Focal basal ganglia lesions are associated with impairments in reward-based reversal learning

Christian Bellebaum,1 Benno Koch,2 Michael Schwarz2 and Irene Daum1

1Department of Neuropsychology, Institute of Cognitive Neuroscience, Ruhr-University Bochum and
2Department of Neurology, Klinikum Dortmund, Germany

Correspondence to: Dr Christian Bellebaum, Institute of Cognitive Neuroscience, Department of Neuropsychology, Ruhr-University of Bochum, Universitaetsstrasse 150, D-44780 Bochum, Germany
E-mail: christian.bellebaum@rub.de

The basal ganglia (BG) are thought to play a key role in learning from feedback, with mesencephalic dopamine neurons coding errors in reward prediction, thereby mediating information processing in the BG and the prefrontal cortex. In the present study, reward-based learning was assessed in patients with focal BG lesions, by studying outcome-based acquisition and reversal of stimulus–stimulus associations with different reward magnitudes in two probabilistic learning tasks. Eleven patients with selective BG lesions (three females) and 18 healthy control subjects (six females) participated in this study. Two cognitive transfer tasks provided a measure of declarative learning strategy application. On the group level, BG patients showed deficits in reversal learning, with dorsal striatum lesion patients being most severely affected. While basic mechanisms of learning from feedback such as the processing of different reward magnitudes appeared to be intact, patients needed more trials than controls to learn a second reward-based task, suggesting reduced carry-over effects in learning. A patient with a bilateral BG lesion showed better performance than controls on most learning tasks, applying a compensatory declarative learning strategy. The results are discussed in terms of the implication of different BG subregions in different aspects of learning from feedback.

Keywords: reward; basal ganglia; striatum; learning; reversal

Abbreviations: BA = basal ganglia; DA = dopamine; MRI = magnetic resonance imaging; MTL = medial temporal lobe; Put = putamen; RT = reaction time; SD = standard deviation; TIA = transient ischaemic attack


Introduction

The dopaminergic system is critically involved in learning from feedback. Dopamine (DA) neurons in the monkey mesencephalic ventral tegmental area and the substantia nigra code errors in reward prediction, i.e. their firing rate is increased, when reward is unexpected and firing is reduced, when expected reward is omitted (see Schultz, 2001). Receiving dopaminergic projections from the mid-brain, the basal ganglia (BG), in particular the caudate nucleus and the putamen (Put) (the striatum) are believed to play a key role in the reward-based acquisition of new stimulus–stimulus or stimulus–response associations. In accordance with this assumption, single cells in the monkey striatum show a variety of different reward-related activations (Apicella et al., 1991; Schultz et al., 1992; Cromwell and Schultz, 2003; Cromwell et al., 2005), integrating information about rewards with visual information, memory or motor behaviour (Kawagoe et al., 1998; Schultz et al., 2000).

In human subjects, the study of Parkinson’s disease patients has provided initial evidence for BG involvement in learning. Disruption of BG and dopaminergic function in these patients has been related to a range of learning deficits (Daum et al., 1995, 1996), including impairments in learning from feedback (Knowlton et al., 1996; Myers et al., 2003; Frank et al., 2004; Cools et al., 2006a). BG activations have frequently been observed in functional neuroimaging studies of feedback-based learning (e.g. Poldrack et al., 2001; Delgado et al., 2003; Haruno et al., 2004; Tanaka et al., 2004; Knutson and Cooper, 2005), but the results are so far inconsistent with respect to the critical involvement of different BG subregions in the learning process. In some studies, strongest activations were observed in the ventral striatum, including the nucleus.
accumbens as well as ventral parts of the caudate nucleus and the Put (Breiter et al., 2001; Pagnoni et al., 2002; Seymour et al., 2007). Other studies reported reward-related activity in both the ventral and dorsal striatum, but only the dorsal activation was modulated by reward valence and magnitude (Delgado et al., 2000, 2003).

In recent years, research has concentrated on functional dissociations within the BG. O’Doherty et al. (2004) reported ventral striatum activations for learning stimulus–outcome association in the context of both classical and instrumental conditioning, whereas dorsal activations were only observed during instrumental conditioning, suggesting that the dorsal striatal region might be more closely related to linking actions and outcomes (O’Doherty et al., 2004; Tricomi et al., 2004, 2006). Another possible functional dissociation within the human striatum refers to the distinct neural correlates of learning from positive and negative feedback (Delgado et al., 2003; Yacubian et al., 2006; Seymour et al., 2007).

In a series of studies, Cools and co-workers (2002) reported accumulating evidence for ventral striatal involvement in feedback-based reversal learning. In healthy human subjects, they found significant activations in the ventral striatum and the ventral prefrontal cortex related to reversal (Cools et al., 2002). In medicated Parkinson’s disease patients, impaired reversal learning was observed, while initial acquisition was intact (Cools et al., 2001, 2006a). This pattern was discussed in relation to ventral striatal dysfunction. In early Parkinson’s disease, dopaminergic dysfunction mainly affects the dorsal striatum (Kish et al., 1988), with L-DOPA medication presumably leading to a DA overdose in the ventral striatum. In fact, L-DOPA was found to selectively affect neural activity in the ventral striatum during a reversal-learning task in Parkinson’s disease patients (Cools et al., 2007).

Taken together, functional neuroimaging studies and studies of Parkinson’s disease patients have provided important insights into the mechanisms of learning from feedback in humans. There is, however, as yet no consistent pattern of associations between different BG regions and learning from feedback. The present study aimed to explore this issue further, based on the study of reward-based learning in patients with focal BG lesions. Given the previously reported findings in Parkinson’s disease patients, one main focus was the assessment of acquisition versus reversal learning.

The BG are closely linked to the prefrontal cortex via three non-motor loops (Alexander et al., 1986; Middleton and Strick, 2000), two of which—providing connections to the orbitofrontal cortex and the anterior cingulate cortex—are of particular interest for reward-based learning. The anterior cingulate cortex has been implicated in the integration of reward-related and motor information with other information relevant for action selection, in particular during processing of negative outcomes (Miltner et al., 1997; Gehring and Willoughby, 2002; Holroyd and Coles, 2002). The orbitofrontal cortex, on the other hand, appears to code reward magnitude (O’Doherty et al., 2001; Roesch and Olson, 2004) and relative reward preference (Tremblay and Schultz, 1999), suggesting that functional interactions between orbitofrontal cortex and striatum are important for stimulus–reward association learning (Rolls, 2000). Although it is difficult to separate different human BG subregions based on anatomical projections, different areas with distinct cortical connectivity patterns have recently been identified (Karachi et al., 2002; Lehericy et al., 2004).

In the present study, two different learning tasks were administered. The first task assessed acquisition and reversal as well as the processing of different reward magnitudes. It was designed to tap striatal function and processing linked to the BG-orbitofrontal cortex circuit (see above). A second reversal with a subsequent cognitive transfer test required subjects to categorize stimuli based on previously learned reward contingencies. The task assessed outcome-based learning of stimulus–stimulus associations and did thus differ from some of the tasks applied in previous functional neuroimaging or event-related potential studies of reward-based learning, which either required learning of stimulus–outcome (O’Doherty et al., 2004) or stimulus–response–outcome associations (Holroyd and Coles, 2002). It did, however, show clear parallels to tasks administered in studies of Parkinson’s disease patients. In the study by Frank et al. (2004), for example, the outcome did not only depend on the chosen stimulus, but also on the context, i.e. on the alternative, not chosen stimulus. Similarly, Myers et al. (2003) assessed the learning of an association between two stimuli based on feedback. The second learning task of the present study, an acquired equivalence task, also shared many features with the procedure used by Myers et al. (2003). Together with the categorization test of the first learning task, the second task aimed to assess processing strategies which have previously been linked to medial temporal lobe (MTL) regions (Poldrack et al., 2001; Myers et al., 2003).

**Methods**

**Participants**

Eleven patients with selective BG lesions (three women and eight men) and 18 healthy control subjects (six women and 12 men) participated in this study. The patients were out-patients of the Department of Neurology at the Klinikum Dortmund, Germany. Lesion aetiologies were haemorrhages in two patients (Patients 3 and 6) and ischaemic events in the remaining patients. BG lesions were documented with MRI using a standard three-dimensional $T_2$-weighted sequence for transverse sections (1 mm × 5 mm × 5 mm voxel size). For five patients, a $T_1$-weighted sequence for coronal sections was also available (1 mm × 5 mm × 5 mm voxel size). In the other six patients, one coronal MR image was available, which illustrated the extent of the lesion in all patients with the exception of Patient 4 (see Fig. 1 for MR images of lesion locations).
In two patients (Patients 9 and 10), BG lesions were incidentally detected within the context of neurological examination because of an epileptic seizure (Patient 9) or a transient ischaemic attack (TIA) (Patient 10). In both patients, MR scanning revealed a BG infarct, but no further abnormality which could have caused the temporary neurological symptoms (TIA-symptoms in Patient 10 affected the left side and could thus not be caused by the left BG lesion). The exact lesion onset in these patients remains unknown. Most probably, the ischaemic event occurred at least a few months before MR examination. For the remaining nine patients, the average lesion test interval was 52 months (SD = 25; see Table 1).

Based on the MR images, lesion locations were determined by two experienced neurologists (M.S. and B.K.), using an established atlas (Mai et al., 1997) and recent reports of anatomical subdivisions of the human BG (Karachi et al., 2002; Lehericy et al., 2004, 2006). Lesions are described in terms of affected BG structure [caudate nucleus, Put or Globus Pallidus] and in addition classified along the anterior–posterior and ventral–dorsal
dimensions (see Table 1), which are of particular relevance for feedback-based learning (O’Doherty et al., 2004; Seymour et al., 2007). In Patient 4, the lesion was not visible on the coronal MR image, but the transverse images suggest a dorsal lesion in this patient. In some patients, lesions affected anterior and posterior parts of the striatum. Similarly, lesions could not clearly be classified as dorsal or ventral in three patients, as both subregions appeared to be affected to some extent. These lesions were classified as ventro-dorsal.

At the time of participation in this study, eight patients did not have any neurological symptoms (Table 1). Patients 1 and 3 still suffered from motor problems affecting the right (Patient 1) and left side (Patient 3), respectively. Arm and hand movements were impaired, and the patients therefore completed the learning tasks with one rather than two hands (see below). Patient 5 still complained of reduced motor strength in her upper extremity, but she did not have any problem performing the required movements with either hand.

Healthy control subjects were recruited from a database in the Department of Neuropsychology of the Institute of Cognitive Neuroscience to match the patients for age and general intellectual ability. Mean age was 61.5 years (SD = 11.0) in the patients and 62.2 years (SD = 9.2) in the control subjects (P > 0.840). An IQ estimate was obtained using the subtests ‘Picture Completion’ and ‘Similarities’ of a German version of the Wechsler Adult Intelligence Scale (Dahl, 1972): the mean IQ estimates of patients and control subjects were 114.1 (SD = 8.3) and 114.2 (SD = 8.0), respectively. The groups did not differ significantly on this measure (P > 0.990).

All participants had normal or corrected-to-normal vision and all subjects were right-handed. Exclusion criteria were history of psychiatric or neurological disease (other than the BG lesion in the patients) and regular use of any medication affecting the central nervous system. The study was approved by the Ethics Committee of the Medical Faculty of the Ruhr-University of Bochum and each subject signed an informed consent form according to the Declaration of Helsinki. The subjects were paid a minimum of €20 for reimbursement of expenses and participation. Depending on the performance of the subjects in the reward learning tasks, they could increase this sum (see below).

**Reward-based acquisition, reversal and single-symbol reversal**

Subjects had to learn associations between four abstract Asian symbols and two colours (red and green) with the help of monetary feedback. On each trial, one of the four symbols was shown on the screen for 2.5 s, followed by two coloured circles (red and green), in left and right positions on a monitor, which were shown for up to 3 s. During presentation, subjects had to select one of the two colours by pressing a left or right response button. Following the button press, the circle of the chosen colour turned to white. After a 1 s interval (black screen), the choice either led to a monetary reward—indicated by a coin in a white circle in the middle of the screen—or to non-reward, indicated by three empty white circles (see Fig. 2A for the sequence of events on a given trial). Unknown to the subjects, each of the four symbols was probabilistically associated with red or green: For each symbol, the choice of one colour (e.g. red) led to reward in 80% of the cases whereas the choice of the other colour (e.g. green) always led to non-reward. Two symbols were associated with red, the other two symbols with green. For two of the symbols (one associated with red, the other with green), correct choices were associated with a 20-cent reward and for the other two symbols with a 5-cent reward.

The learning task consisted of four phases (see Fig. 2B). Acquisition comprised 120 trials, divided into three blocks of 40 trials, i.e. each symbol was presented 30 times, 10 times per block. For each block, the locations of the red and green circles relative to the centre (left or right) were counterbalanced. Similarly, the symbol–colour associations and the reward magnitudes associated with each symbol were counterbalanced across subjects.

In the reversal phase, the symbol–colour associations were reversed: Symbols initially associated with red were now associated with green and vice versa. As acquisition, reversal consisted of three blocks of 40 trials each. In the third phase (single-symbol reversal), the symbol–colour associations were again reversed, back to the initial contingencies of acquisition. However, only two of the symbols were presented (one associated with red, one with green—‘learned symbols’) for new learning. When subjects had
learned the new reversal—as indicated by five correct responses in a row with a minimum of 15 and a maximum of 50 trials—a test phase started, in which all four initial symbols were presented again (the two used for the second reversal learning as well as the two original symbols not used during new learning). In this last learning phase, subjects did not receive feedback on individual trials. To respond correctly, subjects needed to transfer the reversed contingencies to the two symbols which were not presented during the initial part of the single-symbol reversal phase (‘transfer symbols’). To keep up motivation, subjects were informed about the cumulative winnings every five trials. The test phase consisted of 40 trials, 10 for each symbol.

**Reward-based acquired equivalence learning**

The probabilistic acquired equivalence task also entailed associations between four Asian symbols and two colours. Each symbol was again associated with one particular colour, and two symbols were associated with the same colour. A correct choice led to reward in 80% of the trials, whereas the alternative choice was never rewarded. The structure of each trial was identical to the first experiment (see Fig. 2A); reward magnitude was, however, kept constant (5 cents), and new symbols and colours (brown and pink) were used (Fig. 2C). Symbols were randomized across trials and the symbol–colour associations were counterbalanced across subjects. The acquired equivalence task involved three phases (see Fig. 2C).

During initial acquisition, participants had to learn that two symbols were associated with the colour pink and the other two symbols were associated with the colour brown. Acquisition was completed when the subject had reached a criterion of eight correct responses in a row (minimum of 38 trials; maximum of 80 trials). In the second acquisition phase, new colours were presented (blue and yellow) and only two of the four symbols (one initially associated with pink, the other with brown) were used. One of the symbols was now associated with yellow, the other one with blue. In the test phase, the newly learned associations between the new colours and the symbols had to be transferred to the other two symbols (used during initial learning), applying the rule that two symbols are equivalent in the sense that they are always associated with the same colour. The test phase started, when subjects had correctly responded five times in a row in the second acquisition phase (minimum of 13 trials, maximum of 80 trials). The test phase was comparable to the test stage of the single-symbol reversal task. It consisted of 40 trials and there was no feedback on individual trials. Subjects were informed about their cumulative winnings every five trials.

**Post-experimental interview**

After subjects had completed the first reward learning task and the acquired equivalence learning task, respectively, structured interviews were conducted to assess, if subjects had gained insight into reward contingencies and the structure of the task. The first part of each interview consisted of open questions. For the first task, subjects were asked ‘Were there different stages of the task?’ ‘What rule did you apply when making your choices?’, ‘Did the rule change during the course of the experiment?’ and ‘Did you know how to make your decision in the final part of the test, when there was no feedback?’. In the second part of each interview, subjects were asked questions with several response alternatives. In response to the question ‘If the rule changed during the task, how often did it change’, they could, for example, choose between ‘never’, ‘once’, ‘twice’, ‘three times’ and ‘more than three times’.

To obtain an overall insight score, the responses to the free and the structured questions were combined with respect to four different components of the task: (i) the symbol-colour association; (ii) the initial change of contingencies; (iii) the second change of contingencies; and (iv) application of the correct rule in the test phase of single-symbol reversal. The responses were scored on a 0 to 2 point scale (ranging from no insight to partial to full insight), yielding a maximum score of eight for the complete interview.

For the acquired equivalence task, a set of similar questions was administered. The maximum score was six, maximally two each for (i) insight into symbol–colour associations at the beginning of the task; (ii) symbol–colour associations after the colour change; and (iii) insight into symbol equivalence. For both tasks, subjects were also asked, if they had difficulty in discriminating between the different symbols. There was no feedback about the scores.

**Working memory**

To control for a possible mediation of reward-based learning by working memory problems, the working memory scores of the BG patients from a clinical neuropsychological screening were analysed. The task was a 2-back task, in which subjects had to...
press a response button, when a digit appearing on a computer screen matched the digit on the penultimate trial (Working Memory subtest of an established German attention test battery, Zimmermann and Fimm, 1993). The number of misses and the reaction times (RTs) were recorded. As working memory was not assessed in the control subjects of the present study, the BG patients’ scores were compared to the scores of 13 healthy age- and IQ-matched control subjects (mean age = 64.0 years; SD = 8.52 years; mean IQ = 107.3; SD = 12.59), selected from our registry of healthy volunteers who underwent extensive neuropsychological assessment and serve as a reference sample for clinical studies. To further examine the relationship between working memory performance and reward-based learning in the patients, working memory scores were correlated with a range of learning variables.

**Stimulus presentation and data recording**

To control the timing of all stimuli and to record subjects’ responses, Presentation software was used (Neurobehavioral Systems Inc.; http://www.neuro-bs.com)

**Procedure**

Following general information about the aims and tasks of the investigation, a semi-structured interview assessed educational status, previous and current health status and medication. In the patients, subjective symptoms associated with the BG lesion were assessed. Then the first reward-learning task was administered, followed by the post-experimental interview. After completion of the IQ subtests, the acquired equivalence task was administered, followed by the post-experimental interview. The testing session lasted ~1 hour and 45 min.

**Statistical analysis**

Statistical analyses were performed using SPSS 15.0. Data sets were analysed by means of analysis of variance (ANOVA) or t-tests, where appropriate. The significance level was 0.05 (two-tailed).

**Results**

**Reward-based acquisition, reversal and single-symbol reversal**

**Group analysis**

Figure 3 illustrates the performance of patients and controls for acquisition and reversal of symbol–colour associations. For acquisition, ANOVA with factors GROUP (patients versus controls), BLOCK (1 to 3) and REWARD MAGNITUDE (RM; 5 versus 20 cent) yielded a main effect of BLOCK [linear trend: F(1,27) = 43.975, P < 0.001] as well as a significant interaction between BLOCK (linear trend) and RM [F(1,27) = 5.642, P = 0.025]. All other main effects and interactions, including GROUP effects, did not reach significance (all P > 0.275). The BLOCK × RM interaction was due to a highly significant linear trend on trials rewarded with 20 cent [linear trend: F(2,27) = 29.798, P < 0.001], while the trend did not reach significance for 5-cent trials (P > 0.160).

In the reversal phase, a significant GROUP effect reflected a significantly higher number of correct responses in the control subjects compared to the patients F(1,27) = 6.121, P = 0.020). Apart from a near-significant BLOCK effect (P = 0.078), none of the other effects approached significance (all P > 0.110).

To explore GROUP effects on cognitive transfer in the test phase of single-symbol reversal, an ANOVA with factors SYMBOL TYPE (learned versus transfer) and GROUP was conducted (see Fig. 4). The significant SYMBOL TYPE effect indicated higher scores on trials with symbols for which subjects had learned the new reversal [F(1,27) = 4.462, P = 0.044]. None of the effects involving the GROUP factor reached significance (both P > 0.200). For both patients and controls, the performance for the transfer symbols was not above chance level (i.e. 10 correct responses; both P > 0.160).

**Single case analysis**

For single-case analysis of acquisition performance, individual patients’ z-scores were computed for the total number of correct responses in the acquisition phase, relative to
the scores of the 18 control subjects. Interestingly, Patient 7 showed the best acquisition performance of all subjects. With 91 correct responses (maximum 120) and a $z$-score of 2.34, Patient 7 outperformed the control subjects. None of the patients showed a significantly reduced score.

For the reversal phase, the pattern was different. Three individual patients—Patient 1, Patient 4 and Patient 11—showed a significantly lower number of correct responses relative to controls, ($z$-scores of $1.77$, $1.85$ and $1.93$, respectively).

For the transfer task, two patients had significantly higher scores than control subjects for the transfer symbols. Their scores were 20 (Patient 3) and 19 (Patient 7) correct responses (maximum score: 20), yielding $z$-scores of 2.10 and 1.92, respectively. See Table 2 for lesion locations and significant alterations of learning-parameters in individual patients.

Post-experimental interview

On the group level, BG patients and control subjects did not differ with respect to the level of insight into the rules underlying the task ($P = 0.986$). On average, the insight score of the control subjects was 3.17 (SD = 1.69), the mean score of the patients was 3.18 (SD = 1.96). Three patients (Patients 3, 10 and 7) reached very high insight scores (6 to 8). In the controls, the maximum score was six, which was only achieved by one subject. Patient 7 was thus the only participant who gained full insight into all aspects of the first reward-learning task, including the rules underlying reward during acquisition, reversal and single-symbol reversal.

Reward-based acquired equivalence

Group analysis

Figure 5 illustrates the mean number of trials to reach a learning criterion (eight correct responses in the last 10 trials) during acquisition of the acquired equivalence task. Control subjects were significantly faster in reaching this criterion than patients ($t(27) = -2.611, P = 0.015$).

For the analysis of transfer performance in the test phase, only those subjects were considered, who (i) reached this learning criterion within 80 trials; and (ii) also reached the learning criterion in the second acquisition phase, in which associations between two symbols and two new colours had to be learned (five correct responses in a row). All 18 control subjects and nine patients fulfilled these criteria. Figure 6 illustrates performance for the test phase. ANOVA with factors GROUP and SYMBOL TYPE yielded a significant effect of SYMBOL TYPE, with more correct responses for symbols used to learn the new associations in the second acquisition phase compared to transfer symbols ($F(1,27) = 4.558, P = 0.043$). The main GROUP and the interaction effects did not reach significance.
of 1.81 (see Table 2). There were no significantly reduced maximum number of 20 correct responses and a significantly above the control subjects, reaching the all correct responses to transfer symbols. Patient 7 scored respectively (see Table 2).

For initial acquisition, the total number of correct responses in the patients showed a near-significant negative correlation with the number of misses on the 2-back task ($r = -0.556; P = 0.076$). The correlation between the RTs and acquisition performance was not significant ($r = -0.234; P = 0.489$).

For the reversal phase, neither accuracy nor speed of working memory performance correlated significantly with working memory performance (number of misses: $r = -0.215; P = 0.525$; RTs: $r = -0.208; P = 0.541$). Similarly, there were no significant correlations of misses or RTs with the number of trials to reach the learning criterion during acquisition of the acquired equivalence task (misses: $r = 0.214; P = 0.528$; RTs: $r = 0.301; P = 0.369$) or with transfer following single-symbol reversal (misses: $r = -0.451; P = 0.161$; RTs: $r = -0.152 P = 0.655$) or acquired equivalence (misses: $r = 0.197; P = 0.612$; RTs: $r = 0.031; P = 0.937$).

**Discussion**

Errors in reward prediction play a key role in learning from feedback. While prediction errors are coded by dopaminergic neurons in the midbrain, the BG, which receive dopaminergic projections, appear to be critically involved in the learning process. BG activations have been observed in a variety of different feedback-learning paradigms (see Delgado, 2007), and learning deficits in Parkinson’s disease patients were reported to correlate with BG dysfunction (Cools et al., 2007). It is, however, as yet unclear, which parts of the BG and striatal regions in particular are responsible for different cognitive processes involved in feedback-based learning.

The present study addressed this issue by administering different reward-based learning tasks to patients with focal BG lesions and by comparing their performance to healthy control subjects. In the first task, acquisition and reversal of outcome-based stimulus–stimulus associations and the processing of different reward magnitudes were studied. Patients and control subjects showed comparable learning performance during acquisition, with evidence of faster learning with larger reward magnitudes. On the single case level, there was no evidence of a significant impairment in any of the patients. In fact, one patient with a lesion in the left ventro-dorsal posterior Put, extending into the external Globus Pallidus and the tail of the caudate, and a small lesion in the right ventral Put showed evidence of better learning than control subjects.

When contingencies changed during reversal, there was clear evidence of impairment in the BG patients. Three patients with left-sided, dorsal striatal lesions were most severely impaired, with two of the lesions affecting the posterior Put and one the anterior Put. In the transfer test following a further reversal, neither healthy control subjects nor patients performed above chance level. A recent study in our lab yielded consistent above-chance level performance and thus significant transfer in young healthy patients.

(both $P > 0.870$). As in the single-symbol reversal test phase, patients as well as controls did not achieve above-chance performance for transfer symbols (both $P > 0.190$).

**Single case analysis**

In the acquired equivalence task, acquisition performance was assessed in terms of number of trials to reach a learning criterion. Z-score analysis yielded highly significant impairments in six patients. Patients 2, 5, 6, 9, 10 and 11 had $z$-scores of 5.40, 3.22, 7.10, 4.55, 2.98 and 3.10, respectively (see Table 2).

Cognitive transfer was evaluated by analysing the sum of all correct responses to transfer symbols. Patient 7 scored significantly above the control subjects, reaching the maximum number of 20 correct responses and a $z$-score of 1.81 (see Table 2). There were no significantly reduced $z$-scores in any of the patients.

**Post-experimental interview**

With a mean score of 3.50 ($SD = 1.62$), control subjects showed a slightly higher level of insight in the acquired equivalence task than the patients (mean score = 2.55; $SD = 1.97$). This difference was, however, not significant ($P = 0.167$). Only three control subjects and one patient (Patient 7) had full insight into all aspects of the task, reaching the maximum score of six.

**Working memory performance in the patients**

The patients missed only a small number of stimuli in the working memory task (1.18, $SD = 1.08$), which did not differ significantly from the matched reference sample (0.69, $SD = 1.11$; $P = 0.287$). Similarly, the RTs did not differ between the two groups [mean RT patients: 651 ms ($SD = 181$ ms); mean RT controls: 637 ms ($SD = 157$ ms); $P = 0.835$].
subjects on the same task, while older subjects showed some problems with cognitive transfer. The surprisingly poor performance of the control subjects of the present study (mean age 61.4 years) may thus be attributed to age effects. On the level of individual performance, two patients showed clear evidence of good transfer [one bilateral lesion patient (see above) and one patient with a lesion in the right ventral posterior Put].

In a second reward-based learning task involving acquired equivalence of stimuli, BG patients showed deficits relative to control subjects in the acquisition of stimulus–stimulus associations based on reward and non-reward. As a group, they needed a larger number of trials to reach a learning criterion. Six patients, with lesions affecting different parts of the Put, appeared to be disproportionately impaired. In contrast to young healthy subjects, who showed robust and significant transfer effects on the same task in our lab, patients and controls performed relatively poorly on the transfer task (i.e. the test of acquired equivalence), not being able to reach above-chance performance levels. The bilateral lesion patient (see above) was the only subject with a 100% correct response rate, reflecting significantly better performance than controls.

Post-experimental interviews revealed that the vast majority of subjects did not gain insight into reward contingencies and other aspects of the tasks such as reversal and stimulus equivalence. This is surprising, given that only four different symbols and two colours were involved in the associations. The probabilistic nature of the associations was clearly the main problem which prevented full insight into the structure of the tasks. Some subjects reported that they initially attempted to apply explicit strategies such as specific symbol–colour associations. After experiencing that this ‘didn’t work’ (i.e. it did not work 100% of the time), they gave up searching for explicit verbalizable rules. In the following sections, the patients’ deficits in the reversal phase of the first learning task and in the acquisition phase of the second learning task will be addressed, followed by a discussion of the relationship between learning performance and other neuropsychological measures and the issue of the role of implicit versus explicit learning strategies.

Striatal involvement in reversal learning

The most striking deficit observed in the BG patients was impaired reversal learning, despite fully intact acquisition. Reversal learning deficits have consistently been observed in Parkinson’s disease patients on L-DOPA medication (Swainson et al., 2000; Cools et al., 2001) and—in accordance with the present findings in focal BG lesion patients—Cools et al. (2001) observed intact acquisition of probabilistic response outcome associations. Patients off medication, on the other hand, were neither impaired in acquisition nor reversal. Cools et al. (2001) discussed the differential effects of medication on the two types of learning in terms of the ‘DA-overdose’ hypothesis, according to which L-DOPA leads to an overdose of DA in the relatively spared ventral striatum in the initial stages of the disease (Kish et al., 1988; Seger, 2006). In accordance with this hypothesis, DA alters reversal-related activity in the nucleus accumbens of the ventral striatum in Parkinson’s disease patients, but not in the dorsal striatum (Cools et al., 2007).

In contrast to these findings and although the ventral striatum is connected with the orbitofrontal and the ventromedial prefrontal cortices (Karachi et al., 2002; Lehericy et al., 2004), which have been implicated in reversal learning (see Clark et al., 2004; Kringelbach, 2005), the significant reversal deficit in the patients of the present study appears to be mainly driven by the performance of patients with dorsal lesions. Three dorsal lesion patients showed significant deficits relative to controls (Patients 1, 4 and 11, see Results section). Three further patients with dorsal striatal involvement scored more than 1 SD below the controls (z-scores of $-1.05$, $-1.45$ and $-1.13$ for patients 5, 8 and 9, respectively), thereby contributing to the group result. The scores of at least two of the three patients with exclusive ventral striatal lesions (Patients 3, 6 and 10) did not follow this pattern ($z$-scores of 0.77, $-0.26$, $-0.90$). This interpretation needs, however, approval by future research based on larger samples and more selective damage to the dorsal or ventral striatum. The present data do not allow any firm conclusions as to why some dorsal lesion patients were more impaired than others, and why one dorsal lesion patient did not show a reversal impairment at all (Patient 2).

Rather than reflecting reversal as such, poor reversal performance in the context of reward-based learning might reflect impaired negative feedback processing. Consistent with this hypothesis, Seymour et al. (2007) reported a functional dissociation within the human striatum, with activity in the ventral anterior (including the nucleus accumbens) and mid- to dorsal posterior striatum (i.e. Put) reflecting positive and negative prediction errors, respectively. The BG lesions of at least two of the three patients with significant reversal deficits most likely comprise the dorsal striatal region described by Seymour et al. (2007). Although it appears plausible that reversal learning following successful acquisition is more closely related to learning from negative than positive feedback, additional studies are clearly needed to clarify this issue, as our tasks did not allow a distinction between impairments related to learning from positive or negative feedback.

Another possible explanation of impaired reversal learning in patients with dorsal striatal lesions might be related to the functional dissociation within the striatum described by O’Doherty et al. (2004). Based on the actor–critic model of reward-based learning (see Sutton and Barto, 1998), they studied striatal activations during the learning of associations between stimuli and outcomes with and without the need for an active response (instrumental
versus classical conditioning). Dorsal striatal activity was only seen during instrumental conditioning, suggesting that this striatal region might correspond to the actor responsible for the application of a response policy, which maximizes future reward. In the present study, the necessity to change the policy in the reversal phase might have critically depended on the dorsal striatum, leading to significant reversal deficits in patients with selective dorsal lesions. Although stimulus–stimulus–outcome associations had to be learnt, the second stimulus (i.e. the colour) had to be chosen with an active response, linking the learning task to the actor component of reward-based learning.

In the animal literature, specific reversal learning impairments have frequently been observed in relation to striatal dysfunction (see Ragozzino, 2007 for a review of the rat literature). Paralleling the performance patterns of medicated Parkinson’s disease patients and the focal BG lesion patients of the present study, reversal learning deficits were associated with elevated DA levels, e.g. following the administration of endogenous DA in DA deficient mice (Robinson et al., 2005), or following the administration of a DA antagonist in monkeys (Lee et al., 2007). Robinson et al. (2005) concluded that reversal learning depends on a fine-tuning of striatal function, implying that both striatal hyper- as well as hypofunctioning may lead to disruption of learning. It should be noted that most reversal learning studies in rats involved associations between a specific response or behaviour and an outcome (e.g. Divac, 1971), which differs from the features of the tasks administered in the present study. Reversal deficits have, however, also been observed for stimulus–reward associations, e.g. in monkeys (Lee et al., 2007). In rats, lesions of the dorso-medial striatum have been related to reversal learning in different contexts, e.g. associations between reward on the one hand and spatial arrangements or egocentrically defined responses on the other hand (see Ragozzino, 2007).

### Acquisition deficits in BG lesion patients?

The results for initial acquisition in the first task indicate that basic mechanisms of feedback-based learning were fully intact in BG lesion patients. In accordance with the model of Rescorla and Wagner (Rescorla and Wagner, 1972) and with striatal activation patterns in humans (Delgado et al., 2003), both patients and control subjects showed faster learning for larger rewards, i.e. larger positive prediction errors.

It should be noted, however, that acquisition on the first task was generally quite slow, with both patients and controls needing more than 50 trials to reach a criterion of eight correct responses in a sequence of 10 trials (not reported in the Results section). With a mean age of more than 60 years, this finding is probably linked to age-related changes in learning from feedback (Mell et al., 2005), which is presumably linked to age-associated degenerative changes in the striatum or the PFC (see Raz and Rodrigue, 2006). Frontostriatal dysfunction of reward-related processing in healthy elderly subjects shows some resemblance to the pattern observed in Parkinson’s disease patients (Schott et al., 2007).

During acquisition of acquired equivalence, learning was considerably faster. Control subjects needed only about 20 trials to reach the learning criterion. This effect might be mediated by an implicit learning effect across the tasks, since subjects did not receive any explicit information about the underlying rule determining reward frequency or feedback about their own hypotheses before the end of the testing session. The BG patients, on the other hand, did not show evidence of a considerable carry-over effect. On average, they still required 37 trials to reach the criterion. Potential acquisition problems did not occur in initial acquisition of reward-based learning in the first task, presumably because they might have been masked by the somewhat reduced age-related learning performance in the controls. It is interesting to note that (with the exception of Patient 11) there was no overlap in the subgroups of patients who showed acquisition deficits in the second learning task and reversal deficits in the first task, suggesting that different mechanisms may have been involved in the two types of impairments. The reduction of implicit across-task learning can, however, neither be attributed to learning from negative or positive feedback, because subjects most probably used a combination of both, nor can it be related to a certain lesion location in the striatum, as the lesions affected ventral areas in three of the six patients, and dorsal areas in the other three. The underlying problem might be an unspecific deficit in reward-based learning, which only manifests itself during later learning stages and which is related to both the ventral and dorsal striatum, since both are believed to contribute to the association of stimuli, responses and outcomes (see O’Doherty et al., 2004).

Although patients and controls did not differ significantly on the insight scores derived from the post-experimental interview, it cannot be excluded that superior acquisition performance of the controls in the second learning task was somehow mediated by (explicit) knowledge about the underlying rules. On average, there was evidence of a higher level of insight in the controls and, more importantly, three patients but only one control subject did not show any evidence of knowledge (insight scores of 0). Two of these patients were significantly impaired on acquisition on the second task (Patients 5 and 9) and two further patients with significant acquisition impairments also had very low insight scores (Patients 10 and 11; scores of 2 each).

### The effect of reward magnitude

The processing of different reward magnitudes appears to be intact in the patients, a finding which might argue against a specific impairment of reward-related processes.
In the acquisition phase of the first learning task, learning is clearly faster for large compared to small rewards in both patients and controls, and the learning curves for the reversal phase appear to also suggest faster learning for larger rewards, at least on the descriptive level.

On the other hand, faster learning for larger prediction errors does not necessarily exclude deficits in reward processing. It is conceivable that the signals related to prediction errors are generally attenuated in the patients, but still reflect the relative size of the prediction errors.

As the most pronounced deficit of the patients was observed for reversal learning, it is likely that processes specific for this learning phase were disproportionately impaired in many patients, one possibility among others being problems in learning from negative feedback (see section on reversal deficits). Although this explanation must remain speculative at this stage, it is possible that the processing of different prediction errors was spared for positive but not for negative feedback. As pointed out above, the learning tasks administered in the present study do not allow a clear distinction between learning from positive and negative feedback. The present results do, however, suggest future research issues, in terms of efforts to specifically address the role of several BG regions in the processing of positive and negative prediction errors of different magnitudes.

**Relationship to other neuropsychological deficits**

Given the structure of the learning tasks, working memory might make a significant contribution to learning performance. For the acquisition phase of the first learning task, a near-significant correlation between the patients’ working memory accuracy scores and learning provides some support for this idea. Importantly, however, the BG patients did not differ from controls during initial acquisition. Speed and accuracy of working memory performance did not correlate significantly with learning measures during reversal or acquisition of the second task, i.e. with those measures which yielded impairments in the patient group. Taken together, this pattern indicates that it is unlikely that impaired working memory can account for the observed learning deficits in the BG patient group. Furthermore, the comparison of their working memory scores from clinical neuropsychological assessment with the performance of a matched healthy reference group did not provide evidence of a significant impairment of the BG patients who completed the learning tasks.

With respect to the observed reversal problems in the patients, the question arises as to whether the deficit is specific for reward-based learning. Instead, it might simply reflect a more general impairment of executive processing.

**Implicit versus explicit learning**

Acquired equivalence tasks are frequently used to distinguish between implicit (or procedural) and explicit (or declarative) forms of learning. While acquisition critically involves the BG and does not require verbalizable knowledge, cognitive transfer involves explicit learning and critically depends on MTL processing including the hippocampus (Myers et al., 2003). Hippocampal activity and BG activity are negatively correlated, suggesting an inhibitory link between the two systems (Poldrack et al., 2001). Similarly, the acquisition and reversal phases of the first learning task used in the present study are presumably related to implicit types of learning, whereas the transfer test required stimulus categorization and explicit processing. Post-experimental interviews suggested that most patients and controls did not apply explicit learning strategies, since they did not gain insight into stimulus contingencies. They also did not show
above-chance-level transfer for single symbol reversal or acquired equivalence, which would also argue against explicit processing of the relevant rules. Interestingly, Patient 7, the only patient with a bilateral BG lesion, was the only participant who had become fully aware of stimulus contingencies and who showed (nearly) perfect performance in both tasks, with superior performance compared to control subjects. It is therefore likely that he applied an explicit hippocampus-based strategy in associative reward-based learning. BG-based learning mechanisms might be severely disrupted in this patient, leading to a recruitment of MTL-based alternative strategies. Compensatory mechanisms have also been reported to be responsible for intact error correction in patients with focal BG lesions who showed evidence of altered error processing (Ullsperger and von Cramon, 2006).

Conclusion

To our knowledge, the present study is the first to assess reward-based learning in patients with focal BG lesions. As a group, patients showed deficits in reversal learning, which were most pronounced in patients with dorsal striatal lesions. Their impairments might be related to deficits in learning from negative feedback or to a deficit in changing the action policy. Mild acquisition deficits on a subsequent task might be linked to reduced carry-over learning effects in the striatal lesion patients. A patient with a bilateral BG lesion who used an explicit learning strategy showed superior performance on a range of learning tasks, presumably using this strategy to compensate for deficits in implicit feedback-based learning.

Acknowledgement

We gratefully acknowledge support by the German Ministry for Education and Research (Bundesministerium für Bildung und Forschung, BMBF, Grant Nr. 01 GW0541).

References


The basal ganglia and reward learning

Brain (2008), 131, 829–841


