

# Striatal Dopamine Predicts Outcome-Specific Reversal Learning and Its Sensitivity to Dopaminergic Drug Administration

Roshan Cools,<sup>1</sup> Michael J. Frank,<sup>3,4</sup> Sasha E. Gibbs,<sup>2,5</sup> Asako Miyakawa,<sup>2</sup> William Jagust,<sup>2,6</sup> and Mark D'Esposito<sup>2</sup>

<sup>1</sup>Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen Medical Centre, 6500 HB Nijmegen, The Netherlands, <sup>2</sup>Helen Wills Neuroscience Institute, University of California, Berkeley, California 94720-3192, Departments of <sup>3</sup>Cognitive and Linguistic Sciences and <sup>4</sup>Psychology, Brown University, Providence, Rhode Island 02912, <sup>5</sup>Department of Psychology, Stanford University, Stanford, California 94305, and <sup>6</sup>Lawrence Berkeley National Laboratory, Berkeley, California 94720

Individual variability in reward-based learning has been ascribed to quantitative variation in baseline levels of striatal dopamine. However, direct evidence for this pervasive hypothesis has hitherto been unavailable. We demonstrate that individual differences in reward-based reversal learning reflect variation in baseline striatal dopamine synthesis capacity, as measured with neurochemical positron emission tomography. Subjects with high baseline dopamine synthesis in the striatum showed relatively better reversal learning from unexpected rewards than from unexpected punishments, whereas subjects with low baseline dopamine synthesis in the striatum showed the reverse pattern. In addition, baseline dopamine synthesis predicted the direction of dopaminergic drug effects. The D<sub>2</sub> receptor agonist bromocriptine improved reward-based relative to punishment-based reversal learning in subjects with low baseline dopamine synthesis capacity, while impairing it in subjects with high baseline dopamine synthesis capacity in the striatum. Finally, this pattern of drug effects was outcome-specific, and driven primarily by drug effects on punishment-, but not reward-based reversal learning. These data demonstrate that the effects of D<sub>2</sub> receptor stimulation on reversal learning in humans depend on task demands and baseline striatal dopamine synthesis capacity.

**Key words:** dopamine; reward; punishment; striatum; PET; learning

## Introduction

Adaptation to our environment requires the anticipation of biologically relevant events by learning signals of their occurrence, i.e., reward-based learning. Models of reward-based learning use a prediction error signal, representing the difference between expected and obtained events, to update their predictions based on the environment (Sutton and Barto, 1998). A putative mechanism of the prediction error signal for reward is the phasic firing of dopamine neurons in the midbrain (Montague et al., 1996; Schultz et al., 1997). These neurons innervate large parts of the brain, including the striatum, the major input structure of the basal ganglia. In keeping with this anatomical arrangement, the striatum has often been implicated in reward-based learning and

its modulation by dopamine (Cools and Robbins, 2004; Frank, 2005; Pessiglione et al., 2006; Schönberg et al., 2007) and reward-based learning is modulated by agonists of D<sub>2</sub>/D<sub>3</sub> receptors that are abundant in the striatum (Frank and O'Reilly, 2006; Pizzagalli et al., 2008). Schönberg et al. (2007) have recently proposed that individual differences in reward-based learning may reflect differences in striatal dopamine function. However, there is no direct evidence for this hypothesis. Here we demonstrate a significant positive relationship between reward-based learning and baseline striatal dopamine synthesis capacity, as measured with uptake of the positron emission tomography (PET) tracer [<sup>18</sup>F]fluorometatyrosine (FMT).

We further establish the link between dopamine in the striatum and reward-based learning by showing that effects of dopamine D<sub>2</sub> receptor stimulation also depend on baseline striatal dopamine synthesis capacity. Evidence from studies with experimental animals (Williams and Goldman-Rakic, 1995; Zahrt et al., 1997; Arnsten, 1998) has revealed an “inverted U”-shaped relationship between D1 receptor stimulation in the prefrontal cortex and cognitive performance. This relationship has been related to baseline-dependency of drug effects, so that low baseline dopamine levels are remedied, while high baseline dopamine levels are detrimentally over-dosed by the same dopamine D1 receptor agonist (Phillips et al., 2004). Although recent studies with humans, which have made use of genetic variation in the D<sub>2</sub>

Received Sept. 18, 2008; revised Jan. 7, 2009; accepted Jan. 8, 2009.

This work was supported by National Institutes of Health Grants MH63901, NS40813, DA02060, and AG027984. R.C. and M.D. conceived the study. A.M. and S.E.G. collected and analyzed the PET data, while R.C. analyzed and interpreted the psychopharmacological data in relation to the PET data. W.J. developed methods for acquisition and analysis of PET data. M.F. performed the model-based analyses and helped R.C. interpret the data and write the study. All authors discussed the results and commented on the study. We thank Lee Altamirano, Elizabeth Kelley, George Elliott Wimmer, and Emily Jacobs for assistance with data collection and Cindee Madison for assistance with data analysis.

The authors declare no competing financial interests.

Correspondence should be addressed to Roshan Cools, Donders Centre for Cognitive Neuroimaging, Kapittelweg 29, 6500 HB Nijmegen, The Netherlands. E-mail: roshan.cools@donders.ru.nl.

DOI:10.1523/JNEUROSCI.4467-08.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/291538-06\$15.00/0

receptor gene, have suggested that a similar mechanism might underlie contrasting effects of  $D_2$  receptor stimulation in the striatum on reward-based learning (Frank and O'Reilly, 2006; Cohen et al., 2007), direct evidence for baseline-dependency of dopaminergic drug effects on reward-learning in the striatum is lacking. We combined neurochemical PET imaging with behavioral psychopharmacology to test this hypothesis. We studied the effects of a single oral dose (1.25 mg) of the dopamine  $D_2$  receptor agonist bromocriptine on reversal learning in young healthy volunteers, who also, on a separate occasion, underwent a PET scan with the tracer FMT. Subjects with low synthesis capacity were predicted to benefit from  $D_2$  receptor stimulation with bromocriptine, while subjects with high synthesis capacity were predicted to be detrimentally overdosed by the same drug.

We used a reversal learning paradigm that enabled the separate assessment of reward- and punishment-based reversal learning (Cools et al., 2008a). Based on prior data indicating that dopaminergic drug effects are outcome-specific (Cools et al., 2006), we anticipated contrasting effects of bromocriptine on reward- and punishment-based reversal learning.

## Materials and Methods

**General procedure.** The University of California Berkeley Committee for the Protection of Human Subjects approved the procedures, which were in accord with the Helsinki Declaration of 1975.

Eleven subjects [all female; mean (SD) age = 22.2 (2.0)] underwent a single PET scan with the tracer 6-[ $^{18}$ F]fluoro-L-m-tyrosine (FMT). The PET data from these subjects were previously reported in relation to their working memory capacity as measured with the listening span test (Cools et al., 2008b). Detailed neuropsychological characteristics of the subjects are presented in that previous paper. All subjects were screened for psychiatric and neurological disorders; exclusion criteria were any history of cardiac, hepatic, renal, pulmonary, neurological, psychiatric or gastrointestinal disorders, an episode of loss of consciousness, use of psychotropic drugs, sleeping pills and heavy marijuana use (>10 times in a lifetime).

Subjects were selected from a sample of subjects that had also participated in a psychopharmacological study on the effects of bromocriptine (Cools et al., 2007). Initial selection of subjects for this study was based on their high or low scores on the Barratt Impulsiveness Inventory (BIS-11) (Patton et al., 1995). However, there was no relationship between dopamine synthesis capacity and trait impulsivity, as reported in our previous report on the PET data from these subjects (Cools et al., 2008b) (all Pearson correlations <0.1). As part of this study, subjects completed an established observational reversal learning task (Fig. 1) (Cools et al., 2006, 2008a) on two occasions, once after intake of placebo and once after intake of bromocriptine, in a double-blind, placebo-controlled cross-over design. The dose of bromocriptine (1.25 mg) was selected based on previously observed changes in cognitive performance (Kimberg et al., 1997; Gibbs and D'Esposito, 2005) and minimal side effects (Luciana et al., 1992; Luciana and Collins, 1997). The order of bromocriptine and placebo testing was approximately counterbalanced (six subjects received bromocriptine on the first session). One bromocriptine dataset from the punishment condition was missing. The reversal learning task was administered ~3.5 h after capsule intake, coinciding with the time-window of maximal drug effects (Drewe et al., 1988; Lynch, 1997). Subjects were instructed to have a light meal ~1 h before arrival and they were asked to refrain from caffeine and cigarettes on the days of testing.

**PET imaging and analysis.** FMT is a substrate of L-aromatic amino acid



**Figure 1.** Schematic of the observational reversal learning task. The task was administered as reported previously (Cools et al., 2006) with deterministic stimulus-outcome contingencies, as indicated here by the %-labels. Subjects pressed one of two colored buttons on a keyboard to indicate whether they predicted that the highlighted stimulus (face or scene) would lead to reward or punishment. The outcome-response mappings were approximately counterbalanced between subjects (6 subjects pressed the right button for punishment in both drug sessions). Each subject performed two practice blocks and four experimental blocks (120 trials per block) in each session.

decarboxylase and an index of presynaptic dopamine synthesis capacity, i.e., processes that occur in striatal terminals of midbrain dopamine neurons. The tracer was synthesized with a modification of the procedure as previously reported (Namavari et al., 1993). Scanning procedures, data analysis and region of interest procedures were also as previously reported (Cools et al., 2008b).

All subjects were scanned ~60 min after administration of an oral dose of 2.5 mg/kg of the peripheral decarboxylase inhibitor carbidopa to increase brain uptake of the tracer. Participants were positioned on the scanner bed with a pillow and an elastic band to comfortably restrict head motion. Images (voxel size was 3.6 × 3.6 mm in-plane with 4 mm slice thickness) were obtained on a Siemens ECAT EXACT HR scanner in 3D acquisition mode. A 10 min transmission scan was obtained for attenuation correction, then ~2.5 mCi of FMT were subsequently injected as a bolus in an antecubital vein and a dynamic acquisition sequence in 3D mode was obtained: 4 × 1 min, 3 × 2 min, 3 × 3 min, 14 × 5 min for a total of 89 min of scan time.

Data were reconstructed using an ordered subset expectation maximization (OSEM) algorithm with weighted attenuation, an image size of 256 × 256, and 6 iterations with 16 subsets. A Gaussian filter with 6 mm FWHM was applied, with a scatter correction. Images were evaluated for subject motion and realigned as necessary using algorithms implemented in SPM2.

Structural scans (high-resolution MP-FLASH; 0.875 × 0.875 × 1.54 mm) were obtained during the prior magnetic resonance imaging (MRI) study (structural scans were not available for two out of 11 subjects). This permitted the use of the high-resolution MRI image for anatomical verification of the localization of functional PET regions of interest (ROIs). Bilateral cerebellar ROIs (10 mm spheres) were used as a reference ROI in conjunction with ROIs in striatum and a simplified reference tissue model with a graphical analysis approach (Patlak and Blasberg, 1985; Lammertsma and Hume, 1996). This leads to an influx constant  $K_1$ , which reflects regional FMT uptake scaled to the volume of distribution in the reference region.

To test hypotheses about differences in subregions of the striatum, we defined ROIs in the right and left caudate and putamen. An axial image representing the sum of the last four emission scans of the PET scanning session (4 × 5 min frames) was coregistered to the high-resolution MR scan using a 12-parameter affine algorithm implemented in SPM2. ROIs were drawn on these images (Wang et al., 1996; Volkow et al., 1998), using the atlas of Talairach and Tournoux (Talairach and Tournoux, 1988) to delineate the caudate and putamen. Regions were drawn on data in native space to preserve differences in tracer uptake due to anatomical variability between subjects. We have previously demonstrated the ability to draw ROIs with high inter-rater reliability (Klein et al., 1997). The Patlak model was fitted with dynamic data from each ROI from 24 to 89 min, when the regression is highly linear ( $r > 0.99$ ).

It might be noted that there was a delay between the acquisition of the PET data and that of the behavioral data (mean 16.7 months; SD 2.4 months). Uptake of the tracer [ $^{18}$ F]fluorometatyrrosine is thought to

reflect a relatively stable process (synthesis capacity) that is not particularly sensitive to small state-related changes, much like uptake of the tracer F-DOPA. Thus, a study by Vingerhoets et al. (1994) demonstrated that striatal  $K_i$  is a reliable measurement, with it having a 95% chance of lying within 18% of its value within an individual normal subject. We argue that the delay does not confound our results, but rather renders them more striking. Any instability in the PET measurement across the delay would have reduced rather than enhanced the likelihood of obtaining the results, which were statistically controlled for noise by an  $\alpha$  level of 0.05. Data analyses supported this hypothesis, as the effects were stronger when effects of interest were corrected for the delay between the two sessions than when they were not. Here we report only those analyses in which we corrected for the delay, by entering it as a covariate, although all effects were also significant when they were not corrected for the delay. Furthermore, in the supplemental Results C (available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material), we report an additional analysis, explicitly addressing in a quantitative manner the possibility that the effect reflects noise.

**Experimental paradigm.** The task required the learning and reversal of predictions of reward and punishment. On each trial two vertically adjacent stimuli were presented: one face and one scene; location randomized; about 19 inch viewing distance; subtending  $\sim 3^\circ$  horizontally and  $3.5^\circ$  vertically. One of the stimuli was associated with reward, while the other was associated with punishment. On each trial, one of the two stimuli was highlighted with a black border surrounding the stimulus and subjects had to predict, based on trial and error learning, whether the highlighted stimulus would lead to reward or punishment. Subjects indicated their predictions by pressing, with the index or middle finger, one of two colored buttons (corresponding to keys “b” and “n” depending on the response-outcome mapping) on a laptop keyboard. They pressed the green button for reward and the red button for punishment. The outcome-response mappings were counterbalanced between subjects. The (self-paced) response was followed by an interval of 1000 ms, after which the outcome was presented for 500 ms. Note that this outcome did not provide direct performance feedback. Reward consisted of a green smiley face, a “+\$100” sign and a high-frequency jingle tone. Punishment consisted of a red sad face, a “-\$100” sign and a single low-frequency tone. After the outcome, the screen was cleared for 500 ms, after which the next two stimuli were presented. The stimulus-outcome contingencies reversed multiple times provided learning criteria were met.

Each subject performed one practice block and four experimental blocks. Each practice block consisted of one acquisition stage and one reversal stage (learning criterion was 20 [not necessarily consecutive] correct trials). Each experimental block consisted of one acquisition stage and a variable number of reversal stages. The task proceeded from one stage to the next following a number of consecutive correct trials, as determined by a pre-set learning criterion. Learning criteria (i.e., the number of consecutive correct trials following which the contingencies changed) varied between stages (mean = 6.9; SD = 1.8; range from 5 to 9), to prevent predictability of reversals. The maximum number of reversal stages per experimental block was 16, although the block terminated automatically after completion of 120 trials ( $\sim 6.6$  min), so that each subject performed 480 trials (4 blocks) per experimental session. The mean number of stages completed is reported in supplemental Table 1, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material.

The task consisted of two conditions (two blocks per condition). In the unexpected reward condition, reversals were signaled by unexpected reward occurring after the previously punished stimulus was highlighted. Conversely, in the unexpected punishment condition, reversals were signaled by unexpected punishment following the previously rewarded stimulus. Accuracy on the trial following the unexpected outcome (“switch trials”) reflected the degree to which subjects updated their predictions based on unexpected outcomes. The stimulus that was highlighted on the first trial of each reversal stage (on which the unexpected outcome was presented) was always highlighted again on the second trial of that stage (i.e., the switch trial on which the subject had to implement the reversed contingencies and switch their predictions).

Based on prior data (Frank et al., 2004; Cools et al., 2006; Frank and

O’Reilly, 2006), we predicted that bromocriptine would have contrasting effects on reward- and punishment-based reversal learning. Following this prior work, we were specifically interested in the drug effect on the balance between (reversal) learning from reward and from punishment. Therefore, relative reversal learning scores were calculated, by which accuracy scores on punishment-based switch trials were subtracted from accuracy scores on reward-based switch trials. The additional advantage of this method is that it controls for within-subject variability due to other factors such as arousal, attention and motivation, which would have affected each measure in the same direction. Thus, general effects of the drug that were not specific to the ability to learn from reward or punishment were subtracted out. Further, we report drug effects (differences between the placebo and the bromocriptine session), because it is these drug effects that are of primary interest in the present study. Finally, we also report reward- and punishment-based reversal learning under placebo and under bromocriptine separately.

**Statistical analysis.** Mean proportions of correct responses on the learning task were calculated as reported previously (Cools et al., 2006, 2008a). Repeated measures ANOVAs were conducted using SPSS 15.0 with drug and valence as within-subject factors and dopamine synthesis capacity as a covariate. The delay between acquisition of the drug and PET data was also entered as a covariate. Pearson product-moment correlation coefficients were calculated between  $K_i$ -values extracted from ROIs and behavioral data. All correlations represent partial correlations, correcting for the delay between PET and drug data acquisition. All correlations were also significant without this correction. The distribution of none of the parameters assessed here deviated from normality as indicated by Kolmogorov–Smirnov tests (all  $p = 0.2$ ).

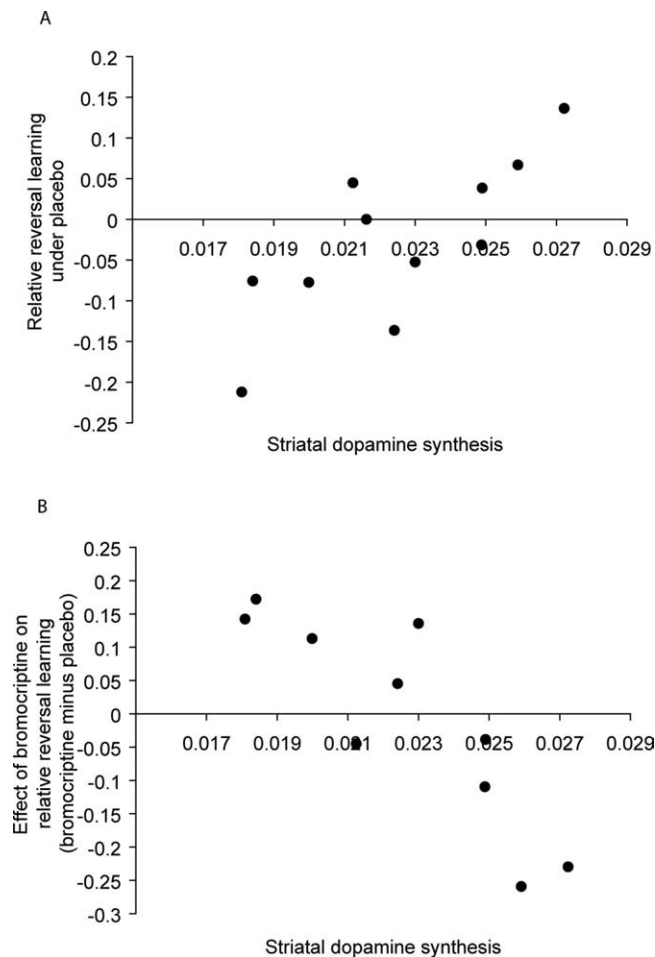
All reported  $p$  values are two-tailed.

## Results

In our sample of young healthy volunteers, influx constant  $K_i$  values varied between 0.018 and 0.027, falling well within the range of “normal” values observed previously (Eberling et al., 2007). Subjects performed well on the reversal learning task, with an average accuracy rate on trials after the unexpected outcomes  $>90\%$  (supplemental Table 2, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material).

First, we analyzed the data from the placebo session. A repeated measures ANOVA was conducted with valence as the within-subject factor and synthesis capacity and acquisition delay as covariates. Consistent with neurophysiological evidence from nonhuman primates (Hollerman and Schultz, 1998), this analysis revealed a highly significant interaction between valence and synthesis capacity ( $F_{(1,8)} = 19.0$ ,  $p = 0.002$ ) and no effects of acquisition delay. This interaction reflected a positive correlation between dopamine synthesis capacity in the striatum and reversal learning from reward relative to punishment under placebo (Fig. 2a). It was present across the entire striatum (averaged across right and left caudate nucleus and putamen;  $r_8 = 0.84$ ,  $p = 0.004$ ), and also within striatal subregions (bilateral caudate nucleus:  $r_8 = 0.8$ ,  $p = 0.007$ ; bilateral putamen:  $r_8 = 0.87$ ,  $p = 0.001$ ). The correlation between dopamine synthesis capacity and relative performance on non-switch trials (trials requiring reward-prediction minus trials requiring punishment-prediction) was also positive, albeit non-significant (entire striatum:  $r_8 = 0.54$ ,  $p = 0.1$ ).

The positive correlation between synthesis capacity and relative reversal learning scores (i.e., the difference between reward and punishment) was driven by a positive correlation between synthesis capacity and reward-based reversal under placebo (accuracy:  $r_8 = 0.79$ ,  $p = 0.007$ ), indicating that greater dopamine synthesis capacity predicted better reward-based reversal. Conversely, punishment-based reversal under placebo did not depend on baseline dopamine synthesis capacity (accuracy:  $r_8 =$



**Figure 2.** Baseline-dependency of relative reversal learning scores and their sensitivity to  $D_2$  receptor stimulation. Relative reversal learning scores represent the proportion of correct responses on switch trials after unexpected reward minus the proportion of correct responses on switch trials after unexpected punishment. **A**, Positive correlation between relative reversal learning and striatal dopamine synthesis capacity ( $K_i$ -values) under placebo. **B**, Negative correlation between the effect of bromocriptine on relative reversal learning and striatal dopamine synthesis capacity.

$-0.2, p = 0.6$ ) (Fig. 3*a,b*). This finding is remarkably consistent with neurophysiological evidence for associations between phasic dopamine burst firing and reward prediction error (Hollerman and Schultz, 1998), given that synthesis capacity likely influences the efficacy of impulse-dependent phasic dopamine release. The finding that the effect did not extend to learning from unexpected punishment suggests that synthesis capacity did not influence the efficacy of the impulse-dependent pause in dopamine firing that accompanies unexpected reward omission (Hollerman and Schultz, 1998).

Next we assessed whether baseline striatal dopamine synthesis capacity also predicted the effects of the dopamine  $D_2$  receptor agonist bromocriptine on reversal learning. To this end, we conducted a repeated measures ANOVA with drug and valence as within-subject factors and dopamine synthesis capacity and acquisition delay as covariates. As predicted, this analysis revealed highly significant two-way drug by valence ( $F_{(1,7)} = 17.4, p = 0.004$ ) and three-way drug by valence by synthesis capacity interactions ( $F_{(1,7)} = 29.4, p = 0.001$ ). There were no main effects (acquisition delay:  $F_{(1,7)} = 0.2, p = 0.7$ ; synthesis capacity:  $F_{(1,7)} = 3.0, p = 0.1$ ; valence:  $F_{(1,7)} = 0.8, p = 0.4$ ; drug:  $F_{(1,7)} = 0.03, p = 0.9$ ) and no other interaction effects (valence by synthesis

capacity:  $F_{(1,7)} = 0.6, p = 0.5$ ; valence by delay:  $F_{(1,7)} = 0.3, p = 0.6$ ; drug by synthesis capacity:  $F_{(1,7)} = 0.2, p = 0.7$ ; drug by delay:  $F_{(1,7)} = 0.003, p = 0.96$ ; drug by valence by delay:  $F_{(1,7)} = 2.4, p = 0.2$ ). The significant three-way interaction reflected a significant negative correlation between synthesis capacity and drug-induced improvement on relative reversal learning scores ( $r_7 = -0.9, p = 0.001$ ) (Fig. 2*b*). Consistent with an ‘inverted u’-shaped dose–response curve, bromocriptine improved reward-based reversal relative to punishment-based reversal in subjects with low baseline levels of striatal dopamine synthesis capacity, but had the reverse effect in subjects with high baseline levels. Again the effect extended across striatal subregions (bilateral caudate nucleus:  $r_7 = -0.89, p = 0.001$ ; bilateral putamen:  $r_7 = -0.89, p = 0.001$ ). The correlation with (relative) performance on non-switch trials was not significant ( $r_7 = -0.45, p = 0.2$ ).

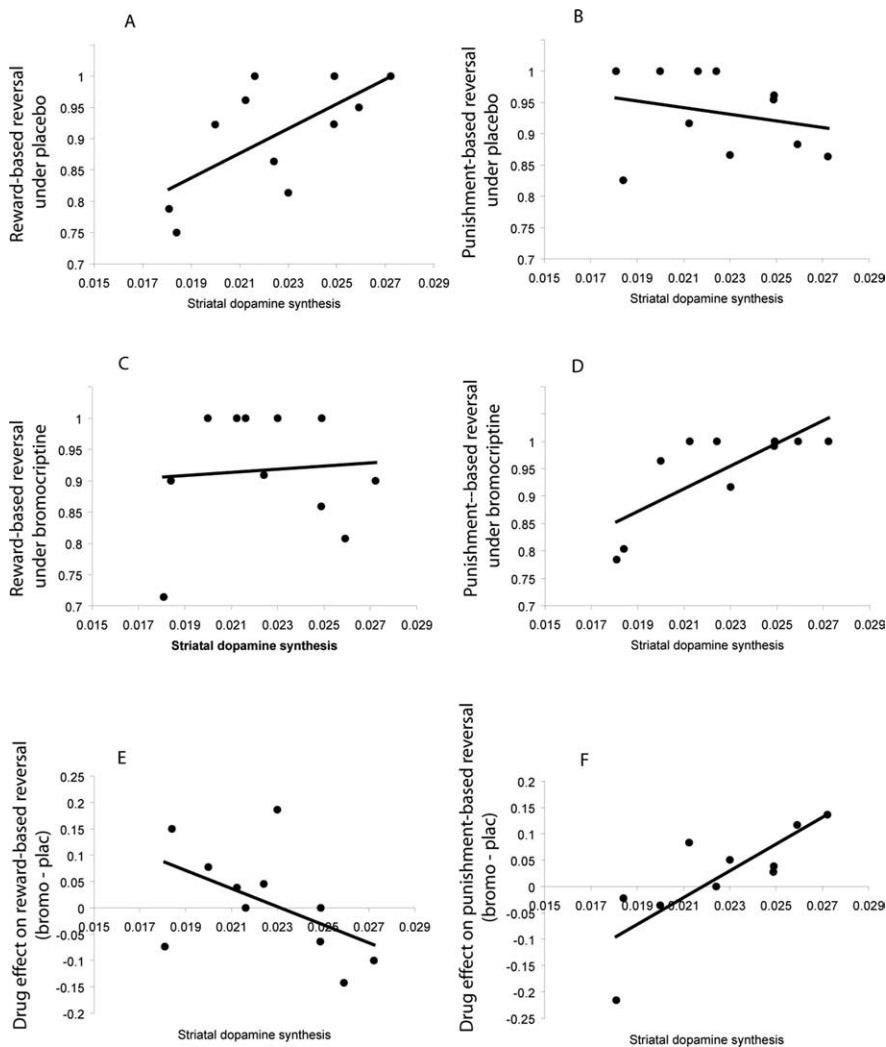
Breakdown of the three-way interaction effect into simple interaction effects for reward and punishment separately revealed a significant interaction between drug and synthesis capacity for punishment-based reversal ( $F_{(1,7)} = 14.2, p = 0.007$ ), as well as a near-significant interaction between drug and synthesis capacity for reward-based reversal ( $F_{(1,8)} = 3.4, p = 0.1$ ). These interactions reflected a highly significant positive correlation between striatal dopamine synthesis and drug-induced improvement in punishment-based reversal ( $r_7 = 0.8, p = 0.007$ ) (Fig. 3*b,d,f*), while the negative correlation between dopamine synthesis and drug effects on reward-based reversal was less convincing ( $r_8 = -0.55, p = 0.1$ ) (Fig. 3*a,c,e*).

In supplementary analyses, we aimed to disentangle two alternative hypotheses regarding dopaminergic modulation. Specifically, to establish whether the here described effects reflect a modulation of learning or switching, we applied computational reinforcement learning algorithms to fit individual subjects’ trial-by-trial sequence of choices (Sutton and Barto, 1998; Frank et al., 2007*b*). These algorithms allowed us to generate learning-rate parameters (separately for reward and punishment) that were not directly observable in the behavioral data. Detailed methods and results are presented in the supplemental Materials, available at [www.jneurosci.org](http://www.jneurosci.org). Critically, a significant relationship was obtained between dopamine synthesis and the drug effect on reward learning rate ( $r_8 = -0.71, p = 0.02$ ), as well as between dopamine synthesis and the drug effect on punishment learning rate ( $r_{10} = 0.78, p = 0.01$ ) (supplemental Figure and Table 3, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material).

In summary, higher dopamine synthesis capacity in the striatum was associated with better reward-based reversal learning under placebo. Furthermore, bromocriptine improved reward-based reversal learning in subjects with low synthesis capacity, while impairing it in subjects with high synthesis capacity. Conversely, the same drug dose impaired punishment-based reversal learning in subjects with low synthesis capacity, while improving it in subjects with high synthesis capacity.

## Discussion

Baseline dopamine measures predicted reversal learning due to reward prediction errors relative to punishment prediction errors. The result provides the first empirical evidence for the pervasive, but hitherto untested hypothesis that individual variation in reward-based learning reflects quantitative variation in baseline levels of striatal dopamine function, as indexed by uptake of a PET dopamine synthesis tracer. Critically, the effect was outcome-specific, so that high dopamine synthesis was associated with a bias toward reward- relative to punishment-based reversal learning. This observation concurs with recent theoretical mod-



**Figure 3.** Baseline-dependency of absolute reversal learning scores and their sensitivity to  $D_2$  receptor stimulation. **A**, Significant positive correlation between striatal dopamine synthesis capacity ( $K_i$ -values) and reward-based reversal learning under placebo. **B**, Nonsignificant correlation between striatal dopamine synthesis capacity and punishment-based reversal learning under placebo. **C**, Nonsignificant correlation between striatal dopamine synthesis capacity and reward-based reversal learning under bromocriptine ( $r_8 = -0.15$ ). **D**, Significant positive correlation between striatal dopamine synthesis capacity and punishment-based reversal learning under bromocriptine ( $r_7 = 0.78, p = 0.015$ ). **E**, Nonsignificant negative correlation between striatal dopamine synthesis capacity and the effect of bromocriptine on reward-based reversal learning (bromocriptine minus placebo). **F**, Significant positive correlation between striatal dopamine synthesis capacity and the effect of bromocriptine on punishment-based reversal learning (bromocriptine minus placebo). For statistics, see Results.

els and empirical work (Frank et al., 2004, 2005). For example, patients with Parkinson's disease, which is characterized by severe dopamine depletion in the striatum, exhibit difficulty with learning from reward relative to punishment, as measured with the present paradigm (Cools et al., 2006) as well as with a probabilistic selection task (Frank et al., 2004). Furthermore, the common dopamine-enhancing antiparkinson medication reversed this bias, leading to difficulty with learning from punishment relative to reward (Cools et al., 2006; Frank et al., 2004). The present data demonstrate that individual differences in baseline levels of dopamine in the healthy population also predict reward relative to punishment-learning biases.

Perhaps most critically, this work provides the first direct evidence for the existence of an inverted u-shaped relationship between striatal function and dopamine  $D_2$  receptor stimulation. Based on evidence from experimental animals (Skirboll et al., 1979; Torstenson et al., 1998), we argue that this curvilinear

dose–response curve might reflect differential stimulation of pre- versus postsynaptic  $D_2$  receptors in high- versus low-baseline subjects respectively. Specifically, we hypothesize that the established self-regulatory mechanism of presynaptic action of bromocriptine, by which dopamine cell firing, release and/or synthesis are inhibited, is more pronounced in subjects with already high baseline levels of synaptic dopamine than in subjects with low baseline levels of dopamine. Furthermore, disproportionate efficacy of self-regulatory (presynaptic) mechanisms in high-baseline subjects might be accompanied by desensitization of postsynaptic  $D_2$  receptors, thereby further reducing the postsynaptic efficacy of bromocriptine. Thus bromocriptine might have paradoxically reduced synaptic dopamine levels, thereby impairing reward-based learning, via a presynaptic mechanism of action in high-baseline subjects. Conversely, we hypothesize that the same drug enhanced dopamine transmission via a postsynaptic mechanism of action in low-baseline subjects.

The effects likely reflect general associative learning from unexpected outcomes rather than switching per se, as demonstrated by the supplementary model-based analyses. Although the rapid updating required in the current task, on which subjects expressed high learning rates, is different from the slower types of incremental probabilistic learning (Cools et al., 2001; Frank et al., 2004), which require the integration of outcomes across more distant histories, we hypothesize that similar associations will be observed between striatal dopamine synthesis and updating during slow learning. Consistent with this hypothesis is the observation that effects of dopaminergic manipulations on incremental positive reinforcement learning

correlate with effects on rapid working memory updating (Frank and O'Reilly, 2006; Frank et al., 2007a), suggestive of similar dopaminergic influences on parallel neurobiological circuits.

One potential caveat of the present study is the considerable delay between the acquisition of the PET data and that of the behavioral data. We argue that this delay does not confound our results for the following reasons. First, there is evidence that measures of dopamine synthesis capacity are relatively stable in healthy volunteers, even across many years. Specifically, Vingerhoets et al. (1994b) have shown that the fluorodopa PET index decreased non-significantly over seven years by 0.3% per year. In addition, the reliability of change coefficient was 96%, confirming their previous study showing that striatal activity measured with PET is a highly reproducible measurement (Vingerhoets et al., 1994a), although we acknowledge that our age group was younger than the one studied by Vingerhoets et al. Second, the effects were statistically controlled for noise by an  $\alpha$  level of 0.05

and remained highly significant after statistical correction for the acquisition delay. Finally, any instability in the PET measurement across the delay would have reduced rather than enhanced the likelihood of obtaining the result. There is a possibility that, if there had been no delay, the correlation between synthesis capacity and behavioral data might have been even stronger. Therefore, the reported correlations might represent noisy versions of the true correlations.

Our data elucidate not only a fundamental mechanism underlying the behavioral role of striatal dopamine, but also identify an important neurobiological factor, i.e., baseline striatal dopamine synthesis, that contributes to the large variability in dopaminergic drug efficacy. This finding should have far-reaching implications for individualized drug development in neuropsychiatry, where variable drug efficacy provides a major problem for the treatment of patients with heterogeneous spectrum disorders like schizophrenia, attention deficit/hyperactivity disorder and drug addiction.

## References

- Arnsten AFT (1998) Catecholamine modulation of prefrontal cortical cognitive function. *Trends Cogn Sci* 2:436–446.
- Cohen MX, Krohn-Grimbergha A, Elger CE, Weber B (2007) Dopamine gene predicts the brain's response to dopaminergic drug. *Eur J Neurosci* 26:3652–3660.
- Cools R, Robbins TW (2004) Chemistry of the adaptive mind. *Philos Transact A Math Phys Eng Sci* 362:2871–2888.
- Cools R, Barker RA, Sahakian BJ, Robbins TW (2001) Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cereb Cortex* 11:1136–1143.
- Cools R, Altamirano L, D'Esposito M (2006) Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia* 44:1663–1673.
- Cools R, Sheridan M, Jacobs E, D'Esposito M (2007) Impulsive personality predicts dopamine-dependent changes in frontostriatal activity during component processes of working memory. *J Neurosci* 27:5506–5514.
- Cools R, Robinson OJ, Sahakian B (2008a) Acute tryptophan depletion in healthy volunteers enhances punishment prediction but does not affect reward prediction. *Neuropsychopharmacology* 33:2291–2299.
- Cools R, Gibbs SE, Miyakawa A, Jagust W, D'Esposito M (2008b) Working memory capacity predicts dopamine synthesis capacity in the human striatum. *J Neurosci* 28:1208–1212.
- Drewe J, Mazer N, Abisch E, Krummen K, Keck M (1988) Differential effect of food on kinetics of bromocriptine in a modified release capsule and a conventional formulation. *Eur J Clin Pharmacol* 35:535–541.
- Eberling JL, Bankiewicz KS, O'Neil JP, Jagust WJ (2007) PET 6-[F]fluoro-L-m-tyrosine studies of dopaminergic function in human and nonhuman primates. *Front Hum Neurosci* 1:9.
- Frank MJ (2005) Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and non-medicated Parkinsonism. *J Cogn Neurosci* 17:51–72.
- Frank MJ, O'Reilly RC (2006) A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. *Behav Neurosci* 120:497–517.
- Frank MJ, Seeberger LC, O'Reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306:1940–1943.
- Frank MJ, Santamaria A, O'Reilly RC, Willcutt E (2007a) Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. *Neuropsychopharmacology* 32:1583–1599.
- Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE (2007b) Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc Natl Acad Sci U S A* 104:16311–16316.
- Gibbs SE, D'Esposito M (2005) A functional MRI study of the effects of bromocriptine, a dopamine receptor agonist, on component processes of working memory. *Psychopharmacology* 180:644–653.
- Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* 1:304–309.
- Kimberg DY, D'Esposito M, Farah MJ (1997) Effects of bromocriptine on human subjects depend on working memory capacity. *Neuroreport* 8:3581–3585.
- Klein GJ, Teng X, Jagust WJ, Eberling JL, Acharya A, Reutter BW, Huesman RH (1997) A methodology for specifying PET VOF's using multimodality techniques. *IEEE Trans Med Imaging* 16:405–415.
- Lammertsma AA, Hume SP (1996) Simplified reference tissue model for PET receptor studies. *Neuroimage* 4:153–158.
- Luciana M, Collins P (1997) Dopaminergic modulation of working memory for spatial but not object cues in normal volunteers. *J Cogn Neurosci* 9:330–347.
- Luciana M, Depue RA, Arbisi P, Leon A (1992) Facilitation of working memory in humans by a D2 dopamine receptor agonist. *J Cogn Neurosci* 4:58–68.
- Lynch MR (1997) Selective effects on prefrontal cortex serotonin by dopamine D3 receptor agonism: interaction with low-dose haloperidol. *Prog Neuropsychopharmacol Biol Psychiatry* 21:1141–1153.
- Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16:1936–1947.
- Namavari M, Satyamurthy N, Phelps ME, Barrio JR (1993) Synthesis of 6-[18F] and 4-[18F]fluoro-L-m-tyrosines via regioselective radiofluorodestannylation. *Appl Radiat Isot* 44:527–536.
- Patlak CS, Blasberg RG (1985) Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. Generalizations. *J Cereb Blood Flow Metab* 5:584–590.
- Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006) Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442:1042–1045.
- Phillips A, Ahn S, Floresco S (2004) Magnitude of dopamine release in medial prefrontal cortex predicts accuracy of memory on a delayed response task. *J Neurosci* 14:547–553.
- Pizzagalli D, Evins AE, Schetter EC, Frank MJ, Patjas PE, Santesso DL, Cullhane M (2008) Single dose of a dopamine agonist impairs reinforcement learning in humans: behavioral evidence from a laboratory-based measure of reward responsiveness. *Psychopharmacology* 196:221–232.
- Schönberg T, Daw ND, Joel D, O'Doherty JP (2007) Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *J Neurosci* 27:12860–12867.
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Skirboll LR, Grace AA, Bunney BS (1979) Dopamine auto- and postsynaptic receptors: Electrophysiological evidence for differential sensitivity to dopamine agonists. *Science* 206:80–82.
- Sutton R, Barto A (1998) Reinforcement learning. Cambridge, MA: MIT.
- Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging. Stuttgart: Thieme.
- Torstenson R, Hartvig P, Långström B, Bastami S, Antoni G, Tedroff J (1998) Effect of apomorphine infusion on dopamine synthesis rate relates to dopaminergic tone. *Neuropharmacology* 37:989–995.
- Vingerhoets F, Snow BJ, Schulzer M, Morrison S, Ruth TJ, Holden JE, Cooper S, Calne DB (1994) Reproducibility of fluorine-18-6-fluorodopa positron emission tomography in normal human subjects. *J Nucl Med* 35:18–24.
- Volkow N, Gur RC, Wang GJ, Fowler JS, Moberg PJ, Ding YS, Hitzemann R, Smith G, Logan J (1998) Association between decline in brain dopamine activity with age and cognitive and motor impairment in healthy individuals. *Am J Psychiatr* 155:344–349.
- Wang GJ, Volkow ND, Levy AV, Fowler JS, Logan J, Alexoff D, Hitzemann RJ, Schlyer DJ (1996) MR-PET image coregistration for quantitation of striatal dopamine D2 receptors. *J Comput Assist Tomogr* 20:423–428.
- Williams GV, Goldman-Rakic PS (1995) Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. *Nature* 376:572–575.
- Zahrt J, Taylor JR, Mathew RG, Arnsten AFT (1997) Supranormal stimulation of D<sub>1</sub> dopamine receptors in the rodent prefrontal cortex impairs spatial working memory performance. *J Neurosci* 17:8528–8535.