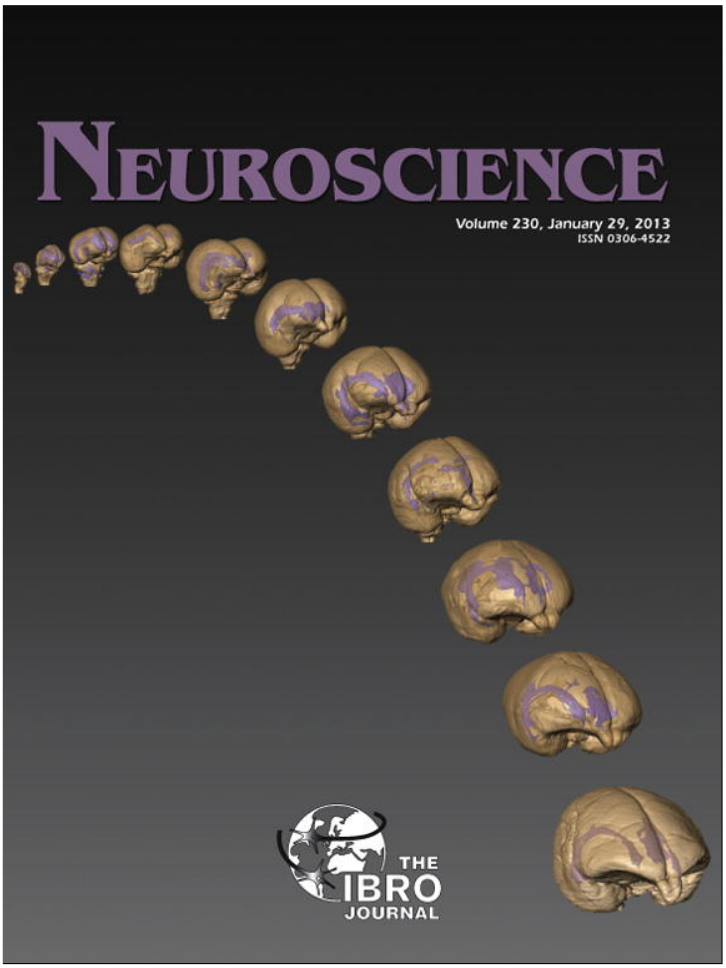

Neuroscience 230 (2013) 132–138

STRIATAL DOPAMINE D1 RECEPTORS ARE INVOLVED IN THE DISSOCIATION OF LEARNING BASED ON REWARD-MAGNITUDE

J. ROSE,* A.-M. SCHIFFER† AND O. GÜNTÜRKÜN

Institute of Cognitive Neuroscience, Biopsychology, Department of Psychology, Ruhr-University of Bochum, GAFO 05/618, D-44780 Bochum, Germany

Abstract—Here we investigate the contribution of striatal dopamine receptors (D1) to the influence of reward-magnitude on learning. Pigeons (*Columba livia*) were trained on a discrimination-task with two pairs of stimuli; correct discrimination resulted in a large reward in one pair of stimuli and in a small reward in the other pair. Acquisition of the discrimination-task was accompanied by intracranial injections to the medial striatum, either of a dopamine-antagonist (Sch23390) or of vehicle. In the control-condition the rate of learning was modulated by the magnitude of the reward; discrimination was learned faster if contingent rewards were large and learning was slower if contingent rewards were small. Following injections of D1 antagonist this effect vanished even though the ability to discriminate between the rewards was unaffected. Interestingly, the mean rate of learning was indistinguishable between the control and antagonist conditions. Consequently, it appears that not learning per se but the effect of reward-magnitude on learning is mediated through D1 receptors in the striatum. We argue that the injections of dopamine-antagonist cause a shift in strategy underlying learning. In the control-condition animals rely on positive feedback and thus learning is affected by the magnitude of the contingent reward; in the antagonist-condition, however, learning might rely on negative feedback and is thus insensitive to reward-magnitude.

Key words: reward, punishment, reinforcement learning, pigeon, behavioral contrast.

INTRODUCTION

The adaptation to environmental needs is crucial for all animals and it often depends on learning from feedback (reinforcement). This feedback can come in two qualities either as positive (reward) or as negative feedback (punishment), but it can also come in different quantities. A long tradition of psychological research was concerned with the effects of reward magnitude on behavior and in particular on learning (Guttman, 1953; Maher and Wickens, 1954; Denny and King, 1955; Collier and Marx, 1959; Neuringer, 1967). Several studies reported the 'Crespi-Effect', also called effect of 'behavioral contrast'. In trained animals, an increase in the amount of reinforcement delivered for a response leads to an increase in response-strength. Interestingly, this effect strongly depends on subjective experience rather than on the objective amount of the reinforcer. In other words, the animal needs to experience the increase in reward-magnitude to exhibit an increase in response-strength (Crespi, 1942). These behavioral data are now complemented by electrophysiological recordings from dopaminergic neurons demonstrating that cellular responses are scaled to the maximal reward available (Tobler et al., 2005) and to the probability of reward (Fiorillo et al., 2003).

The temporal difference (TD) model, today the most widely accepted model of animal reinforcement learning, builds on the activity of dopaminergic neurons. The TD model was originally devised for machine-learning but was then successfully used to formalize the role of dopamine in animal learning (Sutton and Barto, 1981). Early in conditioning dopaminergic neurons respond to reward-delivery and with progression of learning this response shifts to the preceding cue (Schultz et al., 1992, 1997). In the terminology of TD, the cells encode the error of reward-prediction, a signal used by the model to instruct new learning (Schultz, 2002, 2007; Bayer and Glimcher, 2005). Numerous studies involve dopamine in such a learning process (Wickens, 1990; Schultz, 2002; Wickens et al., 2007) and it was recently demonstrated that the activation of dopaminergic neurons alone is sufficient to elicit conditioning (Tsai et al., 2009).

Dopaminergic neurons project to several cortical and subcortical sites, most commonly noted in the context of TD-learning are the projections to striatum and to cortex. The striatum and in particular the 'corticostriatal loop' is crucial for action-outcome learning (Graybiel, 2005; Wickens et al., 2007) and is a site of dopamine-mediated plasticity (Calabresi et al., 1992, 1997; Kerr and Wickens, 2001; Reynolds et al., 2001; Wang et al., 2006). Lesions to the striatum or the disruption of striatal dopaminergic function cause deficits in learning, both in mammals (Clarke et al., 2008) and in birds

*Corresponding author. Present address: The Picower Institute for Learning and Memory, Massachusetts Institute of Technology, 43 Vassar Street, 46-6241 Cambridge, MA 02139, USA. Tel: +1-617-252-1469; fax: +1-617-258-7978.

E-mail addresses: jonasr@mit.edu (J. Rose), anne-marike.schiffer@psy.ox.ac.uk (A.-M. Schiffer), Onur.guentuerkuen@rub.de (O. Güntürkün).

† Present address: Department of Experimental Psychology, University of Oxford, South Parks Road, OX1 3UD Oxford, UK.

Abbreviations: GPe, globus pallidus external; GPi, globus pallidus internal; MSt, medial striatum; SNr, substantia nigra pars reticulata; SpL, spiriformis lateralis; TD, temporal difference.

0306-4522/12 $36.00
http://dx.doi.org/10.1016/j.neuroscience.2012.10.064

(Watanabe, 2001; Kabai et al., 2004). In addition, expression levels of striatal D1-receptors are altered when learning stimulus–response tasks in pigeons (Herold et al., 2012).

In the present study we investigate the contribution of D1 receptors in the medial striatum (MSt) of pigeons to reward-based learning. The dopaminergic system in birds and mammals has comparable organization and function (Reiner et al., 2005; Herold et al., 2008; Güntürkün, 2012). Also, the basal ganglia of birds and mammals show a similar organization with regard to cell-type, neurochemistry and connectivity (Reiner et al., 2005). Comparable to mammals, the avian striatum gives rise to a direct and an indirect pathway. The indirect pathway projects via globus pallidus external-like (GPe) neurons to the subthalamic nucleus which in turn projects to globus pallidus internal-like (GPi) neurons, to the substantia nigra pars reticulata (SNr) and to a pretectal nucleus called spiriformis lateralis (SpL). The projections to GPi and SNr increase inhibition on the motor thalamus and therefore have a net inhibitory effect on movements (Jiao et al., 2000). In the direct pathway, the striatum directly projects to GPi, SNr and SpL. These projections cause disinhibition of the motor thalamus and thereby promote movements (Reiner et al., 1998). There are three notable differences between the basal ganglia of birds and mammals: The projection to SpL is found in birds but not in mammals (Reiner et al., 1998); while separated in mammals, GPi and GPe neurons are intermingled in birds (Reiner et al., 2005); while intermingled in mammals, striato-pallidal and striato-nigral efferents are separated in birds, arising from the lateral and medial striatum respectively (Veenman and Reiner, 1994; Reiner et al., 2005; Sun et al., 2005).

In a previous study we reported that the quantity of reward delivered during the acquisition of simple discrimination greatly affects the learning rate; color-discrimination is learned faster if contingent rewards are large than if contingent rewards are small (Rose et al., 2009). Here we investigate the contribution of striatal dopamine to this effect. Pigeons (*Columba livia*) were first trained to discriminate between two reward sizes and then they were trained on a simple discrimination-task with two pairs of colored textures. In each pair, correct discrimination was rewarded; in pair one with a large reward, in pair two with a small reward. Training on the discrimination-task was accompanied by intracranial injections to the MSt, either of a dopamine receptor (D1) antagonist (Sch23390) or of vehicle. Injections were conducted prior to each training-session, and the condition (antagonist or vehicle) was maintained over the entire acquisition-period of ten days.

EXPERIMENTAL PROCEDURES

Subjects

The present study used seven adult homing pigeons (*C. livia)* with body weights between 330 g and 490 g. The animals were individually housed in wire-mesh cages inside a colony room, had free access to water and grit and during experiments they were maintained on 80% of their free-feeding body weight. The colony room provided a 12-h dark-light cycle with lights on at 8 a.m. and lights off at 8 p.m. The experiment and all experimental procedures were in accordance with the National Institute of Health guidelines for the care and use of laboratory animals and were approved by a national committee (North Rhine-Westphalia, Germany).

Apparatus and stimuli

All training and testing were conducted in an operant chamber, controlled via PC and parallel-port interface by Matlab (The Mathworks Inc., Natick, MA, USA) and the Biopsychology Toolbox (Rose et al., 2008). Matlab was also used for all data analysis.

Four transparent pecking-keys, situated on the front panel of the box were used to record behavioral responses and to present the stimuli. A TFT-Monitor was placed behind the front panel so that images could be displayed behind the pecking-keys to provide a full back-illumination of the pecking-keys; for a detailed description of the setup, refer to Rose et al. (2009). The stimuli consisted of white-light illumination of two lateral pecking-keys on reward-discrimination trials and colored textures, presented on two vertically aligned central pecking keys on discrimination learning trials. During each ten-day acquisition-period a new set of four textures, consisting of two S+ and two S−, was introduced to the animals so that each animal experienced a total of eight stimuli. For each bird, one S+/S− combination was paired with the chance of gaining a large reward, the other combination with the chance of gaining a small reward. Two feeders were situated below the pecking-keys. One feeder gave access to grain for 4.0 s and the other for 1.5 s and these served as large and small rewards, respectively. Mixed grain was used as the reward. All contingencies (the identity of the S+, stimulus-pair and reward-size, reward-size and side of the reward) were balanced between the animals.

Behavioral task

The birds were trained on two distinct tasks, on a simple discrimination between a large and a small reward and on a simple discrimination of colored textures.

Following training on an autoshaping procedure, the animals were first trained to choose the large reward over the small reward. On these trials an inter-trial interval (ITI) of ten seconds was followed by a presentation of the white-light stimulus on both lateral pecking-keys. A response to either side resulted in reward-delivery from the corresponding feeder below the pecking-key. Since each pecking-key was associated with a feeder and the feeder with a given reward-magnitude, the animals thus chose a reward-magnitude by choosing a pecking-key (Fig. 1B). Omission of a response was mildly punished with a 10-s lights-off. After the animals learned to reliably choose the large reward over the small reward (three consecutive days of at least 80% choice of large reward), they were trained on the color-discrimination task. For a more detailed description of the initial training procedure, refer to Rose et al. (2009). This initial training was performed for two reasons. On one hand, it was aimed at ensuring that all subjects understood the difference between the two feeders (reward-magnitudes) and therefore at reducing inter-individual differences during the main experiment. On the other hand it was intended to reduce potential differences between the first and the second acquisition-period that might result from a difference in familiarity with the setup (feeders and reward-magnitudes).

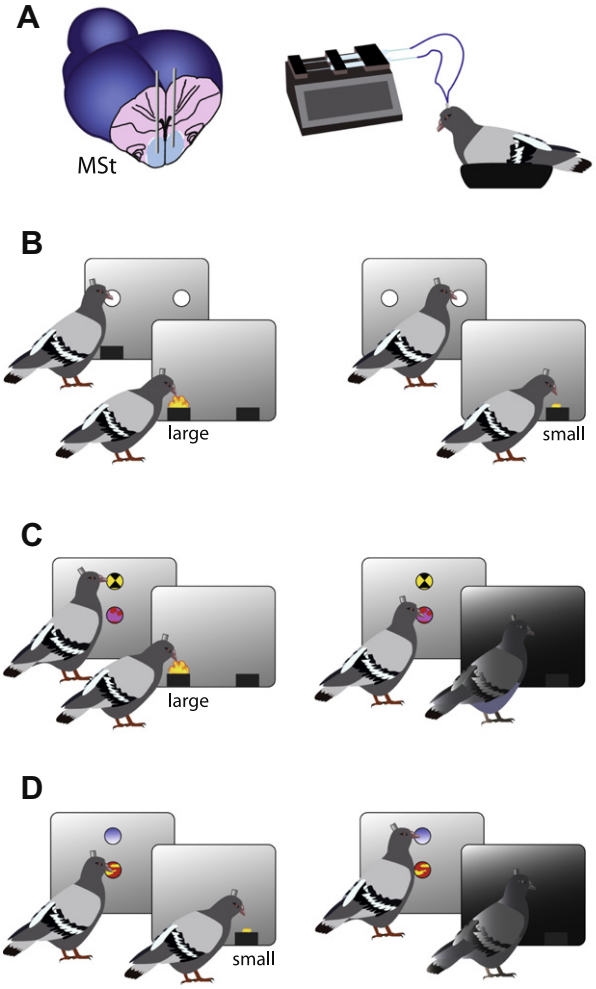

Fig. 1. The behavioral procedure. (A) Prior to each training session the animals received bilateral injections of either Sch23390 or vehicle to the MSt. Injections followed by a within-subject design. Each animal was tested on two blocks of 10 days and during each block one drug was injected and one set of stimuli was acquired. Each test session consisted of reward-choice and discrimination learning trials. (B) Reward-choice trial. The animals choose between a large and a small reward. A response on the left pecking-key was associated with a large reward; a response on the right pecking-key was associated with a small reward. (C) A large reward color-choice trial. One pair of stimuli consisted of an S++, associated with the large reward and an S−, associated with mild punishment. (D) A small reward color-choice trial. One pair of stimuli consisted of an S+, associated with the small reward and an S−, associated with mild punishment. All contingencies were balanced between the animals.

There were 60 trials to each discrimination-learning session, organized in blocks of reward-choice and of learning trials (10 trials of reward-choice, 20 trials of learning, 10 trials of reward-choice, 20 trials of learning). The order of trials was randomized but ensured a balanced presentation of all stimuli and both reward-magnitudes within a given block of 10/20 trials. On discrimination-learning trials, an ITI of ten seconds was followed by the presentation of one pair of stimuli (colored textures) on the two central pecking keys (Fig. 1C). Each stimulus-pair consisted of one S− and either a S+ or a S++. A single peck to the S++ was followed by a large reward and a peck to the S+ followed by a small reward; a peck to either of the S− or response omission was mildly punished with a 10-s lights off. Each animal was trained on two sets of four stimuli, consisting of an S++/S− and an S+/S− that served as learning pairs for the drug and vehicle conditions. Training on each set of stimuli was conducted for 10 days.

Surgery and histology

All animals were chronically implanted with guiding-cannulas (Plastics One Inc., Roanoke, VA, USA, C235G-2.0) and the corresponding dummy-cannulas (Plastics One Inc., C235DC-2.0 and 303DCFT). These cannulas were bilaterally placed in MSt (AP: 12.0, ML: ±1.0, DV: 6.5) according to the stereotactic atlas of the pigeon brain (Karten and Hodos, 1967), for current nomenclature refer to Reiner et al. (2004). Prior to surgery, the animals were deeply anesthetized with 1 ml/kg of a 3:7 mixture of ketamine (Pharmacia & Upjohn Inc., New York, NY, USA, Ketavet) and Xylazine (Bayer GmbH, Leverkusen, Germany, Rompun). After the scalp was cleared of feathers and swabbed with alcohol, the animals were placed in a stereotactic frame. An incision was made to expose the skull, the skull was cleaned and four stainless steel screws were placed in the skull. A small craniotomy was opened above MSt and the cannulas were carefully lowered into the brain. The craniotomy surrounding the cannulas was sealed with medical silicone (Dreve Otoplastik GmbH, Unna, Germany, Biopor AB) and the implant was attached to the screws with dental acrylic. After surgery the animals were allowed to recover for at least five days. Once the experiments were concluded, the animals were sacrificed and the brains fixed in paraformaldehyde. To assess the location of the cannulas, brains were cut and stained with Cresyl Violet.

Injections

Training on the color-discrimination was always preceded by pharmacological intervention; that is intracranial injection of Sch23390 (3 μg/1 μl saline with 1% dimethyl sulfoxide for solubility) or vehicle (physiological saline with 1% dimethyl sulfoxide) to MSt. These injections were always performed in blocks of ten days, during which each animal received injections of one agent and learned one set of four stimuli. Ten days after completion of one block a new set of stimuli and a new drug were used. In other words, each animal was tested on two sets of stimuli so that it received injections of the drug during the acquisition of one set of stimuli and injections of vehicle on during the acquisition of the other set of stimuli. The animals were divided into two groups, one group started with vehicle injections, the other with antagonist injections.

The dopamine antagonist we used, Sch23390 is a selective D1 antagonist with negligible affinity to D2-type dopamine receptors (Barnetta et al., 1986). The concentration and injection parameters we used have successfully been used for intracranial injections in birds (Herold et al., 2008; Rose et al., 2010). While no detailed information is available on the spread of Sch23390 following intracranial injection, injections of TTX to the entopallium of pigeons were showed to spread about 1.5 mm in all directions from the injection-site, an area that should cover the entire MSt and possibly touch on the medial ventricle (Freund et al., 2010).

For the injections, the dummy cannula was removed and replaced with the injection cannula (Plastics One Inc., C235I-2.0 and C232) which was connected to two syringes (Hamilton AG, Bonaduz, GR, Swizerland, Microliter 701, 10 μl). We bilaterally injected 1 μl solution per hemisphere at a rate of 0.2 μl per minute. To control the flow-rate an injection pump (Harvard Apparatus, Hollison, MA, USA, PHD 2000) was used. The injection cannulas remained in place for 5 min after the injections to allow the agent to diffuse freely into the target structure; thereafter the injection cannulas were replaced by the dummy cannulas.

RESULTS

Histological analysis revealed that all injection cannulas were placed within the borders of MSt. However, one animal had to be removed from all subsequent analyses since it failed to acquire the discrimination task altogether (under all reward- and all drug-conditions).

On free-choice blocks, animals consistently chose the large reward over the small reward. This preference was not affected by pharmacological intervention; the animals consistently chose the large reward over the small reward, irrespective of the type of injection. Results are depicted in Fig. 2E, F. Analysis of omission-errors on the discrimination-learning sessions showed that there was no increase in response omissions on D1-antagonist sessions compared to the control sessions (Fig. 2D).

Following injections of vehicle, the acquisition of the discrimination-task was modulated by the magnitude of contingent reward (Fig. 2A). This observation is consistent with previously published results (Rose et al., 2009). For the first analysis we evaluated the number of days until the animals reached criterion (first day with over 85% correct choices). Animals learned the discrimination with the large contingent reward after an average of 1.67 days (SEM: 0.67). Learning the discrimination with the small contingent reward took on average 5.17 days (SEM: 1.27), significantly longer than the discrimination with the large contingent reward (t test, df: 10, $t = -2.4314$, $p = 0.0354$). A repeated measures ANOVA performed on the vehicle injection data (factors training-day and reward-magnitude) revealed a significant main effect of training day (df: 9, F: 4.28, $p = 0.0005$) but no main effect of reward-magnitude (df: 1, F: 2.49, $p = 0.1755$). It did, however, reveal a significant interaction between training-day and reward-magnitude (df: 9, F: 2.66, $p = 0.0144$).

Conversely, following injections of D1-antagonist the effect of reward magnitude on task-acquisition vanished (Fig. 2B). Following antagonist injections, animals learned the stimuli contingent on the large reward after an average of 4.17 days (SEM: 1.57). Stimuli contingent on the small reward took on average 4.67 days (SEM: 1.23), this difference did not reach significance (t test, df: 10, $t = -0.2498$, $p = 0.8078$). A repeated measures ANOVA with factors training-day and reward-magnitude performed on the D1-antagonist injection data revealed a significant main effect of training day (df: 9, F: 7.63, $p = 0.0$) but not of reward-magnitude (df: 1, F: 1.06, $p = 0.3507$). The interaction between training-day and reward-magnitude did not reach significance following antagonist injections (df: 9, F: 0.1, $p = 0.9994$). Furthermore, a repeated measures ANOVA with factors training-day, reward-magnitude and injected drug revealed the main effects of training day (df: 9, F: 6.65, $p = 0.0$) and reward-size (df: 1, F: 8.69, $p = 0.032$) but not of drug (df: 1, F: 0.94, $p = 0.3774$). The interactions training-day∗reward-size (df: 9, F: 0.44, $p = 0.9101$) and training-day∗drug (df: 9, F: 0.25, $p = 0.9854$) did not reach significance, however the interaction reward-size∗drug did reach significance (df: 1, F: 36.24, $p = 0.0$). Taken together these results imply an improvement in performance over time, for the injection of vehicle and D1-antagonist. However, a modulation of this improvement by the magnitude of contingent reward could be observed only following the injections of

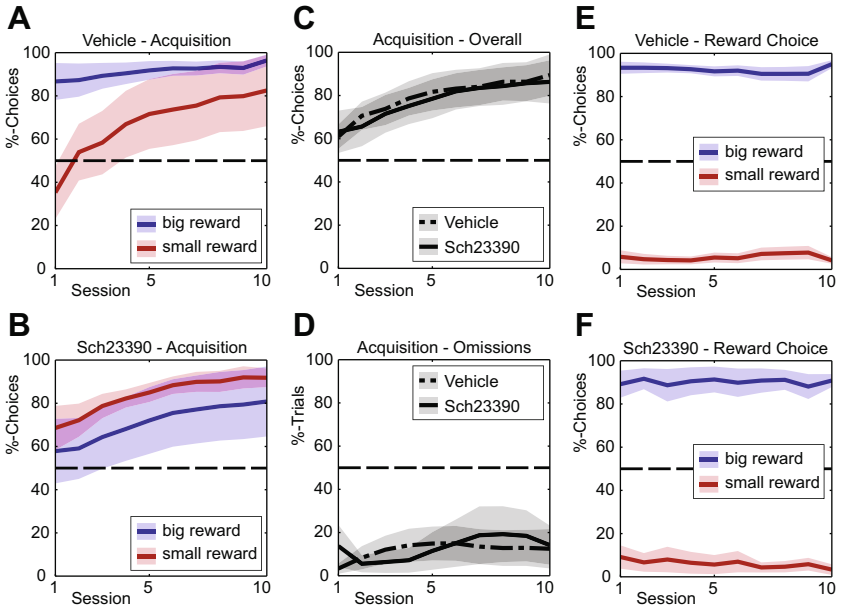

Fig. 2. (A, B) Discrimination-learning is modulated by Sch23390. Following control injections (A), the stimulus-set contingent on a small reward is learned comparably slowly, while stimuli contingent on the large reward are learned very fast, mostly within the first session. This difference in acquisition-speed vanishes following the injections of Sch23390 (B). In this condition both stimulus-sets are learned over the course of several days. (C) Average learning on large-reward and small-reward discrimination trials. The overall acquisition is unimpaired by the injections of Sch23390. (D) Omission-errors on discrimination learning trials. Injections of Sch23390 do not lead to an increase in trial omission-errors. (E, F) The behavior on reward-choice trials following injections of vehicle (E) and injections of Sch23390 (F). Following injection of either drug, animals are unimpaired on the reward-discrimination.

vehicle but not following the injections of the D1-antagonist.

The effect of reward magnitude on task acquisition observed in the control condition vanished following the injection of the D1-antagonist, however, task-acquisition overall (averaged across reward-magnitude) remained undistinguishable between the two drug conditions (Fig. 2C). This result suggests a decrease in performance on the large-reward following injections of D1-antagonist along with an increase in performance on the small reward. Repeated measures ANOVA performed only on large-reward trials (df: 1, F: 28.61, $p = 0.0$) and a separate analysis performed only on small-reward trials (df: 1, F: 11.83, $p = 0.0008$) confirmed this effect.

DISCUSSION

We recently showed that the effectiveness of reinforcement learning is modulated by the magnitude of contingent reinforcer (Rose et al., 2009). Here we further these results and show that this differentiation is dependent on striatal dopamine. Following the injections of the D1-antagonist, the animals learned a color-discrimination on large-reward and small-reward trials equally fast. Interestingly, the animals were unimpaired in learning per se and their preference for the large reward was equally unaffected. Only the ability to dissociate learning based on contingent reward, that is to learn faster on the large-reward condition and slower on the small-reward condition, was impaired.

At a first glance these results seem to contradict the literature. The most prevalent model on reinforcement-learning relies heavily on dopamine-release as a training-signal (Schultz, 2002, 2007) and the striatum is a site of D1-mediated plasticity in the context of learning (Reynolds et al., 2001; Wickens et al., 2007). Yet, we report overall intact learning after the injections of D1-antagonist to the striatum. One possible explanation for this disparity might be an alternative view of the dopaminergic function. In the 'incentive salience' model dopamine is not involved in learning but in the attribution of incentive salience, or 'wanting' (Berridge and Robinson, 1998; Berridge, 2007). Such an account could explain why learning remained intact in our study. The model would also predict a general reduction in motivation (wanting) on all types of trials. We did, however, not find evidence for such impairment. The animals performed normally on reward choice-trials (Fig. 2E, F) and response omissions were not increased during discrimination trials (Fig. 2D).

A closer look at the role of dopamine-release in learning might lead to a more parsimonious explanation of our data. As aforementioned, dopamine-release is often modeled as providing a prediction-error that encodes the difference between obtained and anticipated reward and that is sensitive to the value of the reward. This account thus involves dopaminergic neurons in learning from *positive* feedback. With regards to negative feedback, the role is less clear. Some studies did not find evidence for a role of dopaminergic neurons in learning from negative feedback (Bayer and Glimcher, 2005; Joshua et al., 2008). Others report that the activity of dopamine neurons drops below the baseline when an expected reward is omitted, thus providing a negative prediction-error that is involved in extinction (Pan et al., 2008). Interesting contributions to this discussion come from work with Parkinson patients, a disease caused by the loss of dopaminergic neurons and commonly medicated with dopamine-agonistic drugs. Frank et al. (2004) showed that, in such patients, the effectiveness of the type of reinforcement depends on medication. When off medication, the patients relied heavily on negative feedback, when medicated, however, the patients were relying more on positive feedback. This effect was very pronounced, when off medication, patients even showed enhanced learning from negative feedback compared to elderly controls. The authors explain their results with a dichotomy in dopamine-function (Frank and O'Reilly, 2006). According to their model, positive and negative feedback control plasticity in a 'go-pathway' (direct-pathway) and in a 'nogo-pathway' (indirect pathway). Positive feedback triggers dopamine-release, thereby inducing plasticity into the 'go-pathway' as mediated by D1-receptors and inhibiting the 'nogo-pathway' via D2-receptors. This allows learning to 'act' while blocking learning to 'refrain from action'. Negative feedback (dopamine inhibition) presumably has the opposite effect; it does not induce plasticity in the 'go-pathway' for lack of D1-activation and removes D2-mediated inhibition of the 'no-go-pathway'.

How then, does this model account for our results? The paradigm we used can be solved in two ways, by relying on positive or on negative feedback. When relying on positive feedback the animal learns choosing the stimulus that is associated with reward, when relying on negative feedback it learns choosing the stimulus that is not associated with punishment. In our experiment we use two magnitudes of reward but only one type of punishment. Following control injections, learning-rate was modulated by reward-magnitude, thus we can be certain that the animals relied on positive feedback. Following injections of dopamine-antagonist, however, the animals might have relied on negative feedback. This would explain why learning was overall intact while the effect of reward-magnitude vanished. This account is in line with the results by Frank et al. (2004); patients relied on negative feedback only when dopamine-function was impaired.

A counterintuitive result of the present experiment is that learning, contingent on small-reward seems to benefit from the D1-antagonist. Similarly, Frank et al. (2004) report that Parkinson patients off medication show an improved learning form negative feedback compared to controls. This effect is in accordance with their model; they argue that a reduction in available dopamine induces plasticity in the 'no-go-pathway', thereby fostering learning from negative feedback and inhibiting learning from positive feedback. This accounts for the fact that learning from negative feedback benefits from impaired dopamine-function but it does not

fully account for our results. We observed that learning from negative feedback is more efficient than learning form small positive reward and less efficient than learning from large positive reward.

The subjective nature of reward representations helps to account for this result. If an animal is trained to perform for a reward and is then shifted to a larger or to a smaller reward it will increase or decrease its response–strength respectively (Crespi, 1942; Black, 1968). If the animal never experienced this shift but was initially trained either with the large or with the small reward, response–strength is comparable to that on an intermediate reward. In other words, the difference in behavior is not dependent on the physical quantity of the reinforcer but on the subjective evaluation of the reinforcer. In our study, the animals were trained with two magnitudes of reward and only one punishment. Therefore the small reward was relatively devalued and the value of the large reward relatively increased, the punishment on the other hand reflected an intermediate subjective value. For the present results this implies that an animal with intact striatal dopamine that learned from positive feedback showed relatively good performance on the large reward and relatively poor performance on the small reward since its performance was modulated by subjective value. Conversely, if the animal was deprived of the dopamine-signal and thereby relying on negative feedback it showed intermediate performance in both conditions; thus showing reduced performance on the large reward but also improved performance on the small reward.

To summarize, we demonstrate the involvement of D1 receptors in MSt to the modulation of learning by the magnitude of contingent reward. While the ability to learn discrimination per se was unimpaired by drug injections, the modulation of learning-rate by reward-magnitude vanished. These results are in line with an established model on the role of dopamine in learning from positive and negative feedback. Our results also further the model by demonstrating for the first time that the locus of this effect is in the striatum.

Acknowledgements—This work was supported by the Deutsche Forschungsgemeinschaft through its SFB 874, project B5.

REFERENCES

Barnetta A, Ahna H, Billarda W, Golda EH, Kohli JD, Glock D, Goldberg LI (1986) Relative activities of SCH 23390 and its analogs in three tests for D1/DA1 dopamine receptor antagonism. Eur J Pharmacol 128(3):249–253.

Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 47(1):129–141.

Berridge KC (2007) The debate over dopamine's role in reward: the case for incentive salience. Psychopharmacology 191(3):391–431.

Berridge KC, Robinson TE (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res Brain Res Rev 28(3):309–369.

Black RW (1968) Shifts in magnitude of reward and contrast effects in instrumental and selective learning: a reinterpretation. Psychol Rev 75(2):114–126.

Calabresi P, Maj R, Pisani A, Mercuri NB, Bernardi G (1992) Long-term synaptic depression in the striatum: physiological and pharmacological characterization. J Neurosci 12(11):4224–4233.

Calabresi P, Saiardi A, Pisani A, Baik JH, Centonze D, Mercuri NB, Bernardi G, Borrelli E (1997) Abnormal synaptic plasticity in the striatum of mice lacking dopamine D2 receptors. J Neurosci 17(12):4536–4544.

Clarke HF, Robbins TW, Roberts AC (2008) Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. J Neurosci 28(43):10972–10982.

Collier G, Marx MH (1959) Changes in performance as a function of shifts in the magnitude of reinforcement. J Exp Psychol 57(5):305–309.

Crespi LP (1942) Quantitative variation of incentive and performance in the white rat. Am J Psychol 55(4):467–513.

Denny MR, King GF (1955) Differential response learning on the basis of differential size of reward. J Genet Psychol 87(2):317–320.

Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. Science 299(5614):1898–1902.

Frank MJ, O'Reilly RC (2006) A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. Behav Neurosci 120(3):497–517.

Frank MJ, Seeberger LC, O'Reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science 306(5703):1940–1943.

Freund N, Manns M, Rose J (2010) A method for the evaluation of intracranial tetrodotoxin injections. J Neurosci Methods 186(1):25–28.

Güntürkün O (2012) The convergent evolution of neural substrates for cognition. Psychol Res 76:212–219.

Graybiel AM (2005) The basal ganglia: learning new tricks and loving it. Curr Opin Neurobiol 15(6):638–644.

Guttman N (1953) Operant conditioning, extinction, and periodic reinforcement in relation to concentration of sucrose used as reinforcing agent. J Exp Psychol 46(4):213–224.

Herold C, Diekamp B, Güntürkün O (2008) Stimulation of dopamine D1 receptors in the avian fronto-striatal system adjusts daily cognitive fluctuations. Behav Brain Res 194:223–229.

Herold C, Joshi I, Hollmann M, Güntürkün O (2012) Prolonged cognitive training increases D5 receptor expression in the avian prefrontal cortex. PLoS One 7(5):e36484.

Jiao Y, Medina L, Veenman CL, Toledo C, Puelles L, Reiner A (2000) Identification of the anterior nucleus of the ansa lenticularis in birds as the homolog of the mammalian subthalamic nucleus. J Neurosci 20:6998–7010.

Joshua M, Adler A, Mitelman R, Vaadia E, Bergman H (2008) Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. J Neurosci 28(45):11673–11684.

Kabai P, Stewart MG, Tarcali J, Csillag A (2004) Inhibiting effect of D1, but not D2 antagonist administered to the striatum on retention of passive avoidance in the chick. Neurobiol Learn Mem 81(2):155–158.

Karten HW, Hodos W (1967) A stereotaxic atlas of the brain of the pigeon (*Columba livia*). Baltimore: Johns Hopkins University Press.

Kerr JN, Wickens JR (2001) Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum in vitro. J Neurophysiol 85(1):117–124.

Maher WB, Wickens DD (1954) Effect of differential quantity of reward on acquisition and performance of a maze habit. J Comp Physiol Psychol 47(1):44–46.

Neuringer AJ (1967) Effects of reinforcement magnitude on choice and rate of responding. J Exp Anal Behav 10:417–424.

Pan WX, Schmidt R, Wickens JR, Hyland BI (2008) Tripartite mechanism of extinction suggested by dopamine neuron activity and temporal difference model. J Neurosci 28(39):9619–9631.

Reiner A, Medina L, Veenman CL (1998) Structural and functional evolution of the basal ganglia in vertebrates. Brain Res Brain Res Rev 28:235–288.

Reiner A, Perkel DJ, Bruce LL, Butler AB, Csillag A, Kuenzel W, Paxinos G, Shimizu T, Striedter G, Wild M, Ball GF, Durand S, Güntürkün O, Lee DW, Mello CV, Powers A, White SA, Hough G, Kubikova L, Smulders TV, Wada K, Dugas-Ford J, Husband S, Yamamoto K, Yu J, Siang C, Jarvis ED (2004) Revised nomenclature for avian telencephalon and some related brainstem nuclei. J Comp Neurol 473(3):377–414.

Reiner A, Yamamoto K, Karten HJ (2005) Organization and evolution of the avian forebrain. Anat Rec A 287A:1080–1102.

Reynolds JN, Hyland BI, Wickens JR (2001) A cellular mechanism of reward-related learning. Nature 413(6851):67–70.

Rose J, Otto T, Dittrich L (2008) The Biopsychology-Toolbox: a free, open-source Matlab-toolbox for the control of behavioral experiments. J Neurosci Methods 175(1):104–107.

Rose J, Schmidt R, Grabemann M, Güntürkün O (2009) Theory meets pigeons: The influence of reward-magnitude on discrimination-learning. Behav Brain Res 198(1):125–129.

Rose J, Schiffer AM, Dittrich L, Güntürkün O (2010) The role of dopamine in maintenance and distractability of attention in the in the ''prefrontal cortex'' of pigeons. Neuroscience 167(2):232–237.

Schultz W (2002) Getting formal with dopamine and reward. Neuron 36(2):241–263.

Schultz W (2007) Behavioral dopamine signals. Trends Neurosci 30(5):203–210.

Schultz W, Apicella P, Scarnati E, Ljungberg T (1992) Neuronal activity in monkey ventral striatum related to the expectation of reward. J Neurosci 12(12):4595–4610.

Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275(5306):1593–1599.

Sun Z, Wang HB, Laverghetta AV, Yamamoto K, Reiner A (2005) The distribution and cellular localization of glutamatergic acid decarboxylase-65 (GAD65) mRNA in the forebrain and midbrain of the domestic chick. J Chem Neuroanat 29:265–291.

Sutton RS, Barto AG (1981) Toward a modern theory of adaptive networks: expectation and prediction. Psychol Rev 88(2):135–170.

Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine neurons. Science 307(5715):1642–1645.

Tsai HC, Zhang F, Adamantidis A, Stuber GD, Bonci A, de Lecea L, Deisseroth K (2009) Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. Science 324(5930):1080–1084.

Veenman CL, Reiner A (1994) The distribution of GABA-containing perikarya, fibers, and terminals in the forebrain and midbrain of pigeons, with particular reference to the basal ganglia and its projection targets. J Comp Neurol 339:209–250.

Wang Z, Kai L, Day M, Ronesi J, Yin HH, Ding J, Tkatch T, Lovinger DM, Surmeier DJ (2006) Dopaminergic control of corticostriatal long-term synaptic depression in medium spiny neurons is mediated by cholinergic interneurons. Neuron 50(3):443–452.

Watanabe S (2001) Effects of lobus parolfactorius lesions on repeated acquisition of spatial discrimination in pigeons. Brain Behav Evol 58(6):333–342.

Wickens J (1990) Striatal dopamine in motor activation and reward-mediated learning: steps towards a unifying model. J Neural Transm Gen Sect 80(1):9–31.

Wickens JR, Horvitz JC, Costa RM, Killcross S (2007) Dopaminergic mechanisms in actions and habits. J Neurosci 27(31):8181–8183.

(Accepted 25 October 2012)
(Available online 14 November 2012)