expected values (in the critic) or for facilitating action execution (in the actor). Taken together, our two-stage model provided a good fit of the data and could account for between-group differences by adjusting the free parameters.

To our knowledge, reinforcement learning was understudied in Gilles de la Tourette syndrome, although some authors reported deficits in probabilistic classification learning (Kéri *et al*., 2002;

Marsh *et al*., 2005). This task can be seen as a form of instrumental learning, but it is cognitively more complex and mixes positive and negative feedbacks. It is crucial, when assessing modulation of dopamine transmission, to distinguish between positive and negative feedbacks, as opposite effects of dopaminergic drugs have been repeatedly found for reward versus punishment learning (Frank *et al*., 2004; Pessiglione *et al*., 2006; Bódi *et al*., 2009; Cools *et al*., 2009; Rutledge *et al*., 2009). Indeed, we showed in a previous study (Palminteri *et al*., 2009) that subliminal reward learning, but not punishment learning, is enhanced in unmedicated patients with Gilles de la Tourette syndrome and impaired by treatment with neuroleptics. This is consistent with the present results, suggesting that, even if encountering difficulties in motor skill performance, patients with Gilles de la Tourette syndrome benefit from positive feedbacks. More generally, our findings give empirical evidence to the intuitive idea that providing reinforcements supports people in acquiring motor skills.

Acknowledgements

We wish to thank Helen Bates, for piloting the behavioural task, the staff of the Centre de Neuro-Imagerie de Recherche (CENIR), where the experiments were conducted, and Alina Strasser, who checked the English. We are also grateful to the anonymous referees, whose comments significantly helped us to improve the quality of this work.

Funding

The study was funded by the Ecole de Neurosciences de Paris (ENP). S.P. received a PhD fellowship from the Neuropôle de Recherche Francilien (NERF).

Supplementary material

Supplementary material is available at *Brain* online.

References

Abler B, Erk S, Walter H. Human reward system activation is modulated by a single dose of olanzapine in healthy subjects in an event-related, double-blind, placebo-controlled fMRI study. Psychopharmacology 2007; 191: 823–33.

Agostino R, Berardelli A, Formica A, Accornero N, Manfredi M. Sequential arm movements in patients with Parkinson's disease, Huntington's disease and dystonia. Brain 1992; 115 (Pt 5): 1481–95.

Albin RL, Mink JW. Recent advances in Tourette syndrome research. Trends Neurosci 2006; 29: 175–82.

Argyelan M, Carbon M, Ghilardi MF, Feigin A, Mattis P, Tang C, et al. Dopaminergic suppression of brain deactivation responses during sequence learning. J Neurosci 2008; 28: 10687–95.

Ashe J, Lungu OV, Basford AT, Lu X. Cortical control of motor sequences. Curr Opin Neurobiol 2006; 16: 213–21.

Badgaiyan RD, Fischman AJ, Alpert NM. Striatal dopamine release in sequential learning. Neuroimage 2007; 38: 549–56.

Bar-Gad I, Bergman H. Stepping out of the box: information processing in the neural networks of the basal ganglia. Curr Opin Neurobiol 2001; 11: 689–95.

Bayer HM, Glimcher PW. Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 2005; 47: 129–41.

Berns GS, Sejnowski TJ. A computational model of how the basal ganglia produce sequences. J Cogn Neurosci 1998; 10: 108–21.

Bódi N, Keri S, Nagy H, Moustafa A, Myers CE, Daw N, et al. Reward-learning and the novelty-seeking personality: a between- and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. Brain 2009; 132: 2385–95.

Carbon M, Eidelberg D. Functional imaging of sequence learning in Parkinson's disease. J Neurol Sci 2006; 288: 72–7.

Cools R, Frank MJ, Gibbs SE, Miyakawa A, Jagust W, D'Esposito M. Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. J Neurosci 2009; 29: 1538–43.

Doya K. Complementary roles of basal ganglia and cerebellum in learning and motor control. Curr Opin Neurobiol 2000; 10: 732–9.

Doya K. Metalearning and neuromodulation. Neural Netw 2002; 15: 495–506.

Doya K. Modulators for decision making. Nat Neurosci 2008; 11: 410–6.

Doyon J. Motor sequence learning and movement disorders. Curr Opin Neurol 2008; 21: 478–83.

Doyon J, Bellec P, Amsel R, Penhune V, Monchi O, Carrier J, et al. Contributions of the basal ganglia and functionally related brain structures to motor learning. Behav Brain Res 2009; 199: 61–75.

Frank MJ. Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. Neural Netw 2006; 19: 1120–36.

Frank MJ, Seeberger LC, O'Reilly R C. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science 2004; 306: 1940–3.

Garraux G, Peigneux P, Carson RE, Hallett M. Task-related interaction between basal ganglia and cortical dopamine release. J Neurosci 2007; 27: 14434–41.

Gilbert DL, Christian BT, Gelfand MJ, Shi B, Mantil J, Sallee FR. Altered mesolimbocortical and thalamic dopamine in Tourette syndrome. Neurology 2006; 67: 1695–7.

Graybiel AM, Canales JJ. The neurobiology of repetitive behaviors: clues to the neurobiology of Tourette syndrome. Adv Neurol 2001; 85: 123–31.

Groenewegen HJ, van den Heuvel OA, Cath DC, Voorn P, Veltman DJ. Does an imbalance between the dorsal and ventral striatopallidal systems play a role in Tourette's syndrome? A neuronal circuit approach. Brain Dev 2003; 25 (Suppl 1): S3–S14.

Heathcote A, Brown S, Mewhort DJ. The power law repealed: the case for an exponential law of practice. Psychon Bull Rev 2000; 7: 185–207.

Hikosaka O, Nakamura K, Sakai K, Nakahara H. Central mechanisms of motor skill learning. Curr Opin Neurobiol 2002; 12: 217–22.

Houk JC, Adams JL, Barto AG. A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: Houk JC, Davis JL, Beiser DG, editors. Models of information processing in the basal ganglia. Cambridge, Massachusetts: The MIT press; 1995. p. 249–70.

Jankovic J. Treatment of dystonia. Lancet Neurol 2006; 5: 864–72.

Joel D, Niv Y, Ruppin E. Actor-critic models of the basal ganglia: new anatomical and computational perspectives. Neural Netw 2002; 15: 535–47.

Karabanov A, Cervenka S, de Manzano O, Forssberg H, Farde L, Ullen F. Dopamine D2 receptor density in the limbic striatum is related to implicit but not explicit movement sequence learning. Proc Natl Acad Sci USA 2010; 107: 7574–9.

Kwak Y, Müller ML, Bohnen NI, Dayalu P, Seidler RD. Effect of dopaminergic medications on the time course of explicit motor sequence learning in Parkinson's disease. J Neurophysiol 2010; 103: 942–9.

Kéri S, Szlobodnyik C, Benedek G, Janka Z, Gadoros J. Probabilistic classification learning in Tourette syndrome. Neuropsychologia 2002; 40: 1356–62.

Kumari V, Corr PJ, Mulligan OF, Cotter PA, Checkley SA, Gray JA. Effects of acute administration of d-amphetamine and haloperidol on procedural learning in man. Psychopharmacology 1997; 129: 271–6.

Leckman JF. Tourette's syndrome. Lancet 2002; 360: 1577–86.

Maia TV. Two-factor theory, the actor-critic model, and conditioned avoidance. Learn Behav 2010; 38: 50–67.

Maia TV, Frank MJ. From reinforcement learning models to psychiatric and neurological disorders. Nat Neurosci 2011; 14: 154–62.

Malison RT, McDougle CJ, van Dyck CH, Scahill L, Baldwin RM, Seibyl JP, et al. [123I]beta-CIT SPECT imaging of striatal dopamine transporter binding in Tourette's disorder. Am J Psychiatry 1995; 152: 1359–61.

Marsh R, Alexander GM, Packard MG, Zhu H, Peterson BS. Perceptual-motor skill learning in Gilles de la Tourette syndrome. Evidence for multiple procedural learning and memory systems. Neuropsychologia 2005; 43: 1456–65.

Menon M, Jensen J, Vitcu I, Graff-Guerrero A, Crawley A, Smith MA, et al. Temporal difference modeling of the blood-oxygen level dependent response during aversive conditioning in humans: effects of dopaminergic modulation. Biol Psychiatry 2007; 62: 765–72.

Mink JW. The basal ganglia and involuntary movements: impaired inhibition of competing motor patterns. Arch Neurol 2003; 60: 1365–8.

Minzer K, Lee O, Hong JJ, Singer HS. Increased prefrontal D2 protein in Tourette syndrome: a postmortem analysis of frontal cortex and striatum. J Neurol Sci 2004; 219: 55–61.

O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 2004; 304: 452–4.

Palminteri S, Lebreton M, Worbe Y, Grabli D, Hartmann A, Pessiglione M. Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes. Proc Natl Acad Sci USA 2009; 106: 19179–84.

Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 2006; 442: 1042–5.

Rescorla RA, Wagner AR. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF, editors. Classical conditioning II: current research and theory. New-York: Appleton-Century-Crofts; 1972. p. 64–99.

Rhodes BJ, Bullock D, Verwey WB, Averbeck BB, Page MP. Learning and production of movement sequences: behavioral, neurophysiological, and modeling perspectives. Hum Mov Sci 2004; 23: 699–746.

Rutledge RB, Lazzaro SC, Lau B, Myers CE, Gluck MA, Glimcher PW. Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. J Neurosci 2009; 29: 15104–14.

Saka E, Graybiel AM. Pathophysiology of Tourette's syndrome: striatal pathways revisited. Brain Dev 2003; 25 (Suppl 1): S15–9.

Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science 1997; 275: 1593–9.

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998; 59 (Suppl 20): 22–33;quiz 34–57.

Singer HS. Tourette's syndrome: from behaviour to biology. Lancet Neurol 2005; 4: 149–59.

Suri RE, Schultz W. Learning of sequential movements by neural network model with dopamine-like reinforcement signal. Exp Brain Res 1998; 121: 350–4.

Sutton RS, Barto AG. Reinforcement learning. Cambridge, MA: MIT Press; 1998.

Vidailhet M, Grabli D, Roze E. Pathophysiology of dystonia. Curr Opin Neurol 2009; 22: 406–13.

Wachter T, Lungu OV, Liu T, Willingham DT, Ashe J. Differential effect of reward and punishment on procedural learning. J Neurosci 2009; 29: 436–43.

Willingham DB. A neuropsychological theory of motor skill learning. Psychol Rev 1998; 105: 558–84.

Wong DF, Brasic JR, Singer HS, Schretlen DJ, Kuwabara H, Zhou Y, et al. Mechanisms of dopaminergic and serotonergic neurotransmission in Tourette syndrome: clues from an in vivo neurochemistry study with PET. Neuropsychopharmacology 2008; 33: 1239–51.

Yoon DY, Gause CD, Leckman JF, Singer HS. Frontal dopaminergic abnormality in Tourette syndrome: a postmortem analysis. J Neurol Sci 2007; 255: 50–6.

Zaghloul KA, Blanco JA, Weidemann CT, McGill K, Jaggi JL, Baltuch GH, et al. Human substantia nigra neurons encode unexpected financial rewards. Science 2009; 323: 1496–9.