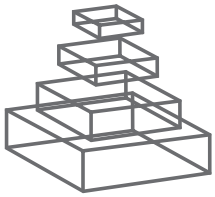
frontiers
RESEARCH TOPICS

# DOPAMINE AND BEHAVIORAL FLEXIBILITY: THE PROBLEM OF MODIFYING ESTABLISHED BEHAVIOR

Topic Editors
Jeff A. Beeler, Roshan Cools, Monica Luciana, Sean B. Ostlund and Giselle Petzinger,

frontiers in NEUROSCIENCE



# frontiers




Cover image provided by Ibbl sarl, Lausanne CH

**ISSN** 1664-8714
**ISBN** 978-2-88919-225-0
**DOI** 10.3389/978-2-88919-225-0


## ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.
All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!
Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

# DOPAMINE AND BEHAVIORAL FLEXIBILITY: THE PROBLEM OF MODIFYING ESTABLISHED BEHAVIOR

Topic Editors:
**Jeff A. Beeler,** Department of Psychology, Queens College and the Graduate Center, CUNY, Queens, NY, USA
**Roshan Cools,** Center for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, Netherlands
**Monica Luciana,** Department of Psychology, University of Minnesota, Minneapolis, MN, USA
**Sean B. Ostlund**, Department of Psychiatry and Biobehavioral Sciences, University of California Los Angeles, Los Angeles, CA, USA
**Giselle Petzinger,** Department of Neurology, University of Southern California, Los Angeles, CA, USA

Dopamine has been associated with many functions critical to adaptive behavior, including regulating synaptic plasticity, signaling the outcomes of behavior, reinforcement learning, regulating appetitive motivation, modulating energy and activity levels, gating short-term memory and facilitating movement, to name a few. In research and theories of dopamine, however, a distinction is rarely made between its role in the initial acquisition and expression of a behavior and its role in modifying an established behavior. For example, though dopamine is widely believed to modulate corticostriatal synaptic plasticity and play a crucial role in motor, habit and procedural learning, the role that dopamine may play in modifying an established behavior or skill, particularly once habitual or automatic, has received little attention. Similarly, though dopamine has been associated with appetitive motivation and reinforcement learning, much of this work has focused on the contribution of dopamine to the development and maintenance of maladaptive, inflexible behaviors, such as in addiction, with less attention to how dopamine may contribute to changing those behaviors.

Dopamine has been associated with both acquiring and promoting learned behaviors. When an established behavior needs to be modified, however, these two aspects of dopamine seem to be at odds: facilitating new learning and modifying established behaviors on one hand, while promoting prior learning on the other. How are these two aspects of dopamine, inherently contradictory, reconciled?

We start from the premise that a crucial, highly conserved neurotransmitter system integral to multiple critical functions, from motivation to movement, likely evolved through a selection process that favored flexible adaptation to changing environments and conditions. The contributions to this special topic explore themes related to how dopamine, often associated with inflexible, compulsive behavior, may contribute to behavioral flexibility. The introductory editorial, highlighting the contributions included here, proposes a shift from conceptualizing dopamine as the 'reward transmitter' to rethinking it as fundamentally a 'flexibility' transmitter.

# Table of Contents


**05** ***A Kinder, Gentler Dopamine... Highlighting Dopamine's Role in Behavioral Flexibility***
Jeff A. Beeler, Roshan Cools, Monica Luciana, Sean B. Ostlund and Giselle Petzinger

**07** ***Accelerated Habit Formation Following Amphetamine Exposure is Reversed by $D_1$, but enhanced by $D_2$, Receptor antagonists***
Andrew J. D. Nelson and Simon Killcross

**20** ***Bidirectional Modulation of Infralimbic Dopamine D1 and D2 Receptor Activity Regulates Flexible Reward Seeking***
Jacqueline M. Barker, Mary M. Torregrossa and Jane R. Taylor

**27** ***Dopaminergic Control of Cognitive Flexibility in Humans and Animals***
Marianne Klanker, Matthijs Feenstra and Damiaan Denys

**51** ***Of Goals and Habits: Age-Related and Individual Differences in Goal-Directed Decision-Making***
Ben Eppinger, Maik Walter, Hauke R. Heekeren and Shu-Chen Li

**65** ***Prefrontal Dopamine and Behavioral Flexibility: Shifting From an "Inverted-U" Toward a Family of Functions***
Stan B. Floresco

**77** ***Dopaminergic Drug Effects During Reversal Learning Depend on Anatomical Connections Between the Orbitofrontal Cortex and the Amygdala***
Marieke E. van der Schaaf, Marcel P. Zwiers, Martine R. van Schouwenburg, Dirk E. M. Geurts, Arnt F. A. Schellekens, Jan K. Buitelaar, Robbert Jan Verkes and Roshan Cools

**87** ***Heads for Learning, Tails for Memory: Reward, Reinforcement and a Role of Dopamine in Determining Behavioral Relevance Across Multiple Timescales***
Mathieu Baudonnat, Anna Huber, Vincent David and Mark Edwin Walton

**101** ***On the Role of Subsecond Dopamine Release in Conditioned Avoidance***
Erik B. Oleson and Joseph F. Cheer

**110** ***Dopamine Imbalance in Huntington's Disease: A Mechanism for the Lack of Behavioral Flexibility***
Jane Y. Chen, Elizabeth A. Wang, Carlos Cepeda and Michael S. Levine

frontiers in
NEUROSCIENCE

EDITORIAL
published: 24 January 2014
doi: 10.3389/fnins.2014.00004

# A kinder, gentler dopamine. . . highlighting dopamine's role in behavioral flexibility

***Jeff A. Beeler[1]*, Roshan Cools[2], Monica Luciana[3], Sean B. Ostlund[4] and Giselle Petzinger[5]***

[1] *Department of Psychology, Queens College and the Graduate Center, CUNY, Flushing, NY, USA*
[2] *Center for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, Netherlands*
[3] *Department of Psychology, University of Minnesota, Minneapolis, MN, USA*
[4] *Psychiatry and Biobehavioral Science, UCLA Semel Institute for Neuroscience, University of California Los Angeles, Los Angeles, CA, USA*
[5] *Department of Neurology, University of Southern California, Los Angeles, CA, USA*
**Correspondence: jbeeler@qc.cuny.edu*

***Edited by:***
*Scott A. Huettel, Duke University, USA*

**Keywords: behavioral flexibility, explore-exploit, dopamine, dopamine D2 receptor, decision-making**

If dopamine were a politician, it might have an image problem. The organizing metaphor for dopamine function is *reward*. Implicit in this formulation is that a system that drives appetitive pursuit toward needed resources is essential for survival. Dopamine is the go-and-get-it neurotransmitter. In it's association with addiction and compulsive behavior— a "hijacked reward system"— dopamine has become neuroscience's version of Freud's *id*, driving appetitive pursuit without regard to consequences.

However, some evidence suggests that dopamine is not essential for basic reward related behaviors. Even with diminished dopamine transmission, animals still like food, still eat, and can still learn about rewards. Dopamine is not at the root of these functions; rather, it is *modulatory*. That is, dopamine *adjusts* these functions, which begs the question *to what end*? The impulse behind this research topic is that the "root" function of dopamine is not to drive reward behavior, but rather to *adjust* reward oriented behavior in order to achieve *adaptive behavioral flexibility*. That is, dopamine evolved to *adapt* reward pursuit, not blindly drive it like a catecholaminergic id. From this perspective, compulsive pathology arises from the *loss* of dopamine's capacities to adapt appetitive behavior to the environment. Instead of thinking of dopamine as a "reward neurotransmitter," we might just as well consider it a "flexibility neurotransmitter." This is not merely semantic but can reframe the questions we ask, how experiments are designed and how data are interpreted. However, conceptual frameworks for thinking about dopamine and behavioral flexibility are not as well developed as those that emphasize dopamine and reward.

The papers in this research topic highlight several themes relevant to viewing dopamine as a neurotransmitter that has evolved to promote flexibility in the service of appetitive pursuit. Fundamental to behavioral flexibility is the ability to strike the right balance between exploiting acquired knowledge and exploring to update one's knowledge, the so-called explore-exploit dilemma, which dopamine may regulate (Beeler et al., 2010; Humphries et al., 2012). Nelson and Killcross (2013), investigating habit formation—the exploit end of this continuum—find that blocking D2 receptors enhances accelerated habit formation associated with amphetamine administration while D1 blockade, in contrast, reverses this effect. These data suggest that diminished D2 signaling may favor exploitation of prior learning and imply, conversely, that D2 promotes behavioral flexibility. This theme is repeated in Barker et al. (2013) study of the role of dopamine in the prefrontal cortex where they observe that *activation* of D2 receptors facilitates flexible behavioral responding and appears necessary for restoring sensitivity to changes in outcome value; that is, necessary for updating value associated with actions and subsequent behavior. Similar to Nelson, Barker found, conversely, that *antagonism* of D1 facilitated flexibility. These studies, together, suggest that D1 may facilitate exploitation of prior learning while D2 may promote behavioral flexibility. Klanker et al. (2013) systematically review the role of dopamine in different aspects of cognitive flexibility. Consistent with Nelson and Barker, D2 signaling again emerges as a key substrate mediating adaptive flexibility. Behavioral flexibility depends upon how information is stored and how it can be accessed and utilized. Eppinger et al. (2013) report a study examining differences between younger and older adults in the utilization of model-based and model-free value representation, reporting that older adults exhibit greater reliance on less flexible model-free strategies.

Of course, the dopamine system is not monolithic. Although we may abstract general themes, as above, other papers in this topic remind us that dopamine contributes to a mosaic of functions. In a review of its role in prefrontal function, Floresco (2013) challenges the widely held notion that the effects of dopamine can be characterized uniformly as an inverted U dose-response function. Instead, he suggests, the interaction between D1 and D2 varies depending upon what aspect of behavioral flexibility is being examined, suggesting that realistic dose-response characterization of dopamine requires a family of functions in which different aspects of behavioral flexibility respond differently to changes in dopamine signaling. While we often focus on dopaminergic modulation of specific targets, van der Schaaf et al. (2013) remind us that these targets do not operate in isolation. As above, the authors found that D2 played a key role in mediating behavioral flexibility (reversal learning), but that it did so in a way that depended on individual differences in anatomical connections between the orbitofrontal cortex and the amygdala. These data highlight that dopamine's effects on behavioral flexibility involve changes in communication between structures.

Baudonnat et al. (2013) take a step back and consider the problem of behavioral flexibility in a more complex, naturalistic context in which part of the challenge faced by the animal is determining which stimuli may or may not be relevant and when to use prior learning, or memory, versus when to pursue new learning. In their review, they focus on elaborating a dopaminergic mechanism—incorporating the different timescales of dopamine action and its modulation of corticostriatal plasticity—by which the animal determines "how surprising" the world is and adjust behavior according to the degree of uncertainty.

Associated primarily with signaling positive, rewarding outcomes, what role does dopamine play in altering behavior in response to deleterious outcomes? Oleson and Cheer (2013) review recent, elegant work in their lab using cyclic voltammetry to characterize changes in dopamine signaling across the course of aversive avoidance conditioning and show that dopamine release in response to cues preceding an aversive stimulus changes as that cue shifts from signaling *fear* to signaling *safety,* once the animal learns to avoid the aversive stimulus, engaging dopamine incentive learning.

Finally, returning to pathology in the dopamine system, Cepeda and Levine (Chen et al., 2013) examine the time course of dopaminergic pathophysiology in Huntington's disease where both hyper- and hypo- dopaminergic pathologies emerge sequentially over time contributing to shifting pathologies in behavioral flexibility. Their detailed review highlights mechanistic questions, consistent with a recurring theme, on the differential contribution of D1 and D2 across this progression. Their analysis reframes critical mechanistic questions and highlights how a more nuanced understanding of dopamine pathophysiology may lead to better therapeutics.

The traditional view of dopamine as the "reward neurotransmitter" has, in recent years, been gradually evolving. Through the work of many, dopamine is increasingly framed in neuroeconomic, decision-making terms (e.g., Gan et al., 2010; Schultz, 2010), shifting emphasis from driving reward pursuit to learning about value associated with stimuli and actions and adapting behavior accordingly. Implicit in this shift is an increasing emphasis on how dopamine mediates adaptive flexibility: rather than being the *id* of neurotransmitters, dopamine emerges more as the brain's chief comptroller in energy allocation (Beeler et al., 2012). We believe much is to be gained from explicitly reframing of dopamine's "root function" not as mediating reward, but as mediating behavioral flexibility in the allocation and pursuit of resources. We hope the contributions in this research topic stimulate thinking about dopamine as the "flexibility neurotransmitter" given its critical role in appetitive motivation.

## REFERENCES


Barker, J. M., Torregrossa, M. M., and Taylor, J. R. (2013). Bidirectional modulation of infralimbic dopamine D1 and D2 receptor activity regulates flexible reward seeking. *Front. Neurosci.* 7:126. doi: 10.3389/fnins.2013.00126

Baudonnat, M., Huber, A., David, V., and Walton, M. E. (2013). Heads for learning, tails for memory: reward, reinforcement and a role of dopamine in determining behavioral relevance across multiple timescales. *Front. Neurosci.* 7:175. doi: 10.3389/fnins.2013.00175

Beeler, J. A., Daw, N., Frazier, C. R. M., and Zhuang, X. (2010). Tonic dopamine modulates exploitation of reward learning. *Front. Behav. Neurosci.* 4:170. doi: 10.3389/fnbeh.2010.00170

Beeler, J. A., Frazier, C. R. M., and Zhuang, X. (2012). Putting desire on a budget: dopamine and energy expenditure, reconciling reward and resources. *Front. Integr. Neurosci.* 6:49. doi: 10.3389/fnint.2012.00049

Chen, J. Y., Wang, E. A., Cepeda, C., and Levine, M. S. (2013). Dopamine imbalance in Huntington's disease: a mechanism for the lack of behavioral flexibility. *Front. Neurosci.* 7:114. doi: 10.3389/fnins.2013.00114

Eppinger, B., Walter, M., Heekeren, H.R., and Li, S. (2013). Of habits and goals: age-related and individual differences in goal-directed decision-making. *Front. Neurosci.* 7:253. doi: 10.3389/fnins.2013.00253

Floresco, S. B. (2013). Prefrontal dopamine and behavioral flexibility: shifting from an "inverted-U" toward a family of functions. *Front. Neurosci.* 7:62. doi: 10.3389/fnins.2013.00062

Gan, J. O., Walton, M. E., and Phillips, P. E. M. (2010). Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nat. Neurosci.* 13, 25–27. doi: 10.1038/nn.2460

Humphries, M. D., Khamassi, M., and Gurney, K. (2012). Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia. *Front. Neurosci.* 6:9. doi: 10.3389/fnins.2012.00009

Klanker, M., Feenstra, M., and Denys, D. (2013). Dopaminergic control of cognitive flexibility in humans and animals. *Front. Neurosci.* 7:201. doi: 10.3389/fnins.2013.00201

Nelson, A. J. D., and Killcross, S. (2013). Accelerated habit formation following amphetamine exposure is reversed by D1, but enhanced by D2, receptor antagonists. *Front. Neurosci.* 7:76. doi: 10.3389/fnins.2013.00076

Oleson, E. B., and Cheer, J. F. (2013). On the role of subsecond dopamine release in conditioned avoidance. *Front. Neurosci.* 7:96. doi: 10.3389/fnins.2013.00096

Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behav. Brain Funct.* 6, 24. doi: 10.1186/1744-9081-6-24

van der Schaaf, M. E., Zwiers, M. P., van Schouwenburg, M. R., Geurts, D. E. M., Schellekens, A. F. A., Buitelaar, J. K., et al. (2013). Dopaminergic drug effects during reversal learning depend on anatomical connections between the orbitofrontal cortex and the amygdala. *Front. Neurosci.* 7:142. doi: 10.3389/fnins.2013.00142


*Received: 31 December 2013; accepted: 08 January 2014; published online: 24 January 2014.*

*Citation: Beeler JA, Cools R, Luciana M, Ostlund SB and Petzinger G (2014) A kinder, gentler dopamine... highlighting dopamine's role in behavioral flexibility. Front. Neurosci.* ***8****:4. doi: 10.3389/fnins.2014.00004*

*This article was submitted to Decision Neuroscience, a section of the journal Frontiers in Neuroscience.*

frontiers in
NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 15 May 2013
doi: 10.3389/fnins.2013.00076

# Accelerated habit formation following amphetamine exposure is reversed by $D_1$, but enhanced by $D_2$, receptor antagonists

***Andrew J. D. Nelson*[1]* and *Simon Killcross*[2]**

[1] School of Psychology, Cardiff University, Cardiff, UK
[2] School of Psychology, University of New South Wales, Sydney, NSW, Australia

***Edited by:***
*Sean B. Ostlund, University of California, Los Angeles, USA*

***Reviewed by:***
*Xinying Cai, Washington University in St Louis, USA*
*Kimberly H. LeBlanc, University of California, Los Angeles, USA*

***Correspondence:**
*Andrew J. D. Nelson, School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff, CF10 3AT, UK.*
*e-mail: nelsona5@cf.ac.uk*

Repeated exposure to the psychostimulant amphetamine has been shown to disrupt goal-directed instrumental actions and promote the early and abnormal development of goal-insensitive habitual responding (Nelson and Killcross, 2006). To investigate the neuropharmacological specificity of this effect as well as restore goal-directed responding in animals with pre-training amphetamine exposure, animals were treated with the non-selective dopamine antagonist α-flupenthixol, the selective $D_1$ antagonist SCH 23390 or the selective $D_2$ antagonist eticlopride, prior to instrumental training (three sessions). Subsequently, the reinforcer was paired with LiCL-induced gastric-malaise and animals were given a test of goal-sensitivity both in extinction and reacquisition. The effect of these dopaminergic antagonists on the sensitivity of lever press performance to outcome devaluation was assessed in animals with pre-training exposure to amphetamine (Experiments 1A–C) or in non-sensitized animals (Experiment 2). Both α-flupenthixol and SCH23390 reversed accelerated habit formation following amphetamine sensitization. However, eticlopride appeared to enhance this effect and render instrumental performance compulsive as these animals were unable to inhibit responding both in extinction and reacquisition, even though a consumption test confirmed they had acquired an aversion to the reinforcer. These findings demonstrate that amphetamine induced-disruption of goal-directed behavior is mediated by activity at distinct dopamine receptor subtypes and may represent a putative model of the neurochemical processes involved in the loss of voluntary control over behavior.

**Keywords: amphetamine, sensitization, $D_1$ and $D_2$ receptor subtypes, habits, goal-directed**

Repeated administration of psychostimulants such as amphetamine leads to behavioral sensitization and induces the appearance of repetitive and stereotyped behaviors that become more exaggerated following additional drug exposure (Kalivas et al., 1993; Canales et al., 2002; Capper-Loup et al., 2002). Consequently, behavioral sensitization has been used to investigate the neural basis of neuropsychiatric disorders that manifest as inflexible and repetitive patterns of behavior such as drug addiction, OCD and Tourette's (e.g., Canales and Graybiel, 2000; Graybiel and Rauch, 2000; Nestler, 2001; Saka et al., 2004).

Significantly, it has been reported that psychostimulant sensitization also disrupts goal-directed behavior leading to the early and abnormal onset of behaviorally-inflexible habitual responses that are not controlled by their consequences but rather by antecedent stimuli (Nelson and Killcross, 2006; Nordquist et al., 2007). During the early stages of acquisition, instrumental performance is normally sensitive to post-conditioning changes in reward value (Adams and Dickinson, 1981; Dickinson and Balleine, 1994) but as training proceeds, response control is ceded to the habit system and as a consequence becomes less sensitive to changes in reward value (Adams, 1982). However, following pre-training exposure to amphetamine, animals display habit-based instrumental performance that is insensitive to outcome devaluation even with only limited amounts of training (Nelson and Killcross, 2006; Nordquist et al., 2007). This finding supports evidence from lesion studies for a dissociation of neural systems that subserve the performance of voluntary goal-directed actions and reflexive, stimulus-bound habitual responding respectively (Coutureau and Killcross, 2003; Killcross and Coutureau, 2003; Yin et al., 2004, 2005; Naneix et al., 2009; Balleine and O'Doherty, 2010) and suggests that the balance between these two systems is acutely sensitive to manipulations of forebrain dopamine (e.g., Faure et al., 2005; Belin and Everitt, 2008). As imbalances in systems that control instrumental behavior are likely to be involved in the development of habitual drug-taking (Everitt and Robbins, 2005; Hogarth et al., 2013) as well as contribute to the production of involuntary and repetitive behaviors associated with OCD and Tourette's Syndrome (Ridley, 1994; Graybiel and Rauch, 2000; Leckman and Riddle, 2000; Gillan et al., 2011), it is of critical importance to understand the neural mechanisms that underpin the transfer of response control from the goal-directed to the habit system.

The neuropharmacological specificity of this effect, however, remains to be elucidated. Indeed, there is evidence that $D_1$ and

$D_2$ receptor subtypes may have dissociable effects on learning (Beninger and Miller, 1998). For example, it has been shown that Pavlovian approach behavior is attenuated by $D_1$ but facilitated by $D_2$ antagonists (e.g., Eyny and Horvitz, 2003). These dissociable effects on learning mirror findings that $D_1$ and $D_2$ receptor subtypes are differentially involved in long-term potentiation (LTP) and depression (LTD) within the striatum: LTP is blocked by $D_1$ antagonists (Kerr and Wickens, 2001) but is enhanced by $D_2$ antagonists and in $D_2$ receptor knock-out mice (Calabresi et al., 1997; Yamamoto et al., 1999). Thus, there are good reasons to assume that the enhancement of S-R habits by prior amphetamine exposure may be mediated by activity at distinct dopamine receptor subtypes.

In the current experiments we sought to reverse amphetamine-induced disruption of goal-directed responding as well as explore the neuropharmacological specificity of this effect by administering both non-selective and selective dopamine antagonists during training. Animals were treated with the non-selective dopamine antagonist α-flupenthixol, the selective $D_1$ antagonist SCH 23390 or the selective $D_2$ antagonist eticlopride during the acquisition of a moderately trained instrumental response. Subsequently, the reinforcer was devalued by LiCl-induced gastric-malaise and animals' propensity to press the lever was indexed both in extinction and reacquisition. In Experiments 1A–C the dopamine antagonists were administered to animals that had received pre-training exposure to amphetamine and in Experiment 2 the effect of these dopamine antagonists was assessed in non-sensitized animals.

## MATERIALS AND METHODS

### SUBJECTS

Male Lister hooded rats were used in these Experiments (Experiment 1A $n = 32$; Experiment 1B $n = 32$; Experiment 1C $n = 32$; Experiment 2 $n = 64$; Harlan UK Ltd., Bicester, Oxon, UK). At the start of behavioral test animals weighted between 263 and 389 g. Rats were housed in pairs in a climate-controlled vivarium (lights on 8:00 A.M. to 8:00 P.M.) and were tested during the light phase of the cycle. All experimental procedures involving animals and their care were carried out in accordance with the UK Animals Scientific Procedures Act (1986) and were subject to Home Office approval (Project License PPL 30/2158).

### DRUGS

For the sensitizing injections (Experiments 1A–C see below) and activity assay (all experiments) *d*-amphetamine sulphate was dissolved in sterile phosphate buffered saline (PBS). Doses of *d*-amphetamine sulphate, 2 mg/kg (sensitizing treatment) and 0.5 mg/kg (activity assay), were calculated as the salt. Alpha-flupenthixol (Experiment 1A) was dissolved in 0.9% physiological saline and administered intraperitoneally (i.p.) 20 min prior to instrumental conditioning at a dose of 0.3 mg/kg. SCH23390 was dissolved in 0.9% physiological saline and administered i.p. 15 min prior to instrumental conditioning at a dose of 0.005 mg/kg. Eticlopride was dissolved in 0.9% physiological saline and administered i.p., 15 min prior to instrumental conditioning at a dose of 0.05 mg/kg (Experiment 1C) and a lower dose of 0.02 mg/kg (Experiment 2). For all drugs, 0.9% saline served as control vehicle solution. All drugs were purchased from Sigma-Aldrich, UK.

### APPARATUS

The training apparatus comprised eight chambers (Paul Fray Ltd, Cambridge, UK) measuring 25 × 25 × 22 cm. The chambers were individually housed within sound-attenuating cabinets and were ventilated by low noise fans. Each chamber had three aluminum walls and a clear Perspex front wall. The roof was made of clear Perspex and the floor consisted of 18, 5 mm diameter steel bars spaced 1.5 mm apart centre-to-centre, parallel to the back of the chamber. A recessed magazine that provided access to rewards via a hinged Plexiglas panel was located in the centre of the left-hand wall. The liquid rewards (0.1 ml) could be delivered into the magazine via a peristaltic pump. The reinforcers used were 20% w/v sucrose solution flavored with grape Kool-Aid (0.05% w/v) and 20% w/v maltodextrin solution flavored with cherry Kool-Aid (0.05% w/v) (Cybercandy Ltd., London, UK). Pilot studies indicated that in normal rats these reinforcers were well matched for motivational value but could be easily discriminated. Levers could be inserted to the left and the right of the magazine. A houselight (3W) mounted in the roof provided general illumination. The apparatus and on-line data collection were controlled by means of an IBM-compatible microcomputer equipped with MED-PC software (Med Associates Inc., VT).

### SENSITIZATION

In Experiments 1A–C all rats received *i.p.* injections of 2 mg/kg *d*-amphetamine sulphate once per day for 7 consecutive days. Rats were returned to their home cages immediately after each injection. Over a seven-day injection-free period, animals were reduced to 80% of their *ad libitum* weight, prior to the start of behavioral training. In Experiment 2, animals underwent the same procedure but received i.p. injections of the equivalent volume of saline.

### BEHAVIORAL TRAINING

Following the sensitization procedure each animal was assigned to one of the eight conditioning chambers, and thereafter was always trained in that chamber. At the start of each session, the house light came on and remained on throughout the session. The house light went out at the end of each session. Behavioral training consisted of three stages: magazine training, instrumental training and devaluation by LiCl.

#### *Magazine training*

All rats were trained to collect food rewards during two, 30 min magazine training sessions. Half the animals were trained to collect the sucrose solution and the other half the maltodextrin solution (counter-balanced across treatment and devaluation groups). The rewards were delivered on a random time (RT) 60 s schedule by which rewards were delivered, on average, every 60 s.

#### *Lever press training and administration of dopamine antagonists*

The rats were initially trained to lever press during two sessions on a continuous schedule of reinforcement, with each press producing reward. One lever was inserted into the chamber at the beginning of the session and retracted at the end of the session.

Each session continued until the rat had earned 25 reinforcers. In the next three sessions of training, rewards were delivered according to a random interval (RI) 30 s schedule (reward available on average every 30 s and delivered following the next lever press). As current evidence suggests that the critical determinant of sensitivity to outcome devaluation is the degree of exposure to the reinforcer rather than the number of responses made, the number of reinforcers earned during acquisition was strictly controlled (Adams, 1982). Thus In each session, animals earned a total of 40 reinforcers so by the end of training animals had earned a total of 120 rewards on this schedule. This protocol has been shown previously to produce goal-directed responding in controls but accelerated habit formation in amphetamine sensitized animals (Nelson and Killcross, 2006). Prior to each of these lever press training sessions, animals received an i.p. injection of a dopamine antagonist (Drug groups) or the equivalent volume of control vehicle solution (Control group). In Experiments 1A–C half the animals (group Drug) received injections of a dopamine antagonist (α-flupenthixol in 1A, SCH23390 in 1B and eticlopride in 1C) and the other half (Controls) injections of saline. In Experiment 2, 16 animals were administered with α-flupenthixol, 16 with SCH23390, 16 with eticlopride and 16 served as vehicle-injected controls. The experimental conditions are summarized in **Table 1**.

## DEVALUATION BY LITHIUM CHLORIDE

### *Taste aversion training*

After the final day of instrumental lever press training, animals received three days of reward devaluation training with LiCl. On each day the rats were placed in the operant chambers and were given 40 free presentations of the instrumental outcome on an RT 30-s schedule. Immediately after the cessation of each session, the devalued group received a 0.15 M, 10 ml/kg (*i.p.*) injection of LiCl solution (Sigma–Aldrich, UK) and the non-devalued group an injection of the equivalent volume of saline. Taste aversion training was conducted drug-free.

### *Extinction test*

24 h after the final session of taste aversion training, animals were placed in the conditioning chambers and received a 10-min, drug-free extinction test conducted in the absence of reward delivery. During this test, lever press performance and magazine entry behavior were assessed.

### *Reacquisition test*

In order to confirm that the taste aversion procedure had successfully devalued the outcome for the devalued groups, all animals underwent a 15-min, drug-free reacquisition test. One day after the extinction tests, the animals were placed in the conditioning chambers and lever pressed to earn the instrumental outcome on an RI 30-s schedule.

### *Consumption test (Experiment 1C only)*

One day after the reacquisition test, animals were placed in feeding cages and given unrestricted access to the instrumental outcome for 15 min. The test was conducted drug-free.

## ACTIVITY ASSAY

To confirm sensitization, all animals were administered a 0.5 mg/kg (*i.p.*) amphetamine challenge before assessment of levels of locomotor activity. These tests occurred immediately following the re-acquisition tests. Activity was monitored using eight chambers (56 cm wide × 39 cm deep × 19 cm high). Activity within each chamber was recorded with pairs of photobeams situated 20 cm apart and 18 cm from the end of the cage connected to a control box (Paul Fray, Cambridge, UK). Each beam break resulted in an incremental count for that chamber and was recorded by an Acorn computer programmed in BBC Basic. Locomotor activity was measured (total number of photobeam breaks) for 30 min.

## DATA ANALYSIS

Statistical analysis was performed using analysis of variance (ANOVA) with between subject factors of devaluation (devalued versus non-devalued) and drug treatment (either dopamine antagonist or saline). As the standard deviation was proportional to the mean, the extinction data were subject to logarithmic transformations (Howell, 2002). Significant main effects with more than two levels were explored with Tukey *post-hoc* tests.

**Table 1 | Summary of main experimental findings.**

| Pre-treatment | Drug | Acquisition | | Extinction | | Reacquisition | |
|---|---|---|---|---|---|---|---|
| | | LP | Mag | LP | Mag | LP | Mag |
| Amphetamine | α-flupenthixol | ↓ | – | ✓ | ✓ | ✓ | ✓ |
| Amphetamine | SCH23390 | ↓ | – | ✓ | ✓ | ✓ | ✓ |
| Amphetamine | Eticlopride | ↓ | – | ✗ | ✓ | ✗ | ✓ |
| Saline | α-flupenthixol | – | – | ✓ | ✓ | ✓ | ✓ |
| Saline | SCH23390 | ↓ | – | ✓ | ✓ | ✓ | ✓ |
| Saline | Eticlopride | ↓ | – | ? | ✓ | ✓ | ✓ |

*In Experiment 1, animals underwent amphetamine sensitization before receiving injections of different dopaminergic antagonists while acquiring an instrumental response. Sensitivity to outcome devaluation was subsequently indexed both in extinction and reacquisition. In Experiment 2, non-sensitized animals underwent identical pharmacological and behavioral procedures. ↓ denotes reduced response rates during acquisition relative to saline injected-controls. ✓ denotes sensitivity and ✗ denotes insensitivity to outcome devaluation during extinction and reacquisition tests. LP denotes lever press behavior and Mag denotes magazine entry behavior.*

## RESULTS

### EXPERIMENT 1A. THE EFFECT OF α-FLUPENTHIXOL ON SENSITIVITY TO OUTCOME DEVALUATION AFTER LIMITED TRAINING IN ANIMALS PRE-TREATED WITH AMPHETAMINE

#### *Instrumental training*

By the end of the three days of RI30 training, all animals had acquired the instrumental response and achieved a stable level of responding. However, α-flupenthixol treatment produced overall lower rates of responding compared to sensitized animals treated with saline. This was confirmed statistically by a main effect of drug [$F_{(1, 28)} = 7.982$, $p < 0.01$] [Mean lever presses per minute (±SEM) AMP + saline group = 10.959 (±1.195); AMP + α-flupenthixol = 7.902 (±0.895)]. However, as the length of each session was determined by the number of reinforcers earned (40 in each) and not time, α-flupenthixol treated animals obtained the same number of reinforcers (120) as controls and hence any differential sensitivity to outcome devaluation observed in the subsequent extinction test cannot be accounted for in terms of differential exposure to the reinforcer. As the critical comparisons at test are between devalued and non-devalued groups within each drug group, it is unlikely that any differences in sensitivity to outcome devaluation are due to these baseline effects. Significantly in this respect, there was neither an effect of intended devaluation ($F < 1$) nor an interaction between drug and devaluation ($F < 1$). In contrast to the depressive effects of α-flupenthixol on lever press acquisition, there was no effect of drug on magazine entry behavior [mean magazine entries per minute (±SEM): AMP + saline group = 5.478 (±1.399); AMP + α-flupenthixol group = 4.642 (±0.974)]. ANOVA yielded no effect of drug ($F < 1$) or devaluation [$F_{(1, 28)} = 2.224$, $p = 0.145$], and no interaction ($F < 1$).

#### *Extinction test—lever press performance*

In order to take account of baseline differences and reduce within subject variability in ANOVA, lever press performance in the extinction test is presented as a proportion of baseline responding. These are presented in the left-hand panel of **Figure 1**. The suggestion from this figure is that administration of α-flupenthixol during training (group AMP + α-flupenthixol) restored goal-sensitivity as the animals in the devalued group (black bars) showed a selective depression in lever press rates compared to animals in the non-devalued group (white bars). On the other hand the responding of animals exposed to amphetamine before training but administered saline during training (group AMP + saline) appeared to be impervious to the current value of the reinforcer as shown by equivalent rates across the two devaluation groups. There was no effect of drug ($F < 1$) but there was a main effect of devaluation [$F_{(1, 28)} = 6.598$, $p < 0.05$] and critically ANOVA revealed a significant drug × devaluation interaction [$F_{(1, 28)} = 4.296$, $p < 0.05$]. Simple effects analysis of this interaction showed that pre-training amphetamine exposure rendered instrumental performance independent of reward value as there was no devaluation effect in these animals ($F < 1$), but there was an effect of devaluation in the animals treated with α-flupenthixol [$F_{(1, 14)} = 7.147$, $p < 0.05$]. The higher rates of responding in the non-devalued α-flupenthixol-treated rats relative to the non-devalued saline-treated rats may have contributed to the this devaluation effect but simple effects analysis revealed no effect of drug in the non-devalued condition [$F_{(1, 14)} = 1.98$, $p = 0.17$]. As such these results replicate previous findings (Nelson and Killcross, 2006) that pre-training amphetamine exposure leads to accelerated habit formation and suggest that this is an effect reversed by the non-selective dopamine antagonist α-flupenthixol.

#### *Extinction test—magazine entry behavior*

Analysis of magazine entry behavior during the extinction test revealed a main effect of devaluation [$F_{(1, 28)} = 12.836$, $p < 0.001$] but no effect of drug or interaction between these factors (both $F$s $< 1$). [Mean magazine entries as a proportion

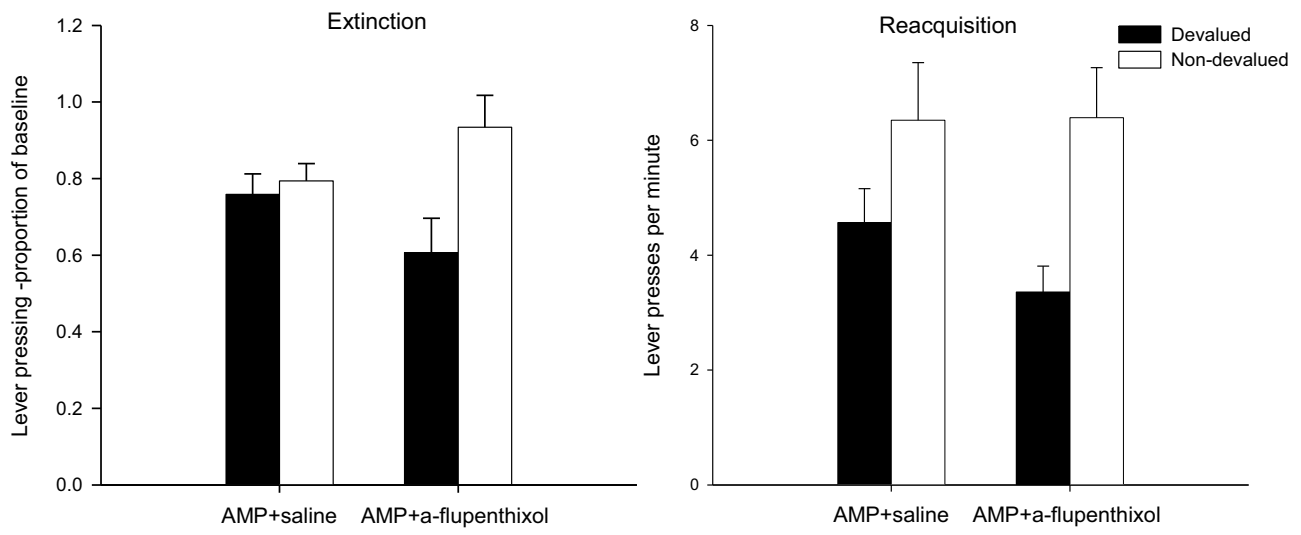



**FIGURE 1 | Effect of α-flupenthixol following amphetamine sensitization on sensitivity of lever pressing to reward devaluation by LiCl-induced nausea.** Mean lever presses per minute as a proportion of baseline (±SEM) in the extinction test (left-hand panel) and lever presses per minute (±SEM) in the rewarded reacquisition test after devaluation by LiCl (devalued—black bars) or no devaluation (non-devalued—white bars).

of baseline (±SEM) Devalued group = 0.389 (±0.106); Non-devalued group = 1.09 (±0.249)]. Thus in contrast to lever press performance, magazine entry behavior was sensitive to outcome value irrespective of drug and suggests that the LiCl treatment successfully devalued the value of the instrumental outcome.

#### *Reacquisition test—lever press performance*

The results of the rewarded reacquisition test confirmed that animals in both devaluation groups had developed an aversion to the reinforcer, as shown in right-hand panel of **Figure 1** by reduced lever press rates compared to non-devalued controls. ANOVA revealed an overall effect of devaluation [$F_{(1, 28)} = 10.036$, $p < 0.01$] but this was unaffected by drug group as there was no effect of drug or interaction (both $F$s $< 1$). Thus, the insensitivity to outcome devaluation in the extinction test observed in the AMP + saline group cannot be attributed to any differential impact of taste aversion training.

#### *Reacquisition test—magazine entry behavior*

Similarly, magazine entry behavior during the 15-min reacquisition test was sensitive to the changed value of the reinforcer. Both devalued groups performed considerably fewer magazine entries during the test compared to the non-devalued controls [$F_{(1, 28)} = 11.569$ $p < 0.01$] (Mean magazine entries per minute (±SEM) Devalued group = 1.912 (±0.377); Non-devalued group = 3.098 (±0.734). There was no effect of drug [$F_{(1, 28)} = 1.690$, $p = 0.204$] nor a drug × devaluation interaction ($F < 1$).

### EXPERIMENT 1B. THE EFFECT OF SCH23390 ON SENSITIVITY TO OUTCOME DEVALUATION AFTER LIMITED TRAINING IN ANIMALS PRE-TREATED WITH AMPHETAMINE

#### *Instrumental training*

All animals acquired the instrumental response but SCH23390 markedly attenuated the rate of responding in animals administered the drug prior to instrumental training. ANOVA yielded a highly significant main effect of drug [$F_{(1, 28)} = 36.392$, $p < 0.001$] [mean lever presses per minute (±SEM) AMP + saline group = 13.967 (±1.435); AMP + SCH23390 group = 6.605 (±0.896)] but no effect of intended devaluation or an interaction between these two factors (both $F$s $< 1$). However, all animals treated with SCH23390 earned all 120 rewards across the three training sessions and hence had the same exposure to the reinforcer as the animals administered saline during instrumental training. The depressive effects of SCH23390 on responding were restricted to lever pressing, as magazine approach behavior was unaffected by the drug [mean magazine entries per minute (±SEM) AMP + saline group = 5.176 (±0.807); AMP + SCH23390 group = 4.219 (±0.768)]. Statistically, there was no effect of drug [$F_{(1, 28)} = 1.434$, $p = 0.241$], intended devaluation nor an interaction (both $F$s $< 1$).

#### *Extinction test—lever press performance*

The lever press performance of saline injected and SCH23390-treated group during the 10-min extinction as a proportion of their baseline responding is presented in the left-hand panel of **Figure 2**. Inspection of this figure suggests that the instrumental performance of animals treated with SCH23390 during training was guided by outcome expectancy as the devalued group (black bars) performed fewer lever presses as a proportion of baseline compared to the non-devalued group (white bars). Conversely, the responding of the AMP + saline group in this test was not goal-directed as demonstrated by their failure to show sensitivity to the change in reward value. This description of the data was confirmed statistically by ANOVA which revealed a main effect of devaluation [$F_{(1, 28)} = 9.157$, $p < 0.01$], no effect of drug ($F < 1$) and significantly, a devaluation × drug interaction [$F_{(1, 28)} = 7.146$, $p < 0.05$]. Subsequent analysis of this interaction yielded no effect of devaluation in the AMP + saline group ($F < 1$) but devalued and non-devalued performance did differ statistically significantly in animals treated with SCH23390 [$F_{(1, 14)} = 8.821$, $p < 0.01$]. It is possible that the higher rates of responding in the SCH23390 non-devalued group may have contributed to the devaluation × drug interaction but simple effects found no evidence that there was an effect of drug in the non-devalued condition [$F_{(1, 14)} = 2.47$, $p = 0.13$]. These findings suggest that the $D_1$ receptor antagonist SCH23390 disrupted the more rapid onset of behavioral autonomy seen after sensitization with amphetamine.

#### *Extinction test—magazine entry behavior*

In contrast, magazine performance during the extinction test was sensitive to the changed value of the reinforcer in both drug groups [mean magazine entries as a proportion of baseline (±SEM) Devalued group = 0.495 (±0.117); non-devalued group = 1.411 (±0.275)]. Indeed, ANOVA revealed only a main effect of devaluation [$F_{(1, 28)} = 18.521$, $p < 0.001$], no effect of drug ($F < 1$) nor an interaction [$F_{(1, 28)} = 1.587$, $p = 0.218$].

#### *Reacquisition test—lever press performance*

The effectiveness of the taste aversion training in devaluing the instrumental outcome is further supported by analysis of lever press rates performed in the rewarded reacquisition test shown in the right-hand panel of **Figure 2**. The impression from this figure is that all animals were averted from the reinforcer, irrespective of drug treatment, and hence pressed the lever at lower rates compared to the non-devalued controls. Statistical analysis by ANOVA revealed a highly significant main effect of devaluation [$F_{(1, 28)} = 25.112$, $p < 0.001$] as well as a main effect of drug [$F_{(1, 28)} = 6.031$, $p < 0.05$] reflecting overall lower response rates in the SCH23390 group, but the level of devaluation in these animals was comparable to that of the AMP + saline animals as there was no drug × devaluation interaction ($F < 1$).

#### *Reacquisition test—magazine entry behavior*

Magazine entry behavior was equally sensitive to outcome value in both drug groups during the reacquisition test [Mean magazine entries per minute (±SEM): Devalued group = 3.983 (±1.408); Non-devalued group = 8.036 (±1.348)]. Statistically, there was an overall effect of devaluation [$F_{(1, 28)} = 10.524$, $p < 0.01$] but no effect of drug ($F < 1$) nor an interaction [$F_{(1, 28)} = 1.322$, $p = 0.26$]. Thus in contrast to lever press performance in the reacquisition test, magazine approach behavior was unaffected by SCH23390.

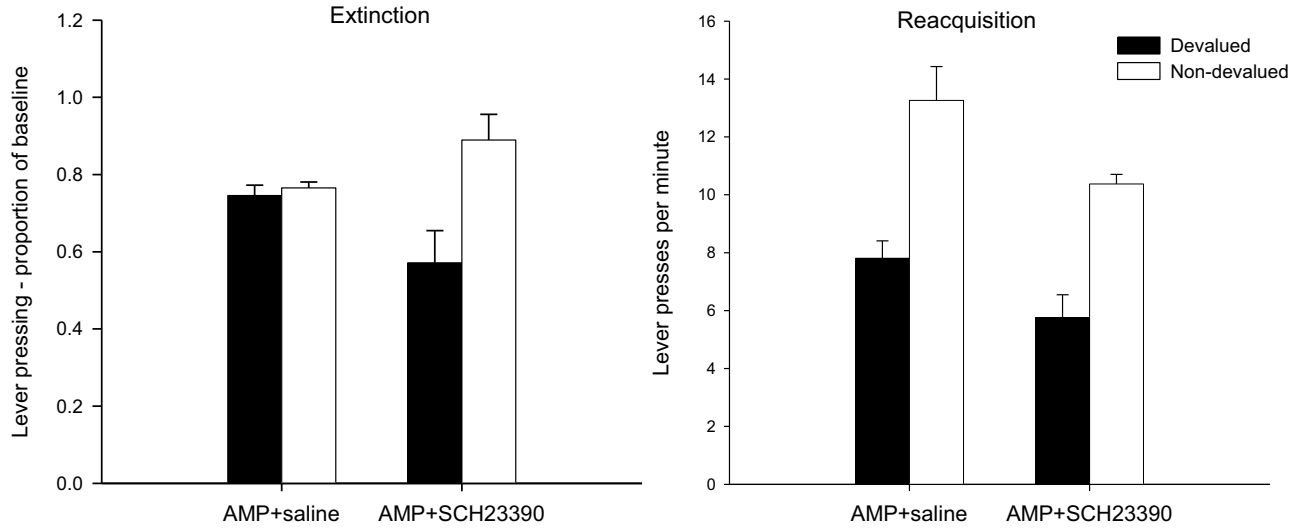



**FIGURE 2 | Effect of SCH23390 following amphetamine sensitization on sensitivity of lever pressing to reward devaluation by LiCl-induced nausea.** Mean lever presses per minute as a proportion of baseline (±SEM) in the extinction test (left-hand panel) and lever presses per minute (±SEM) in the rewarded reacquisition test after devaluation by LiCl (devalued—black bars) or no devaluation (non-devalued—white bars).

## EXPERIMENT 1C. THE EFFECT OF ETICLOPRIDE ON SENSITIVITY TO OUTCOME DEVALUATION AFTER LIMITED TRAINING IN ANIMALS PRE-TREATED WITH AMPHETAMINE

### *Instrumental training*

Both drug groups acquired the instrumental response, albeit at different rates. Eticlopride greatly reduced the rate of responding compared to animals given saline during training. Statistically, ANOVA revealed a highly significant main effect of drug [$F_{(1,\ 28)} = 34.205$, $p < 0.001$] [mean lever presses per minute (±SEM) AMP + saline group = 12.411 (±1.005); AMP + eticlopride group = 7.0122 (±0.795)] but no effect of devaluation group or an interaction (both $F$s < 1). As session length was determined by number of rewards earned (40 per session) rather than time, all the animals in the Eticlopride group earned the 120 rewards over the three sessions. Conversely, eticlopride had no impact on magazine entry behavior as there was no effect of drug, devaluation or an interaction (all $F$s < 1) [mean magazine entries per minute (±SEM): AMP + saline group = 4.311 (±0.610); AMP + eticlopride group = 5.055 (±0.834)].

### *Extinction test—lever press performance*

The mean lever presses per minute in the critical extinction test are presented in the left-hand panel of **Figure 3**. It is clear from this figure that none of the animals, irrespective of drug group, was sensitive to the changed value of the reinforcer as both devalued groups responded at equivalent rates to the non-devalued controls. This was confirmed statistically as there was no effect of devaluation ($F < 1$) and no interaction between drug and devaluation factors ($F < 1$). Eticlopride therefore failed to reverse the effect of pre-training amphetamine exposure on goal-sensitivity after limited training and responding in both groups was habitual even after limited training. However, ANOVA did reveal a highly significant main effect of drug [$F_{(1,\ 28)} = 15.578$, $p < 0.001$], reflecting overall higher rates of responding as a proportion of baseline in the eticlopride group. As the extinction test was conducted drug-free and the data were analyzed as a proportion of baseline, the effect of drug at test may in part reflect the lower rates of responding seen during acquisition under eticlopride. However, the finding that eticlopride-treatment led to reduced responding during acquisition but failed to abolish the enhancement of S-R habits by amphetamine sensitization suggests that the restoration of goal-sensitivity by α-flupenthixol and SCH23390 (Experiments 1A,B, see above) cannot be attributed to their depressive effects on response rates during acquisition alone.

### *Extinction test—magazine entry behavior*

Despite the insensitivity of lever pressing to outcome devaluation, it is clear from the right-hand panel of **Figure 3** that magazine entry behavior in both devalued groups was reduced compared to non-devalued controls. Statistical analysis revealed only an effect of devaluation [$F_{(1,\ 28)} = 10.576$, $p < 0.01$] and no effect of drug nor an interaction (both $F$'s < 1). Thus the demonstration that lever press performance in the extinction test was under the control of S-R habits, whereas magazine approach behavior was guided by outcome value, indicates that the LiCl treatments successfully devalued the instrumental outcome.

### *Reacquisition test—lever press performance*

The results of the rewarded reacquisition test revealed an intriguing dissociation in performance between the two drug groups. The saline treated animals averted from the reinforcer showed a clear devaluation effect: this is consistent with the direct punishment of S-R habits by the presentation of the nausea-inducing reinforcer and with previous findings that pre-training amphetamine exposure promotes lever press performance that is insensitive to outcome devaluation in extinction but not in reacquisition (see Nelson and Killcross, 2006). However, as is clear from the left-hand panel of **Figure 4**, the devalued animals

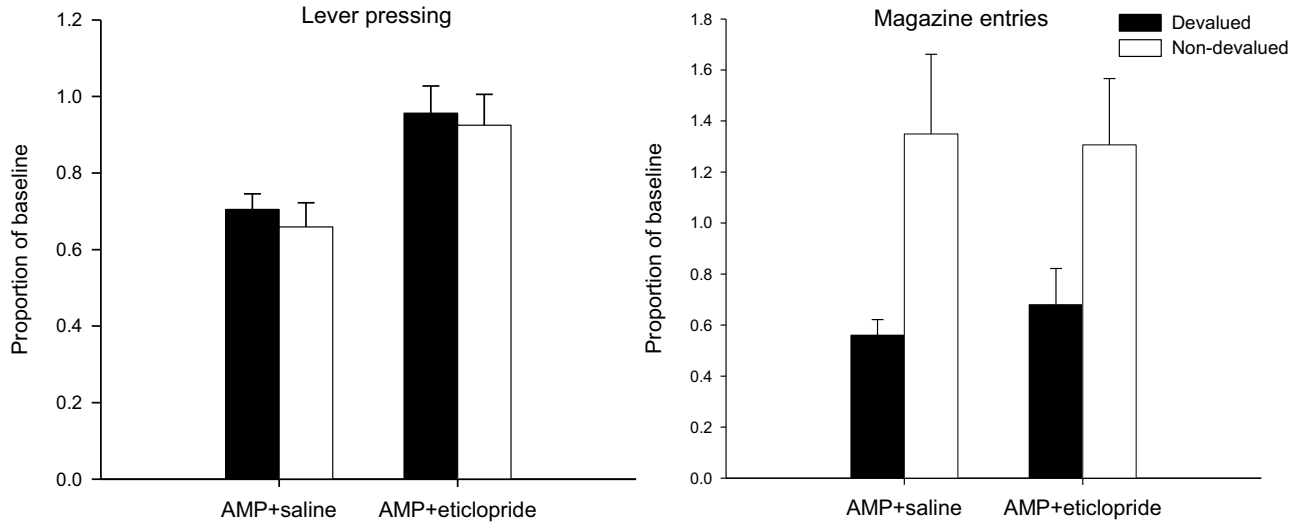



**FIGURE 3 | Effect of eticlopride following amphetamine sensitization on sensitivity of lever pressing (left-hand panel) and magazine entry (right-hand panel) to reward devaluation by LiCl-induced nausea.** Mean lever presses per minute and mean magazine entries per minutes as a proportion of baseline (±SEM) in the extinction test after devaluation by LiCl (devalued—black bars) or no devaluation (non-devalued—white bars).

in the eticlopride group pressed the lever at comparable rates to the non-devalued controls even though responding was reinforced with the reward that had been previously paired in these animals with gastric malaise. This description of the data was supported statistically by ANOVA which revealed a main effect of devaluation [$F_{(1,\ 28)} = 10.384$, $p < 0.01$], no effect of drug ($F < 1$), but crucially a significant interaction between these two factors [$F_{(1,\ 28)} = 5.472$, $p < 0.05$). Subsequent analysis of this interaction with simple effects confirmed that saline-treated animals had acquired an aversion to the reinforcer and could use this representation to guide instrumental performance when presented with the consequences of their actions in reacquisition as there was a highly significant effect of devaluation in these animals [$F_{(1,\ 14)} = 12.171$, $p < 0.01$]. There was no such effect in the eticlopride-treated animals ($F < 1$). This can be taken as evidence that instrumental performance in eticlopride treated animals was completely impervious to reward value and had become compulsive. However, it is possible that this insensitivity arose from a failure of the taste aversion training.

#### *Reacquisition test—magazine entry behavior*

Significantly, analysis of magazine entry behavior during the rewarded reacquisition test suggests that all animals, regardless of drug treatment, had acquired an aversion to the reinforcer. The mean magazine entries per minute in this test are displayed in the right-hand panel of **Figure 4** and in stark contrast to the lever press data reviewed above, magazine approach behavior was sensitive to reward value in both drug groups. ANOVA yielded no effect of drug ($F < 1$) and a highly significant effect of devaluation [$F_{(1,\ 28)} = 45.598, p < 0.001$). The suggestion from the right hand-panel of **Figure 4** is that the devaluation effect may have been slightly attenuated in the eticlopride group but there was no statistical evidence for this as the interaction failed to reach significance [$F_{(1,\ 28)} = 2.743, p = 0.109$].

#### *Consumption test*

In order to confirm that the differential sensitivity of lever press to reward value observed in the reacquisition test could not be explained in terms of any failure of eticlopride-treated animals to acquire an aversion to the reinforcer, all animals were given free access to the instrumental outcome and consumption was measured over a 15-min period. Results of this consumption test revealed that all animals averted from the reinforcer consumed statistically significantly less of the instrumental outcome compared to the non-devalued controls (mean consumption in ml (±SEM): Devalued AMP + saline group = 3.263 (±0.549); Non-devalued AMP + saline group = 6.175 (±0.966); Devalued AMP + eticlopride group = 4.15 (±0.711); Non-devalued AMP + eticlopride group = 7.00 (±0.603). ANOVA revealed a main effect of devaluation [$F_{(1,\ 28)} = 15.776$, $p < 0.001$] and a non-significant trend toward marginally higher overall consumption in eticlopride-treated animals [$F_{(1,\ 28)} = 1.393$, $p = 0.248$]. Critically, the devaluation effect was unaffected by drug as there was no interaction between these two factors ($F < 1$). Coupled with evidence that magazine entry behavior was sensitive to outcome value in both the extinction and reacquisition tests, the results of the consumption test confirm that all animals had acquired an aversion to the reinforcer and hence the effects of eticlopride on the sensitivity of lever pressing to reward value cannot be accounted for in terms of any ineffectiveness of the LiCl devaluation treatments.

### EXPERIMENT 2. THE EFFECT OF α-FLUPENTHIXOL, SCH23390, AND ETICLOPRIDE ON THE SENSITIVITY TO OUTCOME DEVALUATION AFTER LIMITED TRAINING IN NON-SENSITIZED ANIMALS

#### *Instrumental training*

As expected, the dopamine antagonists reduced the rate of responding and this effect was particularly marked in animals treated with SCH23390 and eticlopride [mean lever presses

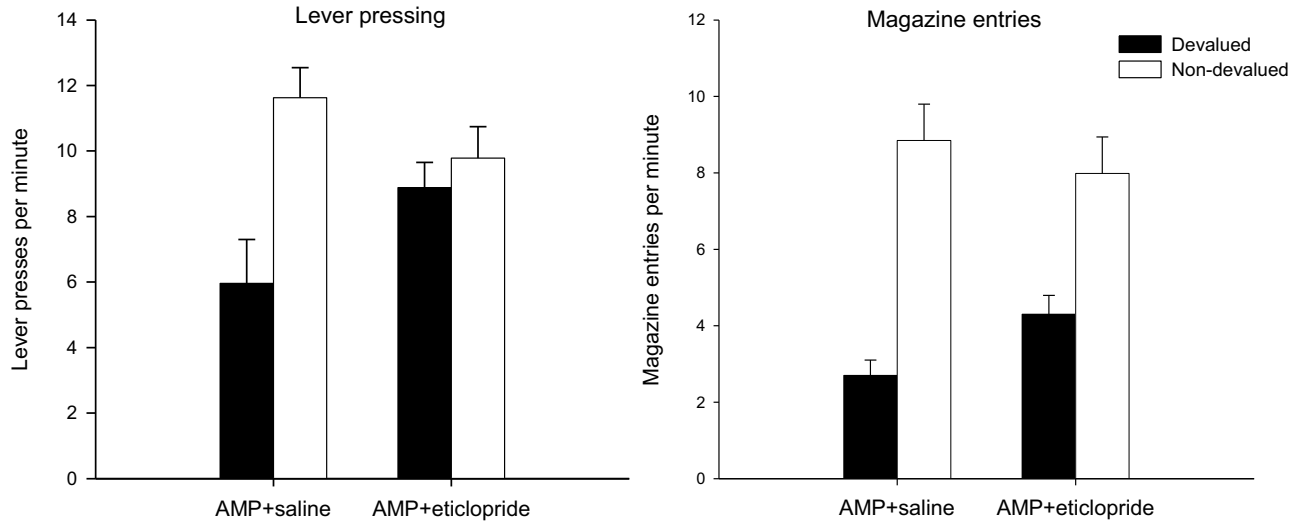



**FIGURE 4 | Effect of eticlopride following amphetamine sensitization on sensitivity of lever pressing (left-hand panel) and magazine entry (right-hand panel) reacquisition after reward devaluation by LiCl-induced nausea.** Mean lever presses per minute and mean magazine entries per minutes as a proportion of baseline (±SEM) in the rewarded reacquisition test after devaluation by LiCl (devalued—black bars) or no devaluation (non-devalued—white bars).

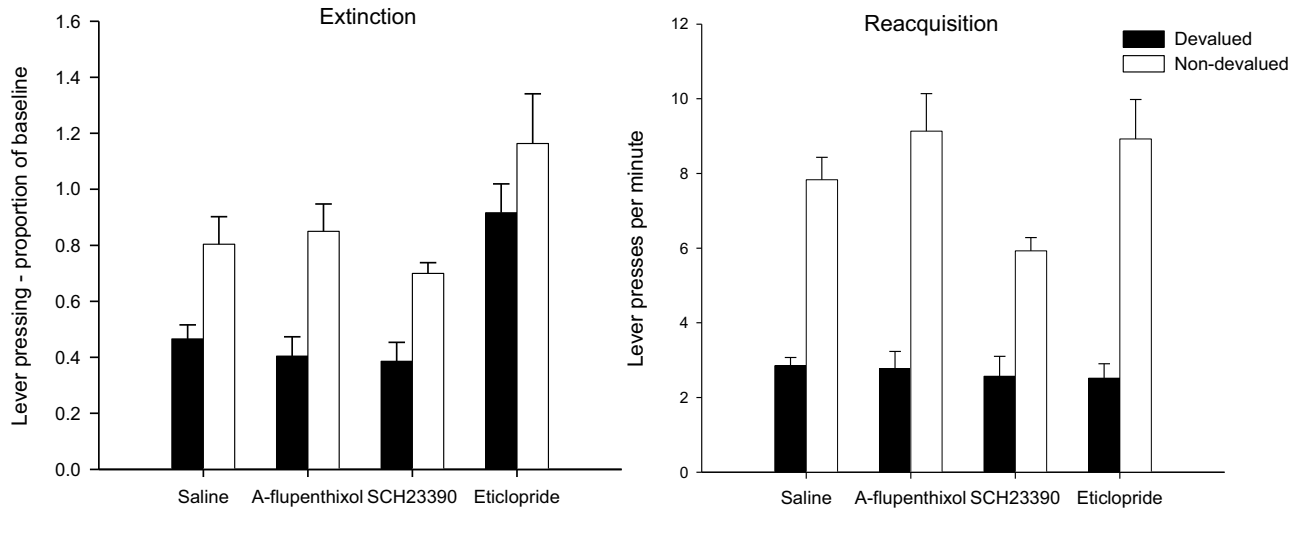



**FIGURE 5 | Effect of α-flupenthixol, SCH23390 and eticlopride on sensitivity of lever pressing to reward devaluation by LiCl-induced nausea.** Mean lever presses per minute as a proportion of baseline (±SEM) in the extinction test (left-hand panel) and lever presses per minute (±SEM) in the rewarded reacquisition test after devaluation by LiCl (devalued—black bars) or no devaluation (non-devalued—white bars).

per minute (±SEM) Saline group = 10.693 (±1.033); α-flupenthixol group = 8.504 (±1.249); SCH23390 group = 6.551 (±0.899); Eticlopride group = 5.819 (±1.245)]. Despite the reduction in the rate of responding, all animals earned 120 reinforcers across the three sessions. This description of the data was confirmed by ANOVA which revealed a main effect of drug [$F_{(3, 56)} = 7.397$, $p < 0.001$] but no effect of intended devaluation nor an interaction between these factors (both $F$'s $< 1$). Subsequent *post-hoc* analysis with Tukey tests confirmed that both SCH23390- ($p < 0.01$) and eticlopride- ($p < 0.001$) treated animals responded at lower rates than saline treated animals. However, magazine entry behavior was unaffected by any of these factors as there was no effect of drug, intended devaluation or interaction [highest $F_{(1, 56)} = 1.947$, $p = 0.168$] [Mean magazine entries per minute (±SEM): Saline group = 4.906 (±0.576); α-flupenthixol group = 4.575 (±0.634); SCH23390 group = 4.777 (±0.807); Eticlopride group = 3.809 (±0.678)].

### *Extinction test—lever press performance*

The left-hand panel of **Figure 5** displays the lever press performance in the extinction test following devaluation by LiCl.

Inspection of this figure suggests that saline controls and animals given SCH23390 and α-flupenthixol during training were goal-directed as animals averted from the reinforcer showed a marked suppression in lever press performance compared to non-devalued control animals. The suggestion from this figure is that the devaluation effect may have been attenuated in animals treated with eticlopride. However, ANOVA only revealed a main effect of devaluation [$F_{(1, 56)} = 24.317$, $p < 0.001$] and no interaction between drug and devaluation ($F < 1$). There was an effect of drug [$F_{(3, 56)} = 10.708$, $p < 0.001$] due to overall higher rates of responding in the eticlopride treated animals. *Post-hoc* Tukey tests revealed that the eticlopride-treated animals pressed at significantly higher rates than all other animals (all *ps* $< 0.01$).

#### *Extinction test—magazine entry behavior*

Analysis of magazine entry behavior during the 10-min extinction test suggests that the LiCl treatment successfully devalued the outcome for all animals as there was a main effect of devaluation [$F_{(1, 56)} = 9.661$, $p < 0.01$] [mean magazine entries as a proportion of baseline (±SEM) Devalued group = 0.443 (±0.059); Non-devalued group = 0.769 (±0.083)]. However, this effect was unaffected by drug group as there was no main effect of drug or an interaction (both *F*s $< 1$).

#### *Reacquisition test—lever press performance*

The results of the rewarded reacquisition test presented in the right-hand panel of **Figure 5** confirmed that all animals had acquired an aversion to the reinforcer. ANOVA yielded a highly significant effect of devaluation [$F_{(1, 56)} = 138.828$, $p < 0.001$] as well as an effect of drug [$F_{(3, 56)} = 2.774$, $p < 0.05$] reflecting lower responding in the SCH23390 group. *Post-hoc* Tukey tests showed that the rate of responding in SCH23390 treated animals differed only from that of α-flupenthixol group ($p < 0.05$). The overall lower responding in the SCH23390-treated animals and in particular the non-devalued SCH23390-treated animals, would account for a marginal significant drug × devaluation interaction [$F_{(3, 56)} = 2.512$, $p = 0.068$]. Nevertheless it is evident from the right-hand panel of **Figure 5** that all devalued groups had acquired a robust aversion to the instrumental outcome and consequently suppressed lever press responding during the rewarded test.

#### *Reacquisition test—magazine entry behavior*

This impression was also confirmed by analysis of magazine approach behavior during the rewarded reacquisition test, with all animals in the devalued groups performing fewer magazine entries compared to the non-devalued controls [$F_{(1, 56)} = 28.010$, $p < 0.001$]. There was also a main effect of drug [$F_{(3, 56)} = 6.521$, $p < 0.001$] as the α-flupenthixol treated animals had higher rates of magazine approach behavior ($p < 0.01$) but this heightened responding did not impact on sensitivity of magazine entry behavior to outcome devaluation as there was no drug x devaluation interaction ($F < 1$).

### ACTIVITY ASSAY

In order to confirm the presence of sensitization in amphetamine pre-treated animals, all animals were administered a 0.5 mg/kg amphetamine challenge allowing between subject comparisons of the locomotor activating effects of amphetamine in sensitized (Experiments 1A–C) and non-sensitized animals (Experiment 2). As is clear from **Figure 6** animals with prior experience of amphetamine showed elevated levels of locomotor activity compared to drug-naïve animals. ANOVA with between-subject factors of sensitization (sensitized with amphetamine or non-sensitized drug-naïve animals) and drug administered during training (saline, α-flupenthixol, SCH23390, or eticlopride) yielded a highly significant effect of sensitization [$F_{(1, 144)} = 48.909$, $p < 0.001$] but also an effect of drug [$F_{(3, 144)} = 4.798$, $p < 0.01$] due to higher locomotor activity in response to the amphetamine challenge in all animals treated with eticlopride during training. There was, however, no interaction between sensitization and drug [$F_{(3, 144)} = 1.702$, $p = 0.169$]. These results confirm that the amphetamine pre-treatment had successfully sensitized animals to amphetamine and provide indirect evidence that antagonism with the $D_2$ antagonist eticlopride enhances the locomotor activating effects of amphetamine irrespective of prior experience with amphetamine.

## DISCUSSION

The experiments reported here examined the effects of both non-selective and selective dopamine antagonists on instrumental performance in a reinforcer devaluation task either in animals given pre-training exposure to amphetamine (Experiments 1A–C) or in non-sensitized animals. Significantly, the experiments replicated our previous finding that pre-training exposure to amphetamine renders instrumental performance autonomous of the current value of the reinforcer even after limited training (Nelson and Killcross, 2006). The results demonstrated that accelerated habit formation seen after amphetamine sensitization is reversed by $D_1$, but enhanced by $D_2$ receptor antagonists. Furthermore, these experiments provided considerable insights into the role of $D_1$

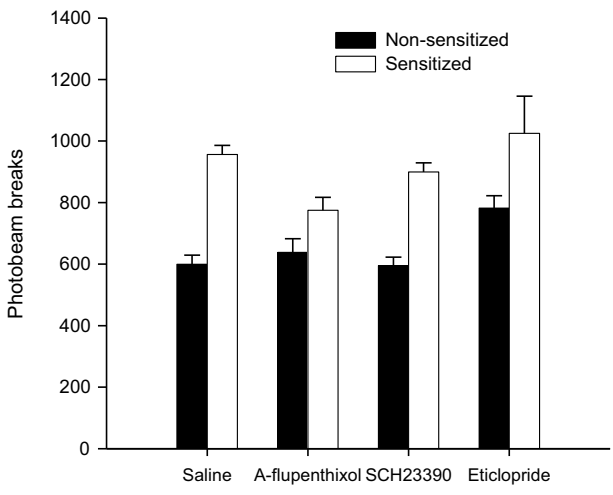



**FIGURE 6 | Activity assay—total mean photobeam breaks following a 0.5 mg/kg amphetamine challenge in sensitized (white bars) and non-sensitized (black bars) animals treated with saline, α-flupenthixol, SCH23390 or eticlopride during training.**

and $D_2$ receptor subtypes in mediating instrumental learning generally as well as susceptibility to devaluation procedures in sensitized and non-sensitized animals.

Before considering test performance, it is important to address the effect of the various dopamine antagonists on acquisition of the instrumental response. Consistent with previous reports from operant procedures (e.g., Tombaugh et al., 1979; Wise and Schwartz, 1981) the administration of dopamine antagonists during training severely retarded the rate of responding in both sensitized and non-sensitized animals. It is important to note that, despite these reduced rates, all animals earned the same number of reinforcers during training as session length was limited by the number of reinforcers earned and not time (Adams, 1982). In the current experiments there was some evidence of a dissociation at the receptor subtype level between the performance of instrumental responses under drug and their expression in drug-free tests. As dopamine has been implicated in various non-associative factors such as motivation, attention and sensorimotor control that contribute to learning, any effects of dopamine antagonism that are restricted to the performance of an instrumental response under drug can be attributed to these non-associative factors. However, if effects of dopaminergic manipulations are seen on the drug-free expression of learned instrumental responses, for example in the current experiments in the extinction and reacquisition tests, then this can be taken as evidence to suggest that a dopaminergic agent may have modulated the course of associative learning.

Here, the non-selective antagonist α-flupenthixol and the selective $D_2$ antagonist eticlopride reduced the rate of instrumental responding during acquisition, but at test the response rates, expressed as a proportion of these reduced baseline rates, were actually higher relative to saline controls. This recovery of responding in the drug-free extinction text indicates that antagonism of $D_2$ receptors may have disrupted the performance of that response during acquisition and not the expression of that learning in the drug-free extinction test. Moreover, these animals showed comparable rates of responding to saline controls in the drug-free reacquisition test. This could similarly be taken as evidence to suggest that these drugs disrupted the performance i.e., led to reduced rates of responding during acquisition but not the subsequent drug-free expression of instrumental conditioning in reacquisition. Dopamine activity in the nucleus accumbens, a structure containing the highest concentration of $D_2$ receptors in the rat brain (Bentivoglio and Morelli, 2005), has been widely implicated in the reinforcing and motivational properties of both natural rewards and drugs of abuse (e.g., Hernandez and Hoebel, 1988; Mark et al., 1994; Wyvell and Berridge, 2000). Thus the disruptive effects of agents selectively and non-selectively targeting $D_2$ receptors on the performance but not the drug-free expression of the instrumental response may have been due to decreased motivation associated with these drugs. However, it is equally possible, the reduced rate of responding could have arisen as a result of the profound motor impairments typically produced by $D_2$ antagonists (e.g., Fowler and Liou, 1994). Whether the disruption was caused by motivational or motor factors or a combination of the two, $D_2$ receptor antagonism appeared to impair the performance but not the drug-free expression of the instrumental response in the current experiments. Conversely, $D_1$ antagonism by SCH23390 not only affected the performance during training but it also reduced the expression of learned instrumental responses at reacquisition. The test was conducted five days after the last SCH23390 treatment and hence the reduced rate of responding may not be accounted for solely in terms of drug induced motivational or sensorimotor deficits but of course these factors cannot be entirely discounted. The results are consistent with previous reports of disruption to operant responding by SCH23390 (e.g., Nakajima, 1986; Sharf et al., 2005) and suggest that $D_1$ receptors may be involved in the associative learning as well as other processes underpinning instrumental responding.

In stark contrast to the effects of dopaminergic drugs on instrumental performance, antagonism of dopaminergic systems failed to impact on magazine approach behavior (but see Choi et al., 2005). Both during acquisition and at test there was no effect of the various dopaminergic agents used in the current experiments on magazine entry behavior. Furthermore in line with previous evidence, magazine approach behavior remained sensitive to outcome devaluation even when instrumental performance (see below) was impervious to changes in reward value (Nelson and Killcross, 2006). Thus the deficits in instrumental performance observed cannot *simply* be attributable to motoric dysfunction as any drug induced motor impairment would presumably impact on magazine approach behavior as well as lever pressing. To the extent that magazine approach behavior in a free operant procedure depends on Pavlovian contingencies, these findings provide yet further evidence that Pavlovian and instrumental conditioning can be subserved by distinct psychological and neural processes (e.g., Holland, 1998; Dickinson et al., 2000; Corbit et al., 2001).

As expected, animals that were not exposed to amphetamine prior to training showed normal sensitivity to outcome devaluation after limited training. The administration of the dopamine antagonists α-flupenthixol and SCH23390 during training had no impact on this sensitivity; it was neither enhanced nor attenuated by these drugs. Eticlopride treatment, however, appeared to reduce sensitivity to the changed value of the reinforcer after taste aversion as evidenced by comparable rates of responding across the two devaluation groups. As there was no statistical evidence for this effect, any inferences from Experiment 1C about the role of $D_2$ receptors in the control of goal-directed behavior in normal animals would be premature.

Nevertheless, the results from the reinforcer devaluation task in animals with prior exposure to amphetamine (Experiments 1A–C) furnish unequivocal evidence for distinct roles of $D_1$ and $D_2$ receptor subtypes in the control of behavior by goal-directed actions and S-R habits. In a replication of our previous findings, animals given pre-training exposure to amphetamine and saline during training showed accelerated habit formation as they failed to alter lever press performance in response to the changed value of the reinforcer. The performance in the reinforcer devaluation task of sensitized animals treated with either the non-selective dopamine antagonist α-flupenthixol or the $D_1$ antagonist SCH23390 during training was not autonomous of the current value of the reinforcer as these animals showed

a selective depression in lever press rates compared to non-devalued controls. Thus the instrumental performance of these animals mirrors that of normal animals after limited training and suggests response control was by goal-directed A-O associations. Given that α-flupenthixol is a non-selective dopamine antagonist that acts at both $D_1$ and $D_2$ receptors it is perhaps noteworthy that its effects in the current study were comparable to those seen with the selective $D_1$ antagonist SCH23390 and not the selective $D_2$ antagonist eticlopride. This would suggest that blockade of $D_1$ receptors by α-flupenthixol was sufficient to reverse amphetamine-induced disruption of goal-directed behavior. Consistent with this profile of action, amphetamine-induced disruption of conditional discrimination performance is attenuated by acute treatment with both selective $D_1$ antagonists and α-flupenthixol but not $D_2$ antagonists (Dunn et al., 2005; Dunn and Killcross, 2006).

The finding, however, that instrumental responding in animals given eticlopride was impervious to the current value of the reinforcer suggests these animals' instrumental performance remained stimulus-bound and governed by S-R habits. The differential sensitivity to outcome devaluation procedures cannot be attributed to impaired acquisition, as responding in all animals was depressed during acquisition irrespective of the antagonist administered. Similarly, all animals acquired an aversion to the reinforcer as evidenced by the marked sensitivity of magazine approach to outcome value in the extinction tests. The consumption test in Experiment 1C similarly confirmed that eticlopride-treated animals had acquired an aversion to the reinforcer and were able to inhibit consummatory behavior. Furthermore, the magazine entry data suggest that the eticlopride treated animals were able, under certain circumstances, to inhibit specific responses. It is not entirely possible to preclude response perseveration as an explanation of the results but the sensitivity of magazine approach to changes in reward value suggests that the insensitivity of lever pressing to outcome devaluation in these animals is unlikely to be entirely attributable to general response perseveration. Although the results of the activity test indicated elevated locomotor activity in eticlopride-treated animals in response to an amphetamine challenge compared to other animals, the sensitivity of magazine entry behavior renders any interpretation of lever press performance in terms of hyperactivity unlikely. The results are therefore specific to an effect on lever pressing and demonstrate that the accelerated habit formation following amphetamine exposure is prevented by $D_1$ but not $D_2$ receptor antagonism. Indeed, this parallels good evidence that the development of sensitization to the locomotor activating effects of amphetamine is also blocked by $D_1$ antagonists. These effects have been observed systemically (Vezina and Stewart, 1989) and with local infusions of SCH23390 into both the VTA and substantia nigra pars reticulata (Stewart and Vezina, 1989; Vezina, 1996). Similarly, $D_1$ receptor knock-out mice fail to develop behavioral sensitivity to amphetamine (Karper et al., 2002; McDougall et al., 2005) and a fMRI study supports the suggestion that $D_1$ receptors are responsible for amphetamine-mediated neurochemical changes and that $D_1$ antagonists inhibit this response to amphetamine (Dixon et al., 2005). Thus the current findings concur with reports of $D_1$ receptor modulation of the neurochemical and locomotor response to amphetamine and extend them to include a further behavioral response; enhanced habit formation.

However, eticlopride administered during training failed to reverse the accelerated formation of S-R habits induced by pre-training amphetamine exposure. This finding is consistent with evidence that $D_2$ antagonism can actually enhance the behavioral and neurochemical effects of amphetamine. For example, the blockade of $D_2$ receptors in the VTA produces persistent elevation of the locomotor activating effects of amphetamine (Tanabe et al., 2004). Indeed, in the current experiments systemic administration of eticlopride during training appeared to heighten the potentiation of locomotor activity by amphetamine in both sensitized and non-sensitized animals in the activity test following a drug challenge. Sulpiride, which has high affinity for $D_2$ receptors, has been shown to enhance the augmentative effects of amphetamine on extracellular striatal dopamine levels measured by *in vivo* microdialysis (Jaworski et al., 2001). Similarly, fMRI measurement of changes in rat brain activation following amphetamine administration shows that pre-treatment with sulpiride facilitates the response elicited by amphetamine (Dixon et al., 2005). Furthermore, the finding that the instrumental performance of animals treated with eticlopride was completely independent of goal-value during the reacquisition test also suggests that antagonism of $D_2$ receptors enhanced the effect of pre-training exposure to amphetamine on the sensitivity of a moderately trained instrumental response to outcome devaluation. The amphetamine-sensitized animals treated with eticlopride clearly had a representation of the devalued outcome as they inhibited magazine entry responses and when given the opportunity consumed less of the outcome compared to controls, but they failed to use this representation to guide instrumental responding. Instrumental performance under the control of S-R habits, whether engendered by overtraining or amphetamine exposure, is normally sensitive to outcome value in re-acquisition and thus the insensitivity of eticlopride-treated animals in the reacquistion test in Experiment 1C is novel and can be taken as evidence of dysfunctional habit learning characteristic of compulsions. By definition, compulsive behavior is carried out repetitively and persists despite adverse consequences. Significantly, there is evidence that abnormal $D_2$ receptor binding may be involved in psychopathologies characterized by compulsive behavior. For example, PET scans have revealed low $D_2$ receptor availability in drug abusers (Wang et al., 1999; Volkow et al., 1999, 2001, 2007) and single photon emission computerized tomography (SPECT) has shown reduced $D_2$ receptor binding in OCD patients (Denys et al., 2004). The current results are consistent with these reports and suggest that sensitization of dopaminergic systems coupled with antagonism of $D_2$ receptors may lead to maladaptive habitual behavior that is compulsive. As such the paradigm developed here could serve as model of the neurochemical changes that accompany the loss of voluntary control over behavior associated with drug addiction and neuropsychiatric disorders such as OCD and Tourette's Syndrome.

The finding of opposing roles of $D_1$ and $D_2$ receptors in the transition from action to habit and compulsion in the experiments presented here is consistent with previous reports

that antagonism of $D_1$ receptors disrupts, but $D_2$ receptor blockade facilitates, learning in a variety of Pavlovian conditioning paradigms (Smith et al., 1997; Horvitz, 2001; Eyny and Horvitz, 2003; Yue et al., 2004; Cassaday et al., 2005). The demonstration here of dissociable effects of $D_1$ and $D_2$ receptor antagonism on instrumental learning and the sensitivity of that learning to outcome devaluation is, however, novel. Moreover, it is consistent with evidence that activity at $D_1$and $D_2$ receptor subtypes can exert opposing effects on dendritic excitability and neuroplasticity within the striatum that in turn may facilitate or inhibit appropriate action selection (Surmeier et al., 2007; Gerfen and Surmeier, 2011). This differential involvement in striatal synaptic plasticity may therefore underlie the effects on learning seen here and more generally accelerated habit formation after sensitization (Gerdeman et al., 2003).

More broadly, the current findings have implications for our understanding of the role of dopamine and activity at different dopamine receptor subtypes in modulating behavioral flexibility. These data provide evidence of $D_1$ receptor involvement in the transition from flexible goal-directed action to inflexible stimulus-driven habits and raise the possibility that antagonism of $D_1$ receptors would reinstate goal-directed behavior in overtrained rats. Similarly, $D_1$ receptor knock-out mice may fail to develop goal-insensitive habitual responding. Conversely, antagonism of $D_2$ receptors appears to exert the opposite effect and render instrumental behavior completely insensitive to changes in outcome value. Thus, antagonism of $D_1$ but not $D_2$ receptors can produce flexible goal-directed behavior when it would otherwise be inflexible.

## ACKNOWLEDGMENTS

This work was supported by a Tourette Syndrome Association grant to Simon Killcross, and a BBSRC studentship award to Andrew J. D. Nelson.

## REFERENCES

Adams, C. D. (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q. J. Exp. Psychol.* 34B, 77–98.

Adams, C. D., and Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Q. J. Exp. Psychol.* 33B, 109–121.

Balleine, B. W., and O'Doherty, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35, 48–69.

Belin, D., and Everitt, B. J. (2008). Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral with the dorsal striatum. *Neuron* 57, 432–441.

Beninger, R. J., and Miller, R. (1998). Doapmine $D_1$-like receptors and reward related incentive learning. *Neurosci. Biobehav. Rev.* 22, 335–345.

Bentivoglio, M., and Morelli, M. (2005). "The organisation and circuits of mesenchepahlic dopaminergic neurons and the distribution of dopamine receptors in the brain," in *Handbook of Chemical Neuroanatomy, Volume 21: Dopamine*, eds S. B. Dunnett, M. Bentivoglio, A. Björklund, and T. Hökfelt (Amsterdam: Elsevier), 1–90.

Calabresi, P., Saiardi, A., Pisani, A., Baik, J. H., Centonze, D., Mercuri, N. B., et al. (1997). Abnormal synaptic plasticity in the striatum of mice lacking dopamine $D_2$ receptors. *J. Neurosci.* 17, 4536–4544.

Canales, J. J., and Graybiel, A. M. (2000). A measure of striatal function predicts motor stereotypy. *Nat. Neurosci.* 3, 377–383.

Canales, J. J., Capper-Loup, C., Hu, D., Choe, E. S., Upadhyay, U., and Graybiel, A. M. (2002). Shifts in striatal responsivity evoked by chronic stimulation of dopamine and glutamate systems, *Brain* 125, 2353–2363.

Capper-Loup, C., Canales, J. J., Kadaba, N., and Graybiel, A. M. (2002). Concurrent activation of dopamine $D_1$ and $D_2$ receptors is required to evoke neural and behavioural phenotypes of cocaine sensitization. *J. Neurosci.* 22, 6218–6227.

Cassaday, H. J., Nelson, A. J. D., and Norman, C. (2005). Haloperidol can increase responding to both discrete and contextual cues in trace conditioned rats. *Behav. Brain Res.* 158, 31–42.

Choi, W. Y., Balsam, P. D., and Horvitz, J. C. (2005). Extended habit training reduces dopamine mediation of appetitive response expression. *J. Neurosci.* 25, 6729–6733.

Corbit, L. H., Muir, J. L., and Balleine, B. W. (2001). The role of the nucleus accumbens in instrumental conditioning: evidence of a functional dissociation between accumbens core and shell. *J. Neurosci.* 21, 3251–3260.

Coutureau, E., and Killcross, S. (2003). Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behav. Brain Res.* 146, 167–174.

Denys, D., van der Wee, N., Janssen, J., De Geus, F., and Westenberg, H. G. (2004). Low level of dopaminergic $D_2$ receptor binding in obsessive-compulsive disorder. *Biol. Psychiatry* 55, 1041–1045.

Dickinson, A., and Balleine, B. (1994). Motivational control of goal-directed action. *Anim. Learn. Behav.* 22, 1–18.

Dickinson, A., Smith, J., and Mirenowicz, J. (2000). Dissociation of pavlovian and instrumental incentive learning under dopamine antagonists. *Behav. Neurosci.* 114, 468–483.

Dunn, M. J., Futter, D., Bonardi, C., and Killcross, S. (2005). Attenuation of d-amphetamine-induced disruption of conditional discrimination performance by alpha-flupenthixol. *Psychopharmacology* 117, 296–306.

Dunn, M. J., and Killcross, S. (2006). Differential attenuation of d-amphetamine-induced disruption of conditioned discrimination performance by dopamine and serotonin antagonists. *Psychopharmacology* 188, 183–192.

Dixon, A. L., Prior, M., Morris, P. M., Shah, Y. B., Joseph, M. H., and Young, A. M. (2005). Dopamine antagonist modulation of amphetamine response as detected using pharmacological MRI. *Neuropharmacology* 48, 236–245.

Eyny, Y. S., and Horvitz, J. C. (2003). Opposing roles of $D_1$ and $D_2$ receptors in appetitive conditioning. *J. Neurosci.* 23, 1584–1587.

Everitt, B. J., and Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* 8, 1481–1489.

Faure, A., Haberland, U., and Conde, F., and Massioui, N. (2005). Lesion to the nigrostriatal dopamine system disrupts stimulus-response habit formation. *J. Neurosci.* 25, 2771–2780.

Fowler, S. C., and Liou, J. R. (1994). Microcatalepsy and disruption of forelimb usage during operant behavior: differences between dopamine $D_1$ (SCH-23390) and $D_2$ (raclopride) antagonists. *Psychopharmacology* 115, 24–30.

Gerdeman, G. L., Partridge, J. G., Lupica, C. R., and Lovinger, D. M. (2003). It could be habit forming: drugs of abuse and striatal synaptic plasticity. *Trends Neurosci.* 26, 184–192.

Gerfen, C. R., and Surmeier, D. J. (2011). Modulation of striatal projection systems by dopamine. *Annu. Rev. Neurosci.* 34, 441–466.

Gillan, C. M., Papmeyer, M., Morein-Zamir, S., Sahakian, B. J., Fineberg, N. A., Robbins, T. W., et al. (2011). Disruption in the balance between goal-directed and habit learning in obsessive-compulsive disorder. *Am. J. Psychiatry* 168, 718–726.

Graybiel, A. M., and Rauch, S. L. (2000). Toward a neurobiology of obsessive-compulsive disorder. *Neuron* 28, 343–347.

Hernandez, L., and Hoebel, B. G. (1988). Food reward and cocaine increase extra-cellular dopamine in the nucleus accumbens as measured by dialysis. *Life Sci.* 42, 1705–1712.

Hogarth, L., Balleine, B. W., Corbit, L. H., and Killcross, S. (2013). Associative learning mechanisms underpinning the transition from recreational drug use to addiction. *Ann. Acad. N.Y. Sci.* 1282, 12–24.

Holland, P. C. (1998). Amount of training affects associatively-activated

event representation. *Neuropharmacology* 37, 461–469.
Horvitz, J. C. (2001). The effects of $D_1$ and $D_2$ receptor blockade on the acquisition and expression of a conditioned appetitive response. *Appetite* 37, 119–120.
Howell, D. C. (2002). *Statistical Methods for Psychology*. Pacific Grove, CA: Duxbury.
Jaworski, J. N., Gonzales, R. A., and Randall, P. K. (2001). Effect of dopamine $D_2/D_3$ receptor antagonist sulpiride on amphetamine-induced changes in striatal extracellular dopamine. *Eur. J. Pharmacol.* 418, 201–216.
Kalivas, P. W., Sorg, B. A., and Hooks, M. S. (1993). The pharmacology and neural circuitry of sensitization to psychostimulants. *Behav. Pharmacol.* 4, 315–334.
Karper, P. E., De la Rosa, H., Newman, E. R., Krall, C. M., Nazarian, A., McDougall, S. A., et al. (2002). Role of $D_1$-like receptors in amphetamine-induced behavioral sensitization: a study using $D_{1A}$ receptor knockout mice. *Psychopharmacology* 159, 407–414.
Kerr, J. N., and Wickens, J. R. (2001). Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum *in vitro*. *J. Neurophysiol.* 85, 117–124.
Killcross, S., and Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex* 13, 400–408.
Leckman, J. F., and Riddle, M. A. (2000). Tourette's Syndrome: when habit-forming systems form habits of their own? *Neuron* 28, 349–354.
Mark, G. P., Smith, S. E., Rada, P. V., and Hoebel, B. G. (1994). An appetitively conditioned taste elicits a preferential increase in mesolimbic dopamine release. *Pharmacol. Biochem. Behav.* 48, 651–660.
McDougall, S. A., Reichel, C. M., Cyr, M. C., Karper, P. E., Nazarian, A., and Crawford, C. A. (2005). Importance of D(1) receptors for associative components of amphetamine-induced behavioral sensitization and conditioned activity: a study using D(1) receptor knockout mice. *Psychopharmacology* 183, 20–30.
Naneix, F., Marchand, A. R., Di Scala, G., Pape, J. R., and Coutureau, E. (2009). A role for medial prefrontal dopaminergic innervation in instrumental conditioning. *J. Neurosci.* 29, 6599–6606.
Nakajima, S. (1986). Suppression of operant responding in the rat by dopamine $D_1$ receptor blockade with SCH23390. *Physiol. Psychol.* 14, 111–114.
Nelson, A., and Killcross, S. (2006). Amphetamine exposure enhances habit formation. *J. Neurosci.* 26, 3805–3812.
Nestler, E. J. (2001). Molecular basis of long-term plasticity underlying drug addiction. *Nat. Rev. Neurosci.* 2, 119–128.
Nordquist, R. E., Voorn, P., de Mooij-van Malsen, J. G., Joosten, R. N., Pennartz, C. M., and Vanderschuren, L. J. (2007). Augmented reinforce value and accelerated habit formation after repeated amphetamine treatement. *Eur. Neuropsychophamacol.* 17, 532–540.
Ridley, R. M. (1994). The psychology of preservative and stereotyped behaviour. *Prog. Neurobiol.* 44, 221–231.
Saka, E., Goodrich, C., Harlan, P., Madras, B., and Graybiel, A. M. (2004). Repetitive behaviours in monkeys are linked to specific striatal activation patterns. *J. Neurosci.* 24, 7557–7565.
Sharf, R., Lee, D. Y., and Ranaldi, R. (2005). Microinjections of SCH 23390 in the ventral tegmental area reduce operant responding under a progressive ratio schedule of food reinforcement in rats. *Brain Res.* 1033, 179–185.
Smith, J. K., Neill, J. C., and Costall, B. (1997). Bidirectional effects of dopamine $D_2$ receptor antagonists on responding for a conditioned reinforcer. *Pharmacol. Biochem. Behav.* 57, 843–849.
Stewart, J., and. Vezina, P. (1989). Microinjections of Sch-23390 into the ventral tegmental area and substantia nigra pars reticulata attenuate the development of sensitization to the locomotor activating effects of systemic amphetamine. *Brain Res.* 495, 401–406.
Surmeier, D. J., Ding, J., Day, M., Wang, Z., and Shen, W. (2007). $D_1$ and $D_2$ dopamine- receptor modulation of striatal glutamerigic signalling in striatal medium spiny neurons. *Trends Neurosci.* 30, 228–235.
Tanabe, L. M., Suto, N., Creekmore, E., Steinmiller, C. L., and Vezina, P. (2004). Blockade of $D_2$ dopamine receptors in the VTA induces a long-lasting enhancement of the locomotor activating effects of amphetamine. *Behav. Pharmacol.* 15, 387–395.
Tombaugh, T. N., Tombaugh, J., and Anisman, H. (1979). Effects of dopamine receptor blockade on alimentary behaviors: home cage food consumption, magazine training, operant acquisition, and performance. *Psychopharmacology* 66, 219–225.
Vezina, P. (1996). $D_1$ dopamine receptor activation is necessary for the induction of sensitization by amphetamine in the ventral tegmental area. *J. Neurosci.* 16, 2411–2420.
Vezina, P., and Stewart, J. (1989). The effect of dopamine receptor blockade on the development of sensitization to the locomotor activating effects of amphetamine and morphine. *Brain Res.* 499, 108–120.
Volkow, N. D., Chang, L., Wang, G. J., Fowler, J. S., Ding, Y. S., Sedler, M., et al. (2001). Low level of brain dopamine $D_2$ receptors in methamphetamine abusers: association with metabolism in the orbitofrontal cortex. *Am. J. Psychiatry* 158, 2015–2021.
Volkow, N. D., Fowler, J. S., Wang, G. J., Swanson, J. M., and Telang, F. (2007). Dopamine in drug abuse and addiction: results of imaging studies and treatment implications. *Arch. Neurol.* 64, 1575–1579.
Volkow, N. D., Wang, G. J., Fowler, J. S., Logan, J., Gatley, S. J., Gifford, A., et al. (1999). Prediction of reinforcing responses to psychostimulants in humans by brain dopamine $D_2$ receptor levels. *Am. J. Psychiatry* 156, 1440–1443.
Wang, G. J., Volkow, N. D., Fowler, J. S., Cervany, P., Hitzemann, R. J., Pappas, N. R., et al. (1999). Regional brain metabolic activation during craving elicited by recall of previous drug experiences. *Life Sci.* 64, 775–784.
Wyvell, C. L., and Berridge, K. C. (2000). Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: enhancement of reward 'wanting' without enhanced 'liking' or response reinforcement. *J. Neurosci.* 20, 8122–8130.
Wise, R. A., and Schwartz, H. V. (1981). Pimozide attenuates acquisition of lever-pressing for food in rats. *Pharmacol. Biochem. Behav.* 15, 655–656.
Yamamoto, Y., Nakanishi, H., Takai, N., Shimazoe, T., Watanabe, S., and Kita, H. (1999). Expression of N-methyl-D-aspartate receptor-dependent long-term potentiation in the neostriatal neurons in an in vitro slice after ethanol withdrawal of the rat. *Neuroscience* 91, 59–68.
Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* 19, 181–189.
Yin, H. H., Ostlund, S. B., Knowlton, B. J., and Balleine, B. W. (2005). The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* 22, 513–523.
Yue, E. L., Cleland, T. A., Pavlis, M., and Linster, C. (2004). Opposing effects of D-1 and D-2 receptor activation on odor discrimination learning. *Behav. Neurosci.* 118, 184–190.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 31 January 2013; paper pending published: 04 March 2013; accepted: 26 April 2013; published online: 15 May 2013.*

*Citation: Nelson AJD and Killcross S (2013) Accelerated habit formation following amphetamine exposure is reversed by $D_1$, but enhanced by $D_2$, receptor antagonists. Front. Neurosci. 7:76. doi: 10.3389/fnins.2013.00076*

*This article was submitted to Frontiers in Decision Neuroscience, a specialty of Frontiers in Neuroscience.*

frontiers in
NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 18 July 2013
doi: 10.3389/fnins.2013.00126

# Bidirectional modulation of infralimbic dopamine D1 and D2 receptor activity regulates flexible reward seeking

***Jacqueline M. Barker*[1,2]*, Mary M. Torregrossa*[1] *and Jane R. Taylor*[1]****

[1] *Department of Psychiatry, Division of Molecular Psychiatry, Center for Genes and Behavior, Yale University School of Medicine, Abraham Ribicoff Research Facilities, Connecticut Mental Health Center, New Haven, CT, USA*
[2] *Interdepartmental Neuroscience Program, Yale University, New Haven, CT, USA*

***Edited by:***
*Roshan Cools, University of Cambridge, UK*

***Reviewed by:***
*Bruno B. Averbeck, National Insitute of Mental Health, USA*
*Erie D. Boorman, University of Oxford, UK*

****Correspondence:***
*Jane R. Taylor, Department of Psychiatry, Division of Molecular Psychiatry, Center for Genes and Behavior, Yale University School of Medicine, Connecticut Mental Health Center, 34 Park Street, Ribicoff Labs, New Haven, CT 06519, USA*
*e-mail: jane.taylor@yale.edu*

The development of addictive behavior is marked by a loss of behavioral flexibility. In part, this is due to an increase in the ability of environmental stimuli to elicit responding and decreased importance of the action-outcome relationship in behavioral control. It has previously been demonstrated that both inactivation of and dopamine (DA) infusions in the infralimbic prefrontal cortex (PFC) can restore behavioral flexibility in paradigms measuring habitual reward seeking. Here, we investigated the mechanism by which cortical DA would act to enable goal-directed actions after the transition to habitual behavior has been established. Further, we extended this work to include a novel mouse model of compulsive-like behavior in which we assessed reward seeking despite the possibility of adverse consequences. Our data show that DA receptor D1 inhibition or D2 activation both promote the expression of a flexible responding after the development of habitual or compulsive-like behavior, and we suggest that the ability of DA infusions in the infralimbic PFC to restore sensitivity to changes in outcome value depends on activation of DA D2 receptors.

**Keywords: habit, dopamine, prefrontal cortex, mouse, behavioral flexibility**

## INTRODUCTION

The transition from casual drug use to addiction is characterized by increasing loss of control over reward seeking. When a behavior is first learned, performance of the action is guided by its relationship to its outcome—*i.e.*, a response is made in order to gain access to a reinforcer. Over time and after repeated execution, behavior transitions from goal-directed action to stimulus-driven habitual behavior (Dickinson, 1985). Habitual reward seeking is no longer mediated by action-outcome relationships or by a representation of the value of an outcome; rather, habitual behavior is automatically elicited by environmental cues and stimuli (c.f., Yin et al., 2008). In addition to habits, addictive behavior also involves the development of compulsive reward seeking that occurs despite adverse consequences (e.g., Everitt et al., 2008; Heyne et al., 2009). Successful treatment of addiction may require restoration of the ability to update behavior in accordance with changed contingencies and in the face of negative outcomes.

The shift in response strategy away from flexible, contingency-mediated behavior to one in which stimulus-response relationships guide behavior is paralleled by a change in the neuroanatomical substrates that mediate behavior from a prefrontal-striatal circuit in which the prefrontal cortex (PFC) monitors the action-outcome relationship, to a more dorsal circuit involving dorsolateral striatum (e.g., Yin and Knowlton, 2006; Balleine and Dickinson, 1998). However, a role for the infralimbic PFC (IL) in the expression of habitual behavior has been demonstrated. When the IL, which projects to the nucleus accumbens shell (e.g., McGeorge and Faull, 1989) and amygdala (Sesack et al., 1989), is lesioned prior to response acquisition, animals are unable to express stimulus-response habits (Killcross and Coutureau, 2003). After extended training, IL lesioned animals remain sensitive to changes in outcome value. Importantly, later research expanded on this finding to show that inactivation of the IL after extended training, at a time point where intact animals are habitual, resulted in the restoration of flexible behavior (Coutureau and Killcross, 2003). More recent work has expanded upon these findings using optogenetic manipulations to investigate online regulation of the IL in the expression of habitual behavior (Smith et al., 2012). Together, these data suggest that the IL is critically involved in the selection of response strategy in situations of conflict between automatic, habitual behaviors and flexible goal-directed actions.

Dopamine (DA) signaling within corticostriatal circuitry has been shown to play a unique role in both the formation and expression of goal-directed vs. habitual instrumental behavior (e.g., Nelson and Killcross, 2006). Our lab has shown that infusions of exogenous DA in the IL, but not the more dorsal prelimbic PFC (PL), restored sensitivity to outcome devaluation after extended training (Hitchcott et al., 2007). While a majority of these studies were performed in rats, we have found using lesion studies that the neuroanatomical mechanisms underlying habit learning are preserved in mice (Quinn et al., 2013). The mechanism by which both inactivation of and DA infusion into the IL can restore sensitivity to the action-outcome relationship is unknown in rodents. Here, we assessed the ability of DA D1 and D2-family specific manipulations in the IL to restore flexible behavior as measured by either sensitivity to changes in action-outcome contingency or reduction of compulsive-like reward-seeking behavior in mice.

## MATERIALS AND METHODS

### SUBJECTS

Male C57b/6 mice were supplied from Charles River and delivered to the Yale University/Connecticut Mental Health Center mouse vivarium between 56 and 70 days of age. These mice were allowed to acclimate for 2 weeks with *ad libitum* access to food and water. All behavioral procedures were approved by the Yale University IACUC and experiments were performed in accordance with the National Institute of Health Guide for Care and Use of Laboratory Animals. After acclimation, mice were food restricted to 90–92% of free feeding weight for all experiments. They had limited access to standard chow in their homecage each day, several hours after training. The amount of food provided was adjusted to maintain weights. Homecage chow was distinct from the purified grain pellets used in both the habitual and compulsive-like food-seeking experiments. There were approximately 5–12 animals in each experimental group after exclusion of mice with inaccurate cannula placement or loss/clogging of cannula during the course of the experiments. Saline groups had large *n*'s (>12) as a cohort of control (saline) animals was included in each testing session to ensure baseline effects were consistent.

### INSTRUMENTAL CONDITIONING CHAMBERS

Instrumental chambers were identical to those described by (Barker et al., 2012). Briefly, 12 mouse instrumental chambers housed within a sound-attenuating box, were used for these experiments (Med-Associates; Georgia, VT). Each chamber was equipped with a 28 V house light located at the top of the middle panel on the left side wall, three adjacent nosepoke apertures located at the bottom of the left side wall, and a magazine located at the bottom of the middle panel on the right side wall. Grain pellets were delivered to a magazine on the opposite wall. Nosepoke apertures and reinforcement magazine were equipped with a light and photobeam sensor. A fan provided background noise and ventilation.

### STEREOTAXIC SURGERY

Mice were anesthetized using ketamine/xylazine. Bilateral cannula (Plastics One; Roanoke, VA) were implanted and mounted to the skull using standard stereotaxic techniques. Cannula were targeted to the IL at AP + 1.7, ML ± 0.25, DV-3.0 from bregma based on coordinates from Wall et al. (2004). For compulsive-like food-seeking experiments, surgeries were performed prior to any training. For instrumental habit experiments, surgeries were performed after 3 days of fixed ratio (FR) 1 training to reduce the amount of time between cannula placement and testing.

### DRUGS AND INFUSIONS

For tests of habitual and compulsive-like food-seeking, mice received two infusions of the same drug prior to a control and experimental session. Infusions were 0.2 uL over 2 min; internal cannula were left in place for an additional 2 min to allow for diffusion. This volume and diffusion duration were chosen based on the literature and our pilot data using thianin which suggested minimal spread to surrounding tissues at this volume and after the delayed removal of cannula. Drugs used were the D1 agonist dihydrexidine HCl (DHX; Tocris; Minneapolis, MN), D1 antagonist SCH23390 (Sigma; St. Louis, MO), D2 agonist quinpirole (Tocris) in saline, and the D2 antagonist sulpiride (Tocris) in acidified saline, each dissolved at 5 nmol per 1 ul.

### INSTRUMENTAL TRAINING

During training, one nosepoke was assigned as the active nosepoke, where a response resulted in reinforcer delivery, and the others designated as inactive nosepokes. Training consisted of 1 day magazine training, 3-days fixed ratio (FR 1) training (in which each active response resulted in reinforcer delivery) and 3-days random interval (RI) 30-s training and 6 RI60 sessions. In RI sessions, reinforcement could be earned every 30 (RI30) or 60 (RI60) s on average. The actual duration of each interval was randomly determined so that reinforcement availability was not predictable. The first active response (nosepoke) after the interval ended resulted in reinforcer delivery; the duration of the next interval was then generated automatically. During each daily training session, the house light and fan were on. All sessions were 30 min in duration.

### CONTINGENCY DEGRADATION TEST

During degradation sessions, conditions were identical to training except that the grain pellet reinforcer was delivered on a non-contingent schedule determined by each individual animal's reinforcement rate on the day prior. Reinforcer delivery was spaced equivalently across the 30-min session. Responses on the active and inactive nosepokes were recorded, but did not result in reinforcer delivery. Infusions of drugs occurred 5 min prior to the start of the degradation session. Mice were assigned to infusion groups by matching baseline response rates, and received a 0.2 ul infusion of either saline ($n = 17$), DHX ($n = 11$), SCH23390 ($n = 6$), quinpirole ($n = 12$), or sulpiride ($n = 9$). More animals were in the saline groups as a cohort of saline animals was included with each behavioral test session to confirm baseline effects were replicated. Data were compared to a non-degraded session in which the animals received the same drug; the order of these sessions was counterbalanced and animals received one normal RI 60 training session between both test sessions where no drug was administered.

### COMPULSIVE-LIKE FOOD-SEEKING TRAINING AND TEST

Additionally, we assessed the effects of IL DA receptor modulation on compulsive-like behavior in mice using a modification of traditional conditioned place preference/aversion testing. Conditioning chambers were standard three chamber boxes with retractable doors (Med Associates; Georgia, VT). Chambers had distinct walls (vertical black and white stripes or diagonal marble and black stripes) and floors (wire mesh or grid). The two conditioning chambers were separated by a neutral, gray chamber. Photocell beam breaks were used to calculate time spent in each chamber, latency to enter the chamber and number of entries by Med-PC IV software. During a single habituation session, mice were placed in the neutral chamber with both doors retracted such that mice could freely explore all chambers. During conditioning, mice were confined to the "paired" chamber for 30 min with access to 30-grain pellets on days 1, 3, and 5. On days 2, 4,

and 6, mice were confined to the opposite chamber for 30 min with an empty food dish.

On day 7, mice received an infusion of either a DA D1 or D2-like receptor agonist or antagonist 5 min prior to being placed in the neutral chamber with both doors retracted and were allowed to freely explore all chambers for 5 min. This duration was chosen because we were able to examine entry into both chambers and latency to enter, but no extinction was expected to occur based on our preliminary data. Mice received a 0.2ul infusion of either saline ($n = 20$), DHX ($n = 8$), SCH23390 ($n = 7$), quinpirole ($n = 8$), or sulpiride ($n = 7$). Latency to enter the chambers was the primary outcome measure.

On the following day, mice were confined to the food-paired chamber. Two minutes after placement, mice received a 2 s, 0.8 mA foot shock. Mice remained in the chamber for 60 s after the shock was terminated and were then returned to their homecage. On day 9, mice received a second infusion of the same drug as day 7. Five min after the infusion, they were returned to the gray chamber and allowed to freely enter both chambers and latency to enter the chambers was assessed in this 20 min session. Latency was selected as the primary measure of compulsive-like behavior because it was not expected to be impacted by the extinction of either the association of the chamber with footshock or the association with the food reward which may be differentially impacted by prefrontal DA manipulations. Importantly, a change in the parameters of the training conditions might have an impact on the expression of reward-seeking under conflict between reward seeking and avoidance of negative consequences, either by increasing the aversive component (e.g., through increasing the shock intensity), the value of the reward, or the extent of learning (e.g., through extended training).

### CONFIRMATION OF PLACEMENT

After behavioral assessment was complete, mice were sacrificed and tissue was fixed in paraformaldehyde for confirmation of cannula placement and location of the infusion tip using standard histological techniques. If cannula were not clogged at the time of sacrifice, thianin was infused at the volume and rate used for testing (0.2 ul over 2 min). If cannula had become clogged, cannula tracts, and tips were confirmed. Mice were excluded if placement could not be confirmed to be in the IL through the use of neuroanatomical landmarks, including white matter tracts.

### STATISTICS

Data were analyzed with JMP Software (SAS Institute) using repeated measures analysis of variance (ANOVA). Significant interactions were further analyzed using Tukey's HSD *post-hoc* tests.

## RESULTS

### CONTINGENCY DEGRADATION

Data were square root transformed to maintain homogeneity of variance. To determine whether agonism and/or antagonism of DA D1 or D2 receptors influenced sensitivity to changes in action-outcome relation, active responding during a degraded session was compared to responding during a non-degraded session; during both test sessions the experimental drug was on board. Importantly, no differences in baseline response rates were seen in animals to-be assigned to groups [$F_{(4,\ 50)} = 1.122$, $p = 0.356$]. Additionally, rmANOVA [drug × non-degraded session ("no drug" vs. "drug")] revealed no differences were observed in response rates between the "drug" and "no drug" non-degraded session ($p > 0.5$ for main effects, $p = 0.185$ for session × drug interaction). Repeated measures ANOVA revealed a significant session (degraded vs. non-degraded) × drug interaction on active responding [$F_{(4,\ 46)} = 2.92$, $p < 0.05$]. *Post-hoc* analyses indicated that responding of the saline-injected animals did not differ significantly between the degraded and non-degraded session, indicating that under basal conditions animals were insensitive to the change in action-outcome relations, consistent with the formation of habit. Critically, responding during the degraded session differed significantly from the non-degraded session only for mice receiving the DA D1 receptor antagonist SCH23390 ($p < 0.05$) or the DA D2 receptor agonist quinpirole ($p < 0.05$; **Figure 1**). Together these data demonstrate that only antagonism of the D1 receptor or agonism of the D2 receptor in the IL are sufficient to restore sensitivity to changes in the action-outcome relationship, indicative of goal-directed instrumental behavior. Mice receiving the DA D1 receptor agonist DHX or DA D2 receptor antagonist sulpiride did not show differential responding between the degraded and non-degraded sessions, confirming that these opposing DA receptor manipulations do not impact sensitivity to changes in contingency after extended training.

Because animals received infusions of the same drug during both test sessions and we used a within subjects analysis to assess responding, we are confident that the marked differences seen between the degraded and non-degraded sessions with either the SCH23390 or the quinpirole infusions reflected a change in response strategy. We do not believe this reduction in responding in the degraded session, which is evidence for goal-directed instrumental action is related to non-specific alterations in task engagement, motivation, or locomotor effects as this would have been reflected as behavioral changes in both the degraded and non-degraded test conditions.

### COMPULSIVE-LIKE FOOD SEEKING

To assess the effect of DA receptor manipulations on a novel measure of compulsive-like reward-seeking behavior, we compared the latency to enter the food reward-paired chamber after training, but prior to shock (pre-shock) with the latency after the animals had received a foot shock in the reward-paired chamber (post-shock). A repeated measures ANOVA revealed a significant session (pre-shock vs. post-shock) × drug interaction on latency to enter the reward paired chamber [$F_{(4,\ 46)} = 2.8205$, $p < 0.05$]. *Post-hoc* analyses indicated that only animals that received SCH23390 or quinpirole infusions had post-shock latencies that were significantly increased compared to saline-infused animals ($p < 0.05$ and $p < 0.01$, respectively; **Figure 2**). Neither of these drugs impacted pre-shock latencies, indicating that DA receptor D1 antagonism or D2 agonism increased the latency to enter the reward-paired chamber *only* after that chamber had been paired with a negative consequence. Additionally,

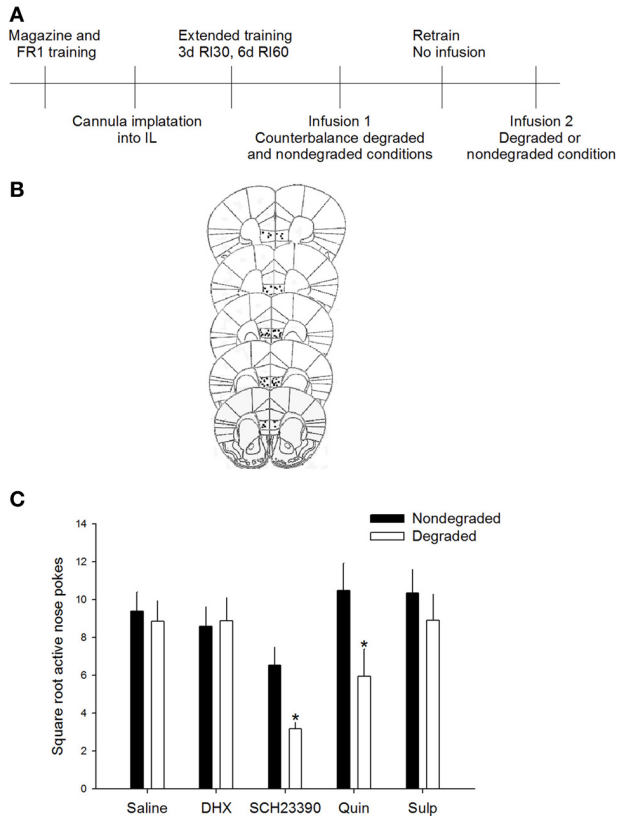



**FIGURE 1 | D1 antagonism or D2 agonism restore goal-directed behavior. (A)** Experimental timeline. Mice only received infusions of drugs during counterbalanced test sessions. One half of mice received the degraded session first while other mice received a non-degraded session first. **(B)** Only mice whose cannulae were placed within the IL were included in analyses. Images modified from Paxinos and Franklin (2001). **(C)** Inhibition of D1 signaling with SCH23390 or agonism of the D2 receptor with quinpirole in the IL resulted in reduced responding *only* during the degraded session, consistent with restoration of goal-directed behavior. Error bars ±SEM. $^{*}p < 0.05$.

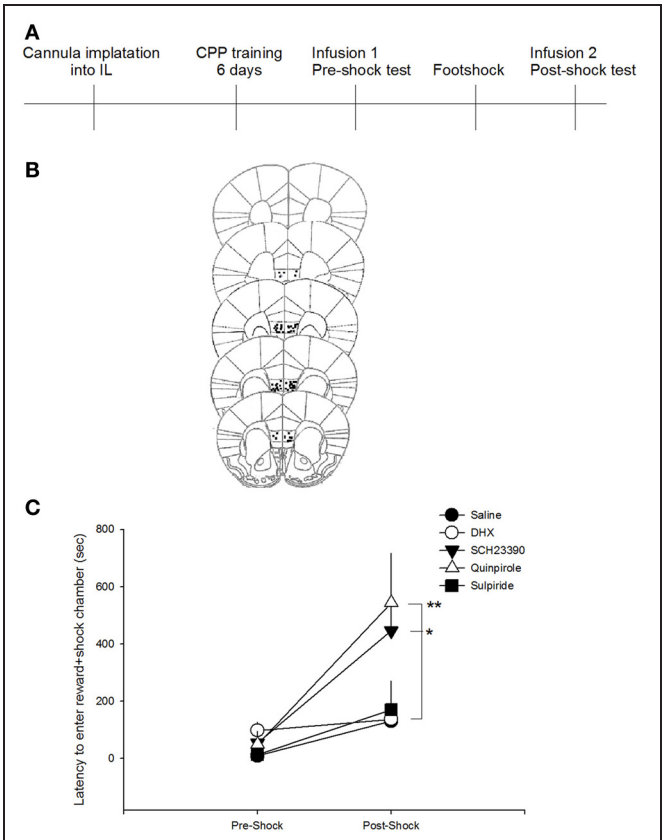



**FIGURE 2 | D1 antagonism or D2 agonism reduce compulsive-like reward seeking. (A)** Experimental timeline. Mice received infusions of drugs into the IL only during test sessions to ensure that effects were on the expression, not acquisition, of compulsive-like behavior. **(B)** Only mice whose cannula could be confirmed to be within the IL were included in behavioral analyses. Images modified from Paxinos and Franklin (2001). **(C)** Antagonism of the D1 receptor with SCH23390 or agonism of the D2 receptor with quinpirole in the IL reduced compulsive-like reward seeking as indicated by an increase in latency to enter the reward-paired chamber *only* after pairing with foot shock (adverse consequence). $^{*}p < 0.05$, $^{**}p < 0.01$.

administration of quinpirole or SCH23390 did not impact the time spent in the reward-paired chamber in either the pre- or post-shock test $[F_{(2,\ 21)} = 0.2022,\ p = 0.8]$, though there was a main effect of session $[F_{(1,\ 21)} = 15.8571,\ p < 0.001]$. These data suggest that inhibition of DA D1 or activation of DA D2 receptors do not impact latency to enter the reward paired chamber in situations where there is no conflict, but decrease compulsive-like reward seeking after the risk of aversive outcome has been learned. Post-shock latencies to enter the reward-paired chamber after infusions of DHX or sulpiride, however, did not differ from saline treated mice ($p > 0.7$), indicating that DA D1 agonism or D2 antagonism did not impact compulsive-like reward seeking.

During the pre-shock interval, only mice receiving DHX infusions showed an increase latency to enter the reward paired chamber as compared to saline treated mice ($p < 0.05$), suggesting that DA D1 agonism impacts latency to enter a reward paired chamber under baseline conditions. Again, because mice receive infusions prior to both the pre-shock and post-shock test sessions, we do not think that the ability of SCH23390 or quinpirole to produce increased latencies to enter the post-shock chamber is reflective of altered activity levels or motivation to enter the chamber. To confirm that these manipulations did not generally increase latencies to enter both the paired and unpaired chambers in the post-shock session, a rmANOVA was performed (shock × drug). The analysis revealed a main effect of drug on latency $[F_{(4,\ 42)} = 3.26,\ p = 0.02]$ and a main effect of session $[F_{(1,\ 42)} = 5.54, p = 0.02]$, but not shock × drug interaction ($p = 0.21$), suggesting that neither the SCH23390 nor the quinpirole interacted with shock exposure to produce a latency to enter both chambers. Further, these data suggest that exposure to these drugs during the pre-shock session did not result in a generalized aversion to both chambers in the post-shock test session. Follow up analyses indicated that SCH23390 administration resulted in an increased latency to enter the unpaired chamber in both the pre- and post-shock sessions.

## DISCUSSION

These experiments investigated the role of specific manipulations of IL DA D1 and D2 receptor signaling in flexible reward seeking. We found that after extended training in an instrumental task, at a time point when control animals were insensitive to changes in contingency, inactivation of DA D1 or activation of DA D2 receptors in the IL was sufficient to render mice sensitive to the change in the relationship between action and outcome. That is, either a decrease in DA D1 activity or an increase in DA D2 signaling resulted in restoration of goal-directed behavior *after* the transition to habit. Conversely, we saw that neither DA D1 agonism nor DA D2 antagonism had any impact on behavior after extended training, indicating that it is not a general change in the ratio of D1 to D2 signaling that produced this increased sensitivity to action-outcome relationship, but rather specific decreases in DA D1 activity or increases in DA D2 signaling allowed alterations in behavior. Importantly, these studies only investigate one form of loss of action-outcome relationship, and future research will be necessary to determine whether selective infralimbic DA manipulations alter flexible responding in paradigms that disrupt contingency through provision of alternative reinforcers, reversal of the action-outcome contingency through selective reinforcement of non-responding, or under conditions of extinction.

In addition to restoration of goal-directed behavior after extended performance of an instrumental response, we similarly showed that D1 antagonism and D2 agonism in the IL reduced compulsive-like reward seeking in a task investigating competition between adverse consequences and reinforcement. Importantly, we again saw no effects of infralimbic D1 agonism or D2 antagonism on the ability to restore behavioral flexibility. The increase in latency to enter the reward-paired chamber in mice receiving IL infusions of the DA D1 antagonist or DA D2 agonist occurred *only* after animals received a foot shock in the same chamber, indicating that these DA manipulations during the test did not impair either the ability to move toward the chamber or motivation to enter the reward paired chamber in the absence of conflict, i.e., prior to foot shock. Notably, IL DA D2 signaling has been shown to be critical for the extinction of conditioned fear (Mueller et al., 2010). However, we do not think this finding in anyway contradicts our conclusion that IL DA D2 activity reduces compulsive reward seeking as infusion of the DA D2 agonist increases latency to enter the shock and reward-paired chamber, indicating that extinction has not occurred. Together, these data suggest that increased DA signaling through D2-like receptors in the IL restores flexible behavior, while DA D1 activity in the IL may be related to reduced sensitivity to action-outcome relationships, including a loss of such relationships through contingency degradation, and the risk of adverse consequences, as loss of signaling at this receptor restores flexible behavior.

Our lab and several others, have long been interested in the role of corticostriatal dysfunction in inflexible, habitual, addiction-related processes (e.g., Jentsch and Taylor, 1999; Robbins and Everitt, 1999). We have previously demonstrated that administration of exogenous DA into the IL restored goal-directed behavior in animals performing habitually (Hitchcott et al., 2007); our current data suggest that this effect was mediated by activity at DA D2 receptors. Importantly, our current work focuses on the ability of DA manipulations to restore sensitivity in changes to action-outcome contingency, without investigating the role of change in outcome value. While in many cases, response strategy selection in these paradigms is consistent, it is possible that the ability to track action-outcome relationships is dependent on IL DA signaling in a way that is separate from the ability to regulate responding for a devalued outcome, and this has yet to be determined. DA has been shown to differentially affect PFC function depending on the task used and the dose tested. For example, DA is thought to impact measures of prefrontal function, such as working memory, in a dose-dependent manner through D1-mediated alterations in the signal-to-noise ratio (e.g., Arnsten, 2007). Our data indicate that in assessments of habit, exogenous DA is primarily acting through DA D2 receptors to decrease infralimbic activity, which is consistent with the ability of both D2 agonists and DA to restore flexible reward seeking. In addition, this finding reconciles the data from studies indicating that both DA infusions (Hitchcott et al., 2007) and inactivation of the IL restore goal-directed behavior (Coutureau and Killcross, 2003). The activation of DA D1 or D2 receptors has distinct and opposing downstream effects. DA D1 receptors are $G\alpha_s$ coupled, and their stimulation results in increased production of cyclic adenosine monophosphate (cAMP) and the cAMP-dependent protein kinase (PKA). Activation of $G\alpha_{i/o}$ coupled DA D2-like receptors, however, inhibits adenylyl cyclase activity, directly opposing DA D1 activity and downstream signaling. In addition to inhibition of pyramidal cells through the above described mechanism, DA D2 activation may further inhibit projection neurons through enhancement of GABAergic interneuron activity (Tseng and O'Donnell, 2007a). Enhanced signaling at infralimbic DA D2-like receptors relative to D1 receptors is likely to result in decreased neuronal activity. Based on the evidence that inactivation or lesion of the IL also impair the expression of stimulus-response habits (Coutureau and Killcross, 2003; Killcross and Coutureau, 2003), we propose that the ability of DA infusions in the IL to reinstate sensitivity to the action-outcome relationship is due to decreased activity and that the balance of D1/D2 activity in the IL is critical to the expression of flexible reward-seeking behavior.

Though a precise role for infralimbic DA D1 and D2 signaling in habitual and compulsive-like reward seeking has not been previously investigated, IL has been implicated in situations of response conflict (Haddon and Killcross, 2011). Further, a role for prefrontal DA signaling has also been investigated in other measures of flexible behavior. Blockade of DA D1 or D2 in the medial PFC has been shown to impair the ability to update behavior to a change in reward value, while not impacting the ability to perceive the change (Winter et al., 2009). Additionally, DA D2 antagonism impaired flexibility in a set-shifting task, though agonism of DA D2 did not promote shifting (Floresco et al., 2006). Inhibition of the DA D4 receptor, a member of the D2-family of receptors, had opposing effects on set shifting. Consistent with these findings, it is possible that the effects of DA D1 inhibition and D2 activation in our experiments result not from a change in infralimbic activity, but rather through changes in PFC network stability. It has been suggested that DA D1 activity can stabilize the existing PFC networks, potentially explaining why loss of DA D1 signaling

can promote flexible behavior through loss of this stabilization (e.g., Seamans and Yang, 2004; Durstewitz and Seamans, 2008). In this model, and consistent with our findings, DA D2 signaling would promote system lability through reduction in signaling in the GABAergic neurons, thus enabling the establishment of new behavioral patterns. The basis for this model, however, is work done in adolescent animals (e.g., Seamans et al., 2001) in which the DA D2 impact on GABAergic signaling may be different (i.e., opposite) from that seen in adult animals (Tseng and O'Donnell, 2007b; O'Donnell, 2010); however, the discrepancy between these findings does not appear to be solely dependent on age (Kroener and Lavin, 2010). It therefore remains unclear whether in adult animals, D2 activation in the IL may act to reduce GABAergic inhibition of pyramidal cells, or perhaps, as described above, to produce a net decrease in IL activity.

As our data suggest that a selective shift in the DA D1:D2 ratio in the IL can enable a shift in response strategy selection, it is important to consider that the observed separation between D1 and D2 effects may result from downstream influences on distinct neuroanatomical targets. It has been well established that in the striatum, DA D1- and D2-receptor containing medium spiny neurons are located in distinct populations of neurons that have separate projection targets. Indeed, striatal D1- and D2-receptor containing neurons that participate in the direct and indirect pathways, respectively, have been shown to differentially contribute to the attribution of value to an action and, therefore, inform response selection in a distinct but complementary fashion (Tai et al., 2012). While there is evidence that PFC neurons may co-express DA D1- and D2-type receptors (Vincent et al., 1995), it has also been demonstrated that D1 and D2 containing neurons are at least in part distinct populations (e.g., Gaspar et al., 1995; Gee et al., 2012). It may be that DA D1- and D2-expressing projection neurons in the IL also have separate targets and that modulation of DA D1 and D2 signaling differentially impacts downstream brain regions, therefore enabling a shift in contribution to response strategy selection between these targets. For example, it has been shown that disconnection of the IL from the nucleus accumbens shell can replicate the effects of IL inactivation on cocaine seeking (Peters et al., 2008). In addition to the nucleus accumbens shell, the IL also projects extensively to amygdalar nuclei (e.g., Vertes, 2004). Though the central nucleus of the amygdala has been shown to interact with the dorsolateral striatum to mediate the expression of goal-directed and habitual behavior (Lingawi and Balleine, 2012), the effect of IL disconnection from its targets on habitual and compulsive reward-seeking behavioral control is still under investigation.

The precise role IL plays in response strategy selection and the mechanism by which decreased activity in the IL would restore goal-directed behavior, remain to be elucidated. Studies by Rich and Shapiro (2009) suggest that infralimbic activity lags behind response switching, while PL activity leads the change, suggesting perhaps that IL is involved in the maintenance of habits while activity in the PL is required to flexibly update responding. Loss of the IL may result in a reversion to the competing memory system that uses knowledge of the action-outcome relationship and outcome value to guide behavior.

### SUMMARY AND IMPLICATIONS

The ability to behave flexibly is critical to the successful control of reward seeking, and a better understanding of the mechanisms by which response strategies shift away from those that are habitual or compulsive to those that are goal-directed, is likely to inform treatment of both drug and food addiction. Here, we show that increased D2 receptor or decreased D1 receptor activity in the IL can restore sensitivity to changes in action-outcome contingency and decrease reward seeking in the face of punishment. Importantly, these data help to explain the apparent discrepancy between the ability of infusions of DA and inactivation of the IL to enable a shift in response strategy, and will help to inform future work investigating the precise role that IL plays in facilitating plastic behavior.

## ACKNOWLEDGMENTS

This research was supported by National Institutes of Health Grants AA012870 (Jane R. Taylor), DA011717 (Jane R. Taylor), AA020135 (Jacqueline M. Barker), DA031745 (Mary M. Torregrossa) and the Connecticut Department of Mental Health and Addiction Services. The authors thank Tiffany L. Wang for her excellent technical assistance.

## REFERENCES

Arnsten, A. F. (2007). Catecholamine and second messenger influences on prefrontal cortical networks of "representational knowledge": a rational bridge between genetics and the symptoms of mental illness. *Cereb. Cortex* 17(Suppl. 1), i6–i15. doi: 10.1093/cercor/bhm033

Balleine, B. W., and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419. doi: 10.1016/S0028-3908(98)00033-1

Barker, J. M., Torregrossa, M. M., and Taylor, J. R. (2012). Low prefrontal PSA-NCAM confers risk for alcoholism-related behavior. *Nat. Neurosci.* 15, 1356–1358. doi: 10.1038/nn.3194

Coutureau, E., and Killcross, S. (2003). Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behav. Brain Res.* 146, 167–174. doi: 10.1016/j.bbr.2003.09.025

Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philos. Trans. R. Soc. B Biol. Sci.* 308, 67–78. doi: 10.1098/rstb.1985.0010

Durstewitz, D., and Seamans, J. K. (2008). The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. *Biol. Psychiatry* 64, 739–749. doi: 10.1016/j.biopsych.2008.05.015

Everitt, B. J., Belin, D., Economidou, D., Pelloux, Y., Dalley, J. W., and Robbins, T. W. (2008). Review. Neural mechanisms underlying the vulnerability to develop compulsive drug-seeking habits and addiction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 3125–3135. doi: 10.1098/rstb.2008.0089

Floresco, S. B., Magyar, O., Ghods-Sharifi, S., Vexelman, C., and Tse, M. T. (2006). Multiple dopamine receptor subtypes in the medial prefrontal cortex of the rat regulate set-shifting. *Neuropsychopharmacology* 31, 297–309. doi: 10.1038/sj.npp.1300825

Gaspar, P., Bloch, B., and Le Moine, C. (1995). D1 and D2 receptor gene expression in the rat frontal cortex: cellular localization in different classes of efferent neurons. *Eur. J. Neurosci.* 7, 1050–1063. doi: 10.1111/j.1460-9568.1995.tb01092.x

Gee, S., Ellwood, I., Patel, T., Luongo, F., Deisseroth, K., and Sohal, V. S. (2012). Synaptic activity unmasks dopamine D2 receptor modulation of a specific class of layer V pyramidal neurons in prefrontal cortex. *J. Neurosci.* 32, 4959–4971. doi: 10.1523/JNEUROSCI.5835-11.2012

Haddon, J. E., and Killcross, S. (2011). Inactivation of the infralimbic prefrontal cortex in rats reduces the influence of inappropriate habitual responding in a response-conflict task. *Neuroscience* 199, 205–212. doi: 10.1016/j.neuroscience.2011.09.065

Heyne, A., Kiesselbach, C., Sahún, I., McDonald, J., Gaiffi, M., Dierssen, M., et al. (2009). An animal model of compulsive food-taking behaviour. *Addict. Biol.* 14, 373–383. doi: 10.1111/j.1369-1600.2009.00175.x

Hitchcott, P. K., Quinn, J. J., and Taylor, J. R. (2007). Bidirectional modulation of goal-directed actions by prefrontal cortical dopamine. *Cereb. Cortex* 17, 2820–2827.

Jentsch, J. D., and Taylor, J. R. (1999). Impulsivity resulting from frontostriatal dysfunction in drug abuse: implications for the control of behavior by reward-related stimuli. *Psychopharmacology* 146, 373–390. doi: 10.1007/PL00005483

Killcross, S., and Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex* 13, 400–408.

Kroener, S., and Lavin, A. (2010). Altered dopamine modulation of inhibition in the prefrontal cortex of cocaine-sensitized rats. *Neuropsychopharmacology* 35, 2292–2304.

Lingawi, N. W., and Balleine, B. W. (2012). Amygdala central nucleus interacts with dorsolateral striatum to regulate the acquisition of habits. *J. Neurosci.* 32, 1073–1081. doi: 10.1523/JNEUROSCI.4806-11.2012

McGeorge, A. J., and Faull, R. L. (1989). The organization of the projection from the cerebral cortex to the striatum in the rat. *Neuroscience* 29, 503–537. doi: 10.1016/0306-4522(89)90128-0

Mueller, D., Bravo-Rivera, C., and Quirk, G. J. (2010). Infralimbic D2 receptors are necessary for fear extinction and extinction-related tone responses. *Biol. Psychiatry* 68, 1055–1060. doi: 10.1016/j.biopsych.2010.08.014

Nelson, A., and Killcross, S. (2006). Amphetamine exposure enhances habit formation. *J. Neurosci.* 26, 3805–3812. doi: 10.1523/JNEUROSCI.4305-05.2006

O'Donnell, P. (2010). Adolescent maturation of cortical dopamine. *Neurotox. Res.* 18, 306–312.

Paxinos, G., and Franklin, K. B. J. (2001). *The Mouse Brain in Stereotaxic Coordinates*. San Diego, CA: Academic Press.

Peters, J. LaLumiere, R. T., and Kalivas, P. W. (2008). Infralimbic prefrontal cortex is responsible for inhibiting cocaine seeking in extinguished rats. *J. Neurosci.* 28, 6046–6053. doi: 10.1523/JNEUROSCI.1045-08.2008

Quinn, J. J., Pittenger, C., Lee, A. S., Pierson, J. L., and Taylor, J. R. (2013). Striatum-dependent habits are insensitive to both increases and decreases in reinforcer value in mice. *Eur. J. Neurosci.* doi: 10.1111/ejn.12106

Rich, E. L., and Shapiro, M. (2009). Rat prefrontal cortical neurons selectively code strategy switches. *J. Neurosci.* 29, 7208–7219. doi: 10.1523/JNEUROSCI.6068-08.2009

Robbins, T. W., and Everitt, B. J. (1999). Drug addiction: bad habits add up. *Nature* 398, 567–570. doi: 10.1038/19208

Seamans, J. K., Gorelova, N., Durstewitz, D., and Yang, C. R. (2001). Bidirectional dopamine modulation of GABAergic inhibition in prefrontal cortical pyramidal neurons. *J. Neurosci.* 21, 3628–3638.

Seamans, J. K., and Yang, C. R. (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog. Neurobiol.* 74, 1–58. doi: 10.1016/j.pneurobio.2004.05.006

Sesack, S. R., Deutch, A. Y., Roth, R. H., and Bunney, B. S. (1989). Topographical organization of the efferent projections of the medial prefrontal cortex in the rat: an anterograde tract-tracing study with *Phaseolus vulgaris* leucoagglutinin. *J. Comp. Neurol.* 290, 213–242. doi: 10.1002/cne.902900205

Smith, K. S., Virkud, A., Diesseroth, K., and Graybiel, A. M. (2012). Reversible online control of habitual behavior by optogenetic perturbation of medial prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18932–18937. doi: 10.1073/pnas.1216264109

Tai, L. H., Lee, A. M., Benavidez, N., Bonci, A., and Wilbrecht, L. (2012). Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat. Neurosci.* 15, 1281–1289. doi: 10.1038/nn.3188

Tseng, K. Y., and O'Donnell, P. (2007a). D2 dopamine receptors recruit a GABA component for their attenuation of excitatory synaptic transmission in the adult rat prefrontal cortex. *Synapse* 61, 843–850.

Tseng, K. Y., and O'Donnell, P. (2007b). Dopamine modulation of prefrontal cortical interneurons changes during adolescence. *Cereb. Cortex* 17, 1235–1240.

Vertes, R. P. (2004). Differential projections of the infralimbic and prelimbic cortex in the rat. *Synapse* 51, 32–58. doi: 10.1002/syn.10279

Vincent, S. L., Khan, T., and Benes, F. M. (1995). Cellular colocalization of dopamine D1 and D2 receptors in rat medial prefrontal cortex. *Synapse* 19, 112–120. doi: 10.1002/syn.890190207

Wall, P. M., Blanchard, R. J., Markham, C., Yang, M., and Blanchard, D. C. (2004). Infralimbic D1 receptor agonist effects on spontaneous novelty exploration and anxiety-like defensive responding in CD-1 mice. *Behav. Brain Res.* 152, 67–79.

Winter, S., Dieckmann, M., and Schwabe, K. (2009). Dopamine in the prefrontal cortex regulates rats behavioral flexibility to changing reward value. *Behav. Brain Res.* 198, 206–213. doi: 10.1016/j.bbr.2008.10.040

Yin, H. H., and Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* 7, 464–476. doi: 10.1038/nrn1919

Yin, H. H., Ostlund, S. B., and Balleine, B. W. (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. *Eur. J. Neurosci.* 28, 1437–1148. doi: 10.1111/j.1460–9568.2008.06422.x

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 30 April 2013; paper pending published: 07 June 2013; accepted: 28 June 2013; published online: 18 July 2013.*

*Citation: Barker JM, Torregrossa MM and Taylor JR (2013) Bidirectional modulation of infralimbic dopamine D1 and D2 receptor activity regulates flexible reward seeking. Front. Neurosci. 7:126. doi: 10.3389/fnins.2013.00126*

*This article was submitted to Frontiers in Decision Neuroscience, a specialty of Frontiers in Neuroscience.*

frontiers in
NEUROSCIENCE

REVIEW ARTICLE
published: 05 November 2013
doi: 10.3389/fnins.2013.00201

# Dopaminergic control of cognitive flexibility in humans and animals

***Marianne Klanker*[1,2]*, *Matthijs Feenstra*[1,2] *and Damiaan Denys*[1,2]**

[1] *Netherlands Institute for Neuroscience, Institute of the Royal Netherlands Academy of Arts and Sciences, Amsterdam, Netherlands*
[2] *Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands*

***Edited by:***
*Roshan Cools, University of Cambridge, UK*

***Reviewed by:***
*Mark Walton, University of Oxford, UK*
*Chung-Chuan Lo, National Tsing Hua University, Taiwan*

***Correspondence:**
*Marianne Klanker, Netherlands Institute for Neuroscience, Meibergdreef 47, 1105 BA Amsterdam, Netherlands*
*e-mail: m.klanker@nin.knaw.nl*

Striatal dopamine (DA) is thought to code for learned associations between cues and reinforcers and to mediate approach behavior toward a reward. Less is known about the contribution of DA to cognitive flexibility—the ability to adapt behavior in response to changes in the environment. Altered reward processing and impairments in cognitive flexibility are observed in psychiatric disorders such as obsessive compulsive disorder (OCD). Patients with this disorder show a disruption of functioning in the frontostriatal circuit and alterations in DA signaling. In this review we summarize findings from animal and human studies that have investigated the involvement of striatal DA in cognitive flexibility. These findings may provide a better understanding of the role of dopaminergic dysfunction in cognitive inflexibility in psychiatric disorders, such as OCD.

**Keywords: dopamine, cognitive flexibility, obsessive-compulsive disorder, reversal learning, set-shifting, task switching**

## INTRODUCTION

In a constantly changing environment behavior has to be adaptive and flexible. Cognitive flexibility is the ability to adapt goal-directed behavior in response to changing situational demands. Cognitive flexibility is one of the cognitive domains that are grouped together as executive functions or executive control (Gilbert and Burgess, 2008). Despite the necessity of cognitive flexibility for everyday functioning there is a substantial variation within the healthy population (Miyake and Friedman, 2012) that can be related to variations in dopamine (DA) related genes in humans (Braver et al., 2010; Barnes et al., 2011) and mice (Laughlin et al., 2011). Specific deficits in the ability to flexibly update behavior are observed in various neurological and psychiatric disorders such as Parkinson's disease, schizophrenia, autism, addiction and obsessive compulsive disorder (OCD) (Cools et al., 2001; Chamberlain et al., 2006; Verdejo-Garcia et al., 2006; Ceaser et al., 2008; Yerys et al., 2009).

Here, we intend to provide an overview of animal and human studies on the relation between cognitive flexibility and DA neurotransmission and relate this to OCD, a psychiatric disease that combines defects in cognitive flexibility and alterations in DA processes.

## TESTING COGNITIVE FLEXIBILITY

The successful adaptation of behavior following changes in the environment encompasses several cognitive processes, such as associative learning, decision making, response selection and inhibition, working memory and attention. Several neuropsychological tests have been constructed to study different types of cognitive flexibility, which may recruit varied cognitive functions and depend on parallel neurobiological substrates. The use and translational applicability of a number of these tasks was discussed by Barch et al. (2009). One set of tasks probes flexibility of choice behavior, where selection of one from two or more options leads to a wanted outcome. For a specific response to be adapted, the behavior has to be acquired first. During discrimination learning, subjects learn to discriminate between a certain rewarded/correct stimulus, strategy or response rule and another one that is not rewarded/correct. When task demands change, the response that has been successful so far no longer yields reward and has to be inhibited, whilst another response/stimulus/strategy has to be chosen, initiated and maintained. This requires extinction of the old association and acquisition of a novel association. Classical reversal learning and intra- and extradimensional attentional set-shifting fall in this category.

### *Reversal learning*

With reversal learning, the ability to adapt behavior in response to a reversal of reinforcement contingencies is studied. This requires a shift in valence between stimuli or locations that have been associated with a specific outcome (e.g., a reward) previously. Depending on the operationalization of the reversal task used, this can be a reversal of all sorts of cues, but the choice options remain the same.

### *Attentional set-shifting and strategy shifting*

Attentional set-shifting requires adaptation of behavior following changes in the relevance of perceptual categories or dimensions. In an *intradimensional* set-shift, new stimulus exemplars (i.e., novel choice options) are presented but the relevant stimulus dimension does not change between trials. Successful shifting requires maintenance of the current rule (attentional set) and adapting behavior accordingly. In an *extradimensional* set-shift, not only are the stimulus exemplars novel, but the reinforced dimension has also changed. This requires a response shift to a

dimension that has previously been irrelevant and bypassing of an acquired attentional bias (Rogers et al., 2000).

In human subjects, the ability to shift cognitive sets is commonly tested with the Wisconsin Card Sorting Test (WCST). The WCST requires matching of a multi-dimensional cue card to one of four reference cards according to a specific stimulus aspect. The attentional set-shifting task has been developed as a non-human primate version of the WCST (Roberts et al., 1988). Because it is a more direct measure of the ability to shift cognitive set and a better measure for frontal lobe impairments (Rogers et al., 2000), it is now often used in human subjects as well.

Both reversal learning and attentional set-shifting paradigms have been developed for humans, non-human primates and rodents. Stimulus dimensions consist of different visual stimulus sets that can be simple or compound in nature (human, non-human primate, rodent) or stimulus sets consisting of multiple sensory dimensions (spatial, odor, touch, visual); rodent bowl digging procedure (Birrell and Brown, 2000; Garner et al., 2006). Discriminations based on stimulus valence have been classified as representing a lower order of abstraction, whereas discriminations based on stimulus components or abstract rules may represent a higher order of abstraction (Wise et al., 1996; Ragozzino, 2007).

Another example of a procedure based on a response rule or strategy and an unannounced switch to a different rule or strategy is response-based versus cue-based responding on a T-maze, often applied in rodents (Packard, 2009).

A general problem with switching responses in these tasks is that several processes occur simultaneously and that incorrect responses may reflect different mechanisms, i.e., resistance to extinction versus learned irrelevance (Maes et al., 2004). Task adaptation (Tait and Brown, 2007) or detailed analysis (e.g., Dias et al., 1996a) lead to more informative outcomes. Three-choice paradigms have been used in non-human primates and may offer superior experimental approaches as they allow testing of more variable conditions and require animals to trace the value of several alternative options, as a change in one option does not automatically imply a change in the other alternative options (Walton et al., 2010).

#### *Task switching*

Task switching is a paradigm that is mostly, but not exclusively (Stoet and Snyder, 2003; Leenaars et al., 2012) used in human subjects and requires the rapid switching between stimulus-response sets that have been acquired previously (Sohn et al., 2000; Monsell, 2003). Presentation of an external cue indicates which task (stimulus-response set) has to be executed in a given trial. This differs fundamentally from reversal learning and set-shifting procedures, where the presentation of altered contingencies (i.e., "the switch") is not cued and subjects have to use the change in reinforcing feedback to adapt behavior accordingly.

#### *Control over prepotent or automatic responses*

Another category incorporates tasks that probe the ability to behave flexibly in conditions that previously allowed automatic or habitual performance. A well-known example is the countermanding or stop-signal task (Logan et al., 1984; Eagle et al., 2008), testing inhibitory control over actions. Another example is the anti-saccade task where a more or less automatic action needs to be suppressed to allow flexible responding (Munoz and Everling, 2004). In the present review we focus on studies using reversal learning, attentional set-shifting (including WCST) and task-switching as these tasks have received most translational interest, have been related to DA function and have been performed in OCD patients.

## NEURAL CIRCUITRY SUPPORTING COGNITIVE FLEXIBILITY

### *Prefrontal cortex*

Within the prefrontal cortex (PFC), damage to different prefrontal areas results in dissociable deficits in separate forms of cognitive flexibility. Damage to the orbitofrontal cortex (OFC) is thought to specifically impair reversal learning, but not attentional set-shifting (Dias et al., 1996a; McAlonan and Brown, 2003; Hornak et al., 2004; Boulougouris et al., 2007). Damage to the lateral PFC [or medial PFC in rodents, suggested to be functionally equivalent; (Uylings et al., 2003)] specifically impairs (extradimensional) shifting of attentional sets but not reversal learning (Owen et al., 1991; Dias et al., 1996a, 1997; Birrell and Brown, 2000; Bissonette et al., 2008). However, the proposed unique role of the OFC in reversal learning is under discussion and alternative views have been presented (Schoenbaum et al., 2009). Recent findings suggest that impaired reversal learning in Rhesus monkeys is only observed following aspiration but not excitotoxic OFC lesions (Rudebeck et al., 2013), suggesting that reversal learning does not depend on an intact OFC but instead on intact communication between other prefrontal areas and more caudal structures. While human brain lesions generally involve passing fibers and brain parenchym, many studies in rodents and new world monkeys report deficits after fiber-sparing lesions. The transient character of impairments in these studies may reflect evolution-related differences in neurobiological and/or anatomical substrates of reversal learning (Rudebeck et al., 2013).

### *Striatum*

Reciprocal projections from PFC to the striatum and thalamus form parallel frontostriatal loops, suggesting striatal regions also contribute to the regulation of cognitive flexibility (Rogers et al., 2000; Floresco et al., 2006a; Ragozzino, 2007; Clarke et al., 2008; Castane et al., 2010). Combined results from lesion and functional imaging studies suggest that different types of cognitive flexibility are regulated by segregated fronto-striatal circuits: OFC and dorsomedial striatum (human/non-human primate: caudate nucleus; functional equivalent rodent area: dorsomedial striatum) are implicated in reversal learning (Divac, 1971; Dias et al., 1996a; Rogers et al., 2000; McAlonan and Brown, 2003; Bellebaum et al., 2008; Clarke et al., 2008; Castane et al., 2010; Ghahremani et al., 2010). Set- and task switching performance relies on connections between the dorsolateral PFC (or the medial PFC in rodents which is in this task functionally equivalent) and striatum (Owen et al., 1991; Dias et al., 1996a,b; Birrell and Brown, 2000; Sohn et al., 2000; Manes et al., 2002; Ragozzino, 2007; Graham et al., 2009). It should be noted that these circuits

are not fully segregated but overlapping. Importantly, these circuits show consistent similarities between primates and rodents (Mailly et al., 2013).

### DOPAMINE

DA is an important neuromodulator in fronto-striatal circuits. A substantial amount of work has described a role for DA in reward-related learning and motivated behavior. More specifically, burst firing of DA neurons (associated with phasic DA release) may code a quantitative prediction error that serves as a teaching signal to guide behavior and is essential for a range of learning situations (Montague et al., 1996; Schultz et al., 1997; Schultz, 2013; Steinberg et al., 2013). Yet not much is known about the contribution of DA to the adaptation of behavior following changing task demands, such as a reversal of contingencies. A common factor in all tests of cognitive flexibility is the expectation of a reward (or absence of punishment) when a correct response is made. The absence of an expected reward and presence of an unexpected reward following a reversal or shift is the archetypal situation for the occurrence of reward prediction errors coded by DA. Therefore, one would expect that DA is in some way involved in the regulation of cognitive flexibility. However, in the past decade the role of the PFC and its serotonergic innervation in cognitive flexibility received most attention (e.g., Robbins and Arnsten, 2009).

In this review, we summarize findings from animal and human studies that investigated whether DA contributes to the regulation of cognitive flexibility. First, we will describe pharmacological manipulations to the DA system in humans and animals, then DA-related genetics in humans and animals. Next, we report on DA changes and cognitive flexibility in OCD, to investigate whether alterations in DA signaling contribute to cognitive inflexibility in this disorder. Previously, OCD has been proposed to be characterized by a hyperdopaminergic state (Denys et al., 2004b) and similar states in animals have repeatedly been described as leading to OCD-like behaviors (see further). This, combined with the suggestion that impairments in the ability to flexibly adapt behavior may be an endophenotype for OCD (Robbins et al., 2012) drove us to review the evidence for a relation between the two.

## PHARMACOLOGICAL MANIPULATIONS AND IMAGING STUDIES IN HUMAN SUBJECTS

### DA SYNTHESIS

DA synthesis capacity in humans is determined after administration of radio labeled F-DOPA or F-tyrosine and imaging the resulting fluorinated amines using PET. The observed variations in DA synthesis capacity may relate to variations in DA neurotransmission, as a significant negative correlation between synthesis capacity and $D_2$- receptor availability was reported (Ito et al., 2011). Decreasing DA synthesis by dietary omission of DA precursors tyrosine and phenylalanine reduces occupation of $D_2$ receptors by endogenous DA, suggesting decreased DA transmission (Montgomery et al., 2003). Administration of the tyrosine hydroxylase inhibitor alpha-methyl-paratyrosine also reduces $D_2$ occupation by endogenous DA (Verhoeff et al., 2003), but affects noradrenergic signaling as well (Krahn et al., 1999).

The small number of studies using these approaches does not support a general relation between DA synthesis and flexible updating of task information: no correlation was observed between DA synthesis capacity and task performance on the WCST (Vernaleken et al., 2007), and reward- and punishment-based reversal learning was not impaired following DA depletion in males (Robinson et al., 2010). In contrast, catecholamine depletion (affecting both DA *and* NA) impaired performance during probabilistic reversal learning (Hasler et al., 2009).

Other studies suggest that when tasks are used that allow more selective approaches, a differential involvement of DA synthesis is observed. Thus, subjects with high DA synthesis capacity perform worse compared to subjects with low DA synthesis capacity when presented with shifts in object features but not in abstract rules in a task-switching paradigm (Dang et al., 2012). Cools et al. (2009) reported that individuals with high DA synthesis capacity perform better when presentation of an unexpected reward signals reversal compared to reversals that are signaled by presentation of an unexpected punishment, whereas the opposite is observed for individuals with low DA synthesis capacity. Females tend to have a higher DA synthesis capacity (Laakso et al., 2002) and this may explain gender-related differences such as the DA depletion-induced improvement of punishment-based but not reward-based reversal learning in females (Robinson et al., 2010).

In conclusion, DA synthesis is differentially associated with task features in cognitive flexibility and variations in synthesis capacity affect performance only in some task conditions, probably depending on specific DA homeostasis parameters in cortical and striatal areas (cf. Cools and D'Esposito, 2011).

### DA RECEPTOR/TRANSPORTER BINDING

Using imaging techniques, baseline availability of DA receptors and transporters can be investigated and related to task performance. Receptor availability in resting conditions provides an index of the number of receptors unoccupied by the endogenous transmitter. Subjects with higher availability of DA transporters in the striatum make less perseverative errors in the WCST (Hsieh et al., 2010) but the interpretation of this finding depends on whether the higher availability reflects the density of the DA innervation or a possible substrate-induced adaptation (Chen et al., 2010).

WCST performance has also been linked to differences in DA receptor availability (see **Table 1**). Decreased striatal $D_2$ availability is associated with impaired performance (Volkow et al., 1998), but $D_2/D_3$ receptor binding in the anterior cingulate cortex correlates positively with the number of errors made in the WCST (Lumme et al., 2007).

For DA transmission through $D_1$ receptors, an optimal level of DA activity is required for best working memory performance (Williams and Goldman-Rakic, 1995; Zahrt et al., 1997; Vijayraghavan et al., 2007). Similar results were obtained for flexible responding in the WCST where impaired performance is observed for both high and low prefrontal $D_1$ (but not $D_2$) binding [(Takahashi et al., 2008), but see Karlsson et al. (2011)].

When receptor availability is assessed during task performance, it provides a measure of task-related release of endogenous DA. Reduced binding to $D_2$ receptors in the dorsal striatum

(Monchi et al., 2006a) and anterior cingulate cortex (Ko et al., 2009) during set-shifting (see Monchi et al., 2006b) suggests that DA is indeed released during tasks requiring flexibility. Transient inactivation of dorsolateral PFC activity impaired striatal DA release as well as task performance, suggesting both are under top–down control by the dorsolateral PFC (Ko et al., 2008).

Taken together, these findings indicate that DA is activated and can influence performance on set-shifting tasks through $D_2$ receptors in the striatum and anterior cingulate cortex, whereas in the PFC, DA activity through $D_1$ receptors can modulate performance. In addition, optimum values may exist for both extracellular DA concentrations and DA receptor numbers. The majority of studies relating performance on cognitive flexibility tasks to DA-receptor binding potential have specifically focused on binding to $D_2$ receptors in specifically delineated brain areas. Therefore, although this provides evidence that $D_2$ receptors modulate performance in these types of tasks, one cannot exclude the involvement of $D_1$ receptors.

## PHARMACOLOGICAL MANIPULATIONS AFFECTING DA SIGNALING

DA neurotransmission during task performance can be influenced by administration of pharmacological agents that directly bind to DA receptors or by drugs that induce DA release. Combining the administration of pharmacological agents with functional imaging during task performance indicates in which brain areas modulation by DA is most pronounced.

### DA ANTAGONIST

Systemic administration of the $D_2$ receptor antagonist sulpiride slows response times during task-switching (Mehta et al., 2004) and impairs performance of an extra-dimensional set-shift, without affecting intra-dimensional set-shifting (Mehta et al., 1999, 2004). Sulpiride enhances performance on reward-based reversal learning (van der Schaaf et al., 2012). This behavioral effect was stronger in subjects with higher working memory capacity [which is assumed to reflect higher striatal DA synthesis capacity (Cools et al., 2008)]. In addition to behavioral effects, sulpiride also increased striatal BOLD signals during unexpected outcomes, irrespective of whether the unexpected outcome was a reward or a punishment (van der Schaaf et al., 2012).

### INDIRECT DA AGONIST

Methylphenidate is a psychostimulant that increases striatal extracellular DA levels (Volkow et al., 2001), but also affects serotonin (5-hydroxytryptamin, 5-HT) and noradrenaline (Kuczenski and Segal, 1997). Administration of methylphenidate leads to displacement of raclopride binding to $D_{2/3}$ receptors (Clatworthy et al., 2009). These changes in the post commissural part of the caudate nucleus were associated with effects on reversal learning, such that a large displacement following methylphenidate was associated with impaired performance and a small displacement with improved performance (Clatworthy et al., 2009). As these effects may depend on individual variation in receptor availability and DA synthesis capacity, behavioral effects of the psychostimulant on measures of flexibility are likely to be averaged out when the individual variation is not taken into account—which may explain the negative results on attentional set-shifting (Elliott et al., 1997).

Administration of methylphenidate influences brain activation in ventral striatal regions during behavioral adaptation and modulates activity in frontal regions during cognitive control. Thus, activation in ventral striatal regions was reduced during reversal errors (even in the absence of behavioral effects), whereas in prefrontal regions, increased activation was observed following correct responses (Dodds et al., 2008). The balance of DA in frontal and striatal regions may therefore be crucial in regulating the balance between cognitive control and cognitive flexibility.

### DA AGONIST

Interestingly, DA synthesis capacity also influences the effect of direct DA agonists on task performance. While (Mehta et al., 2001) originally observed an increase in non-perseverative errors and slowed reaction times during probabilistic reversal learning after administration of the $D_2$ agonist bromocriptine, Cools et al. (2009) later showed that this drug impaired reversal learning from unexpected rewards in subjects with high DA synthesis capacity, but improved the same parameter in subjects with low synthesis capacity in striatal regions.

The beneficial effect of $D_2$ receptor stimulation in subjects with low DA synthesis capacity is not limited to reversal learning. Bromocriptine can also improve performance on the WCST (Kimberg et al., 1997) and task-switching performance (van Holstein et al., 2011) in subjects with low DA synthesis capacity, whereas no effects are observed following administration of pergolide, which differs from bromocriptine in that it also activates $D_1$ receptors (Kimberg and D'Esposito, 2003). That the improvement on task switching after bromocriptine can be specifically related to the function of $D_2$ receptors was shown by (van Holstein et al., 2011), as pre-treatment with the $D_2$ antagonist sulpiride blocked the beneficial effect. Therefore, performance of subjects with high DA synthesis capacity is impaired following administration of bromocriptine, and increases following administration of sulpiride.

## SUMMARY AND CONCLUSION

To conclude (see **Table 1**), flexible updating of behavior in set-shifting tasks (WCST and attentional set-shifting) as well as task switching is associated with increased DA neurotransmission through $D_2$-receptors. In particular, the mediating effects of $D_2$ signaling on task performance have been observed in the dorsal striatum and anterior cingulate cortex, which is in line with observations from imaging and lesion studies suggesting the involvement of the connections between PFC and dorsal striatum in the regulation of these types of flexibility (Owen et al., 1991; Sohn et al., 2000). This also concurs with observations in patients with PD. In the early stages of PD, when DA depletion is largely limited to the dorsal striatum, patients show impairments in task switching whereas reversal learning performance is spared. Administration of levodopa reverses the impairments in task switching, whilst it impairs performance on reversal learning probably due to overstimulation of DA receptors in ventral striatal regions (Cools, 2006; Kehagia et al., 2010). In control subjects increased $D_2$-mediated transmission also impairs reversal learning, although this may turn into an improvement when DA synthesis capacity is low.

**Table 1 | Summary of effects of pharmacological manipulations to the dopamine system on cognitive flexibility in human subjects.**

| Paradigm | Manipulation | | Performance | References |
|---|---|---|---|---|
| **WCST** | | | | |
| | Depletion | | ↓ response times | Nagano-Saito et al., 2008 |
| | Synthesis capacity | | = | Vernaleken et al., 2007 |
| | $D_2$ agonist | | ↑ | Kimberg et al., 1997 |
| | $D_1/D_2$ agonist | | = | Kimberg and D'Esposito, 2003 |
| | DAT availability | Striatum | ↑ reduced errors | Hsieh et al., 2010 |
| | $D_2/D_3$ binding | Anterior cingulate | | Lumme et al., 2007 |
| | $D_1$ binding | Dorsolateral PFC | | Takahashi et al., 2008 |
| **REVERSAL LEARNING** | | | | |
| | DA Depletion | | ↑ punishment based reversal | Robinson et al., 2010 |
| | Catecholamine depletion | | ↓ probabilistic reversal | Hasler et al., 2009 |
| | Synthesis capacity | High<br>Low | ↑ reward based reversal<br>↑punishment based reversal | Cools et al., 2009 |
| | $D_2$ agonist | High DA synthesis Low DA synthesis | ↓ more errors, longer RT<br>↓ reward based reversal<br>↓ reward based reversal | Mehta et al., 2001<br>Cools et al., 2009 |
| | $D_2$ antagonist | | ↑ reward based reversal | van der Schaaf et al., 2012 |
| **TASK SWITCHING** | | | | |
| | Synthesis capacity | | = abstract rule shift<br>↑ object feature shift | Dang et al., 2012 |
| | $D_2$ agonist | Low DA synthesis | ↑ | van Holstein et al., 2011 |
| | $D_2$ antagonist | | ↓ longer RT | |
| **ATTENTIONAL SET-SHIFT** | | | | |
| | $D_2$ binding | Dorsal striatum Anterior Cingulate | Binding reduced during shifts | Monchi et al., 2006a; Ko et al., 2009 |
| | $D_2$ antagonist | | ↓ EDS performance<br>= IDS performance | Mehta et al., 1999, 2004 |
| | $D_2$ agonist | Methylphenidate | = | Elliott et al., 1997 |

*= no effect, ↑ increased performance, ↓ decreased performance.*

*DA, dopamine; RT, reaction time; EDS, extra dimensional set-shift; IDS, intradimensional set-shift.*

Human studies have particularly shown the importance of individual differences in the DA system. Individual differences in DA synthesis capacity influence both task performance and effects of manipulations to the DA system in different types of flexibility. Individual differences in $D_2$ receptor availability also influence stimulation-induced changes in performance during reversal learning. The combined study of manipulations to the DA system with performance on behavioral tasks, indicate that DA transmission in the ventral striatum changes during reversal learning.

These results also indicate that there may be differences in the involvement of DA in reversal learning compared to set-shifting and task switching. As noted before, these paradigms are thought to represent different levels of complexity and may depend on different brain areas. However, studies differ in the task designs used to study one type of cognitive flexibility. Therefore, replication of effects of DAergic manipulations using similar task designs would help in delineating the possible differences in DA contribution to reversal, set-shifting and task switching.

A question remains in what way $D_1$ receptors contribute to behavioral performance during cognitive flexibility tasks. Direct manipulations of $D_1$ signaling or studies relating performance on behavioral task to $D_1$ receptors availability are scarce. Combining the administration of pharmacological agents with functional imaging during performance of different behavioral paradigms may provide more insight on the effects of DA on cognitive flexibility in prefrontal and striatal regions.

## PHARMACOLOGICAL MANIPULATIONS IN ANIMALS

The use of pharmacological imaging in human subjects provides insight into the role of DA in cognitive flexibility, but the use of animals permits direct (and invasive) manipulations and measurements and can extend and specify findings obtained in human subjects. Here, we will discuss animal studies that have used pharmacological manipulations of the DA system or DA depletion to investigate in what way DA in prefrontal and striatal regions contributes to cognitive flexibility.

### DA DEPLETION STUDIES

In rodents, lesioning DAergic projections in the nucleus accumbens core (though DA in the medial PFC was similarly affected) impairs both spatial discrimination and reversal learning on a T-maze (Taghzouti et al., 1985). Selective depletion of DA neurotransmission in the dorsomedial striatum impairs odor guided

reversal learning, without affecting initial discrimination learning (O'Neill and Brown, 2007). A selective deficit in reversal learning following DA depletion in the dorsomedial striatum was observed in primates as well (Clarke et al., 2011). The deficit in reversal learning following DA depletion is not perseverative, suggesting that DA may be particularly important for the learning phase after reversal, rather than mediating response inhibition to the previously rewarded side. The effect was not only shown in the first, but also in subsequent reversals. Importantly, the deficit is neurochemically specific, as depletion of 5-HT neurotransmission in the mediate caudate nucleus does not affect behavioral performance during reversal learning (Clarke et al., 2011). A previous study also found decreased performance on reversal learning (although this did not reach significance) (Collins et al., 2000). Subsequently, Crofts et al. (2001) showed that although acquisition, maintenance and initial shifting of an attentional set are intact, monkeys with DA depletion in the caudate are impaired when they have to make an attentional shift to a stimulus dimension that was learned to be irrelevant in a previous extra dimensional shift (Collins et al., 2000; Crofts et al., 2001). Therefore, DA in the caudate nucleus appears to be involved in situations that require a shift of established cognitive sets (Collins et al., 2000).

In contrast to DA depletion in striatal regions, selective DA depletion in frontal regions is complicated by the accompanied depletion of noradrenaline (Roberts et al., 1994; Crofts et al., 2001). Although Roberts et al. (1994) observed a specific improvement in performance on extra-dimensional set-shifts after prefrontal catecholamine depletion in non-human primates, a later study suggests that this may actually result from an inability to maintain an attentional set (Crofts et al., 2001). Prefrontal catecholamine depletion is associated with long lasting enhancement of striatal DA release, suggesting that it may be the balance between DA levels in prefrontal and striatal regions rather than DA levels in either region that affects behavior (Roberts et al., 1994).

## DA VERSUS 5-HT

Based on data from depletion studies, a neurochemical dissociation between prefrontal and striatal regions in the control of cognitive flexibility during reversal learning has been suggested. In the caudate nucleus, DA, but not 5-HT depletion impairs performance during reversal learning. Previously, it was reported that 5-HT, but not DA neurotransmission in the OFC is required for successful behavioral adaptation in a spatial reversal learning task (Clarke et al., 2004, 2007). Depletion of 5-HT in the OFC specifically impairs reversal learning by increasing perseverative responding, but does not affect attentional set-shifting (Clarke et al., 2005). OFC DA depletion, however, leads to impaired extinction, albeit not in a perseverative manner (Walker et al., 2009). The contributions of 5-HT and DA neurotransmission to cognitive flexibility therefore appear to be confined to separate functions related to regions of the cortico-striatal circuit. Recently, (Groman et al., 2013) suggested that the balance between 5-HT levels in the OFC and DA levels in the dorsal striatum contributes to individual differences in cognitive flexibility. Reduced performance on a reversal learning task is associated with low levels of 5HT in the OFC when DA levels in the putamen are low, but not when DA levels in the putamen are high (Groman et al., 2013). These findings indicate that cognitive flexibility is under control of DA and 5-HT, while other data show involvement of noradrenaline, as well (Bouret and Sara, 2004; Lapiz and Morilak, 2006; Seu et al., 2009).

## EFFECTS OF PSYCHOSTIMULANTS

Psychostimulants such as methylphenidate, (meth)amphetamine and cocaine increase release of DA and other monoamines by blocking catecholamine re-uptake or promoting DA release (Sulzer et al., 2005). Administration of methylphenidate in rodents does not affect reversal learning (Seu and Jentsch, 2009; Cheng and Li, 2013), although the latter authors observed beneficial effects in animals with reversal learning impairments (spontaneously hypertensive rats). Effects of amphetamine and methamphetamine on reversal learning have been variable, but possibly dose-dependent: high doses (5 mg/kg) impair reversal learning (Ridley et al., 1981; Arushanian and Baturin, 1982; Idris et al., 2005; Cheng et al., 2007; White et al., 2009; Izquierdo et al., 2010; Kosheleff et al., 2012; Talpos et al., 2012), while intermediate doses 1–2 mg/kg show no effect or improved learning (Wilpizeski and Hamilton, 1964; Kulig and Calhoun, 1972; Mead, 1974; Weiner and Feldon, 1986; Weiner et al., 1986; Daberkow et al., 2008; Pastuzyn et al., 2012; Soto et al., 2012) and low doses again impair reversal performance (Ridley et al., 1981; Idris et al., 2005). These results are compatible with the general idea that cognitive function depends on DA activity in an inverse U-shaped fashion (Cools and D'Esposito, 2011; Arnsten et al., 2012). However, given the multiple and differential effects of psychostimulants on monoamine release in prefrontal and striatal regions it is often difficult to conclude whether these effects depend on increased DA release. Yet, for methylphenidate Cheng and Li (2013) showed that the beneficial effect were blocked by local injections with haloperidol in the OFC.

## SYSTEMIC EFFECTS OF DA (ANT)AGONISTS

While selective depletion studies indicate specific brain areas where DA modulates flexible behavior, administration of pharmacological agents that are selective for a specific receptor subtype indicate how $D_1$ and $D_2$ receptor subtypes are involved. In primates, both stimulation and inhibition of $D_2/D_3$ receptor function results in difficulties in adapting behavior following changing task demands, but not during acquisition of the original discrimination (Smith et al., 1999; Lee et al., 2007). Administration of the $D_2/D_3$ antagonist raclopride affects performance on reversal learning when administered alone, but only when the reversal is preceded by retention of the originally acquired discrimination (Lee et al., 2007). Performance is also reduced by the $D_3/D_2$ agonist 7-OH-DPAT (Smith et al., 1999) and this deficit is antagonized by co-administration with the $D_2/D_3$ antagonist raclopride, but not the $D_2$-selective antagonist sulpiride, suggesting stimulation of $D_3$ receptors impairs performance (Smith et al., 1999).

In rodents, like in primates, administration of a $D_2/D_3$ agonist (quinpirole) impaired spatial reversal learning in an operant chamber by increasing the number of perseverative errors.

Administration of a $D_2/D_3$ antagonist (raclopride) or selective $D_3$ antagonist (nafadotride) had no effect (Boulougouris et al., 2009). The quinpirole-induced deficit is attenuated when raclopride is co-administered, but worsens after co-administration with nafadotride. Selective stimulation of $D_2$-receptors (co-administration of quinpirole and nafadotride) increased both the number of discrimination errors and of perseverative and learning errors in the reversal phase (Boulougouris et al., 2009). Thus, stimulation of $D_3$ receptors may be important for the acquisition of altered response-reward contingencies during reversal learning whereas $D_2$-receptor activation may cause a more generalized impairment (Boulougouris et al., 2009).

Systemic administration of a $D_1/D_5$ antagonist does not affect reversal learning in primates (Lee et al., 2007), though in rodents systemic administration of a $D_1$ agonist (SKF-812979) impairs early, but not late stages of reversal learning (Izquierdo et al., 2006). Extradimensional set-shifting on the other hand improves following intermediate, but not high or low doses of a $D_1$ agonist (Nikiforuk, 2012).

These findings suggest that $D_2$-like receptors contribute to the regulation of cognitive flexibility, possibly in a dose-dependent manner. System administration of $D_1$-like receptors has received less attention and could affect cognitive flexibility depending on the species or behavioral task used.

### LOCAL EFFECTS IN THE STRIATUM

Local manipulations of DA neurotransmission can elucidate in which way DA neurotransmission in specific subregions of the fronto-striatal circuit can contribute to cognitive flexibility (although see, Arnt, 1985) for the limitations of this approach). Execution or suppression of actions leading to reward are controlled by two parallel cortico-striato-thalamo-cortical pathways (Frank and Claus, 2006). From the striatum, output neurons in the *direct pathway* connect to cortical regions via connections to globus pallidus pars interna (GPi)/substantia nigra pars reticulata (SNr) and thalamus. Output neurons in the *indirect pathway* project via globus pallidus pars externa, subthalamic nucleus to GPi/SNr, thalamus and cortex. Activity in these pathways can be differentially modulated by activation of $D_1$ or $D_2$ receptors in the striatum (Frank and Claus, 2006). Yawata et al. (2012) investigated pathway specific control of reward learning and cognitive flexibility. Blocked neurotransmission in the direct pathway, combined with $D_1$ blockade in the contralateral nucleus accumbens impaired the acquisition phases of the original discrimination as well as the discrimination presented after a reversal or a rule shift, while stimulation of $D_1$ receptors did not influence behavior (Yawata et al., 2012). Application of a $D_2$ agonist combined with contralateral blockade of the indirect pathway induced perseverative responding during reversal learning and also affected rule shifting, without affecting acquisition of the original discrimination problem (Yawata et al., 2012). These findings suggest that within the nucleus accumbens, stimulation of DA $D_1$ receptors (direct pathway) aids the acquisition and relearning of behavioral responses to a particular stimulus, whereas suppression (i.e., a phasic interruption) of $D_2$-mediated transmission (indirect pathway) may be required to allow reorganization of ongoing behavioral patterns. These results are in line with previous findings reporting impaired reversal learning after local stimulation of $D_2$ receptors, while during set-shifting blocking $D_1$ receptors impaired maintenance of the new strategy and stimulation of $D_2$ receptors induced perseverative responding (Haluk and Floresco, 2009).

### LOCAL EFFECTS IN THE PREFRONTAL CORTEX

DA depletion in the OFC did not affect reversal learning (Clarke et al., 2007), but local manipulation of DA receptors in the OFC can influence aspects of cognitive flexibility. Blockage of $D_1$ or $D_2$ receptors in OFC prevents development of discriminative reaction times to high and low rewards under reversal conditions, without affecting accuracy (Calaminus and Hauber, 2008). In a task that required rats to adapt behavior following a change in reward value, by manipulating the amount of lever presses required to obtain a food pellet, local inhibition of $D_1$ but not $D_2$ receptors in the OFC impaired performance (Winter et al., 2009). In the MPFC, local inhibition of both $D_1$ and $D_2$ receptors inhibits performance (Winter et al., 2009). Set-shifting ability in a maze-based shifting task is affected by manipulations of several DA receptors in the MPFC. Local blockade of $D_1$ and $D_2$ receptors as well as stimulation of $D_4$ receptors results in perseverative responding, whereas blockade of the $D_4$ receptor improves performance (Ragozzino, 2002; Floresco et al., 2006b). This contrasts with the findings of $D_1$ blockade in the nucleus accumbens, which did not induce perseverative responding, but affected maintenance of the new strategy.

### *In vivo* DA MEASUREMENTS RELATED TO COGNITIVE FLEXIBILITY

Only a few reports on the measurement of extracellular levels of DA in the brain (reflecting DA release) are available. In the nucleus accumbens, these levels are higher during acquisition of a rule shift compared to simple rule acquisition in a T-maze set-shift paradigm (Stefani and Moghaddam, 2006), clearly suggesting a role for DA in the nucleus accumbens in the regulation of cognitive flexibility, in particular strategy or set-shifting. In the mPFC, both rule acquisition and rule shifting in a T-maze are accompanied by increased DA levels and higher basal mPFC DA levels were associated with rapid shifting between discrimination rules (Stefani and Moghaddam, 2006). After inhibiton of COMT, animals also show increased task-related, but not basal extracellular DA levels in the medial PFC, suggesting that task-induced increases in PFC DA release may contribute to set-shifting performance (Tunbridge et al., 2004).

DA (but not noradrenaline) release in the MPFC is elevated and prolonged during performance of a spatial reversal session in a skinnerbox, compared to release in a discrimination session preceding reversal (van der Meulen et al., 2007). Within the reversal session, the DA elevation was most pronounced during the phase in which rats improved performance.

These findings suggest elevated DA release in both striatal and prefrontal regions during execution of cognitive flexibility tasks.

### SUMMARY AND CONCLUSION

Taken together (see **Table 2**), DA appears to be actively involved in the performance of tasks requiring cognitive flexibility: DA

**Table 2 | Summary of effects of pharmacological manipulations to the dopamine system on cognitive flexibility in animals.**

| Paradigm | Region | | Manipulation | Performance | References |
|---|---|---|---|---|---|
| **SET-SHIFT** | | | | | |
| | Nucleus accumbens | $D_1$ | Agonist | = | Haluk and Floresco, 2009 |
| | | | Antagonist | ↓ Impaired maintenance new strategy | |
| | | $D_2$ | Agonist | ↓ perseveration | Haluk and Floresco, 2009 |
| | | | Antagonist | = | |
| | Dorsomedial striatum | | Depletion | ↓ EDS, only when switching to previously dimension | Collins et al., 2000 |
| | MPFC | $D_1$ | Agonist | = | Floresco et al., 2006b |
| | | | Antagonist | ↓ Perseverative | Ragozzino, 2002 |
| | | $D_2$ | Agonist | = | Floresco et al., 2006b |
| | | | Antagonist | ↓ more trials/errors to criterion. Perseverative | |
| | | D4 | Agonist | ↓ more trials/errors to criterion. Perseverative | Floresco et al., 2006b |
| | | | Antagonist | ↓ more trials/errors to criterion. Perseverative | |
| | Frontal | | Depletion | ↓ Maintenance of set (IDS) | Crofts et al., 2001<br>Roberts et al., 1994 |
| **REVERSAL** | | | | | |
| | Systemic (primate) | $D_1$ | Antagonist | = | Lee et al., 2007 |
| | Systemic (rodent) | | | ↓ | Izquierdo et al., 2006 |
| | Systemic (rodent) | $D_2/D_3$ | Agonist | ↓ perseveration | Boulougouris et al., 2009 |
| | Systemic (primate) | $D_2/D_3$ | Antagonist | ↓ more trials/errors to criterion | Lee et al., 2007 |
| | Systemic (rodent) | | | = | Boulougouris et al., 2009 |
| | Systemic (primate) | $D_3/D_2$ | Agonist | ↓ more trials/errors to criterion | Smith et al., 1999 |
| | Nucleus accumbens | $D_1$ | Agonist | = | Haluk and Floresco, 2009 |
| | | | Antagonist | = | Calaminus and Hauber, 2007 |
| | | $D_2$ | Agonist | ↓ trials to criterion/errors, but not perseveration | Haluk and Floresco, 2009 |
| | | | Antagonist | = | Calaminus and Hauber, 2007 |
| | | | Depletion | ↓ | Taghzouti et al., 1985 |
| | Dorsomedial striatum | | Depletion | ↓ more trials to criterion | O'Neill and Brown, 2007<br>Clarke et al., 2011 |
| | OFC | $D_1$ | Antagonist | ↓ absence discriminative reaction times (high/low reward)<br>↓ impaired maintenance low effort response | Calaminus and Hauber, 2008<br>Winter et al., 2009 |
| | | $D_2$ | Antagonist | ↓ absence discriminative reaction times (high/low reward)<br>= reversal required effort not affected | Calaminus and Hauber, 2008<br>Winter et al., 2009 |
| | MPFC | $D_1$ | Antagonist | ↓ impaired maintenance low effort response | Winter et al., 2009 |
| | | $D_2$ | Antagonist | ↓ impaired maintenance low effort response | Winter et al., 2009 |

*= no effect, ↑ increased performance, ↓ decreased performance.*

*EDS, extra dimensional set-shift; IDS, intradimensional set-shift; MPFC, medial prefrontal cortex; OFC, orbitofrontal cortex.*

release is increased, local DA depletion impairs performance and pharmacological interference alters task execution. Whereas DA depletion studies indicated ventral and dorsomedial striatum as the primary location where DA influences cognitive flexibility, specific DA receptor stimulation/blockade studies and *in vivo* release measurements implicate prefrontal regions as well. A complicating factor is that manipulation of prefrontal DA also affects striatal DA transmission (Roberts et al., 1994).

It is important to note that impairment of reward-related learning and cognitive flexibility following perturbations in DA signaling is almost always of transient nature: subjects eventually do make the switch when sufficient trials are presented, suggesting that DA may facilitate these behaviors, but is not indispensable.

Interestingly, most pharmacological studies investigating the involvement of DA-subtype selective receptors have indicated that striatal blockade of $D_1$-receptors and overactivation of $D_2$-receptors impairs performance. This was most elegantly shown in the study of Yawata et al. (2012): DA signaling through $D_1$ receptors in the nucleus accumbens and the direct basal ganglia pathway contributes to the acquisition of a new reward-directed behavior in a four-armed maze once switching has occurred (i.e., $D_1$ stimulation could contribute to new learning following a behavioral switch), whereas suppression of $D_2$-mediated transmission in the accumbens and the indirect pathway is required for the reorganization of behavioral patterns. A transient elevation in DA potentiates connections in the direct pathway to initiate movement toward reward, whereas a transient dip in DA potentiates connections in the indirect pathway to suppress movements that are no longer rewarded (Hong and Hikosaka, 2011). The findings from animal studies do indicate a role for the DA in the nucleus accumbens mediating cognitive flexibility, both reversal and strategy or set-shifting, whereas less research has focused on local manipulation of $D_1$ or $D_2$ receptors in dorsomedial or dorsolateral striatal regions. However, a role for dorsal striatal regions has been indicated by selective DA depletion studies as well as a significant amount of human data. Moreover, in the primate dorsal striatum (caudate and putamen), availability of $D_2$-receptors can be related to performance during reversal but not discrimination learning (Groman et al., 2011). This warrants further investigation of the effects of manipulating $D_1$ or $D_2$ signaling in striatal regions other than the nucleus accumbens.

In general, these conclusions are similar to those based on human data, as discussed in the previous section. However, unlike what was reported in humans, $D_2$-based manipulations seem to affect lower order (cue reversal) and higher order (rule or task switch) processes in a similar way. It is unclear if $D_2$-mediated effects in animals depend on DA synthesis capacity.

## CONTRIBUTIONS OF DA GENOTYPE TO COGNITIVE FLEXIBILITY IN HUMANS

Individual variability in executive functioning may be subserved by a strong genetic component (Friedman et al., 2008). The expression of complex traits such as cognitive flexibility is likely regulated by multiple genes that each contribute a small effect. Several polymorphisms in genes affecting DA functioning have been investigated to explain individual variability in cognitive flexibility.

### DA RECEPTORS AND INTRACELLULAR SIGNALING

#### $D_1$

DARPP-32 (DA and cAMP regulated phosphoprotein of 32kDA) is strongly expressed in medial spiny neurons in the striatum, where it is stimulated by $D_1$ and inhibited by $D_2$ receptor activation and mediates post-receptor effects of DA (Nishi et al., 1997; Svenningsson et al., 2004). Enhanced performance on several cognitive tasks, including the WCST, was observed for a frequent haplotype in the DARPP-32 gene that is associated with increased post-mortem DARPP-32 expression and affects structural and functional connectivity between PFC and striatum (Meyer-Lindenberg et al., 2007). The polymorphism was also associated with better learning from positive feedback (Frank et al., 2007). This suggests $D_1$ receptors in the striatum could contribute to learning after positive feedback, supporting successful switching of behavior in cognitive flexibility tasks by maintaining responses to the newly rewarded site.

#### $D_2$

The DRD2-TAQ1 polymorphism is located close to the exon coding for the $D_2$ receptor. A1-allele carriers show a reduced number of available $D_2$ receptors [(Thompson et al., 1997; Pohjalainen et al., 1998), but see Lucht and Rosskopf (2008)] and the A1-allele is associated with increased DA synthesis in the striatum (indicating reduced autoreceptor-mediated feedback regulation) (Laakso et al., 2005). In a probabilistic learning task, carriers of the A1-allele showed reduced ability to learn from errors accompanied by functional changes in the frontostriatal circuitry (Klein et al., 2007). A1-carriers showed blunted reward-related activity in the NAC, reduced activity in the posterior medial frontal cortex during negative feedback and reduced interactions between the medial frontal cortex and hippocampus (Klein et al., 2007). The use of feedback is required to adapt responding during reversal learning and, not surprisingly, A1-carriers perform worse (Jocham et al., 2009). Following presentation of a reversal, they were less likely to maintain the newly rewarded response, but kept alternating responses and showed diminished activation of orbitofrontal and ventral striatal regions during reversals (Jocham et al., 2009). Task-switching performance on the other hand is improved in A1-carriers, who show reduced switch costs associated with decreased activity in the lateral PFC and decreased connectivity between PFC and dorsal striatal regions (Stelzel et al., 2010). Switching tasks does not depend on the use of feedback and is supported by different circuits/areas than switching responses based on the use of feedback (Stelzel et al., 2010). This illustrates how impaired DA transmission could have different effects depending on the operationalization of the cognitive flexibility task that is used, i.e., whether on-line feedback-induced response adaptation ("learning") is essential or not.

A second polymorphism affecting availability of striatal $D_2$ receptors is the C957T polymorphism of the DRD2 gene (Hirvonen et al., 2004, 2005). CC-allele carriers show reduced binding potential to striatal $D_2$ receptors (Hirvonen et al., 2004, 2005) and impaired responding in the WCST (Rodriguez-Jimenez et al., 2006). In addition, CC-allele carriers are reduced in their ability to use negative feedback in a probabilistic reinforcement learning task (Frank et al., 2007). These concurrent findings suggest that reduced availability of $D_2$ receptors is associated with impaired cognitive flexibility, resulting from an inability to use negative feedback to adapt behavior.

### DA TRANSPORTER AND METABOLIZING ENZYMES

The DA transporter (DAT) regulates re-uptake of DA from the synaptic cleft in striatal regions, whereas its influence in the PFC is

less pronounced (Sesack et al., 1998). Using a task-switching protocol based on the WCST, Garcia-Garcia et al. (2010) observed impaired performance and electrophysiological differences in 9-repeat allele carriers compared to 10-repeat allele carriers of the DAT gene. During task-switching, manipulation of reward anticipation affects performance and striatal activity depending on DAT genotype, suggesting striatal DA levels mediate the influence of motivational effects on cognitive flexibility (Aarts et al., 2010). However, considering that it is unclear how this polymorphism relates to DAT expression *in vivo* [Heinz et al., 2000; Martinez et al., 2001; van Dyck et al., 2005; van de Giessen et al., 2009; meta-analysis by Costa et al., 2011], these results should be interpreted with caution.

The polymorphism that has received most attention relating DAergic gene function to executive functioning is the Valine (Val)/Methione (Met) polymorphism at codon 158 of the Catechol-*O*-methyltranserase (COMT) gene (Lotta et al., 1995). Activity of COMT is thought to be lower in homozygote Met allele carriers compared to homozygote Val carriers, presumably resulting in higher prefrontal DA levels in Met homozygotes (Lotta et al., 1995; Chen et al., 2004; Meyer-Lindenberg et al., 2005), although striatal DA levels may also be altered (Akil et al., 2003). Most studies investigating the association between the COMT Val/Met polymorphism and cognitive flexibility used perseverative responding or perseverative errors in the WCST as a measure of flexible behavior. Results have not been consistent: although an initial meta-analysis (Barnett et al., 2007) reported a small effect of COMT genotype on performance in the WCST, with reduced perseverative errors for the Met homozygotes, a second meta-analysis could not confirm an association between COMT genotype and perseverative responding on the WCST and several other cognitive measures, suggesting that the COMT polymorphism does not consistently relate to cognitive functioning (Barnett et al., 2008). It has been suggested that the variety of cognitive functions contributing to WCST performance complicate attribution of impaired performance to deficits in cognitive flexibility or deficits in cognitive stability (Bilder et al., 2004). Other test measures of cognitive flexibility might be more sensitive and more selective indicators of alterations in this function.

Despite the inconsistent effects of COMT genotype on perseverative errors in the WCST, the COMT Val/Met genotype is associated with differential activation patterns in the PFC during other cognitive paradigms (Mier et al., 2010). Therefore, it is interesting to relate COMT genotype to neural activation during other tasks that measure separate aspects of cognitive flexibility more specifically, to see whether this genotype influences neural activation in these tasks. Indeed, when (Krugel et al., 2009) studied the influence of COMT gene polymorphisms on performance and neural activity during probabilistic reversal learning, Val homozygotes performed better than Met homozygotes and showed increased striatal BOLD responses during prediction errors. In addition, higher connectivity between frontal and ventral striatal regions could be related to learning rate in Val homozygotes (Krugel et al., 2009). Interestingly, these findings suggest that striatal activity reflecting prediction errors might be modulated by DA levels in the PFC. However, during acquisition of probabilistic reinforcement learning, Val homozygotes show reduced switching of responses following negative outcomes on a trial-by-trial basis (Frank et al., 2007). This suggests that striatal DA function may be differentially regulated by DA levels in the PFC during response acquisition or adaptation of an existing response. In addition to a behavioral advantage during reversal learning, Val homozygotes also have smaller switch costs on a task switching paradigm when trials have short intervals (Colzato et al., 2010). Together these findings indicate a behavioral advantage on both reversal learning and task switching paradigms for Val homozygotes, suggesting that lower baseline levels of prefrontal DA may benefit cognitive flexibility in humans.

### SUMMARY AND CONCLUSION

A substantial amount of studies investigating the influence of genes mediating DA function on cognitive flexibility have limited analysis to a task that likely measures several complex cognitive functions, i.e., the WCST (Friedman et al., 2008). A more promising approach may be to study the effect of DA related genes on well-defined operationalizations of cognitive flexibility, such as initial discrimination learning, reversal learning, attentional set-shifting or task switching. A confound in the study of cognitive effects of genetic polymorphisms is that the effect of a polymorphism on DA transmission or even on gene expression is often not known. This hampers translational approaches, in which effects of increased or decreased expression and/or DA transmission might be studied in a controlled and reproducible manner.

To summarize, the studies reviewed above suggest an association between polymorphisms regulating DA function and cognitive flexibility. Reduced availability of $D_2$ receptors, presumably affecting striatal DA activity, impairs the use of negative feedback and the maintenance of a new response during reversal learning and set-shifting (in the WCST), whereas increased availability of $D_2$ receptors impairs task switching, suggesting different involvement of $D_2$ receptors in these tasks. Striatal $D_1$ signaling, mediated by DARPP-32 function, also contributes to cognitive functioning, although this has not yet been verified using specific measures of cognitive flexibility. Presumed lower levels of prefrontal DA, mediated by COMT-genotype appear to facilitate behavioral adaptation in both reversal learning and task-switching paradigms (see **Table 3**).

To conclude, considering that the genetic underpinnings of complex cognitive functions are likely to be polygenic and not limited to DA, studying additive genetic effects of DA related genes on cognitive flexibility as well as the study of interactions between DA related genes and other genes regulating frontostriatal function could provide a better understanding of the genetic basis of cognitive flexibility (Frank and Fossella, 2011).

## EFFECTS OF GENETIC MANIPULATIONS IN DA RELATED GENES ON COGNITIVE FLEXIBILITY IN ANIMALS

The use of genetically modified animals provides an invaluable tool to study the role of DA related genes in cognitive flexibility. Selectively targeted mutations on a known genetic background can elucidate the genetic and neurobiological basis of complex behavior.

**Table 3 | Effects of polymorphisms in dopamine related genes on cognitive flexibility in human subjects.**

| Paradigm | Gene | | Presumed DA effect | Performance | References |
|---|---|---|---|---|---|
| **REVERSAL** | | | | | |
| | $D_2$ | A1 | ↓ $D_2$ binding striatum | ↓ reversal learning | Jocham et al., 2009 |
| | COMT | Val/Val | ↓ COMT activity PFC | ↑ reversal learning | Krugel et al., 2009 |
| **TASK SWITCH** | | | | | |
| | $D_2$ | Non-A1 | ↑ $D_2$ binding striatum | ↓increased switch cost | Stelzel et al., 2010 |
| | DAT | 9-repeat | Striatum | ↓increased RT cue switch/ = task switch<br>↑ task switch high rewarded trials | Garcia-Garcia et al., 2010<br>Aarts et al., 2010 |
| | COMT | Val/Val | ↓ COMT activity PFC | ↑ reduced switch cost | Colzato et al., 2010 |
| **WCST** | | | | | |
| | $D_2$ | C957T – CC | ↓ $D_2$ binding striatum | ↓ WCST categories completed, perseveration | Rodriguez-Jimenez et al., 2006 |
| | COMT | Val/Val | ↓ COMT activity PFC | = | Barnett et al., 2008 |
| | DARPP-32 | Haplotype | Striatum | ↑ WCST performance | Meyer-Lindenberg et al., 2007 |

*= no effect, ↑ increased performance, ↓ decreased performance.*

*PFC, prefrontal cortex; RT, reaction time; WCST, Wisconsin Card Sorting Task.*

### DA DEFICIENCY

An example of an advanced genetic approach is selective reinstatement of DA signaling in ventral or dorsal striatum of DA-deficient mice (Darvas and Palmiter, 2011). Restoring DA signaling specifically to either dorsal or ventral striatum supports acquisition and reversal of a turn-based escape strategy in a water maze (Darvas and Palmiter, 2011). However, the ability to switch from one escape strategy to another (strategy set-shift) is impaired when DA signaling is limited to the ventral striatum, suggesting DA neurotransmission in the dorsal striatum is required for strategy set-shifting, whereas DA in either ventral or dorsal striatum is sufficient to support reversal learning (Darvas and Palmiter, 2011). It should be noted, however, that the translational value of the tasks used is not established.

### DA RECEPTORS AND INTRACELLULAR SIGNALING

#### $D_1$

Mice lacking functional $D_1$ receptors show attenuated operant responding for reward (El-Ghundi et al., 2003). They show a general deficit in reinforcement learning, impaired motivation to work for a reward, are slow to discriminate between a reinforced and non-reinforced lever and are impaired in reversal learning, during which they maintain responding to both levers. Heterozygote mice are also impaired on reversals, although not as severely (El-Ghundi et al., 2003). The observed general deficits in motivation and reinforcement learning in $D_1$-knockout mice, however, prevent the drawing of conclusions about the contribution of $D_1$ receptors to cognitive flexibility.

Activation of $D_1$ receptors modulates striatal function through phosphorylation of DARPP (Walaas and Greengard, 1984). Next to a minor reduction in performance during discrimination learning, DARPP-32 knockout mice show a pronounced deficit in reversal learning. Although knockout mice eventually were able to switch responding to the newly rewarded side, it took them significantly more sessions to do so (Heyser et al., 2000). This is indirect evidence that $D_1$ receptor activation is needed for reversal learning.

#### $D_2$

Genetic manipulations of $D_2$ receptors also affect performance on cognitive flexibility tasks. Female mice with a complete knock-out of functional $D_2$ receptors make more errors during odor discrimination and reversal learning whereas male $D_2$-knockouts are impaired during reversal learning only; both sexes show perseveration to the previously rewarded side (Kruzich and Grandy, 2004; Kruzich et al., 2006). This was confirmed by De Steno and Schmauss (2009), who also showed a similar impairment with chronic treatment with the $D_2$ antagonist haloperidol. Glickstein et al. (2005) observed a deficit of male $D_2$-knockouts during compound discrimination, but not reversal, whereas $D_3$ receptor knockouts showed increased performance during the reversal. The differences in behavioral performance were paralleled by opposite prefrontal activation patterns following the task sequence: activity dependent gene expression in the MPFC is increased for $D_3$ mutants and decreased for $D_2$ mutants (Glickstein et al., 2005; De Steno and Schmauss, 2009). Interestingly, knockout of neither $D_2$ nor $D_3$ receptors

affects performance on intra- or extradimensional set-shifts (De Steno and Schmauss, 2009), suggesting differential contribution of $D_2/D_3$receptors to the regulation of reversal learning or set-shifting.

Selective overexpression of $D_2$ receptors in the striatum does not affect learning of a discrimination, a reversal or an intra- or extradimensional set-shift. Response latencies were longer during reversal trials only, suggesting the animals had some difficulties adapting established responses (Kellendonk et al., 2006). Interestingly, these mice also show physiological changes in the medial PFC where DA turnover was decreased and activation of $D_1$ receptors increased (Kellendonk et al., 2006).

### METABOLIZING ENZYMES

Overexpression of the human COMT-Val polymorphism in mice increases COMT enzyme activity (suggesting lower prefrontal extracellular DA) and induces specific deficits in cognitive flexibility. Although discrimination and reversal learning are not affected, these mice make more errors and need more time to complete an extra-dimensional set-shift (Papaleo et al., 2008). In contrast to behavioral impairments observed after increased COMT enzyme activity, pharmacological inhibition of COMT can improve performance (Tunbridge et al., 2004).

### SUMMARY AND CONCLUSION

The studies using selective DA-reinstatement in DA-deficient mice show that higher order flexibility [strategy shifting (Wise et al., 1996)] is associated with dorsal striatal DA, whereas lower order flexibility (reversal learning) may be supported by DA in all striatal areas. Similarly, human studies suggest influence of DA genotype on activity in ventral striatal regions or increased connectivity between PFC and ventral striatum during reversal learning and in dorsal striatal regions during task switching.

The $D_1$ receptor is involved in cognitive flexibility, although this is overshadowed by a general impairment in goal-directed behavior in full knock-outs. DARPP-32 expression (reflecting $D_1$ activity) is associated with cognitive performance in both humans and animals.

The findings described above, and the observation that performance of reversal learning in mice covaries with $D_2$ receptor levels in the ventral midbrain (Laughlin et al., 2011), indicate the importance of $D_2$ receptors for flexible behavior, specifically in a situation where response-reward contingencies are reversed (see **Table 4**). This compares to the influence of polymorphisms in the $D_2$ receptor gene on the ability to learn from negative feedback in human subjects.

Expressing the human COMT-Val polymorphism (increasing COMT-activity and presumably decreasing extracellular prefrontal DA) in mice impairs extra dimensional set-shift. This concurs with the improved set-shifting performance after COMT-inhibition in rats. However, presence of the Val-polymorphism in humans has been associated with a behavioral advantage during reversal learning and task-switching suggesting that confirmation of these studies is needed before we can draw conclusions.

Caution should be exerted when interpreting results from animals in which a receptor is completely knocked out as compensatory mechanisms (such as increased neurotransmitter levels) during development may contribute to the observed deficits. Also, in the case of complete knock-outs it is not possible to locate the neurobiological substrate of the impairment as the knock-out is present throughout the brain. Finally, mice with intermediate expression of specific receptors (heterozygotes) are

**Table 4 | Effects of genetic manipulations to dopamine related genes on cognitive flexibility in animals.**

| Paradigm | Gene | | Performance | References |
|---|---|---|---|---|
| **DISCRIMINATION** | | | | |
| | $D_1$ KO | | ↓ more errors | El-Ghundi et al., 2003 |
| | $D_2$ KO | Female | ↓ more errors | Kruzich and Grandy, 2004 |
| | $D_2$ KO | Male | = | Kruzich et al., 2006 |
| | COMT-Val overexpression | | = | Papaleo et al., 2008 |
| **REVERSAL** | | | | |
| | $D_1$ KO | | ↓ more errors | El-Ghundi et al., 2003 |
| | $D_2$ KO | Male + female | ↓ more errors ↓ increased RT reversal phase set-shift = reversal phase set-shift | Kruzich and Grandy, 2004 Kruzich et al., 2006 |
| | DARPP-32 KO | | ↓ more errors | Heyser et al., 2000 |
| **ATTENTIONAL SET-SHIFT** | | | | |
| | $D_2$ KO | | = | De Steno and Schmauss, 2009 Glickstein et al., 2005 |
| | $D_2$ overexpression | Striatum only | = | Kellendonk et al., 2006 |
| | COMT-Val overexpression | | ↓ impaired EDS | Papaleo et al., 2008 |

*= no effect, ↑ increased performance, ↓ decreased performance. KO, knock out; RT, reaction time; EDS, extradimensional set-shift.*

useful for studying gene-dosage effects on behavior, which could be particularly relevant when compared to differences in receptor expression levels observed in humans.

## OCD

OCD is a psychiatric disorder that is characterized by recurrent intrusive, unwanted thoughts (obsessions) that are often accompanied by repetitive ritualistic behaviors (compulsions). Although the precise neurobiological substrates underlying OCD symptoms are not known, structural and functional imaging studies show alterations in frontal and orbitofrontal cortices and basal ganglia in OCD patients (Pujol et al., 2004; Menzies et al., 2008a,b; van den Heuvel et al., 2009; Rotge et al., 2010). Symptom severity correlates with increased functional connectivity between OFC and striatal regions (Harrison et al., 2009), which normalizes after treatment (Figee et al., 2013).

The repeated performance of ritual-like action sequences has led to the hypothesis that decreased cognitive flexibility or increased habitual behavior (Gillan et al., 2011) is a major underlying factor of OCD and could be a potential endophenotype for the disorder (Robbins et al., 2012). This might be an attractive suggestion considering that associated circuits and neurotransmitters related to these processes are (partly) known. Indications for abnormal flexibility have been described in OCD patients (Chamberlain et al., 2006; Gu et al., 2008) and there is evidence for altered DA signaling (Denys et al., 2004a,b; Moresco et al., 2007; Perani et al., 2008). Therefore, an important question is how DA contributes to this disorder. In the next sections, we will describe studies reporting alterations in the DA system in OCD patients as well as studies investigating cognitive flexibility in OCD.

### DA ALTERATIONS IN OCD

Although there is strong evidence that serotonin plays a role in the treatment of OCD (van Dijk et al., 2010), it is clear that OCD pathophysiology also involves alterations in fronto-striatal circuitry and its neuromodulation by DA. Indirect evidence comes from clinical observations that administration of DA antagonists can improve symptoms in OCD-patients that do not respond to SSRI's alone [(McDougle et al., 2000; Dougherty et al., 2004); see Denys et al. (2004b) for review]. In animals, administration of drugs acting on DAergic receptors and genetic manipulations of DA receptors induces compulsive, stereotypic behaviors similar to the repetitive behaviors of OCD patients (Szechtman et al., 1998; Campbell et al., 1999; Joel and Doljansky, 2003; Denys et al., 2004b; Sesia et al., 2013).

Importantly, direct evidence indicating altered DA signaling in OCD patients is also available. Kim et al. (2003) observed a higher density of the DA transporter (DAT) in the right basal ganglia that normalized after SSRI treatment (Kim et al., 2007). However, these findings were not consistently replicated (Nikolaus et al., 2010): van der Wee et al. (2004) also showed higher binding ratios using OCD patients without co-morbid disorders, but Hesse et al. (2005) observed reduced striatal DAT binding and Pogarell et al. (2003) did not observe differences in DAT availability between OCD patients and healthy controls. The latter authors also reported increased instead of decreased DAT binding after SSRI's.

OCD-patients show reduced binding to $D_1$ receptors in caudate nucleus and putamen (Olver et al., 2009) and in anterior cingulate cortex (Olver et al., 2010), although reduced binding does not correlate with symptom severity.

Studies investigating binding to striatal $D_2$ receptors in OCD patients present a more consistent picture. The original finding by Denys et al. (2004a) of reduced binding to $D_2$ receptors in the caudate nucleus of OCD patients was replicated by others (Perani et al., 2008; Schneier et al., 2008; Denys et al., 2013). In medication-naïve OCD patients, repeated administration of an SSRI increased binding to striatal $D_2$ receptors, illustrating that regulation of DA release can be modulated by 5-HT (Moresco et al., 2007).

Taken together, the studies mentioned here described reduced binding to DA receptors in OCD patients, mainly in, but not limited to striatal regions. The most replicated finding is reduced availability of $D_2$ receptors in striatal regions. It has been hypothesized that reduced availability of DA receptors in OCD patients could be the result of increased DA release in the striatum (Denys et al., 2004a). However, the observed changes in the DA system do not correlate with symptom severity or duration of illness and it is possible that the DAergic alterations are secondary to diminished serotonergic tone.

### COGNITIVE FLEXIBILITY IN OCD

Although the repetitive execution of behavioral patterns that is often observed in OCD patients could be defined as inflexible or perseverative behavior, the question is whether this translates to impaired performance on measurements of cognitive flexibility that are currently used in tests of executive functioning.

Findings using the WCST have been contradictory, with some studies observing impaired performance in OCD patients (Lucey et al., 1997; Lacerda et al., 2003; Bohne et al., 2005; Lawrence et al., 2006; Bucci et al., 2007; de Geus et al., 2007; Cavedini et al., 2010), whilst others do not (Gambini et al., 1993; Abbruzzese et al., 1995, 1997; Cavedini et al., 1998; Moritz et al., 2002; Fenger et al., 2005; Henry, 2006). The former studies often describe an increase in the number of perseverative errors. The observation that deficits in flexibility may persist after remission or use of medication and that unaffected family members also show reduced flexibility, suggests that these deficits are trait-like and independent of OCD-symptomatology (Bannon et al., 2006; Cavedini et al., 2010), supporting the hypothesis that inflexible, rigid and habit-like behavior is an endophenotype in OCD.

#### *Reversal learning*

Alterations in recruitment of fronto-striatal circuitry in the absence of behavioral impairments have been observed in both OCD patients and their unaffected first-degree relatives during reversal learning (Chamberlain et al., 2008). Remijnse et al. (2006) observed attenuated responsiveness of OFC and striatal regions during reward and affective switching in OCD patients with and without comorbidities. In these studies, as well as in others (Valerius et al., 2008; Ersche et al., 2011) no clear evidence for behavioral impairments during task performance was obtained, although OCD patients do show a somewhat slowed response pattern, suggesting they may require more processing time when

faced with altered response-reward contingencies. Altered recruitment of fronto-striatal circuitry during these tests suggests that even though overt behavioral performance (i.e., reaction times, number of errors, number of trials required to reach criterion) may not be impaired, the processing of cognitive information is altered in OCD patients during reversal learning.

#### *Attentional set-shifting*

Performance on tasks that require shifting between different stimulus dimensions does appear to be affected in OCD patients. Behavioral impairments have been observed in OCD patients and unaffected first-degree relatives in an attentional set-shifting task (Veale et al., 1996; Fenger et al., 2005; Watkins et al., 2005; Chamberlain et al., 2006, 2007) but see (Purcell et al., 1998a,b), with some reporting reduced performance on extra-dimensional set-shifts (Veale et al., 1996; Watkins et al., 2005; Chamberlain et al., 2006, 2007) and others on intra-dimensional set-shifts (Veale et al., 1996; Fenger et al., 2005). Response to SSRI-treatment was found to be related to set-shifting ability (Fontenelle et al., 2001).

#### *Task switching*

Increased switch costs (decreased accuracy or increased response times) have been observed in OCD patients during performance of task switching paradigms (Moritz et al., 2004; Gu et al., 2008; Page et al., 2009). Gu et al. (2008) found an increase in the number of errors made during task-switching trials in OCD patients, but others report slowed responding (Moritz et al., 2004; Remijnse et al., 2013) or no effect (Page et al., 2009). However, when task switching is combined with functional imaging, activity in the dorsal fronto-striatal circuit is consistently found to differ between OCD patients and healthy controls. Whereas activation of the dorsal fronto-striatal circuit is observed in healthy controls during task-switching trials, this is not the case in OCD patients (Gu et al., 2008; Page et al., 2009; Remijnse et al., 2013).

### SUMMARY AND CONCLUSION

Several problems arise when interpreting the deficits of OCD patients on cognitive flexibility and the mixed outcomes of the studies investigating these deficits. Next to the influence of medication and the need for careful matching of patient and control groups, the high comorbidity with other psychiatric disorders, in particular depression is an important confounding factor. Although the use of subject groups with OCD as the only clinical diagnosis could be thought of as misrepresentative for the population of OCD patients *because* comorbidity is so common (Olley et al., 2007), the use of well-defined clinical populations in studies combining neuropsychological testing with measurements of brain activity in particular, could contribute to the knowledge about distorted recruitment of frontostriatal circuitry in cognitive flexibility.

As far as we know, studies directly linking measurements of cognitive flexibility to alterations in DA signaling have not been performed in OCD patients. The most consistent alteration in the DA system is changed DA receptor binding, mostly in striatal regions. Replication of these findings, especially of both $D_1$ and $D_2$ receptor binding, in different OCD samples would enhance our understanding of the contribution of DA to OCD. For performance on cognitive flexibility tasks, behavioral performance on lower order cognitive flexibility (reversal learning) is not altered, whilst OCD patients may be impaired on higher order flexibility tasks (attentional set-shift and task switching). Irrespective of the presence of behavioral impairments, activity and connectivity in neural circuits regulating flexible behavior (OFC-ventral striatum for reversal learning, PFC-dorsal striatum for task-switching) are altered in OCD patients during task execution. Considering the modulatory effect of DA in these neural circuits, it is possible that altered striatal DA contributes to different activity in these circuits during task performance.

## OCD ANIMAL MODELS: DOPAMINE AND COGNITIVE FLEXIBILITY

Animal models of psychiatric disorders cannot reflect all aspects of the disease (Nestler and Hyman, 2010). In line with this, OCD models that show a combination of the critical face, predictive and construct validities (Korff and Harvey, 2006; Wang et al., 2009; Fineberg et al., 2011; Albelda and Joel, 2012b) predominantly mirror the compulsive acts of OCD patients. This applies for models based on spontaneous behavior [ethological models, e.g., compulsive dogs, (Vermeire et al., 2012)], behavioral models [e.g., compulsive lever-pressing during signal attenuation in rats (Joel, 2006)], pharmacological models [e.g., quinpirole-induced checking in rats (Szechtman et al., 1998)], and transgenic models [e.g., compulsive grooming in Sapap3-mutant mice, (Welch et al., 2007)]. Compulsive acts are behaviorally and conceptually not always clearly differentiated from simple repetitive behaviors. Repetitive, stereotyped, perseverative, rigid and habitual behavior have been grouped together into (overlapping) clusters of compulsive-like behavior [(Langen et al., 2011; Ting and Feng, 2011; Robbins et al., 2012); for a critical discussion of the distinction between stereotypies and compulsions, see (Lewis et al., 2007)]. These clusters are relevant not only for OCD, but also for other psychiatric disorders and may share a relative DAergic hyperactivity in the basal ganglia (Pitman, 1989). Two recent studies highlight the direct involvement of specific projections from OFC to ventromedial striatum in the regulation of compulsive-like, repetitive behavior in normal mice (Ahmari et al., 2013) and compulsively grooming Sapap-3 mice (Burguiere et al., 2013).

Stereotyped repetitive behavior, in particular, is strongly linked to DA mechanisms (Randrup and Munkvad, 1975; Ridley, 1994). Next to the quinpirole-model (repeated administration of a $D_{2/3}$-selective agonist), the DAT-knockdown mouse that shows stronger and more rigid grooming behavior, has been proposed as an OCD-model based on DA hyperactivation (Berridge et al., 2005). Another model of increased DA-related neuronal activity is the D1CT transgenic mouse, showing repetition of all normal behaviors (Campbell et al., 1999).Most other validated OCD-models also show involvement of DA mechanisms in their compulsive behavior (Joel and Doljansky, 2003; Presti et al., 2003; Albelda and Joel, 2012a; Moreno and Flores, 2012; Vermeire et al., 2012; Sesia et al., 2013), although DA mechanisms were not tested in compulsively grooming transgenic mouse models (Welch et al., 2007; Shmelkov et al., 2010).

The relationship between repetitive behavior and cognitive flexibility as probed in tasks using translationally valid constructs of reversal learning, attentional set-shifting or task switching has received only limited attention. In deer mice, stereotyped jumping was correlated with the number of incorrect responses in a reversal of escape-learning in a water-filled T-maze (Tanimura et al., 2008). BTBR T+ tf.J mice, showing compulsive grooming and increased marble burying, show impaired probabilistic reversal learning (Amodeo et al., 2012). A task probing recurrent perseveration (two-choice task where continuous switching provides the optimal strategy) showed a correlation between stereotyped behavior and recurrent perseveration in farmed minks, but not in ICR CD-1 mice (Gross et al., 2011). Finally, rats compulsively drinking in the schedule-induced polydipsia model displayed increased perseveration during extinction of the 5-choice serial reaction time task and perseveration during extinction of other operant procedures was reported in bank voles (Garner and Mason, 2002) and caged bears (Vickery and Mason, 2005).

However, if we focus on reversal learning, attentional set-shifting or task switching there are no studies available that show task impairments in OCD animal models, let alone impairments related to DA mechanisms. The only possible exception is stereotyped behavior in deer mice, which correlated to the number of incorrect responses during reversal learning and decreased after striatal administration of a $D_1$-selective antagonist (Presti et al., 2003; Tanimura et al., 2008), though the relation between reversal learning and DA was not directly investigated.

In conclusion, a possible relation between compulsive behavior and cognitive flexibility, including the possibility that DA mechanisms might play a role in this, did not receive much attention up to now. One can understand that the introduction of translational valid paradigms for cognitive flexibility in exotic species such as bank voles, mink or bears is not an easy task. But using behavioral testing in reversal learning, attentional set-shifting or task switching in rodent OCD-models should be a priority for researchers who want to study the neurobiological underpinnings of OCD.

## CONCLUSION

Evidence for a role of DA in the control of cognitive flexibility comes from a range of human and animal studies that have been reviewed above. This overview indicates that DA is involved in different facets of cognitive flexibility, including reversal learning, set-shifting and task-switching. Moreover, DA in both cortical and subcortical parts of the corticostriatal circuits seem to be involved in the regulation of these different aspects of cognitive flexibility. The idea that DA facilitates flexibility or switching behavior can be traced back to older studies that used different behavioral paradigms than the studies reviewed here. For example, a role for DA in switching strategies in a swim test was suggested by Cools (1980) and van den Bos and Cools (1989), while the importance of DA in switching (increasing the probability that another behavioral output is chosen) was advocated by Oades (1985).

However, the general picture arises that although DA may facilitate cognitive flexibility, it is not required. Following a variety of manipulations to the DA system the ability to successfully shift behavior following changes in reinforcer contingencies is impaired but not completely absent (in rodents, non-human primates and humans). How does the supportive role of DA in cognitive flexibility (i.e., behavioral adaptation to a change in conditions) compare to its role in initial learning about rewards? The question whether DA is necessary for learning has been addressed by studying acquisition of learning in DA deficient mice—the conclusion was that loss of DA may impair, but does not inhibit reward learning (Berridge, 2005; Robinson et al., 2005; Palmiter, 2008; Darvas and Palmiter, 2010). Animals may become less motivated, but were still able to learn cue-reward associations. Disruption of phasic DA activity by deletion of NMDA-receptors from DA neurons again showed that learning may be retarded, but not inhibited (Zweifel et al., 2009). A recent study using an optogenetics approach showed that phasic DA stimulation may drive associative learning or impair extinction learning, suggesting a causal role for DA (Steinberg et al., 2013). However, DA stimulation could not maintain the original behavior, so that other processes are probably involved as well. During performance of cognitive flexibility tasks, a number of cognitive processes act simultaneously and DA may be especially important to switch behavior rapidly. The contribution of DA to new learning therefore appears to be facilitatory rather than a prerequisite and the supportive role of DA appears to be present both in initial learning and adaptation of learning.

Both pharmacological and genetic studies in human subjects and animals point to a role for $D_2$ receptors in the regulation of cognitive flexibility. However, the regulation is not limited to $D_2$ receptor activity: $D_1$ and $D_2$ receptors both contribute and appear to be cooperatively involved in discrimination learning and the flexible adaptation of behavior. One could argue that successful behavioral switching requires three processes that may partly occur in parallel: extinction of the response that is no longer rewarded, behavioral switch to the newly rewarded side and response maintenance. A complication in delineating the contribution of DA to either process is that these processes occur simultaneously during behavioral adaptation. DA signaling through $D_1$ receptors may not be essential for switching behavior *per se*, but animal studies suggest that activation of $D_1$ receptors contributes to the acquisition and maintenance of a new response, also when acquisition follows a reversal. In contrast, inactivation of $D_2$ receptors may allow switching of behavior patterns. The contributions of $D_1$ versus $D_2$ receptors in the regulation of reward learning and behavior switching has been related to involvement of the direct and indirect pathway of the basal ganglia in these processes, and several models have been put forward to describe the possible components involved in regulating this behavior (Frank and Claus, 2006; Hong and Hikosaka, 2011). In general these models assume the presence of $D_1$ receptors in the direct pathway (direct projections from striatal medium spiny neurons (MSN) to the substantia nigra) and expression of $D_2$ receptors on MSN's of the indirect pathway (projections from MSN to substantia nigra via the globus pallidus) (Deng et al., 2006). Because binding affinity differs for $D_1$ and $D_2$ receptors (Richfield et al., 1989), fluctuations in DA levels during different stages of discrimination and reversal learning may result in different activation of $D_1$ (direct pathway) or $D_2$ (indirect pathway)

expressing neurons. When a reward is presented unexpectedly, or when a stimulus that predicts reward is presented, a transient increase in DA release occupies low affinity $D_1$ receptors and activates the direct pathway, allowing facilitation of response execution and prompting reward-related learning. Switching of behavioral patterns on the other hand might require reduced occupancy of high affinity $D_2$ receptors. Omission of an expected reward following altered reinforcer contingencies results in transient reductions in striatal DA levels and diminished inhibition of the indirect pathway by $D_2$ receptors, resulting in inhibition of the previously successful response. Both facilitation of behavioral adaptation by deactivation of striatal $D_2$ receptors and facilitation of the acquisition of the "new" behavioral response by striatal $D_1$ activation suggests the importance of phasic fluctuations in striatal DA levels during execution of cognitive flexibility. This may be illustrated for the $D_2$-mediated response: both continuously higher and lower tonic $D_2$ activation could impair detection of the transient reduction of DA. As tonic DA may be related to general synaptic factors such as synthesis capacity, uptake activity and metabolic efficiency, all these factors may influence flexible responding through $D_2$ receptor dependent transmission. However, it is difficult to separate tonic from phasic DA signaling with most manipulations used. Tonic prefrontal DA (Seamans and Yang, 2004) probably contributes as well. In addition, activation of $D_1/D_2$ receptors in prefrontal regions may differ from the activation in striatal regions. It has been suggested, for example, that $D_2$ stimulation in prefrontal regions may facilitate flexible behavior (Durstewitz and Seamans, 2008) whereas in striatal regions, deactivation of $D_2$ receptors is suggested to facilitate cognitive flexibility (Yawata et al., 2012). The combined study of genetic effects on behavioral performance and patterns of neural activation also suggests that although DA genotype may primarily affect expression of DA related genes in either striatal or prefrontal areas, functional effects of DA genotype are not limited to either region but are observed throughout the frontostriatal circuit. Genetic and imaging studies suggest that DA in ventral regions of the striatum (or connections between PFC and ventral striatum) contributes to reversal learning (lower order complexity), whereas DA in dorsal regions may be more important for attentional set-shifting and task switching (higher order complexity). However, animal studies have also described effects of DA in the NAC on attentional set-shifts and animals that only have DA signaling in dorsal striatal regions are able to learn a reversal. In addition, in human imaging studies it is not always clear if activation is limited to either ventral or dorsal striatum because analysis was limited to that particular striatal region or because the other striatal region was not activated. Therefore, it appears to be more likely that the relative activation of $D_1/D_2$ in prefrontal and striatal regions as well as the interaction with other neuromodulators (5-HT, NA) determines the control of cognitive flexibility. Considering the complexity of DA modulation in frontostriatal circuitry (Seamans and Yang, 2004), it may not be surprising DA modulation in neither frontal nor striatal regions that exclusively determines behavioral performance on tasks of cognitive flexibility.

So how do these findings relate to altered cognitive flexibility in OCD patients? If cognitive flexibility can indeed be used as an endophenotype for OCD, do the alterations in DA signaling that have been observed in OCD patients comply with the proposed role for DA in cognitive flexibility? The most replicated alteration in the DA system of OCD patients is reduced binding to $D_2$ receptors in the striatum. A questions remains, how reduced $D_2$ receptor binding relates to DAergic activity *in vivo*. A reduction in binding potential to $D_2$ receptors may result from increased striatal DA levels or altered availability of $D_2$ receptors. In both cases, reduced flexibility could be expected. However, behavioral performance (i.e., accuracy) on reversal learning tasks is not impaired in OCD patients. On reversal learning tasks, if any behavioral effect is found, it is a slowing of response times rather than an effect on the amount of errors that are made. Differences in accuracy have been observed in attentional set-shifting and task switching paradigms. It is possible that reversal learning may be a paradigm that is too simple for gross behavioral abnormalities to be observed in OCD patients. Increased reactions times on flexibility tasks, however, do suggest altered cognitive processing in OCD patients during cognitive flexibility and the measurement of reaction times should therefore be included in studies investigating differences in cognitive flexibility between healthy controls and OCD patients. The altered recruitment of frontostriatal circuitry during the execution of reversal learning as well as task switching is another indication for altered cognitive processing in OCD patients. Altered DA signaling is a potential contributor to changes in frontostriatal activity when performing cognitive tasks. Altered activity in the frontostriatal circuit (OFC-ventral striatum) during reversal learning, as observed in OCD patients is also found in subjects with polymorphisms in the $D_2$ gene that result in reduced binding to $D_2$ receptors. Most likely, however, abnormalities in prefrontal regions and 5-HT modulation in OCD patients also contribute.

An important step in investigating the possibility of altered cognitive processing in cognitive flexibility tasks as an endophenotype for OCD would be the replication of studies using cognitive flexibility tasks in OCD patients with the use of strictly defined patient and control groups. Considering that altered neural correlates of OCD could be symptom dimension-specific (van den Heuvel et al., 2009), separate study of the different symptom dimensions contributes to the identification of possible endophenotypes. Preferably, these studies combine behavioral testing with measurements of brain activity and/or DA activity to further investigate the neurobiological basis of altered cognitive processing during cognitive flexibility tests in OCD patients.

## REFERENCES

Aarts, E., Roelofs, A., Franke, B., Rijpkema, M., Fernandez, G., Helmich, R. C., et al. (2010). Striatal dopamine mediates the interface between motivational and cognitive control in humans: evidence from genetic imaging. *Neuropsychopharmacology* 35, 1943–1951. doi: 10.1038/npp.2010.68

Abbruzzese, M., Ferri, S., and Scarone, S. (1995). Wisconsin Card Sorting Test performance in obsessive-compulsive disorder: no evidence for involvement of dorsolateral prefrontal cortex. *Psychiatry Res.* 58, 37–43. doi: 10.1016/0165-1781(95)02670-R

Abbruzzese, M., Ferri, S., and Scarone, S. (1997). The selective breakdown of frontal functions in patients with obsessive-compulsive disorder and in patients with schizophrenia: a double dissociation experimental finding. *Neuropsychologia* 35, 907–912. doi: 10.1016/S0028-3932(96)00095-4

Ahmari, S. E., Spellman, T., Douglass, N. L., Kheirbek, M. A., Simpson, H. B., Deisseroth, K., et al. (2013). Repeated cortico-striatal stimulation generates persistent OCD-like behavior. *Science* 340, 1234–1239. doi: 10.1126/science.1234733

Akil, M., Kolachana, B. S., Rothmond, D. A., Hyde, T. M., Weinberger, D. R., and Kleinman, J. E. (2003). Catechol-O-methyltransferase genotype and dopamine regulation in the human brain. *J. Neurosci.* 23, 2008–2013.

Albelda, N., and Joel, D. (2012a). Animal models of obsessive-compulsive disorder: exploring pharmacology and neural substrates. *Neurosci. Biobehav. Rev.* 36, 47–63. doi: 10.1016/j.neubiorev.2011.04.006

Albelda, N., and Joel, D. (2012b). Current animal models of obsessive compulsive disorder: an update. *Neuroscience* 211, 83–106. doi: 10.1016/j.neuroscience.2011.08.070

Amodeo, D. A., Jones, J. H., Sweeney, J. A., and Ragozzino, M. E. (2012). Differences in BTBR T+ tf/J and C57BL/6J mice on probabilistic reversal learning and stereotyped behaviors. *Behav. Brain Res.* 227, 64–72. doi: 10.1016/j.bbr.2011.10.032

Arnsten, A. F., Wang, M. J., and Paspalas, C. D. (2012). Neuromodulation of thought: flexibilities and vulnerabilities in prefrontal cortical network synapses. *Neuron* 76, 223–239. doi: 10.1016/j.neuron.2012.08.038

Arnt, J. (1985). Antistereotypic effects of dopamine D-1 and D-2 antagonists after intrastriatal injection in rats. Pharmacological and regional specificity. *Naunyn. Schmiedebergs. Arch. Pharmacol.* 330, 97–104. doi: 10.1007/BF00499901

Arushanian, E. B., and Baturin, V. A. (1982). Effect of haloperidol and chlorpromazine on reversal learning of normal and striatectomized rats in a Y-maze. *Pharmacol. Biochem. Behav.* 16, 541–545. doi: 10.1016/0091-3057(82)90411-7

Bannon, S., Gonsalvez, C. J., Croft, R. J., and Boyce, P. M. (2006). Executive functions in obsessive-compulsive disorder: state or trait deficits? *Aust. N. Z. J. Psychiatry* 40, 1031–1038.

Barch, D. M., Carter, C. S., Arnsten, A., Buchanan, R. W., Cohen, J. D., Geyer, M., et al. (2009). Selecting paradigms from cognitive neuroscience for translation into use in clinical trials: proceedings of the third CNTRICS meeting. *Schizophr. Bull.* 35, 109–114. doi: 10.1093/schbul/sbn163

Barnes, J. J., Dean, A. J., Nandam, L. S., O'Connell, R. G., and Bellgrove, M. A. (2011). The molecular genetics of executive function: role of monoamine system genes. *Biol. Psychiatry* 69, e127–e143. doi: 10.1016/j.biopsych.2010.12.040

Barnett, J. H., Jones, P. B., Robbins, T. W., and Muller, U. (2007). Effects of the catechol-O-methyltransferase Val158Met polymorphism on executive function: a meta-analysis of the Wisconsin Card Sort Test in schizophrenia and healthy controls. *Mol. Psychiatry* 12, 502–509.

Barnett, J. H., Scoriels, L., and Munafo, M. R. (2008). Meta-analysis of the cognitive effects of the catechol-O-methyltransferase gene Val158/108Met polymorphism. *Biol. Psychiatry* 64, 137–144. doi: 10.1016/j.biopsych.2008.01.005

Bellebaum, C., Koch, B., Schwarz, M., and Daum, I. (2008). Focal basal ganglia lesions are associated with impairments in reward-based reversal learning. *Brain* 131, 829–841. doi: 10.1093/brain/awn011

Berridge, K. C. (2005). Espresso reward learning, hold the dopamine: theoretical comment on Robinson et al. (2005). *Behav. Neurosci.* 119, 336–341. doi: 10.1037/0735-7044.119.1.336

Berridge, K. C., Aldridge, J. W., Houchard, K. R., and Zhuang, X. (2005). Sequential super-stereotypy of an instinctive fixed action pattern in hyper-dopaminergic mutant mice: a model of obsessive compulsive disorder and Tourette's. *BMC Biol.* 3:4. doi: 10.1186/1741-7007-3-4

Bilder, R. M., Volavka, J., Lachman, H. M., and Grace, A. A. (2004). The catechol-O-methyltransferase polymorphism: relations to the tonic-phasic dopamine hypothesis and neuropsychiatric phenotypes. *Neuropsychopharmacology* 29, 1943–1961. doi: 10.1038/sj.npp.1300542

Birrell, J. M., and Brown, V. J. (2000). Medial frontal cortex mediates perceptual attentional set shifting in the rat. *J. Neurosci.* 20, 4320–4324.

Bissonette, G. B., Martins, G. J., Franz, T. M., Harper, E. S., Schoenbaum, G., and Powell, E. M. (2008). Double dissociation of the effects of medial and orbital prefrontal cortical lesions on attentional and affective shifts in mice. *J. Neurosci.* 28, 11124–11130. doi: 10.1523/JNEUROSCI.2820-08.2008

Bohne, A., Savage, C. R., Deckersbach, T., Keuthen, N. J., Jenike, M. A., Tuschen-Caffier, B., et al. (2005). Visuospatial abilities, memory, and executive functioning in trichotillomania and obsessive-compulsive disorder. *J. Clin. Exp. Neuropsychol.* 27, 385–399. doi: 10.1080/13803390490520418

Boulougouris, V., Castane, A., and Robbins, T. W. (2009). Dopamine D2/D3 receptor agonist quinpirole impairs spatial reversal learning in rats: investigation of D3 receptor involvement in persistent behavior. *Psychopharmacology (Berl.)* 202, 611–620. doi: 10.1007/s00213-008-1341-2

Boulougouris, V., Dalley, J. W., and Robbins, T. W. (2007). Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. *Behav. Brain Res.* 179, 219–228. doi: 10.1016/j.bbr.2007.02.005

Bouret, S., and Sara, S. J. (2004). Reward expectation, orientation of attention and locus coeruleus-medial frontal cortex interplay during learning. *Eur. J. Neurosci.* 20, 791–802. doi: 10.1111/j.1460-9568.2004.03526.x

Braver, T. S., Cole, M. W., and Yarkoni, T. (2010). Vive les differences! Individual variation in neural mechanisms of executive control. *Curr. Opin. Neurobiol.* 20, 242–250. doi: 10.1016/j.conb.2010.03.002

Bucci, P., Galderisi, S., Catapano, F., Di, B. R., Piegari, G., Mucci, A., et al. (2007). Neurocognitive indices of executive hypercontrol in obsessive-compulsive disorder. *Acta Psychiatr. Scand* 115, 380–387. doi: 10.1111/j.1600-0447.2006.00911.x

Burguiere, E., Monteiro, P., Feng, G., and Graybiel, A. M. (2013). Optogenetic stimulation of lateral orbitofronto-striatal pathway suppresses compulsive behaviors. *Science* 340, 1243–1246. doi: 10.1126/science.1232380

Calaminus, C., and Hauber, W. (2007). Intact discrimination reversal learning but slowed responding to reward-predictive cues after dopamine D1 and D2 receptor blockade in the nucleus accumbens of rats. *Psychopharmacology (Berl.)* 191, 551–566. doi: 10.1007/s00213-006-0532-y

Calaminus, C., and Hauber, W. (2008). Guidance of instrumental behavior under reversal conditions requires dopamine D1 and D2 receptor activation in the orbitofrontal cortex. *Neuroscience* 154, 1195–1204. doi: 10.1016/j.neuroscience.2008.04.046

Campbell, K. M., de Lecea, L., Severynse, D. M., Caron, M. G., McGrath, M. J., Sparber, S. B., et al. (1999). OCD-Like behaviors caused by a neuropotentiating transgene targeted to cortical and limbic D1+ neurons. *J. Neurosci.* 19, 5044–5053.

Castane, A., Theobald, D. E., and Robbins, T. W. (2010). Selective lesions of the dorsomedial striatum impair serial spatial reversal learning in rats. *Behav. Brain Res.* 210, 74–83. doi: 10.1016/j.bbr.2010.02.017

Cavedini, P., Ferri, S., Scarone, S., and Bellodi, L. (1998). Frontal lobe dysfunction in obsessive-compulsive disorder and major depression: a clinical-neuropsychological study. *Psychiatry Res.* 78, 21–28. doi: 10.1016/S0165-1781(97)00153-4

Cavedini, P., Zorzi, C., Piccinni, M., Cavallini, M. C., and Bellodi, L. (2010). Executive dysfunctions in obsessive-compulsive patients and unaffected relatives: searching for a new intermediate phenotype. *Biol. Psychiatry* 67, 1178–1184. doi: 10.1016/j.biopsych.2010.02.012

Ceaser, A. E., Goldberg, T. E., Egan, M. F., McMahon, R. P., Weinberger, D. R., and Gold, J. M. (2008). Set-shifting ability and schizophrenia: a marker of clinical illness or an intermediate phenotype? *Biol. Psychiatry* 64, 782–788. doi: 10.1016/j.biopsych.2008.05.009

Chamberlain, S. R., Fineberg, N. A., Blackwell, A. D., Robbins, T. W., and Sahakian, B. J. (2006). Motor inhibition and cognitive flexibility in obsessive-compulsive disorder and trichotillomania. *Am. J. Psychiatry* 163, 1282–1284. doi: 10.1176/appi.ajp.163.7.1282

Chamberlain, S. R., Fineberg, N. A., Menzies, L. A., Blackwell, A. D., Bullmore, E. T., Robbins, T. W., et al. (2007). Impaired cognitive flexibility and motor inhibition in unaffected first-degree relatives of patients with obsessive-compulsive disorder. *Am. J. Psychiatry* 164, 335–338. doi: 10.1176/appi.ajp.164.2.335

Chamberlain, S. R., Menzies, L., Hampshire, A., Suckling, J., Fineberg, N. A., Del, C. N., et al. (2008). Orbitofrontal dysfunction in patients with obsessive-compulsive disorder and their unaffected relatives. *Science* 321, 421–422. doi: 10.1126/science.1154433

Chen, R., Furman, C. A., and Gnegy, M. E. (2010). Dopamine transporter trafficking: rapid response on demand. *Future Neurol.* 5, 123–134. doi: 10.2217/fnl.09.76

Chen, J., Lipska, B. K., Halim, N., Ma, Q. D., Matsumoto, M., Melhem, S., et al. (2004). Functional analysis of genetic variation in catechol-O-methyltransferase (COMT): effects on mRNA. protein, and enzyme activity in postmortem human brain. *Am. J. Hum. Genet.* 75, 807–821. doi: 10.1086/425589

Cheng, J. T., and Li, J. S. (2013). Intra-orbitofrontal cortex injection of haloperidol removes the beneficial effect of methylphenidate on reversal learning of spontaneously hypertensive rats in an attentional set-shifting task. *Behav. Brain Res.* 239, 148–154. doi: 10.1016/j.bbr.2012.11.006

Cheng, R. K., Etchegaray, M., and Meck, W. H. (2007). Impairments in timing, temporal memory, and reversal learning linked to neurotoxic regimens of methamphetamine intoxication. *Brain Res.* 1186, 255–266. doi: 10.1016/j.brainres.2007.10.002

Clarke, H. F., Dalley, J. W., Crofts, H. S., Robbins, T. W., and Roberts, A. C. (2004). Cognitive inflexibility after prefrontal serotonin depletion. *Science* 304, 878–880. doi: 10.1126/science.1094987

Clarke, H. F., Hill, G. J., Robbins, T. W., and Roberts, A. C. (2011). Dopamine, but not serotonin, regulates reversal learning in the marmoset caudate nucleus. *J. Neurosci.* 31, 4290–4297. doi: 10.1523/JNEUROSCI.5066-10.2011

Clarke, H. F., Robbins, T. W., and Roberts, A. C. (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, 10972–10982. doi: 10.1523/JNEUROSCI.1521-08.2008

Clarke, H. F., Walker, S. C., Crofts, H. S., Dalley, J. W., Robbins, T. W., and Roberts, A. C. (2005). Prefrontal serotonin depletion affects reversal learning but not attentional set shifting. *J. Neurosci.* 25, 532–538. doi: 10.1523/JNEUROSCI.3690-04.2005

Clarke, H. F., Walker, S. C., Dalley, J. W., Robbins, T. W., and Roberts, A. C. (2007). Cognitive inflexibility after prefrontal serotonin depletion is behaviorally and neurochemically specific. *Cereb. Cortex* 17, 18–27. doi: 10.1093/cercor/bhj120

Clatworthy, P. L., Lewis, S. J., Brichard, L., Hong, Y. T., Izquierdo, D., Clark, L., et al. (2009). Dopamine release in dissociable striatal subregions predicts the different effects of oral methylphenidate on reversal learning and spatial working memory. *J. Neurosci.* 29, 4690–4696. doi: 10.1523/JNEUROSCI.3266-08.2009

Collins, P., Wilkinson, L. S., Everitt, B. J., Robbins, T. W., and Roberts, A. C. (2000). The effect of dopamine depletion from the caudate nucleus of the common marmoset (*Callithrix jacchus*) on tests of prefrontal cognitive function. *Behav. Neurosci.* 114, 3–17. doi: 10.1037/0735-7044.114.1.3

Colzato, L. S., Waszak, F., Nieuwenhuis, S., Posthuma, D., and Hommel, B. (2010). The flexible mind is associated with the catechol-O-methyltransferase (COMT) Val158Met polymorphism: evidence for a role of dopamine in the control of task-switching. *Neuropsychologia* 48, 2764–2768. doi: 10.1016/j.neuropsychologia.2010.04.023

Cools, A. R. (1980). Role of the neostriatal dopaminergic activity in sequencing and selecting behavioural strategies: facilitation of processes involved in selecting the best strategy in a stressful situation. *Behav. Brain Res.* 1, 361–378. doi: 10.1016/0166-4328(80)90035-2

Cools, R. (2006). Dopaminergic modulation of cognitive function-implications for L-DOPA treatment in Parkinson's disease. *Neurosci. Biobehav. Rev.* 30, 1–23. doi: 10.1016/j.neubiorev.2005.03.024

Cools, R., Barker, R. A., Sahakian, B. J., and Robbins, T. W. (2001). Mechanisms of cognitive set flexibility in Parkinson's disease. *Brain* 124, 2503–2512. doi: 10.1093/brain/124.12.2503

Cools, R., and D'Esposito, M. (2011). Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biol. Psychiatry* 69, e113–e125. doi: 10.1016/j.biopsych.2011.03.028

Cools, R., Frank, M. J., Gibbs, S. E., Miyakawa, A., Jagust, W., and D'Esposito, M. (2009). Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. *J. Neurosci.* 29, 1538–1543. doi: 10.1523/JNEUROSCI.4467-08.2009

Cools, R., Gibbs, S. E., Miyakawa, A., Jagust, W., and D'Esposito, M. (2008). Working memory capacity predicts dopamine synthesis capacity in the human striatum. *J. Neurosci.* 28, 1208–1212. doi: 10.1523/JNEUROSCI.4475-07.2008

Costa, A., Riedel, M., Muller, U., Moller, H. J., and Ettinger, U. (2011). Relationship between SLC6A3 genotype and striatal dopamine transporter availability: a meta-analysis of human single photon emission computed tomography studies. *Synapse* 65, 998–1005. doi: 10.1002/syn.20927

Crofts, H. S., Dalley, J. W., Collins, P., Van Denderen, J. C., Everitt, B. J., Robbins, T. W., et al. (2001). Differential effects of 6-OHDA lesions of the frontal cortex and caudate nucleus on the ability to acquire an attentional set. *Cereb. Cortex* 11, 1015–1026. doi: 10.1093/cercor/11.11.1015

Daberkow, D. P., Riedy, M. D., Kesner, R. P., and Keefe, K. A. (2008). Effect of methamphetamine neurotoxicity on learning-induced Arc mRNA expression in identified striatal efferent neurons. *Neurotox Res.* 14, 307–315. doi: 10.1007/BF03033855

Dang, L. C., Donde, A., Madison, C., O'Neil, J. P., and Jagust, W. J. (2012). Striatal dopamine influences the default mode network to affect shifting between object features. *J. Cogn. Neurosci.* 24, 1960–1970. doi: 10.1162/jocn_a_00252

Darvas, M., and Palmiter, R. D. (2010). Restricting dopaminergic signaling to either dorsolateral or medial striatum facilitates cognition. *J. Neurosci.* 30, 1158–1165. doi: 10.1523/JNEUROSCI.4576-09.2010

Darvas, M., and Palmiter, R. D. (2011). Contributions of striatal dopamine signaling to the modulation of cognitive flexibility. *Biol. Psychiatry* 69, 704–707. doi: 10.1016/j.biopsych.2010.09.033

de Geus, F., Denys, D. A., Sitskoorn, M. M., and Westenberg, H. G. (2007). Attention and cognition in patients with obsessive-compulsive disorder. *Psychiatry Clin. Neurosci.* 61, 45–53. doi: 10.1111/j.1440-1819.2007.01609.x

De Steno, D. A., and Schmauss, C. (2009). A role for dopamine D2 receptors in reversal learning. *Neuroscience* 162, 118–127. doi: 10.1016/j.neuroscience.2009.04.052

Deng, Y. P., Lei, W. L., and Reiner, A. (2006). Differential perikaryal localization in rats of D1 and D2 dopamine receptors on striatal projection neuron types identified by retrograde labeling. *J. Chem. Neuroanat.* 32, 101–116. doi: 10.1016/j.jchemneu.2006.07.001

Denys, D., de Vries, F., Cath, D., Figee, M., Vulink, N., Veltman, D. J., et al. (2013). Dopaminergic activity in Tourette syndrome and obsessive-compulsive disorder. *Eur. Neuropsychopharmacol.* doi: 10.1016/j.euroneuro.2013.05.012. [Epub ahead of print].

Denys, D., van der Wee, N., Janssen, J., de Geus, F., and Westenberg, H. G. (2004a). Low level of dopaminergic D2 receptor binding in obsessive-compulsive disorder. *Biol. Psychiatry* 55, 1041–1045. doi: 10.1016/j.biopsych.2004.01.023

Denys, D., Zohar, J., and Westenberg, H. G. (2004b). The role of dopamine in obsessive-compulsive disorder: preclinical and clinical evidence. *J. Clin. Psychiatry* 65(Suppl. 14), 11–17.

Dias, R., Robbins, T. W., and Roberts, A. C. (1996a). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* 380, 69–72. doi: 10.1038/380069a0

Dias, R., Robbins, T. W., and Roberts, A. C. (1996b). Primate analogue of the Wisconsin Card Sorting Test: effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behav. Neurosci.* 110, 872–886. doi: 10.1037/0735-7044.110.5.872

Dias, R., Robbins, T. W., and Roberts, A. C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: restriction to novel situations and independence from "on-line" processing. *J. Neurosci.* 17, 9285–9297.

Divac, I. (1971). Frontal lobe system and spatial reversal in the rat. *Neuropsychologia* 9, 175–183. doi: 10.1016/0028-3932(71)90041-8

Dodds, C. M., Muller, U., Clark, L., van Loon, A., Cools, R., and Robbins, T. W. (2008). Methylphenidate has differential effects on blood oxygenation level-dependent signal related to cognitive subprocesses of reversal learning. *J. Neurosci.* 28, 5976–5982. doi: 10.1523/JNEUROSCI.1153-08.2008

Dougherty, D. D., Rauch, S. L., and Jenike, M. A. (2004). Pharmacotherapy for obsessive-compulsive disorder. *J. Clin. Psychol.* 60, 1195–1202. doi: 10.1002/jclp.20083

Durstewitz, D., and Seamans, J. K. (2008). The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. *Biol. Psychiatry* 64, 739–749. doi: 10.1016/j.biopsych.2008.05.015

Eagle, D. M., Bari, A., and Robbins, T. W. (2008). The neuropsychopharmacology of action inhibition: cross-species translation of the stop-signal and go/no-go tasks. *Psychopharmacology (Berl.)* 199, 439–456. doi: 10.1007/s00213-008-1127-6

El-Ghundi, M., O'Dowd, B. F., Erclik, M., and George, S. R. (2003). Attenuation of sucrose reinforcement in dopamine D1 receptor deficient mice. *Eur. J. Neurosci.* 17, 851–862. doi: 10.1046/j.1460-9568.2003.02496.x

Elliott, R., Sahakian, B. J., Matthews, K., Bannerjea, A., Rimmer, J., and Robbins, T. W. (1997). Effects of methylphenidate on spatial working memory and planning in healthy young adults. *Psychopharmacology (Berl.)* 131, 196–206. doi: 10.1007/s002130050284

Ersche, K. D., Roiser, J. P., Abbott, S., Craig, K. J., Muller, U., Suckling, J., et al. (2011). Response perseveration in stimulant dependence is associated with striatal dysfunction and can be ameliorated by a D(2/3) receptor agonist. *Biol. Psychiatry* 70, 754–762. doi: 10.1016/j.biopsych.2011.06.033

Fenger, M. M., Gade, A., Adams, K. H., Hansen, E. S., Bolwig, T. G., and Knudsen, G. M. (2005). Cognitive deficits in obsessive-compulsive disorder on tests of

frontal lobe functions. *Nord. J. Psychiatry* 59, 39–44. doi: 10.1080/08039480510018814

Figee, M., Luigjes, J., Smolders, R., Valencia-Alfonso, C. E., van Wingen, G., de Kwaasteniet, B., et al. (2013). Deep brain stimulation restores frontostriatal network activity in obsessive-compulsive disorder. *Nat. Neurosci.* 16, 386–387. doi: 10.1038/nn.3344

Fineberg, N. A., Chamberlain, S. R., Hollander, E., Boulougouris, V., and Robbins, T. W. (2011). Translational approaches to obsessive-compulsive disorder: from animal models to clinical treatment. *Br. J. Pharmacol.* 164, 1044–1061. doi: 10.1111/j.1476-5381.2011.01422.x

Floresco, S. B., Ghods-Sharifi, S., Vexelman, C., and Magyar, O. (2006a). Dissociable roles for the nucleus accumbens core and shell in regulating set shifting. *J. Neurosci.* 26, 2449–2457. doi: 10.1523/JNEUROSCI.4431-05.2006

Floresco, S. B., Magyar, O., Ghods-Sharifi, S., Vexelman, C., and Tse, M. T. (2006b). Multiple dopamine receptor subtypes in the medial prefrontal cortex of the rat regulate set-shifting. *Neuropsychopharmacology* 31, 297–309. doi: 10.1038/sj.npp.1300825

Fontenelle, L., Marques, C., Engelhardt, E., and Versiani, M. (2001). Impaired set-shifting ability and therapeutic response in obsessive-compulsive disorder. *J. Neuropsychiatry Clin. Neurosci.* 13, 508–510. doi: 10.1176/appi.neuropsych.13.4.508

Frank, M. J., and Claus, E. D. (2006). Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol. Rev.* 113, 300–326. doi: 10.1037/0033-295X.113.2.300

Frank, M. J., and Fossella, J. A. (2011). Neurogenetics and pharmacology of learning, motivation, and cognition. *Neuropsychopharmacology* 36, 133–152. doi: 10.1038/npp.2010.96

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., and Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc. Natl. Acad. Sci. U.S.A.* 104, 16311–16316. doi: 10.1073/pnas.0706111104

Friedman, N. P., Miyake, A., Young, S. E., Defries, J. C., Corley, R. P., and Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *J. Exp. Psychol. Gen.* 137, 201–225. doi: 10.1037/0096-3445.137.2.201

Gambini, O., Abbruzzese, M., and Scarone, S. (1993). Smooth pursuit and saccadic eye movements and Wisconsin Card Sorting Test performance in obsessive-compulsive disorder. *Psychiatry Res.* 48, 191–200. doi: 10.1016/0165-1781(93)90071-N

Garcia-Garcia, M., Barcelo, F., Clemente, I. C., and Escera, C. (2010). The role of the dopamine transporter DAT1 genotype on the neural correlates of cognitive flexibility. *Eur. J. Neurosci.* 31, 754–760. doi: 10.1111/j.1460-9568.2010.07102.x

Garner, J. P., and Mason, G. J. (2002). Evidence for a relationship between cage stereotypies and behavioural disinhibition in laboratory rodents. *Behav. Brain Res.* 136, 83–92. doi: 10.1016/S0166-4328(02)00111-0

Garner, J. P., Thogerson, C. M., Wurbel, H., Murray, J. D., and Mench, J. A. (2006). Animal neuropsychology: validation of the Intra-Dimensional Extra-Dimensional set shifting task for mice. *Behav. Brain Res.* 173, 53–61. doi: 10.1016/j.bbr.2006.06.002

Ghahremani, D. G., Monterosso, J., Jentsch, J. D., Bilder, R. M., and Poldrack, R. A. (2010). Neural components underlying behavioral flexibility in human reversal learning. *Cereb. Cortex* 20, 1843–1852. doi: 10.1093/cercor/bhp247

Gilbert, S. J., and Burgess, P. W. (2008). Executive function. *Curr. Biol.* 18, R110–R114. doi: 10.1016/j.cub.2007.12.014

Gillan, C. M., Papmeyer, M., Morein-Zamir, S., Sahakian, B. J., Fineberg, N. A., Robbins, T. W., et al. (2011). Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *Am. J. Psychiatry* 168, 718–726. doi: 10.1176/appi.ajp.2011.10071062

Glickstein, S. B., Desteno, D. A., Hof, P. R., and Schmauss, C. (2005). Mice lacking dopamine D2 and D3 receptors exhibit differential activation of prefrontal cortical neurons during tasks requiring attention. *Cereb. Cortex* 15, 1016–1024. doi: 10.1093/cercor/bhh202

Graham, S., Phua, E., Soon, C. S., Oh, T., Au, C., Shuter, B., et al. (2009). Role of medial cortical, hippocampal and striatal interactions during cognitive set-shifting. *Neuroimage* 45, 1359–1367. doi: 10.1016/j.neuroimage.2008.12.040

Groman, S. M., James, A. S., Seu, E., Crawford, M. A., Harpster, S. N., and Jentsch, J. D. (2013). Monoamine levels within the orbitofrontal cortex and putamen interact to predict reversal learning performance. *Biol. Psychiatry* 73, 756–762. doi: 10.1016/j.biopsych.2012.12.002

Groman, S. M., Lee, B., London, E. D., Mandelkern, M. A., James, A. S., Feiler, K., et al. (2011). Dorsal striatal D2-like receptor availability covaries with sensitivity to positive reinforcement during discrimination learning. *J. Neurosci.* 31, 7291–7299. doi: 10.1523/JNEUROSCI.0363-11.2011

Gross, A. N., Engel, A. K., Richter, S. H., Garner, J. P., and Wurbel, H. (2011). Cage-induced stereotypies in female ICR CD-1 mice do not correlate with recurrent perseveration. *Behav. Brain Res.* 216, 613–620. doi: 10.1016/j.bbr.2010.09.003

Gu, B. M., Park, J. Y., Kang, D. H., Lee, S. J., Yoo, S. Y., Jo, H. J., et al. (2008). Neural correlates of cognitive inflexibility during task-switching in obsessive-compulsive disorder. *Brain* 131, 155–164. doi: 10.1093/brain/awm277

Haluk, D. M., and Floresco, S. B. (2009). Ventral striatal dopamine modulation of different forms of behavioral flexibility. *Neuropsychopharmacology* 34, 2041–2052. doi: 10.1038/npp.2009.21

Harrison, B. J., Soriano-Mas, C., Pujol, J., Ortiz, H., Lopez-Sola, M., Hernandez-Ribas, R., et al. (2009). Altered corticostriatal functional connectivity in obsessive-compulsive disorder. *Arch. Gen. Psychiatry* 66, 1189–1200. doi: 10.1001/archgenpsychiatry.2009.152

Hasler, G., Mondillo, K., Drevets, W. C., and Blair, J. R. (2009). Impairments of probabilistic response reversal and passive avoidance following catecholamine depletion. *Neuropsychopharmacology* 34, 2691–2698. doi: 10.1038/npp.2009.95

Heinz, A., Goldman, D., Jones, D. W., Palmour, R., Hommer, D., Gorey, J. G., et al. (2000). Genotype influences *in vivo* dopamine transporter availability in human striatum. *Neuropsychopharmacology* 22, 133–139. doi: 10.1016/S0893-133X(99)00099-8

Henry, J. D. (2006). A meta-analytic review of Wisconsin Card Sorting Test and verbal fluency performance in obsessive-compulsive disorder. *Cogn. Neuropsychiatry* 11, 156–176. doi: 10.1080/13546800444000227

Hesse, S., Muller, U., Lincke, T., Barthel, H., Villmann, T., Angermeyer, M. C., et al. (2005). Serotonin and dopamine transporter imaging in patients with obsessive-compulsive disorder. *Psychiatry Res.* 140, 63–72. doi: 10.1016/j.pscychresns.2005.07.002

Heyser, C. J., Fienberg, A. A., Greengard, P., and Gold, L. H. (2000). DARPP-32 knockout mice exhibit impaired reversal learning in a discriminated operant task. *Brain Res.* 867, 122–130. doi: 10.1016/S0006-8993(00)02272-1

Hirvonen, M., Laakso, A., Nagren, K., Rinne, J. O., Pohjalainen, T., and Hietala, J. (2004). C957T polymorphism of the dopamine D2 receptor (DRD2) gene affects striatal DRD2 availability *in vivo*. *Mol. Psychiatry* 9, 1060–1061. doi: 10.1038/sj.mp.4001561

Hirvonen, M., Laakso, A., Nagren, K., Rinne, J. O., Pohjalainen, T., and Hietala, J. (2005). C957T polymorphism of the dopamine D2 receptor (DRD2) gene affects striatal DRD2 availability *in vivo*. *Mol. Psychiatry* 10, 889 (Corrigendum). doi: 10.1038/sj.mp.4001707

Hong, S., and Hikosaka, O. (2011). Dopamine-mediated learning and switching in cortico-striatal circuit explain behavioral changes in reinforcement learning. *Front. Behav. Neurosci.* 5:15. doi: 10.3389/fnbeh.2011.00015

Hornak, J., O'Doherty, J., Bramham, J., Rolls, E. T., Morris, R. G., Bullock, P. R., et al. (2004). Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *J. Cogn. Neurosci.* 16, 463–478. doi: 10.1162/089892904322926791

Hsieh, P. C., Yeh, T. L., Lee, I. H., Huang, H. C., Chen, P. S., Yang, Y. K., et al. (2010). Correlation between errors on the Wisconsin Card Sorting Test and the availability of striatal dopamine transporters in healthy volunteers. *J. Psychiatry Neurosci.* 35, 90–94. doi: 10.1503/jpn.090007

Idris, N. F., Repeto, P., Neill, J. C., and Large, C. H. (2005). Investigation of the effects of lamotrigine and clozapine in improving reversal-learning impairments induced by acute phencyclidine and D-amphetamine in the rat. *Psychopharmacology (Berl.)* 179, 336–348. doi: 10.1007/s00213-004-2058-5

Ito, H., Kodaka, F., Takahashi, H., Takano, H., Arakawa, R., Shimada, H., et al. (2011). Relation between presynaptic and postsynaptic dopaminergic functions measured by positron emission tomography: implication of dopaminergic tone. *J. Neurosci.* 31, 7886–7890. doi: 10.1523/JNEUROSCI.6024-10.2011

Izquierdo, A., Belcher, A. M., Scott, L., Cazares, V. A., Chen, J., O'Dell, S. J., et al. (2010). Reversal-specific learning impairments after a binge regimen of methamphetamine in rats: possible involvement of striatal dopamine. *Neuropsychopharmacology* 35, 505–514. doi: 10.1038/npp.2009.155

Izquierdo, A., Wiedholz, L. M., Millstein, R. A., Yang, R. J., Bussey, T. J., Saksida, L. M., et al. (2006). Genetic and dopaminergic modulation of reversal learning in a

touchscreen-based operant procedure for mice. *Behav. Brain Res.* 171, 181–188. doi: 10.1016/j.bbr.2006.03.029

Jocham, G., Klein, T. A., Neumann, J., von Cramon, D. Y., Reuter, M., and Ullsperger, M. (2009). Dopamine DRD2 polymorphism alters reversal learning and associated neural activity. *J. Neurosci.* 29, 3695–3704. doi: 10.1523/JNEUROSCI.5195-08.2009

Joel, D. (2006). The signal attenuation rat model of obsessive-compulsive disorder: a review. *Psychopharmacology (Berl.)* 186, 487–503. doi: 10.1007/s00213-006-0387-2

Joel, D., and Doljansky, J. (2003). Selective alleviation of compulsive lever-pressing in rats by D1, but not D2, blockade: possible implications for the involvement of D1 receptors in obsessive-compulsive disorder. *Neuropsychopharmacology* 28, 77–85. doi: 10.1038/sj.npp.1300010

Karlsson, S., Rieckmann, A., Karlsson, P., Farde, L., Nyberg, L., and Backman, L. (2011). Relationship of dopamine D1 receptor binding in striatal and extrastriatal regions to cognitive functioning in healthy humans. *Neuroimage* 57, 346–351. doi: 10.1016/j.neuroimage.2011.04.047

Kehagia, A. A., Barker, R. A., and Robbins, T. W. (2010). Neuropsychological and clinical heterogeneity of cognitive impairment and dementia in patients with Parkinson's disease. *Lancet Neurol.* 9, 1200–1213. doi: 10.1016/S1474-4422(10)70212-X

Kellendonk, C., Simpson, E. H., Polan, H. J., Malleret, G., Vronskaya, S., Winiger, V., et al. (2006). Transient and selective overexpression of dopamine D2 receptors in the striatum causes persistent abnormalities in prefrontal cortex functioning. *Neuron* 49, 603–615. doi: 10.1016/j.neuron.2006.01.023

Kim, C. H., Cheon, K. A., Koo, M.-S., Ryu, Y. H., Lee, J. D., Chang, J. W., et al. (2007). Dopamine transporter density in the basal ganglia in obsessive-compulsive disorder, measured with [123I]IPT SPECT before and after treatment with serotonin reuptake inhibitors. *Neuropsychobiology* 55, 156–162. doi: 10.1159/000106474

Kim, C. H., Koo, M. S., Cheon, K. A., Ryu, Y. H., Lee, J. D., and Lee, H. S. (2003). Dopamine transporter density of basal ganglia assessed with [123I]IPT SPET in obsessive-compulsive disorder. *Eur. J. Nucl. Med. Mol. Imag.* 30, 1637–1643. doi: 10.1007/s00259-003-1245-7

Kimberg, D. Y., and D'Esposito, M. (2003). Cognitive effects of the dopamine receptor agonist pergolide. *Neuropsychologia* 41, 1020–1027. doi: 10.1016/S0028-3932(02)00317-2

Kimberg, D. Y., D'Esposito, M., and Farah, M. J. (1997). Effects of bromocriptine on human subjects depend on working memory capacity. *Neuroreport* 8, 3581–3585. doi: 10.1097/00001756-199711100-00032

Klein, T. A., Neumann, J., Reuter, M., Hennig, J., von Cramon, D. Y., and Ullsperger, M. (2007). Genetically determined differences in learning from errors. *Science* 318, 1642–1645. doi: 10.1126/science.1145044

Ko, J. H., Monchi, O., Ptito, A., Bloomfield, P., Houle, S., and Strafella, A. P. (2008). Theta burst stimulation-induced inhibition of dorsolateral prefrontal cortex reveals hemispheric asymmetry in striatal dopamine release during a set-shifting task: a TMS-[(11)C]raclopride PET study. *Eur. J. Neurosci.* 28, 2147–2155. doi: 10.1111/j.1460-9568.2008.06501.x

Ko, J. H., Ptito, A., Monchi, O., Cho, S. S., Van, E. T., Pellecchia, G., et al. (2009). Increased dopamine release in the right anterior cingulate cortex during the performance of a sorting task: a [11C]FLB 457 PET study. *Neuroimage* 46, 516–521. doi: 10.1016/j.neuroimage.2009.02.031

Korff, S., and Harvey, B. H. (2006). Animal models of obsessive-compulsive disorder: rationale to understanding psychobiology and pharmacology. *Psychiatry Clin. North Am.* 29, 371–390. doi: 10.1016/j.psc.2006.02.007

Kosheleff, A. R., Rodriguez, D., O'Dell, S. J., Marshall, J. F., and Izquierdo, A. (2012). Comparison of single-dose and extended methamphetamine administration on reversal learning in rats. *Psychopharmacology (Berl.)* 224, 459–467. doi: 10.1007/s00213-012-2774-1

Krahn, L. E., Lin, S. C., Klee, G. G., Lu, P. Y., Ory, S. J., and Zimmermann, R. C. (1999). The effect of presynaptic catecholamine depletion on 6-hydroxymelatonin sulfate: a double blind study of alpha-methyl-para-tyrosine. *Eur. Neuropsychopharmacol.* 9, 61–66. doi: 10.1016/S0924-977X(98)00006-6

Krugel, L. K., Biele, G., Mohr, P. N., Li, S. C., and Heekeren, H. R. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17951–17956. doi: 10.1073/pnas.0905191106

Kruzich, P. J., and Grandy, D. K. (2004). Dopamine D2 receptors mediate two-odor discrimination and reversal learning in C57BL/6 mice. *BMC Neurosci.* 5:12. doi: 10.1186/1471-2202-5-12

Kruzich, P. J., Mitchell, S. H., Younkin, A., and Grandy, D. K. (2006). Dopamine D2 receptors mediate reversal learning in male C57BL/6J mice. *Cogn. Affect. Behav. Neurosci.* 6, 86–90. doi: 10.3758/CABN.6.1.86

Kuczenski, R., and Segal, D. S. (1997). Effects of methylphenidate on extracellular dopamine, serotonin, and norepinephrine: comparison with amphetamine. *J. Neurochem.* 68, 2032–2037. doi: 10.1046/j.1471-4159.1997.68052032.x

Kulig, B. M., and Calhoun, W. H. (1972). Enhancement of successive discrimination reversal learning by methamphetamine. *Psychopharmacologia* 27, 233–240. doi: 10.1007/BF00422803

Laakso, A., Pohjalainen, T., Bergman, J., Kajander, J., Haaparanta, M., Solin, O., et al. (2005). The A1 allele of the human D2 dopamine receptor gene is associated with increased activity of striatal L-amino acid decarboxylase in healthy subjects. *Pharmacogenet. Genom.* 15, 387–391. doi: 10.1097/01213011-200506000-00003

Laakso, A., Vilkman, H., Bergman, J., Haaparanta, M., Solin, O., Syvälahti, E., Salokangas, R. K. R., and Hietala, J. (2002). Sex differences in striatal presynaptic dopamine synthesis capacity in healthy subjects. *Biol. Psychiatry* 52, 759–763. doi: 10.1016/S0006-3223(02)01369-0

Lacerda, A. L., Dalgalarrondo, P., Caetano, D., Haas, G. L., Camargo, E. E., and Keshavan, M. S. (2003). Neuropsychological performance and regional cerebral blood flow in obsessive-compulsive disorder. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 27, 657–665. doi: 10.1016/S0278-5846(03)00076-9

Langen, M., Durston, S., Kas, M. J., van Engel, H., and Staal, W. G. (2011). The neurobiology of repetitive behavior:...and men. *Neurosci. Biobehav. Rev.* 35, 356–365. doi: 10.1016/j.neubiorev.2010.02.005

Lapiz, M. D., and Morilak, D. A. (2006). Noradrenergic modulation of cognitive function in rat medial prefrontal cortex as measured by attentional set shifting capability. *Neuroscience* 137, 1039–1049. doi: 10.1016/j.neuroscience.2005.09.031

Laughlin, R. E., Grant, T. L., Williams, R. W., and Jentsch, J. D. (2011). Genetic dissection of behavioral flexibility: reversal learning in mice. *Biol. Psychiatry* 69, 1109–1116. doi: 10.1016/j.biopsych.2011.01.014

Lawrence, N. S., Wooderson, S., Mataix-Cols, D., David, R., Speckens, A., and Phillips, M. L. (2006). Decision making and set shifting impairments are associated with distinct symptom dimensions in obsessive-compulsive disorder. *Neuropsychology* 20, 409–419. doi: 10.1037/0894-4105.20.4.409

Lee, B., Groman, S., London, E. D., and Jentsch, J. D. (2007). Dopamine D2/D3 receptors play a specific role in the reversal of a learned visual discrimination in monkeys. *Neuropsychopharmacology* 32, 2125–2134. doi: 10.1038/sj.npp.1301337

Leenaars, C. H., Joosten, R. N., Zwart, A., Sandberg, H., Ruimschotel, E., Hanegraaf, M. A., et al. (2012). Switch-task performance in rats is disturbed by 12 h of sleep deprivation but not by 12 h of sleep fragmentation. *Sleep* 35, 211–221.

Lewis, M. H., Tanimura, Y., Lee, L. W., and Bodfish, J. W. (2007). Animal models of restricted repetitive behavior in autism. *Behav. Brain Res.* 176, 66–74. doi: 10.1016/j.bbr.2006.08.023

Logan, G. D., Cowan, W. B., and Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: a model and a method. *J. Exp. Psychol. Hum. Percept. Perform.* 10, 276–291. doi: 10.1037/0096-1523.10.2.276

Lotta, T., Vidgren, J., Tilgmann, C., Ulmanen, I., Melen, K., Julkunen, I., et al. (1995). Kinetics of human soluble and membrane-bound catechol O-methyltransferase: a revised mechanism and description of the thermolabile variant of the enzyme. *Biochemistry* 34, 4202–4210. doi: 10.1021/bi00013a008

Lucey, J. V., Burness, C. E., Costa, D. C., Gacinovic, S., Pilowsky, L. S., Ell, P. J., et al. (1997). Wisconsin Card Sorting Task (WCST) errors and cerebral blood flow in obsessive-compulsive disorder (OCD). *Br. J. Med. Psychol.* 70(Pt 4), 403–411. doi: 10.1111/j.2044-8341.1997.tb01916.x

Lucht, M., and Rosskopf, D. (2008). Comment on "Genetically determined differences in learning from errors". *Science* 321, 200. doi: 10.1126/science.1155372

Lumme, V., Aalto, S., Ilonen, T., Nagren, K., and Hietala, J. (2007). Dopamine D2/D3 receptor binding in the anterior cingulate cortex and executive functioning. *Psychiatry Res.* 156, 69–74. doi: 10.1016/j.pscychresns.2006.12.012

Maes, J. H., Damen, M. D., and Eling, P. A. (2004). More learned irrelevance than perseveration errors in rule shifting in healthy subjects. *Brain Cogn.* 54, 201–211. doi: 10.1016/j.bandc.2004.01.003

Mailly, P., Aliane, V., Groenewegen, H. J., Haber, S. N., and Deniau, J. M. (2013). The rat prefrontostriatal system analyzed in 3D: evidence for multiple interacting functional units. *J. Neurosci.* 33, 5718–5727. doi: 10.1523/JNEUROSCI.5248-12.2013

Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., et al. (2002). Decision-making processes following damage to the prefrontal cortex. *Brain* 125, 624–639. doi: 10.1093/brain/awf049

Martinez, D., Gelernter, J., Abi-Dargham, A., van Dyck, C. H., Kegeles, L., Innis, R. B., et al. (2001). The variable number of tandem repeats polymorphism of the dopamine transporter gene is not associated with significant change in dopamine transporter phenotype in humans. *Neuropsychopharmacology* 24, 553–560. doi: 10.1016/S0893-133X(00)00216-5

McAlonan, K., and Brown, V. J. (2003). Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. *Behav. Brain Res.* 146, 97–103. doi: 10.1016/j.bbr.2003.09.019

McDougle, C. J., Epperson, C. N., Pelton, G. H., Wasylink, S., and Price, L. H. (2000). A double-blind, placebo-controlled study of risperidone addition in serotonin reuptake inhibitor-refractory obsessive-compulsive disorder. *Arch. Gen. Psychiatry* 57, 794–801. doi: 10.1001/archpsyc.57.8.794

Mead, P. G. (1974). Effects of overtraining and pretrial administration of dextroamphetamine on reversal learning in rats. *Percept. Mot. Skills* 38:566. doi: 10.2466/pms.1974.38.2.566

Mehta, M. A., Manes, F. F., Magnolfi, G., Sahakian, B. J., and Robbins, T. W. (2004). Impaired set-shifting and dissociable effects on tests of spatial working memory following the dopamine D2 receptor antagonist sulpiride in human volunteers. *Psychopharmacology (Berl.)* 176, 331–342. doi: 10.1007/s00213-004-1899-2

Mehta, M. A., Sahakian, B. J., McKenna, P. J., and Robbins, T. W. (1999). Systemic sulpiride in young adult volunteers simulates the profile of cognitive deficits in Parkinson's disease. *Psychopharmacology (Berl.)* 146, 162–174. doi: 10.1007/s002130051102

Mehta, M. A., Swainson, R., Ogilvie, A. D., Sahakian, J., and Robbins, T. W. (2001). Improved short-term spatial memory but impaired reversal learning following the dopamine D(2) agonist bromocriptine in human volunteers. *Psychopharmacology (Berl.)* 159, 10–20. doi: 10.1007/s002130100851

Menzies, L., Chamberlain, S. R., Laird, A. R., Thelen, S. M., Sahakian, B. J., and Bullmore, E. T. (2008a). Integrating evidence from neuroimaging and neuropsychological studies of obsessive-compulsive disorder: the orbitofronto-striatal model revisited. *Neurosci. Biobehav. Rev.* 32, 525–549. doi: 10.1016/j.neubiorev.2007.09.005

Menzies, L., Williams, G. B., Chamberlain, S. R., Ooi, C., Fineberg, N., Suckling, J., et al. (2008b). White matter abnormalities in patients with obsessive-compulsive disorder and their first-degree relatives. *Am. J. Psychiatry.* 165, 1308–1315. doi: 10.1176/appi.ajp.2008.07101677

Meyer-Lindenberg, A., Kohn, P. D., Kolachana, B., Kippenhan, S., McInerney-Leo, A., Nussbaum, R., et al. (2005). Midbrain dopamine and prefrontal function in humans: interaction and modulation by COMT genotype. *Nat. Neurosci.* 8, 594–596. doi: 10.1038/nn1438

Meyer-Lindenberg, A., Straub, R. E., Lipska, B. K., Verchinski, B. A., Goldberg, T., Callicott, J. H., et al. (2007). Genetic evidence implicating DARPP-32 in human frontostriatal structure, function, and cognition. *J. Clin. Invest.* 117, 672–682. doi: 10.1172/JCI30413

Mier, D., Kirsch, P., and Meyer-Lindenberg, A. (2010). Neural substrates of pleiotropic action of genetic variation in COMT: a meta-analysis. *Mol. Psychiatry* 15, 918–927. doi: 10.1038/mp.2009.36

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458

Monchi, O., Ko, J. H., and Strafella, A. P. (2006a). Striatal dopamine release during performance of executive functions: a [(11)C] raclopride PET study. *Neuroimage* 33, 907–912. doi: 10.1016/j.neuroimage.2006.06.058

Monchi, O., Petrides, M., Strafella, A. P., Worsley, K. J., and Doyon, J. (2006b). Functional role of the basal ganglia in the planning and execution of actions. *Ann. Neurol.* 59, 257–264. doi: 10.1002/ana.20742

Monsell, S. (2003). Task switching. *Trends Cogn. Sci.* 7, 134–140. doi: 10.1016/S1364-6613(03)00028-7

Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.

Montgomery, A. J., McTavish, S. F., Cowen, P. J., and Grasby, P. M. (2003). Reduction of brain dopamine concentration with dietary tyrosine plus phenylalanine depletion: an [11C]raclopride PET study. *Am. J. Psychiatry* 160, 1887–1889. doi: 10.1176/appi.ajp.160.10.1887

Moreno, M., and Flores, P. (2012). Schedule-induced polydipsia as a model of compulsive behavior: neuropharmacological and neuroendocrine bases. *Psychopharmacology (Berl.)* 219, 647–659. doi: 10.1007/s00213-011-2570-3

Moresco, R. M., Pietra, L., Henin, M., Panzacchi, A., Locatelli, M., Bonaldi, L., et al. (2007). Fluvoxamine treatment and D2 receptors: a pet study on OCD drug-naive patients. *Neuropsychopharmacology* 32, 197–205. doi: 10.1038/sj.npp.1301199

Moritz, S., Birkner, C., Kloss, M., Jahn, H., Hand, I., Haasen, C., et al. (2002). Executive functioning in obsessive-compulsive disorder, unipolar depression, and schizophrenia. *Arch. Clin. Neuropsychol.* 17, 477–483.

Moritz, S., Hubner, M., and Kluwe, R. (2004). Task switching and backward inhibition in obsessive-compulsive disorder. *J. Clin. Exp. Neuropsychol.* 26, 677–683.

Munoz, D. P., and Everling, S. (2004). Look away: the anti-saccade task and the voluntary control of eye movement. *Nat. Rev. Neurosci.* 5, 218–228. doi: 10.1038/nrn1345

Nagano-Saito, A., Leyton, M., Monchi, O., Goldberg, Y. K., He, Y., and Dagher, A. (2008). Dopamine depletion impairs frontostriatal functional connectivity during a set-shifting task. *J. Neurosci.* 28, 3697–3706. doi: 10.1523/JNEUROSCI.3921-07.2008

Nestler, E. J., and Hyman, S. E. (2010). Animal models of neuropsychiatric disorders. *Nat. Neurosci.* 13, 1161–1169. doi: 10.1038/nn.2647

Nikiforuk, A. (2012). Dopamine D1 receptor modulation of set shifting: the role of stress exposure. *Behav. Pharmacol.* 23, 434–438. doi: 10.1097/FBP.0b013e328356522f

Nikolaus, S., Antke, C., Beu, M., and Muller, H. W. (2010). Cortical GABA striatal dopamine and midbrain serotonin as the key players in compulsive and anxiety disorders–results from *in vivo* imaging studies. *Rev. Neurosci.* 21, 119–139. doi: 10.1515/REVNEURO.2010.21.2.119

Nishi, A., Snyder, G. L., and Greengard, P. (1997). Bidirectional regulation of DARPP-32 phosphorylation by dopamine. *J. Neurosci.* 17, 8147–8155.

Oades, R. D. (1985). The role of noradrenaline in tuning and dopamine in switching between signals in the CNS. *Neurosci. Biobehav. Rev.* 9, 261–282. doi: 10.1016/0149-7634(85)90050-8

Olley, A., Malhi, G., and Sachdev, P. (2007). Memory and executive functioning in obsessive-compulsive disorder: a selective review. *J. Affect Disord.* 104, 15–23. doi: 10.1016/j.jad.2007.02.023

Olver, J. S., O'Keefe, G., Jones, G. R., Burrows, G. D., Tochon-Danguy, H. J., Ackermann, U., et al. (2009). Dopamine D1 receptor binding in the striatum of patients with obsessive-compulsive disorder. *J. Affect Disord.* 114, 321–326. doi: 10.1016/j.jad.2008.06.020

Olver, J. S., O'Keefe, G., Jones, G. R., Burrows, G. D., Tochon-Danguy, H. J., Ackermann, U., et al. (2010). Dopamine D(1) receptor binding in the anterior cingulate cortex of patients with obsessive-compulsive disorder. *Psychiatry Res.* 183, 85–88. doi: 10.1016/j.pscychresns.2010.04.004

O'Neill, M., and Brown, V. J. (2007). The effect of striatal dopamine depletion and the adenosine A2A antagonist KW-6002 on reversal learning in rats. *Neurobiol. Learn. Mem.* 88, 75–81. doi: 10.1016/j.nlm.2007.03.003

Owen, A. M., Roberts, A. C., Polkey, C. E., Sahakian, B. J., and Robbins, T. W. (1991). Extra-dimensional versus intra-dimensional set shifting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia* 29, 993–1006. doi: 10.1016/0028-3932(91)90063-E

Packard, M. G. (2009). Exhumed from thought: basal ganglia and response learning in the plus-maze. *Behav. Brain Res.* 199, 24–31. doi: 10.1016/j.bbr.2008.12.013

Page, L. A., Rubia, K., Deeley, Q., Daly, E., Toal, F., Mataix-Cols, D., et al. (2009). A functional magnetic resonance imaging study of inhibitory control in obsessive-compulsive disorder. *Psychiatry Res.* 174, 202–209. doi: 10.1016/j.pscychresns.2009.05.002

Palmiter, R. D. (2008). Dopamine signaling in the dorsal striatum is essential for motivated behaviors: lessons from dopamine-deficient mice. *Ann. N.Y. Acad. Sci.* 1129, 35–46. doi: 10.1196/annals.1417.003

Papaleo, F., Crawley, J. N., Song, J., Lipska, B. K., Pickel, J., Weinberger, D. R., et al. (2008). Genetic dissection of the role of catechol-O-methyltransferase

in cognition and stress reactivity in mice. *J. Neurosci.* 28, 8709–8723. doi: 10.1523/JNEUROSCI.2077-08.2008

Pastuzyn, E. D., Chapman, D. E., Wilcox, K. S., and Keefe, K. A. (2012). Altered learning and Arc-regulated consolidation of learning in striatum by methamphetamine-induced neurotoxicity. *Neuropsychopharmacology* 37, 885–895. doi: 10.1038/npp.2011.265

Perani, D., Garibotto, V., Gorini, A., Moresco, R. M., Henin, M., Panzacchi, A., et al. (2008). *In vivo* PET study of 5HT(2A) serotonin and D(2) dopamine dysfunction in drug-naive obsessive-compulsive disorder. *Neuroimage* 42, 306–314. doi: 10.1016/j.neuroimage.2008.04.233

Pitman, R. K. (1989). Animal models of compulsive behavior. *Biol. Psychiatry* 26, 189–198. doi: 10.1016/0006-3223(89)90022-X

Pogarell, O., Hamann, C., Popperl, G., Juckel, G., Chouker, M., Zaudig, M., et al. (2003). Elevated brain serotonin transporter availability in patients with obsessive-compulsive disorder. *Biol. Psychiatry* 54, 1406–1413. doi: 10.1016/S0006-3223(03)00183-5

Pohjalainen, T., Rinne, J. O., Nagren, K., Lehikoinen, P., Anttila, K., Syvalahti, E. K., et al. (1998). The A1 allele of the human D2 dopamine receptor gene predicts low D2 receptor availability in healthy volunteers. *Mol. Psychiatry* 3, 256–260. doi: 10.1038/sj.mp.4000350

Presti, M. F., Mikes, H. M., and Lewis, M. H. (2003). Selective blockade of spontaneous motor stereotypy via intrastriatal pharmacological manipulation. *Pharmacol. Biochem. Behav.* 74, 833–839. doi: 10.1016/S0091-3057(02)01081-X

Pujol, J., Soriano-Mas, C., Alonso, P., Cardoner, N., Menchon, J. M., Deus, J., et al. (2004). Mapping structural brain alterations in obsessive-compulsive disorder. *Arch. Gen. Psychiatry* 61, 720–730. doi: 10.1001/archpsyc.61.7.720

Purcell, R., Maruff, P., Kyrios, M., and Pantelis, C. (1998a). Cognitive deficits in obsessive-compulsive disorder on tests of frontal-striatal function. *Biol. Psychiatry* 43, 348–357. doi: 10.1016/S0006-3223(97)00201-1

Purcell, R., Maruff, P., Kyrios, M., and Pantelis, C. (1998b). Neuropsychological deficits in obsessive-compulsive disorder: a comparison with unipolar depression, panic disorder, and normal controls. *Arch. Gen. Psychiatry* 55, 415–423. doi: 10.1001/archpsyc.55.5.415

Ragozzino, M. E. (2002). The effects of dopamine D(1) receptor blockade in the prelimbic-infralimbic areas on behavioral flexibility. *Learn. Mem.* 9, 18–28. doi: 10.1101/lm.45802

Ragozzino, M. E. (2007). The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. *Ann. N.Y. Acad. Sci.* 1121, 355–375. doi: 10.1196/annals.1401.013

Randrup, A., and Munkvad, I. (1975). Stereotyped behavior. *Pharmacol. Ther. B* 1, 757–768. doi: 10.1016/0306-039X(75)90027-6

Remijnse, P. L., Nielen, M. M., van Balkom, A. J., Cath, D. C., van Oppen, P., Uylings, H. B., et al. (2006). Reduced orbitofrontal-striatal activity on a reversal learning task in obsessive-compulsive disorder. *Arch. Gen. Psychiatry* 63, 1225–1236. doi: 10.1001/archpsyc.63.11.1225

Remijnse, P. L., van den Heuvel, O. A., Nielen, M. M., Vriend, C., Hendriks, G. J., Hoogendijk, W. J., et al. (2013). Cognitive inflexibility in obsessive-compulsive disorder and major depression is associated with distinct neural correlates. *PLoS ONE* 8:e59600. doi: 10.1371/journal.pone.0059600

Richfield, E. K., Penney, J. B., and Young, A. B. (1989). Anatomical and affinity state comparisons between dopamine D1 and D2 receptors in the rat central nervous system. *Neuroscience* 30, 767–777. doi: 10.1016/0306-4522(89)90168-1

Ridley, R. M. (1994). The psychology of perserverative and stereotyped behaviour. *Prog. Neurobiol.* 44, 221–231. doi: 10.1016/0301-0082(94)90039-6

Ridley, R. M., Haystead, T. A., and Baker, H. F. (1981). An analysis of visual object reversal learning in the marmoset after amphetamine and haloperidol. *Pharmacol. Biochem. Behav.* 14, 345–351. doi: 10.1016/0091-3057(81)90401-9

Robbins, T. W., and Arnsten, A. F. (2009). The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Annu. Rev. Neurosci.* 32, 267–287. doi: 10.1146/annurev.neuro.051508.135535

Robbins, T. W., Gillan, C. M., Smith, D. G., de Wit, S., and Ersche, K. D. (2012). Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends Cogn. Sci.* 16, 81–91. doi: 10.1016/j.tics.2011.11.009

Roberts, A. C., De Salvia, M. A., Wilkinson, L. S., Collins, P., Muir, J. L., Everitt, B. J., et al. (1994). 6-Hydroxydopamine lesions of the prefrontal cortex in monkeys enhance performance on an analog of the Wisconsin Card Sort Test: possible interactions with subcortical dopamine. *J. Neurosci.* 14, 2531–2544.

Roberts, A. C., Robbins, T. W., and Everitt, B. J. (1988). The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates. *Q. J. Exp. Psychol. B* 40, 321–341.

Robinson, O. J., Standing, H. R., DeVito, E. E., Cools, R., and Sahakian, B. J. (2010). Dopamine precursor depletion improves punishment prediction during reversal learning in healthy females but not males. *Psychopharmacology (Berl.)* 211, 187–195. doi: 10.1007/s00213-010-1880-1

Robinson, S., Sandstrom, S. M., Denenberg, V. H., and Palmiter, R. D. (2005). Distinguishing whether dopamine regulates liking, wanting, and/or learning about rewards. *Behav. Neurosci.* 119, 5–15. doi: 10.1037/0735-7044.119.1.5

Rodriguez-Jimenez, R., Hoenicka, J., Jimenez-Arriero, M. A., Ponce, G., Bagney, A., Aragues, M., et al. (2006). Performance in the Wisconsin Card Sorting Test and the C957T polymorphism of the DRD2 gene in healthy volunteers. *Neuropsychobiology* 54, 166–170. doi: 10.1159/000098652

Rogers, R. D., Andrews, T. C., Grasby, P. M., Brooks, D. J., and Robbins, T. W. (2000). Contrasting cortical and subcortical activations produced by attentional-set shifting and reversal learning in humans. *J. Cogn. Neurosci.* 12, 142–162. doi: 10.1162/089892900561931

Rotge, J. Y., Langbour, N., Guehl, D., Bioulac, B., Jaafari, N., Allard, M., et al. (2010). Gray matter alterations in obsessive-compulsive disorder: an anatomic likelihood estimation meta-analysis. *Neuropsychopharmacology* 35, 686–691. doi: 10.1038/npp.2009.175

Rudebeck, P. H., Saunders, R. C., Prescott, A. T., Chau, L. S., and Murray, E. A. (2013). Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nat. Neurosci.* 16, 1140–1145. doi: 10.1038/nn.3440

Schneier, F. R., Martinez, D., Abi-Dargham, A., Zea-Ponce, Y., Simpson, H. B., Liebowitz, M. R., et al. (2008). Striatal dopamine D(2) receptor availability in OCD with and without comorbid social anxiety disorder: preliminary findings. *Depress Anxiety* 25, 1–7. doi: 10.1002/da.20268

Schoenbaum, G., Roesch, M. R., Stalnaker, T. A., and Takahashi, Y. K. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nat. Rev. Neurosci.* 10, 885–892.

Schultz, W. (2013). Updating dopamine reward signals. *Curr. Opin. Neurobiol.* 23, 229–238. doi: 10.1016/j.conb.2012.11.012

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593

Seamans, J. K., and Yang, C. R. (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog. Neurobiol.* 74, 1–58. doi: 10.1016/j.pneurobio.2004.05.006

Sesack, S. R., Hawrylak, V. A., Matus, C., Guido, M. A., and Levey, A. I. (1998). Dopamine axon varicosities in the prelimbic division of the rat prefrontal cortex exhibit sparse immunoreactivity for the dopamine transporter. *J. Neurosci.* 18, 2697–2708.

Sesia, T., Bizup, B., and Grace, A. A. (2013). Evaluation of animal models of obsessive-compulsive disorder: correlation with phasic dopamine neuron activity. *Int. J. Neuropsychopharmacol.* 16, 1295–1307. doi: 10.1017/S146114571200154X.

Seu, E., and Jentsch, J. D. (2009). Effect of acute and repeated treatment with desipramine or methylphenidate on serial reversal learning in rats. *Neuropharmacology* 57, 665–672. doi: 10.1016/j.neuropharm.2009.08.007

Seu, E., Lang, A., Rivera, R. J., and Jentsch, J. D. (2009). Inhibition of the norepinephrine transporter improves behavioral flexibility in rats and monkeys. *Psychopharmacology (Berl.)* 202, 505–519. doi: 10.1007/s00213-008-1250-4

Shmelkov, S. V., Hormigo, A., Jing, D., Proenca, C. C., Bath, K. G., Milde, T., et al. (2010). Slitrk5 deficiency impairs corticostriatal circuitry and leads to obsessive-compulsive-like behaviors in mice. *Nat. Med.* 16, 598–602, 1p.

Smith, A. G., Neill, J. C., and Costall, B. (1999). The dopamine D3/D2 receptor agonist 7-OH-DPAT induces cognitive impairment in the marmoset. *Pharmacol. Biochem. Behav.* 63, 201–211. doi: 10.1016/S0091-3057(98)00230-5

Sohn, M. H., Ursu, S., Anderson, J. R., Stenger, V. A., and Carter, C. S. (2000). The role of prefrontal cortex and posterior parietal cortex in task switching. *Proc. Natl. Acad. Sci. U.S.A.* 97, 13448–13453. doi: 10.1073/pnas.240460497

Soto, P. L., Wilcox, K. M., Zhou, Y., Ator, N. A., Riddle, M. A., Wong, D. F., et al. (2012). Long-term exposure to oral methylphenidate or dl-amphetamine mixture in peri-adolescent rhesus monkeys: effects on physiology, behavior, and dopamine system development. *Neuropsychopharmacology* 37, 2566–2579. doi: 10.1038/npp.2012.119

Stefani, M. R., and Moghaddam, B. (2006). Rule learning and reward contingency are associated with dissociable patterns of dopamine activation in the rat prefrontal cortex, nucleus accumbens, and dorsal striatum. *J. Neurosci.* 26, 8810–8818. doi: 10.1523/JNEUROSCI.1656-06.2006

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., and Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nat. Neurosci.* 16, 966–973. doi: 10.1038/nn.3413

Stelzel, C., Basten, U., Montag, C., Reuter, M., and Fiebach, C. J. (2010). Frontostriatal involvement in task switching depends on genetic differences in d2 receptor density. *J. Neurosci.* 30, 14205–14212. doi: 10.1523/JNEUROSCI.1062-10.2010

Stoet, G., and Snyder, L. H. (2003). Executive control and task-switching in monkeys. *Neuropsychologia* 41, 1357–1364. doi: 10.1016/S0028-3932(03)00048-4

Sulzer, D., Sonders, M. S., Poulsen, N. W., and Galli, A. (2005). Mechanisms of neurotransmitter release by amphetamines: a review. *Prog. Neurobiol.* 75, 406–433. doi: 10.1016/j.pneurobio.2005.04.003

Svenningsson, P., Nishi, A., Fisone, G., Girault, J. A., Nairn, A. C., and Greengard, P. (2004). DARPP-32: an integrator of neurotransmission. *Annu. Rev. Pharmacol. Toxicol.* 44, 269–296. doi: 10.1146/annurev.pharmtox.44.101802.121415

Szechtman, H., Sulis, W., and Eilam, D. (1998). Quinpirole induces compulsive checking behavior in rats: a potential animal model of obsessive-compulsive disorder (OCD). *Behav. Neurosci.* 112, 1475–1485. doi: 10.1037/0735-7044.112.6.1475

Taghzouti, K., Louilot, A., Herman, J. P., Le, M. M., and Simon, H. (1985). Alternation behavior, spatial discrimination, and reversal disturbances following 6-hydroxydopamine lesions in the nucleus accumbens of the rat. *Behav. Neural. Biol.* 44, 354–363. doi: 10.1016/S0163-1047(85)90640-5

Tait, D. S., and Brown, V. J. (2007). Difficulty overcoming learned non-reward during reversal learning in rats with ibotenic acid lesions of orbital prefrontal cortex. *Ann. N.Y. Acad. Sci.* 1121, 407–420. doi: 10.1196/annals.1401.010

Takahashi, H., Kato, M., Takano, H., Arakawa, R., Okumura, M., Otsuka, T., et al. (2008). Differential contributions of prefrontal and hippocampal dopamine D(1) and D(2) receptors in human cognitive functions. *J. Neurosci.* 28, 12032–12038. doi: 10.1523/JNEUROSCI.3446-08.2008

Talpos, J. C., Fletcher, A. C., Circelli, C., Tricklebank, M. D., and Dix, S. L. (2012). The pharmacological sensitivity of a touchscreen-based visual discrimination task in the rat using simple and perceptually challenging stimuli. *Psychopharmacology (Berl.)* 221, 437–449. doi: 10.1007/s00213-011-2590-z

Tanimura, Y., Yang, M. C., and Lewis, M. H. (2008). Procedural learning and cognitive flexibility in a mouse model of restricted, repetitive behaviour. *Behav. Brain Res.* 189, 250–256. doi: 10.1016/j.bbr.2008.01.001

Thompson, J., Thomas, N., Singleton, A., Piggott, M., Lloyd, S., Perry, E. K., et al. (1997). D2 dopamine receptor gene (DRD2) Taq1 A polymorphism: reduced dopamine D2 receptor binding in the human striatum associated with the A1 allele. *Pharmacogenetics* 7, 479–484. doi: 10.1097/00008571-199712000-00006

Ting, J. T., and Feng, G. (2011). Neurobiology of obsessive-compulsive disorder: insights into neural circuitry dysfunction through mouse genetics. *Curr. Opin. Neurobiol.* 21, 842–848. doi: 10.1016/j.conb.2011.04.010

Tunbridge, E. M., Bannerman, D. M., Sharp, T., and Harrison, P. J. (2004). Catechol-o-methyltransferase inhibition improves set-shifting performance and elevates stimulated dopamine release in the rat prefrontal cortex. *J. Neurosci.* 24, 5331–5335. doi: 10.1523/JNEUROSCI.1124-04.2004

Uylings, H. B., Groenewegen, H. J., and Kolb, B. (2003). Do rats have a prefrontal cortex? *Behav. Brain Res.* 146, 3–17. doi: 10.1016/j.bbr.2003.09.028

Valerius, G., Lumpp, A., Kuelz, A. K., Freyer, T., and Voderholzer, U. (2008). Reversal learning as a neuropsychological indicator for the neuropathology of obsessive compulsive disorder? A behavioral study. *J. Neuropsychiatry Clin. Neurosci.* 20, 210–218. doi: 10.1176/appi.neuropsych.20.2.210

van de Giessen, E., de Win, M. M., Tanck, M. W., van den Brink, W., Baas, F., and Booij, J. (2009). Striatal dopamine transporter availability associated with polymorphisms in the dopamine transporter gene SLC6A3. *J. Nucl. Med.* 50, 45–52. doi: 10.2967/jnumed.108.053652

van den Bos, R., and Cools, A. R. (1989). The involvement of the nucleus accumbens in the ability of rats to switch to cue-directed behaviours. *Life Sci.* 44, 1697–1704. doi: 10.1016/0024-3205(89)90486-4

van den Heuvel, O. A., Remijnse, P. L., Mataix-Cols, D., Vrenken, H., Groenewegen, H. J., Uylings, H. B., et al. (2009). The major symptom dimensions of obsessive-compulsive disorder are mediated by partially distinct neural systems. *Brain* 132, 853–868. doi: 10.1093/brain/awn267

van der Meulen, J. A., Joosten, R. N., de Bruin, J. P., and Feenstra, M. G. (2007). Dopamine and noradrenaline efflux in the medial prefrontal cortex during serial reversals and extinction of instrumental goal-directed behavior. *Cereb. Cortex* 17, 1444–1453. doi: 10.1093/cercor/bhl057

van der Schaaf M. E., van Schouwenburg, M. R., Geurts, D. E., Schellekens, A. F., Buitelaar, J. K., Verkes, R. J., et al. (2012). Establishing the dopamine dependency of human striatal signals during reward and punishment reversal learning. *Cereb. Cortex*. doi: 10.1093/cercor/bhs344. [Epub ahead of print].

van der Wee, N. J., Stevens, H., Hardeman, J. A., Mandl, R. C., Denys, D. A., Van Megen, H. J., et al. (2004). Enhanced dopamine transporter density in psychotropic-naive patients with obsessive-compulsive disorder shown by [123I]{beta}-CIT SPECT. *Am. J. Psychiatry* 161, 2201–2206. doi: 10.1176/appi.ajp.161.12.2201

van Dijk, A., Klompmakers, A., and Denys, D. (2010). "The serotonergic system in obsessive compulsive disorder," in *Handbook of the Behavioral Neurobiology of Serotonin*, eds C. P. Mu¨ller and B. L. Jacobs (London : Academic Press), 547–564.

van Dyck, C. H., Malison, R. T., Jacobsen, L. K., Seibyl, J. P., Staley, J. K., Laruelle, M., et al. (2005). Increased dopamine transporter availability associated with the 9-repeat allele of the SLC6A3 gene. *J. Nucl. Med.* 46, 745–751.

van Holstein, M., Aarts, E., van der Schaaf, M. E., Geurts, D. E., Verkes, R. J., Franke, B., et al. (2011). Human cognitive flexibility depends on dopamine D2 receptor signaling. *Psychopharmacology (Berl.)* 218, 567–578. doi: 10.1007/s00213-011-2340-2

Veale, D. M., Sahakian, B. J., Owen, A. M., and Marks, I. M. (1996). Specific cognitive deficits in tests sensitive to frontal lobe dysfunction in obsessive-compulsive disorder. *Psychol. Med.* 26, 1261–1269. doi: 10.1017/S0033291700035984

Verdejo-Garcia, A., Bechara, A., Recknor, E. C., and Perez-Garcia, M. (2006). Executive dysfunction in substance dependent individuals during drug use and abstinence: an examination of the behavioral, cognitive and emotional correlates of addiction. *J. Int. Neuropsychol. Soc.* 12, 405–415. doi: 10.1017/S1355617706060486

Verhoeff, N. P., Christensen, B. K., Hussey, D., Lee, M., Papatheodorou, G., Kopala, L., et al. (2003). Effects of catecholamine depletion on D2 receptor binding, mood, and attentiveness in humans: a replication study. *Pharmacol. Biochem. Behav.* 74, 425–432. doi: 10.1016/S0091-3057(02)01028-6

Vermeire, S., Audenaert, K., De Meester, R., Vandermeulen, E., Waelbers, T., De Spiegeleer, B., et al. (2012). Serotonin 2A receptor, serotonin transporter and dopamine transporter alterations in dogs with compulsive behaviour as a promising model for human obsessive-compulsive disorder. *Psychiatry Res.* 201, 78–87. doi: 10.1016/j.pscychresns.2011.06.006

Vernaleken, I., Buchholz, H. G., Kumakura, Y., Siessmeier, T., Stoeter, P., Bartenstein, P., et al. (2007). 'Prefrontal' cognitive performance of healthy subjects positively correlates with cerebral FDOPA influx: an exploratory [18F]-fluoro-L-DOPA-PET investigation. *Hum. Brain Mapp.* 28, 931–939. doi: 10.1002/hbm.20325

Vickery, S. S., and Mason, G. J. (2005). Stereotypy and perseverative responding in caged bears: further data and analyses. *Appl. Anim. Behav. Sci.* 247–260. doi: 10.1016/j.applanim.2005.01.005

Vijayraghavan, S., Wang, M., Birnbaum, S. G., Williams, G. V., and Arnsten, A. F. (2007). Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nat. Neurosci.* 10, 376–384. doi: 10.1038/nn1846

Volkow, N. D., Gur, R. C., Wang, G. J., Fowler, J. S., Moberg, P. J., Ding, Y. S., et al. (1998). Association between decline in brain dopamine activity with age and cognitive and motor impairment in healthy individuals. *Am. J. Psychiatry* 155, 344–349.

Volkow, N. D., Wang, G., Fowler, J. S., Logan, J., Gerasimov, M., Maynard, L., et al. (2001). Therapeutic doses of oral methylphenidate significantly increase extracellular dopamine in the human brain. *J. Neurosci.* 21:RC121.

Walaas, S. I., and Greengard, P. (1984). DARPP-32, a dopamine- and adenosine 3':5'-monophosphate-regulated phosphoprotein enriched in dopamine-innervated brain regions. I. Regional and cellular distribution in the rat brain. *J. Neurosci.* 4, 84–98.

Walker, S. C., Robbins, T. W., and Roberts, A. C. (2009). Differential contributions of dopamine and serotonin to orbitofrontal cortex function in the marmoset. *Cereb. Cortex* 19, 889–898. doi: 10.1093/cercor/bhn136

Walton, M. E., Behrens, T. E., Buckley, M. J., Rudebeck, P. H., and Rushworth, M. F. (2010). Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron* 65, 927–939. doi: 10.1016/j.neuron.2010.02.027

Wang, L., Simpson, H. B., and Dulawa, S. C. (2009). Assessing the validity of current mouse genetic models of obsessive-compulsive disorder. *Behav. Pharmacol.* 20, 119–133. doi: 10.1097/FBP.0b013e32832a80ad

Watkins, L. H., Sahakian, B. J., Robertson, M. M., Veale, D. M., Rogers, R. D., Pickard, K. M., et al. (2005). Executive function in Tourette's syndrome and obsessive-compulsive disorder. *Psychol. Med.* 35, 571–582. doi: 10.1017/S0033291704003691

Weiner, I., and Feldon, J. (1986). Reversal and nonreversal shifts under amphetamine. *Psychopharmacology (Berl.)* 89, 355–359. doi: 10.1007/BF00174374

Weiner, I., Feldon, J., and Ben-Shahar, O. (1986). Simultaneous brightness discrimination and reversal: the effects of amphetamine administration in the two stages. *Pharmacol. Biochem. Behav.* 25, 939–942. doi: 10.1016/0091-3057(86)90066-3

Welch, J. M., Lu, J., Rodriguiz, R. M., Trotta, N. C., Peca, J., Ding, J. D., et al. (2007). Cortico-striatal synaptic defects and OCD-like behaviours in Sapap3-mutant mice. *Nature* 448, 894–900. doi: 10.1038/nature06104

White, I. M., Minamoto, T., Odell, J. R., Mayhorn, J., and White, W. (2009). Brief exposure to methamphetamine (METH) and phencyclidine (PCP) during late development leads to long-term learning deficits in rats. *Brain Res.* 1266, 72–86. doi: 10.1016/j.brainres.2009.02.024

Williams, G. V., and Goldman-Rakic, P. S. (1995). Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. *Nature* 376, 572–575. doi: 10.1038/376572a0

Wilpizeski, C. R., and Hamilton, C. L. (1964). Drug-induced decrement in spatial reversal learning in the white rat. *Psychopharmacologia* 6, 475–479. doi: 10.1007/BF00429574

Winter, S., Dieckmann, M., and Schwabe, K. (2009). Dopamine in the prefrontal cortex regulates rats behavioral flexibility to changing reward value. *Behav. Brain Res.* 198, 206–213. doi: 10.1016/j.bbr.2008.10.040

Wise, S. P., Murray, E. A., and Gerfen, C. R. (1996). The frontal cortex-basal ganglia system in primates. *Crit. Rev. Neurobiol.* 10, 317–356. doi: 10.1615/CritRevNeurobiol.v10.i3-4.30

Yawata, S., Yamaguchi, T., Danjo, T., Hikida, T., and Nakanishi, S. (2012). Pathway-specific control of reward learning and its flexibility via selective dopamine receptors in the nucleus accumbens. *Proc. Natl. Acad. Sci. U.S.A.* 109, 12764–12769. doi: 10.1073/pnas.1210797109

Yerys, B. E., Wallace, G. L., Harrison, B., Celano, M. J., Giedd, J. N., and Kenworthy, L. E. (2009). Set-shifting in children with autism spectrum disorders: reversal shifting deficits on the Intradimensional/Extradimensional Shift Test correlate with repetitive behaviors. *Autism* 13, 523–538.

Zahrt, J., Taylor, J. R., Mathew, R. G., and Arnsten, A. F. (1997). Supranormal stimulation of D1 dopamine receptors in the rodent prefrontal cortex impairs spatial working memory performance. *J. Neurosci.* 17, 8528–8535.

Zweifel, L. S., Parker, J. G., Lobb, C. J., Rainwater, A., Wall, V. Z., Fadok, J. P., et al. (2009). Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior. *Proc. Natl. Acad. Sci. U.S.A.* 106, 7281–7288. doi: 10.1073/pnas.0813415106

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 22 May 2013; accepted: 11 October 2013; published online: 05 November 2013.*

*Citation: Klanker M, Feenstra M and Denys D (2013) Dopaminergic control of cognitive flexibility in humans and animals. Front. Neurosci. 7:201. doi: 10.3389/fnins.2013.00201*

*This article was submitted to Decision Neuroscience, a section of the journal Frontiers in Neuroscience.*

frontiers in NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 24 December 2013
doi: 10.3389/fnins.2013.00253

# Of goals and habits: age-related and individual differences in goal-directed decision-making

**Ben Eppinger[1,2]*, Maik Walter[3], Hauke R. Heekeren[2,3] and Shu-Chen Li[1,2]**

[1] Chair of Lifespan Developmental Neuroscience, Department of Psychology, TU Dresden, Dresden, Germany
[2] Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany
[3] Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany

**Edited by:**
*Monica Luciana, University of Minnesota, USA*

**Reviewed by:**
*Nathaniel D. Daw, New York University, USA*
*Gregory R. Samanez-Larkin, Yale University, USA*
*Samuel J. Gershman, Massachusetts Institute of Technology, USA*

***Correspondence:**
*Ben Eppinger, Department of Psychology, Technische Universität Dresden, Zellescher Weg 17, 01062 Dresden, Germany*
*e-mail: benjamin.eppinger@tu-dresden.de*

In this study we investigated age-related and individual differences in habitual (model-free) and goal-directed (model-based) decision-making. Specifically, we were interested in three questions. First, does age affect the balance between model-based and model-free decision mechanisms? Second, are these age-related changes due to age differences in working memory (WM) capacity? Third, can model-based behavior be affected by manipulating the distinctiveness of the reward value of choice options? To answer these questions we used a two-stage Markov decision task in in combination with computational modeling to dissociate model-based and model-free decision mechanisms. To affect model-based behavior in this task we manipulated the distinctiveness of reward probabilities of choice options. The results show age-related deficits in model-based decision-making, which are particularly pronounced if unexpected reward indicates the need for a shift in decision strategy. In this situation younger adults explore the task structure, whereas older adults show perseverative behavior. Consistent with previous findings, these results indicate that older adults have deficits in the representation and updating of expected reward value. We also observed substantial individual differences in model-based behavior. In younger adults high WM capacity is associated with greater model-based behavior and this effect is further elevated when reward probabilities are more distinct. However, in older adults we found no effect of WM capacity. Moreover, age differences in model-based behavior remained statistically significant, even after controlling for WM capacity. Thus, factors other than decline in WM, such as deficits in the in the integration of expected reward value into strategic decisions may contribute to the observed impairments in model-based behavior in older adults.

**Keywords: aging, decision-making and learning, dopamine, goal-directed, habitual**

## INTRODUCTION

Many simple everyday decision-making tasks, such as which cereals to take for breakfast or which subway to take to work in the morning, can be solved via habitual decision mechanisms. However, in more complex decision scenarios, such as how to spend annual bonus or how to plan retirement savings, it may be adaptive to anticipate the consequences of future decisions and to choose the options that are likely to yield higher long-term benefits. In the current study we examined age and individual differences in the interplay between habitual and goal-directed decision-making. We had three specific research questions in mind: first, does aging affect the balance between habitual and goal-directed decision mechanisms? Second, are age differences in the interplay of these decision mechanisms related to age differences in working memory (WM) capacity? Third, can model-based choice behavior be affected by manipulating the distinctiveness of the reward value of different choice options? To address these questions we adapted a two-state Markov decision task (Daw et al., 2011; Wunderlich et al., 2012) in combination with computational reinforcement learning (RL) modeling.

### MODEL-FREE AND MODEL-BASED DECISION-MAKING

The dissociation between habitual and goal-directed mechanisms is at the core of many current theories of learning and decision-making (Daw et al., 2005; Balleine and O'Doherty, 2010; Kahneman, 2011). Habitual or model-free learning refers to the acquisition of behavior based on associations between actions and effects: actions that are followed by reward are more likely to reoccur (Thorndike, 1911). Model-free learning is a robust and computationally efficient mechanism. However, it can come at the cost of being inflexible, especially in dynamically changing environments, which constrain the adaptive value of habitual responses (Doll et al., 2012). Computational accounts suggest that habitual learning is driven by the discrepancy between the current reward and the expected plus the (discounted) sum of all future rewards (i.e., the prediction error signal) (Sutton and Barto, 1998; Niv and Schoenbaum, 2008). Results from electrophysiological studies in animals and neuroimaging work in humans show that these reward predictions errors seem to be coded in phasic changes of dopaminergic activity in the midbrain and ventral striatum (Schultz et al., 1997; Montague et al., 2004; D'Ardenne et al., 2008; Niv et al., 2012).

In comparison, goal-directed or model-based decision-making reflects choices that are guided by internal goal representations or "cognitive maps" (Tolman, 1948; Miller and Cohen, 2001). These representations involve knowledge of the structure of the environment that can be used to make adaptive and foresighted decisions (Doll et al., 2012). One way of thinking about these representations is in terms of a decision space that represents the consequences of actions with respect to sequential transitions in the environment and possible future rewards. The advantage of model-based decision mechanisms is that they allow individuals to flexibly adjust behavior to changes in the environment. One downside of model-based decision-making is that it is computationally more expansive and effortful than the relatively more automatic habitual mechanisms. Recent neuroimaging work has started to investigate the neural mechanisms underlying model-based decision-making (Gläscher et al., 2010; Daw et al., 2011). Whereas results from Daw et al. (2011) suggest that model-based and model-free decisions may implicate overlapping neural systems, involving the ventral striatum and ventromedial prefrontal cortex (vmPFC), findings from Gläscher et al. (2010) show that the learning of new task structures may involve cortical areas such the lateral PFC and parietal cortex.

#### RELATIONS BETWEEN WORKING MEMORY CAPACITY AND MODEL-BASED DECISION-MAKING

Support for the idea that model-based decision-making relies on higher-order cognitive control mechanisms comes from a recent study that combined a two-state Markov decisions task with a concurrent WM paradigm (Otto et al., 2013). This study showed that high WM load resulted in a reduced degree of model-based behavior, suggesting that goal-directed decisions rely on WM functions and the associated neural systems. Similar results were obtained by Worthy and Maddox (2012). Using a dynamic decision-making task these authors showed that WM load seems to shift behavior from a heuristics-based win-stay lose-shift (WSLS) strategy toward a model-free (reinforcement-based) strategy (Worthy and Maddox, 2012). Taken together, evidence from these studies suggests that model-based decision-making may partially rely on WM function and that increasing WM demands may lead to a shift from model-based to model-free decision-making. Another implication from these findings is that model-based decision-making abilities can be understood as a limited cognitive resource. However, what remains unclear from these studies is the degree to which age and individual differences in WM capacity may be associated with differences in model-based behavior (Otto et al., 2013).

#### AGE DIFFERENCES IN LEARNING AND DECISION-MAKING

Results from recent studies on age differences in learning and decision-making suggest that older adults are impaired in learning from uncertain and ambiguous reward. This does not seem to be the case in situations in which reward information is fully predictable (deterministic). Electrophysiological results indicate that learning impairments are associated with deficits in error detection as well as less differentiated reward representations (Eppinger et al., 2008; Eppinger and Kray, 2011; Pietschmann et al., 2011; Hämmerer and Eppinger, 2012). Moreover, results from recent fMRI studies show that age-related impairments in RL are associated with a reduced correlation between reward prediction errors and ventral striatal activity in older than younger adults (Chowdury et al., 2013; Eppinger et al., 2013). In line with the idea of age-related changes in striatal prediction error signaling, Samanez-Larkin et al. (2010) found that suboptimal financial decision-making in older adults is associated with increased temporal variability of the ventral striatal BOLD signal. Taken together, these results are consistent with several theoretical proposals, suggesting that age-related impairments in reward-based learning might result from reduced dopaminergic projections from the midbrain to the ventral striatum and vmPFC (Nieuwenhuis et al., 2002; Frank and Kong, 2008; Hämmerer and Eppinger, 2012). However, it should be noted that there is also evidence indicating that age-related deficits in learning are, at least partially, mediated by decreased white matter integrity in fronto-striatal pathways (Samanez-Larkin et al., 2012).

Only a few studies so far have focused on age-related differences in more complex learning and decision-making (Mata et al., 2010; Worthy et al., 2011; Worthy and Maddox, 2012). In a recent study Worthy and Maddox (2012) used a dynamic decision-making task in which reward depended on the choice history. Results showed that older adults performed better on this task than younger adults. Using computational approaches the authors showed that this effect was due to the fact that older adults relied more on decision heuristics such as a win-stay lose-shift, whereas younger adults relied on RL. Findings by Mata et al. (2010) show that older adults perform poorer in a probabilistic inference task than younger adults if the decision environment favors the use of a cognitively demanding strategy. This is consistent with the idea that strategic, planning-related cognitive processes are a constrained resource, especially in older adults. Taken together, these results point to the view that age-related impairments in decision-making may depend on the complexity of the decision environment (Mata et al., 2010). Older adults may do well or even better than younger adults in tasks that favor the use of decision strategies with shorter temporal horizons, such as WSLS, whereas older adults may be impaired in decision-making if they have to use strategic, model-based processes.

Although the results of the previous studies may point to an age-related shift in the balance between model-based and model-free decision processes, the tasks applied in these studies do not allow to formally dissociate between these decision mechanisms. To address this question and to examined age and individual differences in model-free and model-based decision processes, we adapted a two-stage Markov decision task (cf. Daw et al., 2011) that allowed us to separate the contributions of these decisions mechanisms to choice behavior (see **Figure 1A**).

Specifically, we were interested in three major research questions (a) Does aging affect the balance between model-based and model-free decision-making mechanisms? (b) Are age-related changes in decision mechanisms related to age differences in WM capacity?; and (c) Can model-based behavior in older adults be supported by enhancing the distinctiveness of the reward value of the different choice options?

The two-stage Markov decision task consists of two decision stages in each trial (see **Figure 1A**). The first decision stage

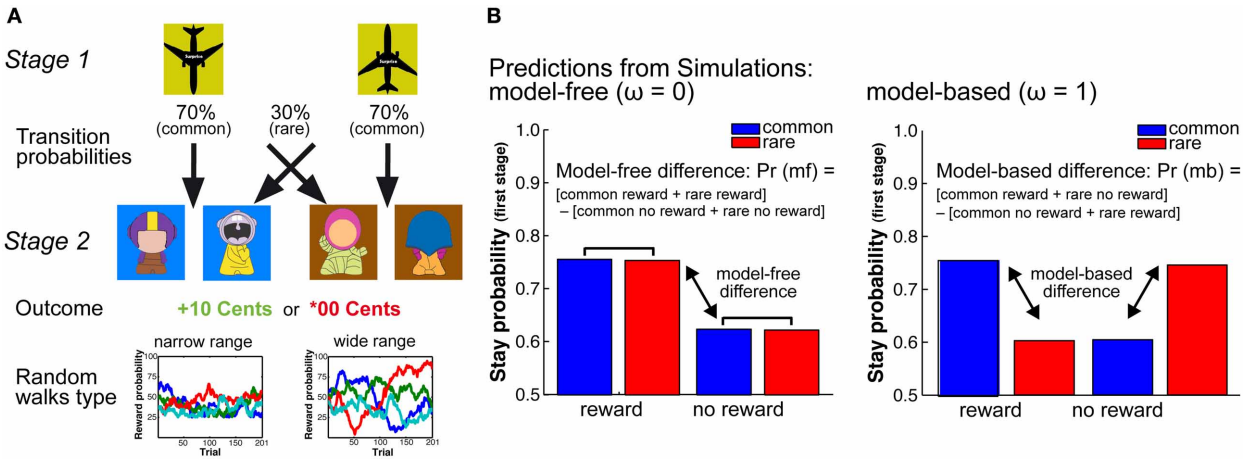



**FIGURE 1 | (A)** Schematic Figure of the 2-stage Markov decision task. In this task participants have to constantly update reward predictions on the second stage (model-free decision-making) and use these reward predictions to make goal-directed decisions on the first stage of the task. To support model-free learning in the two age groups we manipulated the random walks that determine the probability of getting a reward at the second stage. We applied two different probability ranges, one with a narrow range of reward probabilities (25–75% reward probability) and one with a wide range of reward probabilities (0–100% reward probability). **(B)** Model predictions. Left panel: Simulations show that model-free decision-making is reflected in a main effect of reward. That is, stay behavior on the first choice depends on whether behavior on the previous trial was rewarded or not. Model-free behavior is independent of the transition probability structure. Right panel: Model-based behavior is reflected in an interaction between transition on the previous trial and reward on the previous trial. That is, model-based behavior takes the model-free information as well as knowledge of the transition structure into account.

involves two choice options that are associated with different transition probabilities to the second-stage choice options that are then either rewarded or not rewarded. In this task, participants have to constantly update reward predictions at the second stage (model-free decision-making) and use this information prospectively to make goal-directed (model-based) decisions at the first decision stage on the next trial (see **Figure 1A**). To manipulate the demands on the representation and updating of reward value we varied the distinctiveness of the reward probabilities at the second stage (see **Figure 1A**). This was done by increasing the range of the reward probabilities of the choice options at the second stage on each trial. The larger the range of the reward probabilities of the four potential choice options, the easier it should be to differentiate and represent the reward histories associated with these options (cf. Eppinger et al., 2011). More differentiable reward probabilities on the second stage should support the ability to make deliberate, goal-directed decisions on the first stage and may hence be less demanding in terms of the representation of the stage transition structure of the task.

Given previous findings that point to age-related behavioral deficits in complex decision tasks (Mata et al., 2010) we expected older adults to be impaired in model-based decision-making compared to younger adults. Furthermore, given evidence for age-related impairments in learning from probabilistic outcomes (Eppinger et al., 2008; Hämmerer et al., 2011; Pietschmann et al., 2011), we predicted that older adults should benefit from more distinctive reward probabilities at the second stage of the task. That is, we should find enhanced model-based decision-making in the wide compared to the narrow probability range condition. To investigate the association between individual differences in WM and individual differences in model-based decision-making we also acquired a WM capacity using the operation span task (Turner and Engle, 1989; Unsworth et al., 2005). Given results suggesting that WM capacity is critical for model-based behavior we expected that higher WM capacity should be associated with model-based choice patterns (Otto et al., 2013). To the degree that this is the case, age-related deficits in model-based decision-making may be mediated by age-related decline in WM capacity (Salthouse et al., 1989; Salthouse, 1994).

## MATERIAL AND METHODS

### PARTICIPANTS

Sixty younger and sixty two older adults took part in the study. Two older adults had to be excluded because they were unable to perform the experimental tasks. Two younger adults were excluded because they did not return for the second testing session. Two further younger adults needed to be excluded, one due to technical problems during data acquisition and the other due to chance level performance in the WM task. Thus, the effective sample consisted of 56 younger adults (mean age: 24, age range 20–30 years, 27 females) and 60 older adults (mean age: 69, age range: 56–78 years, 27 females). Participants gave written informed consent. The Institutional Review Board of the Max-Planck Institute for Human Development approved the study. Participants completed a biographical and a personality questionnaire (Carver and White, 1994) as well as several psychometric tests: (1) Identical pictures test (Ekstrom et al., 1976); (2) Raven's Progressive matrices (Raven et al., 1998); (3) Spot-the-Word test (Baddeley et al., 1992). As shown in **Table 1** older adults had lower scores on the Identical Pictures test and Raven's matrices than younger adults ($p$'s $< 0.001$, $\eta^2 > 0.46$). In contrast, older adults obtained higher scores than younger adults on the Spot-the-Word test ($p < 0.001$, $\eta^2 = 0.24$). Consistent with previous findings from larger population-based samples (e.g., Li et al.,

**Table 1 | Psychometric variables displayed separately for the two age groups and the two working memory performance groups.**

| Age group | Younger adults | | Performance effect | Older adults | | Performance effect | Age effect |
|---|---|---|---|---|---|---|---|
| WM performance groups | Low (mean, *SE*) | High (mean, *SE*) | *p*-value, effect size | Low (mean, *SE*) | High (mean, *SE*) | *p*-value, effect size | *p*-value, effect size |
| Age | 25.2 (0.6) | 23.8 (0.6) | $p = 0.08$, $\eta^2 = 0.06$ | 69.1 (1.0) | 68.4 (0.9) | $p = 0.60$, $\eta^2 = 0.00$ | $p < 0.001$, $\eta^2 = 0.97$ |
| Raven | 11.3 (0.5) | 12.9 (0.5) | $p = 0.03$, $\eta^2 = 0.08$ | 5.7 (0.5) | 7.4 (0.7) | $p = 0.05$, $\eta^2 = 0.07$ | $p < 0.001$, $\eta^2 = 0.46$ |
| Ospan partial score | 20.7 (1.5) | 49.1 (1.9) | $p < 0.001$, $\eta^2 = 0.72$ | 9.2 (1.1) | 34.6 (2.1) | $p < 0.001$, $\eta^2 = 0.67$ | $p < 0.001$, $\eta^2 = 0.14$ |
| Ospan total score | 45.4 (2.0) | 63.6 (0.9) | $p < 0.001$, $\eta^2 = 0.56$ | 26.3 (2.2) | 53.3 (1.4) | $p < 0.001$, $\eta^2 = 0.66$ | $p < 0.001$, $\eta^2 = 0.20$ |
| Processing speed | 29.0 (1.0) | 33.8 (0.7) | $p < 0.001$, $\eta^2 = 0.25$ | 22.03 (0.6) | 21.5 (0.7) | $p = 0.55$, $\eta^2 = 0.00$ | $p < 0.001$, $\eta^2 = 0.57$ |
| Spot-a-word | 19.5 (1.1) | 20.4 (1.3) | $p = 0.58$, $\eta^2 = 0.00$ | 25.4 (1.0) | 26.7 (0.9) | $p = 0.33$, $\eta^2 = 0.02$ | $p < 0.001$, $\eta^2 = 0.24$ |
| BIS | 19.5 (0.7) | 18.82 (0.8) | $p = 0.50$, $\eta^2 = 0.00$ | 19.86 (0.6) | 17.93 (0.6) | $p = 0.03$, $\eta^2 = 0.08$ | $p = 0.69$, $\eta^2 = 0.00$ |
| BAS | 13.5 (0.3) | 13.6 (0.3) | $p = 0.82$, $\eta^2 = 0.00$ | 13.5 (0.4) | 13.0 (0.3) | $p = 0.29$, $\eta^2 = 0.02$ | $p = 0.48$, $\eta^2 = 0.00$ |

2004), these results suggest age-related reductions in fluid intelligence and age-related improvements in crystallized intelligence. We did not find significant age differences behavioral inhibition or approach (BIS/BAS) scores ($p$'s $> 0.48$) (Carver and White, 1994).

### MEDIAN SPLIT OF GROUPS BASED ON WORKING MEMORY CAPACITY

To examine the associations between individual differences in WM capacity on model-based and model-free decision-making we performed a median split for the operation span total score separately for the two age groups (Unsworth et al., 2005). High and low capacity groups did not differ significantly with respect to mean age (younger adults: $p = 0.08$, older adults: $p = 0.60$, see **Table 1**). However, as expected, given the well-documented positive association between WM capacity and fluid intelligence (Duncan et al., 2012), we found significantly higher Raven scores for high than low WM capacity groups in both age groups ($p$'s $< 0.05$; $\eta^2$'s $> 0.07$). Significant differences between high and low capacity in processing speed were observed for younger ($p < 0.001$; $\eta^2 > 0.25$), but not for older adults ($p = 55$). High and low WM capacity groups did not differ with respect to semantic knowledge in either age group ($p$'s $> 0.33$) (see **Table 1**).

### STIMULI

Stimuli on the first stage were two airplanes that either pointed to the top or the bottom of the screen, indicating the two different choice options. Stimuli on the second stage were 8 colored figures ("GoGos") that we generated using a freeware on the gogos-crazybones.com website and processed for presentation purposes in Photoshop (see **Figure 1A**). Background colors of the second stage stimuli were either blue or brown. Feedback stimuli either indicated a monetary gain of 10 Euro Cents, displayed in green or a neutral outcome of 00 Euro Cents, displayed in red (see **Figure 1A**). Stimuli were presented on a 19-inch computer screen using the program EPrime 2.0 software (PST Inc., Pittsburgh, PA).

### TASK

The task involves two decision stages. At the first stage participants had to make a decision between two choice options (the two airplanes), which occurred randomly either on the left or right side of the screen. This decision determined the transition to the next (second) stage (see **Figure 1A**). We refer to more likely (70%) transitions as common transitions and less likely (30%) transitions as rare transitions. Participant had to indicate their choice within 2 s of stimulus presentation using the "f" or "j" key on a standard computer keyboard. If no response occurred within 2 s the trial was aborted and a new trial started. At the second stage participants had to make another decision between two choice options (the GoGos), which were displayed randomly either on the left or the right side of the screen (see **Figure 1A**). This decision had to be made within 2 s of stimulus presentation using the same keys as in the first decision ("f" and "j"). If no response occurred within 2 s, three white question marks appeared on the screen for 1 s and the trial was aborted. Choices were either rewarded (+10 Cents) or not rewarded (*00 Cents). The probabilities of getting a reward were determined by Gaussian random walks (see **Figure 1A**). The feedback stimuli were displayed for 1 s. Before and after the feedback stimulus a fixation cross was displayed for 500 ms. Reward probabilities were determined by a slowly drifting random walk. At each trial we added Gaussian noise with a mean of 0 and standard deviation of 0.025 to the reward probabilities. To manipulate the demands on the updating of reward value representation we applied two types of random walks with different reflecting boundaries: in the narrow probability range condition the random walks had reflecting boundaries of 0.25 and 0.75 (Daw et al., 2011). In the wide probability range condition we increased the reflecting boundaries from 0.00 to 1.00 (for examples see **Figure 1A**). The broader reflecting boundaries in the wide probability range condition result in more differentiable random walks for the four second-stage options. Participants performed 201 trials with the narrow random walk and 201 trials with the wide random walk in two separate sessions.

To improve subject's understanding of the task structure we designed a cover story for the task. The cover story is about a businessman who has to decide between two airplanes each of which will bring him to one of two islands (see **Figure 1A**). The airline is called "Surprise" and is somewhat unreliable with respect to its destinations (the transition probabilities are made explicit and are practiced by the participants). At each of the islands the businessman can trade with one of two populations of inhabitants (represented by the GoGo Figures). The productivity of the

populations changes across time. The task of the businessman is to make as much money as possible by integrating information about the reward probability on the second stage and the transition structure on the first stage.

### PROCEDURE

Participants performed two sessions, which were separated by a minimum of 1 week and a maximum of 3 weeks. In the first session participants performed a demographic questionnaire, the BIS/BAS personality questionnaire (Carver and White, 1994), Raven's progressive matrices (Raven et al., 1998) and one version of the two-stage Markov decision task (either with narrow or wide probability range condition). In the second session subjects performed an automated version of the Operation Span Task (Unsworth et al., 2005), Spot-a-Word and the Identical pictures test (Li et al., 2004), a version of the two-stage Markov decision task and an additional experimental task, data of which will be presented elsewhere. Half of the participants in each group performed the narrow probability range condition first and vice versa for the second half of the samples. Participants were informed about the nature of the transition probability structure. We also explained (and showed) to the subjects that the likelihood of getting a reward at the second stage varies over time and differs between sessions.

Prior to the task in the first session, participants completed a computerized training session, which was supervised by instructed student research assistants. In the first part of the training participants were introduced to the reward probability structure of the second (model-free) stage of the task. To familiarize participants with probabilistic reward they had to first perform 10 choices between options with a fixed reward probability of 60%. To support the understanding of probabilistic information we always referred to reward probabilities in terms absolute numbers (i.e., getting reward in 6 of 10 cases). Thereafter, participants were given 20 additional trials, in which they had to find the option with the highest reward probability (out of four choice options). After making sure that everyone found the best option we explained that the reward probabilities would change slowly across the experiment. For illustration purposes two examples of the random walks (see **Figure 1**) were shown in a graph.

In the next training phase participants were introduced to the transition probability structure on the first stage. That is, we informed them about the fact that there are common (characteristic) and rare (uncharacteristic) transitions and showed them a graphical picture of the transition structure (similar to **Figure 1**). Then participants performed 20 trials in which they practiced the transitioning from the first stage options to the second stage options (without receiving a reward). Finally, subjects played 30 trials of the experimental task (involving all stages as well as probabilistic rewards) using a different stimulus set [for similar procedures see Daw et al. (2011)]. Before the task in the second session participants performed a short practice session of 20 min. Reward was accumulated across sessions and participants were compensated according to their earnings in the task.

### DATA ANALYSIS

Stay-switch behavior at the first stage was analyzed using Matlab (MATLAB, Mathworks Inc, Natick, MA) and SAS (SAS Institute Inc, Cary, NC). We defined stay-switch behavior as the probability to repeat a choice on the first stage as a function of the transition (common or rare) and the outcome (reward, no reward) on the previous trials. Mean stay probabilities were analyzed using a repeated measures ANOVA with the between subjects factors Age Group (younger, older) and WM capacity (high, low), as well as the within subjects factors probability range (narrow, wide), previous transition type (common, rare) and previous outcome (reward, no-reward). For follow-up analyses we calculated differences measures for model-free behavior [(common reward + rare reward) − (common no reward + rare no reward)] and model-based behavior [(common rewarded + rare unrewarded) − (rare rewarded + common unrewarded)] (see **Figure 1B**). The model-based and model-free difference values were analyzed using an ANOVA with the factors age group and range of reward probability.

### COMPUTATIONAL MODEL

Choice behavior was fit using a hybrid RL algorithm (Daw et al., 2011; Wunderlich et al., 2012). This algorithm assumes that choices on the first stage of the task are driven by a weighted combination of model-based RL, which accounts for the transition structure, and model-free SARSA (λ) TD learning. The weighting of model-based vs. model-free decision mechanism is determined by the free parameter omega, ω, which is held constant across trials and is constrained from 0 to 1. If ω approaches 0 behavior is model-free, which is reflected in a main effect of reward (see **Figure 1B**). In contrast, an omega close to 1 indicates model-based choice behavior, which is reflected in an interaction between transition structure and reward on the previous trial (see **Figure 1B**). Participants are assumed to select actions stochastically according to a softmax function. The choice probabilities were determined by the state-action values. For the model-fit we estimated the free parameters of the hybrid model for each probability range and subject individually via maximum likelihood. We first iterated all parameters individually by using a grid search to get a rough estimate. Subsequently, we extracted the twelve best fitting parameter combinations of both probability ranges and entered them as starting points for a precise parameter estimation, using Matlab routine fMincon.

The task consists of two stages and three states (first stage: $S_A$; Second stage: $S_B$, $S_C$) (see **Figure 1A**). Each state is associated with two actions ($a_A$, $a_B$). At both stages (i) a state-action value function $Q_{Si}(a)$ is learned that maps each state action pair to its expected value. We refer to the model-based value function at the first stage as $Q_{S1^{MB}}$ and to the model-free value function as $Q_{Si^{MF}}$.

### MODEL-FREE STATE ACTION VALUES

Model-free state action values at the second stage were updated using SARSA(λ) temporal difference learning (Rummery and Niranjan, 1994). The state-action pairs were updated in each trial t according to:

$$Q_{S2^{MF}}(a, t+1) = Q_{S2^{MF}}(a,\ t) + \alpha_2(r(t) - Q_{S2^{MF}}(a, t))$$

where $\alpha_i$ is the learning-rate at a given stage (here stage 2) and $r(t)$ is the received reward in that trial.

The state-action value and the reward at the second stage are then used to update the model-free values for the next choice at the *first stage* of the next trial. This updating mechanism followed the same temporal difference learning rule, with an additional parameter, λ allowing eligibility traces:

$$Q_{S1^{MF}}(a, t+1) = Q_{S1^{MF}}(a, t) + \alpha_1 \left(Q_{S2^{MF}}(a_{\text{chosen}}, t) - Q_{S1^{MF}}(a, t)\right) + \alpha_1 \lambda \left((r(t) - Q_{S2^{MF}}(a, t)\right)$$

Note that eligibility traces are not assumed to carry over from trial to trial. The reason for this is the task structure that involved changing reward probabilities (the random walks) for each option across trials.

#### MODEL-BASED STATE-ACTION VALUES

Model-based state-action values are computed using Bellman's equation by taking the model-free state-action values from the second stage and the transition probabilities into account.

$$\begin{aligned} Q_{S1^{MB}}(a_1) &= \text{HighTran} * \max\left[Q^{MF}_{S2-B}(a)\right] \\ &\quad + \text{LowTran} * \max\left[Q^{MF}_{S2-C}(a)\right] \\ Q_{S1^{MB}}(a_2) &= \text{LowTran} * \max\left[Q^{MF}_{S2-B}(a)\right] \\ &\quad + \text{HighTran} * \max\left[Q^{MF}_{S2-C}(a)\right] \end{aligned}$$

In this equation "HighTran" is defined as the highest transition probability of the current condition (0.7) and "LowTran" is defined as the lowest transition probability of that condition (0.3). Before each block participants were explicitly instructed about the nature of the transition probabilities and practiced the transitioning between states.

In the full hybrid model the $Q_{\text{net}}$ state-action value was calculated as the weighted sum of model-based and model-free values:

$$Q_{S1^{\text{net}}} = \omega * Q_{S1^{MB}}(a) + (1-\omega) * Q_{S1^{MF}}(a)$$

where ω is the weighting parameter. At the second stage the $Q_{\text{net}}$ state-action value is equal to the model-free state-action value ($Q_{S2^{\text{net}}} = Q_{S2^{MF}}$).

#### SOFTMAX RULE

Choice probabilities at the first stage were calculated according to a softmax rule:

$$P_{S_i}(a_{1,t}) = \frac{\exp\left(\beta_i * \left[Q_{S_1^{\text{net}}}(a_{1,t}) + \pi * rep(a_1)\right]\right)}{\left(\exp\left(\beta_i * \left[Q_{S_1^{\text{net}}}(a_{1,t}) + \pi * rep(a_1)\right]\right) + \exp\left(\beta_i * \left[Q_{S^{\text{net}}_1}(a_{2,t}) + \pi * rep(a_2)\right]\right)\right)}$$

where $\beta_i$ is the inverse softmax temperature paramter controlling the distinctiveness of the choices. We allowed both learning parameters ($\alpha_1, \alpha_2$) and the softmax temperature parameters ($\beta_1$, $\beta_2$) to differ between the two stages. The indicator function rep(a) is defined as 1 if *a* is a top-stage action and is the same as was chosen on the previous trial, zero otherwise. Taken together, the function rep(a) and the parameter π capture the degree of perseveration at the first-stage ($\pi > 0$) or the switching ($\pi < 0$) at first-stage options (Lau and Glimcher, 2005).

Choice probabilities at the second stage were calculated as follows:

$$P_{S_i}(a_{1,t}) = \frac{\exp\left(\beta_i * Q_{S_2^{\text{net}}}(a_{1,t})\right)}{\left(\exp\left(\beta_i * Q_{S_2^{\text{net}}}(a_{1,t})\right)\right) + \left(\exp\left(\beta_i * Q_{S_2^{\text{net}}}(a_{2,t})\right)\right)}$$

The model contained 7 free parameters ($\alpha_1$, $\alpha_2$, $\beta_1$, $\beta_2$, π, λ, ω). The median parameter values are shown separately for the two age groups, the two performance groups and the two probability ranges in **Table 2**.

#### MODEL FITS

An ANOVA with the factors age group, performance group and probability range on the negative log-likelihoods (-LL) showed no significant age differences in the model-fits ($p = 0.79$) and no significant difference between WM capacity groups ($p = 0.16$) (see **Table 2**). However, we obtained a significant main effect of probability range, indicating better model fits for the wide compared to narrow probability range ($p = 0.004$, $\eta^2 = 0.07$). Taken together, these results show comparable model fits for younger and older adults as well as for high compared to low WM capacity groups. Furthermore, as shown in **Table 2** the model fits as well as the parameter estimates are comparable with those of previous studies (Daw et al., 2011; Wunderlich et al., 2012).

## RESULTS

#### OVERALL TASK PERFORMANCE

To examine age differences in overall task performance we calculated the mean payoffs (earnings), separately for each individual and probability range condition. Mean payoffs were analyzed using an ANOVA with the between subject factors Age group, WM capacity and probability range condition. As shown in **Figure 2C** the analysis showed higher mean payoffs for younger compared to older adults ($p = 0.03$, $\eta^2 = 0.04$). Furthermore, participants earned more money in the wide probability range condition compared to the narrow probability condition ($p < 0.001$, $\eta^2 = 0.26$, see **Figure 2C**). No significant interaction effects were obtained ($p$'s $> 19$).

#### EFFECTS OF AGE GROUP ON MODEL-BASED BEHAVIOR

The overall ANOVA revealed a significant interaction between age group, transition type, and outcome $F_{(1,112)} = 29.66$, $p < 0.001$, $\eta^2 = 0.14$. To test whether there are significant age differences in decision strategies (model-based vs. model-free) we ran an ANOVA with the factors age group, WM capacity, probability range condition and decision strategy (model-based, mb vs. model-free, mf). This analysis revealed a significant interaction between age group and decision strategy $F_{(1,\ 112)} = 17.41$, $p < 0.001$, $\eta^2 = 0.12$. Separate analyses for the two decision strategies revealed significant age differences for model-based decision-making ($t < 4.04$, $p < 0.001$, $\eta^2 = 0.20$), but not for model-free decision-making ($t = -1.31$, $p = 0.19$).

**Table 2 | Computational model parameters (Median, 25th and 75th percentile) as a function of Age group (Younger older adults) and WM Performance group (low performers, high performers) and Transition probability range (narrow, wide).**

| Age/performance | Probability range | Parameter | $\beta_1$ | $\beta_2$ | $\alpha_1$ | $\alpha_2$ | $\lambda$ | $\pi$ | $\omega$ | −LL |
|---|---|---|---|---|---|---|---|---|---|---|
| *Younger adults Low performance group* | Narrow | 25th percentile | 2.33 | 2.64 | 0.24 | 0.30 | 0.05 | 0.10 | 0.09 | 180.04 |
| | | Median | 3.91 | 4.71 | 0.63 | 0.49 | 0.23 | 0.21 | 0.57 | 208.55 |
| | | 75th percentile | 8.90 | 5.81 | 0.85 | 0.69 | 0.52 | 0.42 | 0.69 | 231.83 |
| | Wide | 25th percentile | 2.63 | 2.83 | 0.26 | 0.37 | 0.03 | 0.06 | 0.15 | 170.27 |
| | | Median | 5.12 | 4.02 | 0.52 | 0.56 | 0.17 | 0.16 | 0.57 | 200.94 |
| | | 75th percentile | 10.37 | 5.21 | 0.86 | 0.73 | 0.62 | 0.33 | 0.79 | 220.00 |
| *Younger adults High performance group* | Narrow | 25th percentile | 3.96 | 3.67 | 0.04 | 0.52 | 0.00 | 0.05 | 0.21 | 226.90 |
| | | Median | 7.63 | 4.82 | 0.39 | 0.64 | 0.09 | 0.11 | 0.61 | 191.52 |
| | | 75th percentile | 10.73 | 6.04 | 0.67 | 0.86 | 0.89 | 0.167 | 0.74 | 171.43 |
| | Wide | 25th percentile | 3.46 | 6.57 | 0.07 | 0.39 | 0.03 | 0.09 | 0.34 | 159.39 |
| | | Median | 9.99 | 4.59 | 0.40 | 0.65 | 0.15 | 0.12 | 0.64 | 183.55 |
| | | 75th percentile | 12.39 | 3.78 | 0.58 | 0.85 | 0.36 | 0.24 | 0.78 | 214.80 |
| *Older adults Low performance group* | Narrow | 25th percentile | 2.41 | 1.83 | 0.18 | 0.05 | 0.09 | 0.09 | 0.08 | 249.69 |
| | | Median | 5.83 | 2.92 | 0.65 | 0.33 | 0.33 | 0.18 | 0.24 | 218.33 |
| | | 75th percentile | 9.44 | 5.24 | 0.95 | 0.87 | 0.90 | 0.27 | 0.67 | 188.11 |
| | Wide | 25th percentile | 2.64 | 1.43 | 0.04 | 0.08 | 0.02 | 0.05 | 0.15 | 161.15 |
| | | Median | 6.12 | 3.36 | 0.33 | 0.44 | 0.33 | 0.14 | 0.35 | 208.76 |
| | | 75th percentile | 13.54 | 4.91 | 0.82 | 0.88 | 0.83 | 0.29 | 0.49 | 234.82 |
| *Older adults High performance group* | Narrow | 25th percentile | 2.19 | 1.98 | 0.04 | 0.11 | 0.11 | 0.04 | 0.01 | 170.26 |
| | | Median | 3.62 | 2.69 | 0.40 | 0.56 | 0.38 | 0.18 | 0.15 | 202.09 |
| | | 75th percentile | 6.42 | 6.20 | 0.84 | 0.90 | 0.82 | 0.35 | 0.48 | 246.27 |
| | Wide | 25th percentile | 2.52 | 2.10 | 0.09 | 0.06 | 0.08 | 0.04 | 0.01 | 148.89 |
| | | Median | 6.32 | 3.75 | 0.62 | 0.32 | 0.49 | 0.18 | 0.11 | 187.34 |
| | | 75th percentile | 8.98 | 5.89 | 0.91 | 0.69 | 0.80 | 0.35 | 0.48 | 232.80 |

To further examine age differences in model-based behavior we performed separate analyses for the factors transition type and reward. These analyses showed a significant effect of age group only on rare rewarded trials ($t = 2.80$, $p = 0.006$, $\eta^2 = 0.07$). To confirm that the age effect is specific to rare rewarded trials rather than rare unrewarded trials we performed a *post-hoc* contrast between the two conditions and tested for age differences. We obtained a significant interaction between age group and reward. $F_{(1,\ 112)} = 17.41$, $p < 0.001$, $\eta^2 = 0.12$, indicating that older adults show enhanced stay behavior after rare transition that lead reward than younger adults, whereas age groups don't differ in their behavior after rare transitions that are followed by no reward (see **Figure 2B**).

Thus, age-related deficits in decision-making seem to be particularly pronounced if participants receive an unexpected reward after an uncharacteristic transition. In such a situation younger adults tend to switch to the other first stage choice option because this option is more reliably associated with the stimulus that was rewarded on the previous trial. In contrast, older adults tend to perseverate on options that were rewarded, independently of whether the reward was preceded by a common or rare transition (see **Figure 2A**).

### EFFECTS OF WM CAPACITY AND AGE DIFFERENCES IN MODEL-BASED BEHAVIOR

The overall analysis also revealed a significant interaction between age group, WM capacity, transition type and outcome, $F_{(1,112)} = 10.14$, $p = 0.002$, $\eta^2 = 0.04$. Separate analyses for the two age groups showed a significant interaction between WM capacity, transition type and outcome for younger adults ($p = 0.007$, $\eta^2 = 0.09$), but not for older adults ($p = 0.20$). Analyses of the difference values showed enhanced model-based choice behavior in high WM capacity younger adults compared to low WM capacity younger adults ($t = 2.83$, $p = 0.007$, $\eta^2 = 0.14$) but no effects of WM capacity in older adults ($t = -1.30$, $p = 0.20$) (see **Figure 3A**). Follow-up analyses for the factors WM capacity, transition type and reward revealed greater switching behavior after rare transitions that were followed by reward for high performing younger compared to high performing older adults ($t = 2.59$, $p = 0.01$, $\eta^2 = 0.11$). As shown in **Figure 3B**, younger adults with high WM capacity show enhanced switching behavior after rare transitions that were followed by reward.

### EFFECTS OF PROBABILITY RANGE ON MODEL-BASED BEHAVIOR

Furthermore, the overall ANOVA revealed a significant interaction between probability range, transition type and outcome

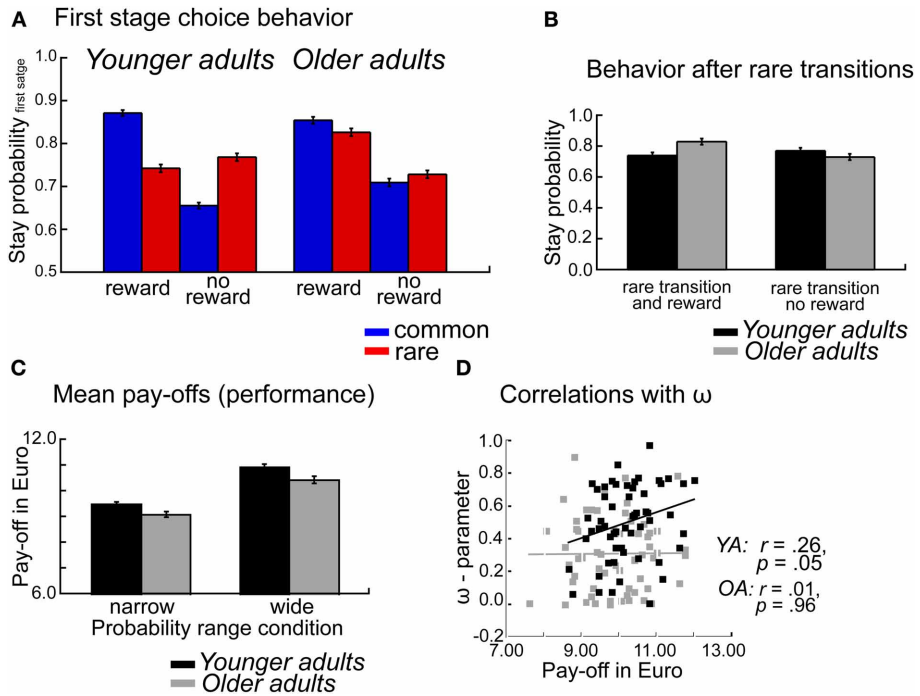



**FIGURE 2 | (A)** Probability of repeating the same first stage choice as a function of the transition on the previous trials (common, rare transition) and the outcome received on the previous trial (reward, no reward). Stay probabilities are displayed separately for younger adults (left panel) and older adults (right panel), error bars reflect the standard error of the mean (s.e.m.). **(B)** Age differences in stay behavior after rare transitions as a function of age group and reward on the previous trial. **(C)** Mean pay-offs in Euro per session, displayed separately for the factors Age group and Probability range condition. **(D)** Correlations between mean pay-offs in Euro and degree of model-based choice behavior (ω-parameter).

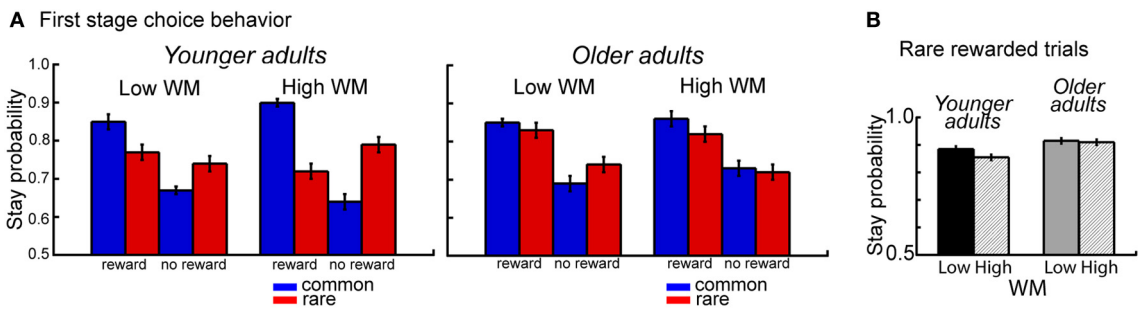



**FIGURE 3 | (A)** Probability of repeating the same first stage choice as a function of the transition on the previous trials (common, rare transition) and the outcome received on the previous trial (reward, no reward). Stay probabilities are displayed separately for younger adults (left panel) and older adults (right panel) and low and high WM capacity groups. Error bars reflect the standard error of the mean (s.e.m.). **(B)** Age differences in stay behavior after rare transitions and reward on the previous trial as a function of age group and WM capacity.

$F_{(1,112)} = 5.16$, $p = 0.02$, $\eta^2 = 0.05$, as well as between WM capacity, probability range, transition type and outcome $F_{(1,\ 112)} = 4.24$, $p = 0.04$, $\eta^2 = 0.04$. Separate analyses for the factor WM capacity showed a significant interaction between probability range, transition type and reward for high WM capacity groups ($p < 0.006$, $\eta^2 = 0.12$) but not for low capacity groups ($p = 0.60$). Analyses of the difference values showed enhanced model-based choice behavior in the wide probability range compared to the narrow probability range for high capacity groups ($t = 2.76$, $p = 0.008$, $\eta^2 = 0.13$) but not for low WM capacity groups ($t = 0.18$, $p = 0.86$) (see **Figure 4**). Separate analyses for the factors transition type and reward showed a significant main effect of probability range only for common rewarded trials ($t = 2.97$, $p < 0.004$, $\eta^2 = 0.08$). These results suggest that the effects of probability range on model-based behavior are primarily driven by enhanced stay behavior after common transitions that were followed by reward.

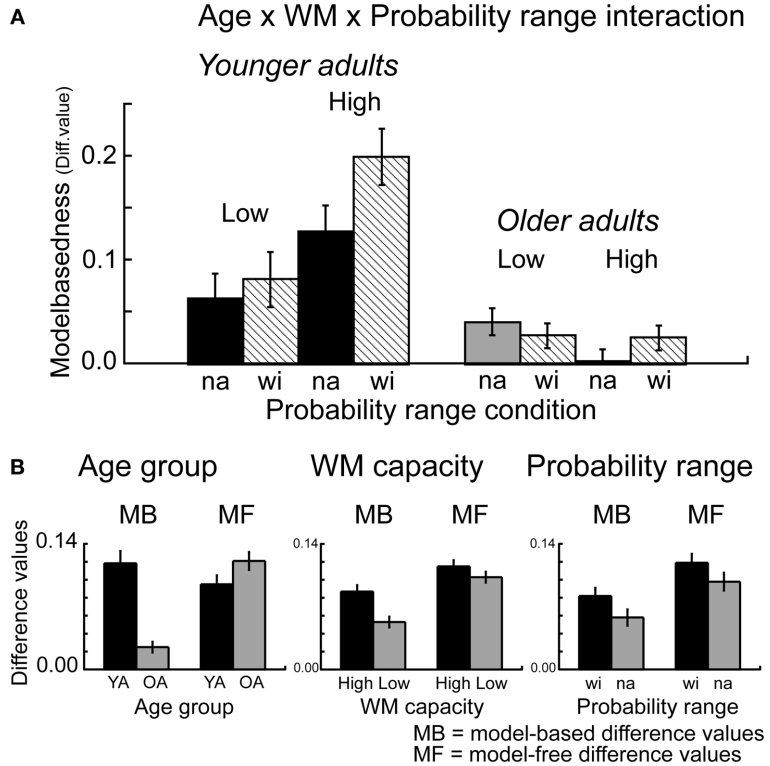



**FIGURE 4 | (A)** Difference values (proportion stay trials) for model-based behavior [(common rewarded + rare unrewarded) − (rare rewarded + common unrewarded)]. Model-based differences values are shown separately for the factors Age group (younger, older adults), WM capacity (high, low capacity) and probability range (narrow, wide probability range). Error bars reflect the standard error of the mean (s.e.m.). **(B)** Model-based (MB) and model-free (MF) differences values, displayed separately for the factors Age group, WM capacity and Probability range condition. Error bars reflect the standard error of the mean (s.e.m.).

#### GENERALIZED LINEAR MIXED MODEL ANALYSIS

Given that the first stage choice proportions are binomial data (and may hence not be normally distributed) we also used a mixed logit model (mixed effects logistic regression) as implemented in the lme4 package (Bates et al., 2013) in the statistical software R (R Development Core and Team, 2010) to fit choice behavior [see also Daw et al., 2011]. The design involved the same factor structure as for the repeated measures ANOVA. The analysis revealed qualitatively similar results as the overall results from ANOVA described above. We obtained a significant interaction between age group, transition type and outcome ($p < 0.001$), reflecting greater model-based choice behavior in younger than older adults (see **Figure 2** and **Table 3**). We also found a significant interaction between age group, WM capacity, transition type and outcome ($p < 0.001$). Separate analyses for the two age groups showed enhanced model-based behavior in high compared to low WM capacity groups in younger adults ($p < 0.001$) but not in older adults ($p = 0.69$), (see **Figure 3A**). Furthermore, we obtained a significant interaction between the factors probability range, transition type and outcome ($p < 0.01$). As shown in **Figure 4**, model-based behavior seems to be more pronounced in the wide compared to the narrow probability range condition. Taken together, the results of the mixed effects logistic regression are qualitatively consistent the results of the repeated measures ANOVA.

#### WM CAPACITY COVARIANCE ANALYSIS

To further analyze the effects of WM capacity on age differences in model-based and model-free decision-making we performed analyses of covariance (ANCOVA) with the between-subjects factor age group, within-subjects factor probability range and the (continuous) covariate WM capacity. For model-based differences values the analysis showed a significant main effect of age group $F_{(1,\ 112)} = 41.96$, $p < 0.001$, $\eta^2 = 0.15$. Importantly, this effect remained significant after controlling for WM capacity $F_{(1,112)} = 4.39$, $p = 0.03$, $\eta^2 = 0.02$, indicating that additional factors contributes to age differences in model-based behavior beyond the effects of WM capacity. Furthermore, we obtained a significant interaction between age group and WM capacity $F_{(1,\ 112)} = 12.19$, $p < 0.001$, $\eta^2 = 0.04$. Separate analyses for the two age groups showed a significant effect of WM capacity for younger adults ($t = 3.06$, $p = 0.002$, $\eta^2 = 0.08$), but not for older adults ($t = 0.69$, $p = 0.87$). No significant effects of age group, WM capacity or probability range were obtained for model-free differences values ($p$'s $> 0.14$). Taken together, these results line up with the findings of the median split analysis and show that in younger adults enhanced WM is associated with

**Table 3 | Results of the logistic regression analysis with the between subjects factors Age group and Working memory capacity and the within subjects factors Probability range condition (walk), transition type and reward.**

| Predictor | Estimate | *p*-value |
|---|---|---|
| (Intercept) | 1.35 | 0.08 |
| Walk | −0.18 | 0.18 |
| Transition | 0.21 | 0.10 |
| Reward | 0.26 | 0.05 |
| Age | 0.08 | 0.87 |
| WM | −0.17 | 0.73 |
| Walk × transition | 0.06 | 0.67 |
| Walk × reward | −0.06 | 0.67 |
| Transition × reward | 1.62 | <0.001 |
| Age × walk | 0.12 | 0.15 |
| Age × transition | 0.00 | 0.96 |
| Age × reward | 0.08 | 0.31 |
| WM × walk | 0.11 | 0.20 |
| WM × transition | −0.07 | 0.43 |
| WM × reward | −0.02 | 0.84 |
| Age × WM | 0.09 | 0.77 |
| Walk × transition × reward | −0.33 | 0.01 |
| Age × walk × transition | −0.02 | 0.85 |
| Age × walk × reward | 0.05 | 0.52 |
| Age × transition × reward | −0.78 | <0.001 |
| WM × walk × transition | −0.05 | 0.51 |
| WM × walk × reward | 0.06 | 0.47 |
| WM × transition × reward | −0.64 | <0.001 |
| Age × WM × walk | −0.10 | 0.06 |
| Age × WM × transition | −0.02 | 0.61 |
| Age × WM × reward | 0.00 | 0.99 |
| Age × walk × transition × reward | 0.12 | 0.16 |
| WM × walk × transition × reward | 0.12 | 0.15 |
| Age × WM × walk × transition | 0.02 | 0.67 |
| Age × WM × walk × reward | −0.06 | 0.20 |
| Age × WM × transition × reward | 0.33 | <0.001 |
| Age × WM × walk × transition × reward | −0.03 | 0.55 |

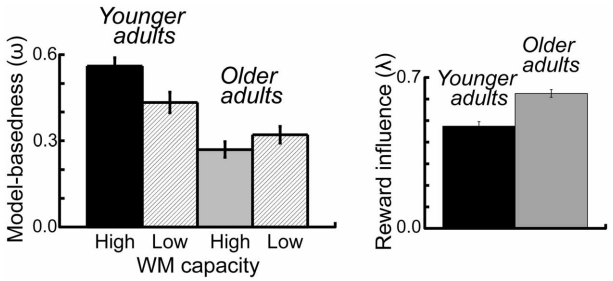



**FIGURE 5 | Left panel:** Mean parameter estimates for the omega (**ω**-) parameter, displayed separately for the two age groups and the two WM capacity groups. The (ω-) parameter reflects the relative contribution model-based and model-free mechanisms to first stage choice behavior. **Right panel**: Mean parameter estimates for the lambda (λ-) parameter, displayed separately for the two age groups. The (λ-) parameter reflects the direct influence of reward on the previous trial on first stage choice behavior. Error bars reflect the standard error of the mean (s.e.m.).

greater model-based behavior, which is not the case in older adults. Importantly, the results also show that age differences in model-based behavior remain, even after controlling for WM capacity.

### MODELING RESULTS

To examine the effects of age group and WM capacity on the model parameters we applied an ANOVA with the between subjects factors age group and performance group and the within subjects factor probability range. For the ω-parameter, we found a significant main effect of age group $F_{(1,112)} = 20.42, p < 0.001$, $\eta^2 = 0.15$, As shown in **Figure 5** younger adults showed a higher degree of model-based decision-making (as reflected in the ω-parameter) than older adults. Furthermore, we obtained a significantly greater λ-parameter for older compared to younger adults, $F_{(1,\ 112)} = 7.72, p = 0.006, \eta^2 = 0.06$ (see **Figure 5**). This finding indicates that in older adults the reward received on the second stage has a greater impact on choice behavior on the first stage than this is the case in younger adults. For the inverse temperature parameter on the first stage ($\beta_1$) we found a significant interaction between age group and WM capacity, $F_{(1,\ 112)} = 5.13$, $p = 0.03$, $\eta^2 = 0.04$. Separate analyses for the two age groups showed a more differentiated choice pattern in high WM capacity compared to low WM capacity younger adults ($t = 2.02$, $p = 0.05$, $\eta^2 = 0.07$) but no effect of WM capacity in older adults ($t = -1.19$, $p = 0.24$), (see **Table 2**). Finally, we obtained a significant main effect of age group on the learning rate $\alpha_1$ on the second stage $F_{(1,\ 112)} = 4.2$, $p = 0.04$, $\eta^2 = 0.04$. As shown in **Table 2**, younger adults had a lower learning rate on the second stage than older adults, indicating that they update value representation less rapidly than older adults.

## DISCUSSION

In this study we examined age-related and individual differences in habitual (model-free) and goal-directed (model-based) decision-making. Specifically, we were interested in three major questions: (a) Does aging affect the balance between model-based and model-free decision mechanisms? (b) Are age-related changes in decision mechanisms related to age differences in WM capacity, and (c) Can model-based behavior be supported by manipulating the distinctiveness of the reward value of the different choice options? To examine these questions, we used a two-stage Markov decision task that allows us to separate the contributions of model-free and model-based decision processes to choice behavior (Daw et al., 2011; Wunderlich et al., 2012). To support model-based behavior in this task we manipulated the range of the reward probabilities associated with the different options on the second stage (see **Figure 1A**). More differentiable reward probabilities on the second stage should support the ability to make deliberate, goal-directed decisions on the first stage and may hence be protective against age-related deficits in model-based decision-making. Furthermore, we acquired a WM capacity measure to investigate the impact of WM capacity on individual differences in model-based decision-making automated operation span, (Unsworth et al., 2005). Based on these WM capacity

scores we separated younger and older samples into high and low WM groups.

### AGE-RELATED IMPAIRMENTS IN MODEL-BASED DECISION-MAKING

The analysis of the stay-switch behavior on the first stage revealed significant impairments in model-based decision-making in older adults (see **Figure 1A**). In contrast, no significant age differences in model-free decision-making were obtained (see **Figures 1A**, **4B**). An analysis of age differences in the model-parameters supports these findings by showing a significant age-related reduction of the ω- parameter, which reflects the relative contribution of model-based compared to model-free decision processes to choice behavior on the first stage of the task (see **Figure 5**). An analysis of overall task performance showed higher mean pay-offs in younger than older adults, indicating that the more model-based strategy in younger adults is beneficial in terms of overall performance (**Figure 2C**). Furthermore, correlation analyses showed that in younger adults greater model-based behavior is associated with higher mean pay-offs. This is not the case in older adults (see **Figure 2D**). Thus, older adults who engage in a more model-based strategy do not seem to benefit from it in terms of overall performance. One interpretation of this effect might be that even though those older adults make strategic decisions on the first stage, they do not consistently choose the option with the highest expected value on the second stage. That is, overall deficits in task performance in older adults may reflect problems in the integration of model-free and model-based information.

Interestingly, age-related deficits in model-based decision-making seem to be particularly pronounced if participants receive an unexpected reward after an uncharacteristic transition and have to revise their decision strategy (see **Figure 2B**). In such a situation younger adults tend to switch to the other first stage option because this option is more reliably associated with the stimulus that was rewarded on the previous trial. This switching behavior can be understood in terms of a model-based exploration in which the younger adults switch to a state that may offer a greater probability of reward than the one they currently exploit. In contrast, older adults tend to perseverate on options that were rewarded, independently of whether the reward was preceded by a common or rare transition. Therefore, the current results suggest that older adults have deficits in applying their knowledge of the task structure if the reward on the previous trial reinforces stay behavior, whereas the fact that it was an uncharacteristic transition indicates the need for a shift in the response strategy on the first stage.

This interpretation is supported by two results of the modeling analysis: first, older adults show a higher λ- parameter than younger adults (see **Figure 5**). The λ- parameter reflects the direct influence of reward on the previous trial on stay-switch behavior on the first stage. That is, a high λ- parameter in older adults indicates that their choice behavior on the first stage is primarily influenced by the outcome on the previous trial rather than their representation of the expected value of the choice options on the previous trial. Second, we found a higher learning rate for older than younger adults on the second stage of the task. This result indicates that older adults are less consistent in their choice behavior on the second stage of the task, which may lead to deficits in building up differentiated reward value representations. Thus, our results are in line with previous findings that point to age-related impairments in the representation and updating of the expected value of choice options during RL (Eppinger et al., 2008; Eppinger and Kray, 2011; Hämmerer et al., 2011; Pietschmann et al., 2011). Furthermore, our findings line up with data from neuroimaging studies, which indicate that impairments RL in older adults are associated with age-related deficits in striatal reward prediction error signaling (Chowdury et al., 2013; Eppinger et al., 2013). However, it seems also plausible that age-related deficits in model-based decision-making are due to more complicated neuromodulatory effects in higher-order cortical areas, particularly the ventromedial and lateral prefrontal cortex. Consistent with such view, recent findings from Samanez-Larkin et al. (2012) suggest that age-related deficits in reward-based learning are, at least partially, mediated by decreased white matter integrity in fronto-striatal pathways (Samanez-Larkin et al., 2012).

Taken together, the current results suggest that age-related impairments in the updating of reward value representations may lead to deficits in goal-directed decision-making in older adults. These deficits are particularly pronounced if reward on the previous trial reinforces stay behavior, whereas the fact that it was an uncharacteristic transition indicates the need for a shift in the response strategy on the first stage. In these situations younger adults use their knowledge of the task structure to engage in strategic exploratory behavior, whereas older adults perseverate on the option they are currently exploiting.

### EFFECTS OF WM CAPACITY AND AGE GROUP ON MODEL-BASED BEHAVIOR

To examine the effects of individual differences in WM capacity on model-based behavior in the two age groups, we acquired a WM measure automated operation span, (Unsworth et al., 2005) and subdivided the younger and older adult samples into high and low WM capacity groups. We found enhanced model-based behavior for high capacity compared to low capacity younger adults, but no effect of WM capacity in older adults (see **Figure 3A**). Moreover, similar to the age-effects on model-based behavior, WM capacity-related differences in younger adults were most pronounced in switching behavior after rare transitions that were followed by reward (see **Figure 3B**). These results suggest that WM capacity is an important determinant of whether individuals engage in a model-based or model-free decision strategy. Furthermore, high WM capacity in younger adults seems to be associated with greater ability for strategic exploratory behavior. The results in younger adults are consistent with recent findings from a study that used the two stage Markov decision task in combination with a concurrent WM manipulation (Otto et al., 2013). Results of this study showed that taxing WM disrupts model-based behavior in younger adults.

What remained unclear from this study is at which decision stage the effects of WM occur. This is an interesting question, because on the one hand, WM may play a role for the representation and maintenance of the state actions values of the different options on the second stage. On the other hand, WM might also play role while trying to integrate model-free information with

information about the transition structure on the first stage of the task (Gershman et al., 2013; Otto et al., 2013). The current findings show that in younger adults the effects of WM capacity are enhanced if the reward probabilities of the different options are more differentiable from each other (in the high probability range condition, see **Figure 4**). Hence, our findings seem more consistent with the first view, suggesting that enhanced WM capacity is associated with a greater ability to maintain model-free value representations and use them for model-based decision-making.

### ABSENCE OF WM EFFECTS ON MODEL-BASED BEHAVIOR IN OLDER ADULTS

As shown in **Figure 6** in older adults we found no significant correlation between WM capacity and model-based behavior. In contrast, in younger adults enhanced WM capacity is associated with a higher degree of model-based behavior, particularly in the high probability range condition. At first sight, one way to interpret these effects would be in terms of a floor effect in WM capacity in older adults. However, as also shown in **Figure 6**, even those older adults with high WM capacity (comparable to young high performing individuals), did not show evidence for enhanced model-based behavior. These findings suggest that factors other than WM might explain the age-related decline in model-based behavior. This interpretation is backed-up by the results of a covariance analysis which show that age differences in model-based behavior remain significant even after controlling for the effects of WM capacity. The question is what those factors might be. Consistent with the interpretation offered above, it could be argued that deficits in the updating of expected reward value might lead to these impairments. However, it could also be argued that that these impairments are due to more complex interactions between areas that represent the expected value of options and areas that are involved in implementing strategic operations, such as the lateral PFC. Results from a recent fMRI study using a three state Markov learning task suggest that age-related impairments in learning of higher order transition structures (models) are associated with a reduced recruitment of the lateral prefrontal cortex (Eppinger et al., 2012). Furthermore, results of that study indicate that model-based learning correlates positively with reasoning abilities but not with WM. Thus, these results point to the view that there is a specific deficit in older adults that relates to the learning and application of higher order associations such as sequential contingencies between events or probabilistic transition structures (such as in the current task).

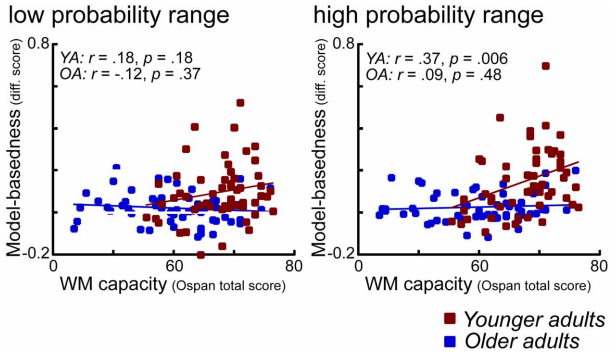



**FIGURE 6 | Scatter plots for the correlations between model-based difference values (see Methods) and WM capacity (OSPAN total scores), displayed separately for the low and the high reward probability range conditions.** Younger adults are shown in red, older adults are shown in blue.

Another interpretation of the absence WM effects in older adults could be that they are less willing (or able) to use an effortful decision strategy that relies on WM and rather fall back on a simpler decision strategy such as win-stay and lose-shift (Mata et al., 2010). This is somewhat supported by the modeling results, which suggest that older adults focus more on the most recent outcome than younger adults. However, given the overall performance deficits in older adults (see **Figure 2C**) such a strategy seems to reflect an adaptation to a behavioral impairment rather than a general difference in their approach to the task.

### EFFECTS OF PROBABILITY RANGE ON MODEL-BASED BEHAVIOR

The analyses of the stay-switch behavior also revealed that model-based behavior is enhanced when reward probabilities on the second stage are more differentiable from each other (in the wide compared to the narrow range probability condition). Furthermore, this effect is more pronounced in high WM capacity groups compared to low WM capacity groups (see **Figures 5**, **6**). The effects of probability range on model-based behavior are interesting for several reasons. First of all, these findings show that a manipulation that seems to primarily affect the second stage of the task can lead to a greater degree of model-based decision-making on the first stage of the task. That is, more differentiated value representations on the second stage seem to support model-based behavior on the first stage. Second, the interaction with WM capacity suggests that enhanced model-based behavior in individuals with high WM capacity may be due to a better ability to maintain and update those value representations in WM. Interestingly, a follow-up analysis of these results showed that greater model-based behavior in the wide probability range was primarily driven by enhanced stay behavior after common transitions that were followed by reward. This finding is in line with the idea that more differentiated reward probabilities on the second stage result in more consistent stay behavior on the first stage options, presumably by reducing uncertainty about the currently best option. The idea here would be that a greater distinctiveness of the value of choice options on the second stage supports the updating of those values in WM, particularly in individuals with high WM capacity. A better representation of the values of the different options on the second stage may then lead to more consistent choice behavior after common transitions that were followed by reward (i.e., in situations in which the available evidence indicates that the best thing to do is to stick to the option that has been chosen on the previous trial). Although such an interpretation seems speculative, it is consistent with theoretical ideas, suggesting that WM updating may be regulated by phasic dopaminergic prediction error signals (Braver and Cohen, 2000; Frank et al., 2001; D'Ardenne et al., 2012). According to the gating theory, it could be argued that the probability range manipulation results in more distinctive prediction error signaling and hence more reliable value representation for the different second-stage

choice options. A more reliable and differentiated representation of state-action values in WM may then support the application of model-based decision strategies on the first stage of the task.

## CONCLUSIONS

Taken together, the current results show impairments in model-based decision-making in older compared to younger adults. These deficits are particularly pronounced in situations in which reward on the previous trial reinforces stay behavior, whereas the fact that it was an uncharacteristic transition indicates the need for a shift in decision strategy. In these situations younger adults engage in a strategic exploration of the task structure, whereas older adults perseverate on the option they are currently exploiting. Analyses of the model parameters showed that decision-making deficits in older adults are associated with less consistent choice patterns on the second stage and a greater direct influence of reward on the previous trial on first stage choice behavior. Thus, the current findings are consistent with the idea that age-related deficits in model-based decision-making reflect impairments in the representation and updating of expected reward value (Eppinger et al., 2011; Chowdury et al., 2013; Eppinger et al., 2013). As a consequence of those deficits, older adults rely more on the most recent outcome rather than their (impoverished) representation of the expected value of choice options on the second stage.

In addition to age-related changes in goal-directed decision-making our findings also point to substantial individual differences in model-based behavior. In younger adults high WM capacity is associated with enhanced model-based behavior. Moreover, this effect is further elevated when reward probabilities on the second stage are more differentiable from each other. The implications of these effects are two-fold: first, these findings suggest that model-based behavior is particularly prevalent in younger individuals with high WM capacity. Second, these results indicate that high WM capacity supports the ability to maintain and update (model-free) value representations and use them for strategic exploration. It could be argued that the absence of a WM effect on model-based behavior in older adults reflects a floor effect in WM capacity. However, the fact that age-related deficits in model-based behavior remain significant even after controlling for the effects of WM capacity indicates that additional factors might play a role. Based on recent fMRI findings (Eppinger et al., 2012) we argue that an under-recruitment of the lateral PFC during the integration of expected reward value into model-based decisions might be one possible explanation for these effects.

## AUTHORS CONTRIBUTION

Ben Eppinger, Maik Walter, Shu-Chen Li, Hauke R. Heekeren designed the study. Ben Eppinger, Maik Walter, acquired and analyzed the data. Ben Eppinger, Maik Walter, Shu-Chen Li; Hauke R. Heekeren wrote the manuscript.

## ACKNOWLEDGMENTS

We would like to thank Rasmus Bruckner and Katharina Wermuth for help during data acquisition, as well as our participants for their contribution to the study. We are grateful to Nathanial Daw and Robert C. Wilson for helpful advice regarding RL modeling. This research was funded by the German Federal Ministry of Education and Research (BMBF) grant Bernstein Focus Neuronal Basis of Learning (FKZ 01Gq0913).

## REFERENCES

Baddeley, A. D., Emslie, H., and Nimmo-Smith, I. (eds.). (1992). *The Speed and Capacity of Language Processing (SCOLP) Test*. Bury St. Edmunds, Suffolk, England: Thames Valley Test Company.

Balleine, B. W., and O'Doherty, J. P. (2010). Human and rodent homologies in action control: cortico-striatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35, 48–69. doi: 10.1038/npp.2009.131

Bates, D., Maechler, M., and Bolker, B. (2013). lme4: linear mized-effects models using S4 classes. R version 3.0.2.

Braver, T. S., and Cohen, J. D. (2000). "On the control of control: the role of dopamine in regulating prefrontal function and working memory," in *Attention and Performance XVIII,* eds. S. Monsell and J. Driver (Cambridge, MA: MIT Press), 713–737.

Carver, C. S., and White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *J. Pers. Soc. Psychol.* 67, 319–333. doi: 10.1037/0022-3514.67.2.319

Chowdury, R., Guitart-Masip, M., Lambert, C., Dayan, P., Huys, Q., Duezel, E., et al. (2013). Dopamine restores reward prediction errors in old age. *Nat. Neurosci.* 16, 648–653. doi: 10.1038/nn.3364

D'Ardenne, K., Eshel, N., Luka, J., Lenartowicz, A., Nystrom, L. E., and Cohen, J. D. (2012). Role of prefrontal cortex and the midbrain dopamine system in working memory updating. *Proc. Natl. Acad. Sci. U.S.A.*109, 19900–19909. doi: 10.1073/pnas.1116727109

D'Ardenne, K., Mcclure, S. M., Nystrom, L., and Cohen, J. D. (2008). BOLD Responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* 319, 1264–1267. doi: 10.1126/science.1150605

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., and Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69, 1204–1215. doi: 10.1016/j.neuron.2011.02.027

Daw, N. D., Niv, Y., and Dayan, P. (2005). Human and rodent homologies in action control: cortico-striatal determinants of goal-directed and habitual action. *Nat. Neurosci.* 8, 1704–1711. doi: 10.1038/nn1560

Doll, B. B., Simon, D. A., and Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. *Curr. Opin. Neurobiol.* 22, 1075–1081. doi: 10.1016/j.conb.2012.08.003

Duncan, J., Schramm, M., Thompson, R., and Dumontheil, I. (2012). Task rules, working memory and fluid intelligence. *Psychon. Bull. Rev.* 19, 864–870. doi: 10.3758/s13423-012-0225-y

Ekstrom, R. B., French, J. W., and Harman, H. H. (1976). *Kit of Factor-Referenced Cognitive Tests.* Princeton, NJ: E.T. Service.

Eppinger, B., Haemmerer, D., and Li, S.-C. (2011). Neuromodulation of reward-based learning and decision making in human aging. *Ann. N.Y. Acad. Sci.* 1235, 1–17. doi: 10.1111/j.1749-6632.2011.06230.x

Eppinger, B., Heekeren, H. R., and Li, S.-C. (2012). "When two birds in the bush are better than one in the hand: age-related impairments in learning to predict future rewards," in *Annual Meeting of the Society for Neuroscience*, (New Orleans, LA).

Eppinger, B., and Kray, J. (2011). To choose or to avoid: age differences in learning form positive and negative feedback. *J. Cogn. Neurosci.* 23, 41–52. doi: 10.1162/jocn.2009.21364

Eppinger, B., Kray, J., Mock, B., and Mecklinger, A. (2008). Better or worse than expected? Aging, Learning, and the ERN. *Neuropsychologia* 46, 521–539. doi: 10.1016/j.neuropsychologia.2007.09.001

Eppinger, B., Schuck, N. W., Nystrom, L. E., and Cohen, J. D. (2013). Reduced striatal responses to reward prediction errors in older compared to younger adults. *J. Neurosci.* 33, 9905–9912. doi: 10.1523/JNEUROSCI.2942-12.2013

Frank, M. J., and Kong, L. (2008). Learning to avoid in older age. *Psychol. Aging* 23, 392–398. doi: 10.1037/0882-7974.23.2.392

Frank, M. J., Loughry, B., and O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cogn. Affect. Behav. Neurosci.* 1, 137–160. doi: 10.3758/CABN.1.2.137

Gershman, S. J., Markman, A. B., and Otto, A. R. (2013). Retrospective revaluation in sequential decision making: a tale of two systems. *J. Exp. Psychol. Gen.* doi: 10.1037/a0030844. [Epub ahead of print].

Gläscher, J., Daw, N. D., Dayan, P., and O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* 66, 585–595. doi: 10.1016/j.neuron.2010.04.016

Hämmerer, D., and Eppinger, B. (2012). Dopaminergic and prefrontal contributions to reward-based learning and outcome monitoring during child development and aging. *Dev. Psychol.* 48, 862–874. doi: 10.1037/a0027342

Hämmerer, D., Li, S.-C., Mueller, V., and Lindenberger, U. (2011). Lifespan differences in electrophysiological correlates of monitoring gains and losses during probabilistic reinforcement learning. *J. Cogn. Neurosci.* 23, 579–592. doi: 10.1162/jocn.2010.21475

Kahneman, D. (2011). *Thinking, Fast and Slow.* New York. NY: Farrar, Strauss, Giroux.

Lau, B., and Glimcher, P. W. (2005). Dynamic response-by-response models of matching behavior in Rhesus monkeys. *J. Exp. Anal. Behav.* 84, 555–579. doi: 10.1901/jeab.2005.110-04

Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., and Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the lifespan. *Psychol. Sci.* 15, 155–163. doi: 10.1111/j.0956-7976.2004.01503003.x

Mata, R., Von Helversen, B., and Rieskamp, J. (2010). Learning to choose: cognitive aging and strategy selection learning in decision making. *Psychol. Aging* 25, 299–309. doi: 10.1037/a0018923

Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167

Montague, P. R., Hyman, S. E., and Cohen, J. D. (2004). Computational roles for dopamine in behavioral control. *Nature* 431, 760–767. doi: 10.1038/nature03015

Nieuwenhuis, S., Ridderinkhof, K. R., Talsma, D., Coles, M. G. H., Holroyd, C. B., Kok, A., et al. (2002). A computational account of altered error processing in older age: dopamine and the error-related negativity. *Cogn. Affect. Behav. Neurosci.* 2, 19–36. doi: 10.3758/CABN.2.1.19

Niv, Y., Edlund, J. A., Dayan, P., and O'Doherty, J. P. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *J. Neurosci.* 32, 551–562. doi: 10.1523/JNEUROSCI.5498-10.2012

Niv, Y., and Schoenbaum, G. (2008). Dialogues on prediction errors. *Trends Cogn. Sci.* 12, 265–272. doi: 10.1016/j.tics.2008.03.006

Otto, A. R., Gershman, S. J., Markman, A. B., and Daw, N. D. (2013). The curse of planning: dissecting multiple reinforcement learning systems by taxing the central executive. *Psychol. Sci.* 24, 751–761. doi: 10.1177/0956797612463080

Pietschmann, M., Endrass, T., Czerwon, B., and Kathmann, N. (2011). Aging, probabilistic learning and performance monitoring. *Biol. Psychol.* 86, 74–82. doi: 10.1016/j.biopsycho.2010.10.009

Raven, J. C., Raven, J. E., and Court, J. H. (eds.). (1998). *Progressive Matrices*. Oxford: Oxford Psychologists Press.

R Development Core and Team (2010). *R: A Language and Environment for Statistical Computing.* Vienna: R.F.F.S. Computing.

Rummery, G., and Niranjan, M. (1994). *On-line Q-learning Using Connectionist Systems.* Cambridge: Cambridge University.

Salthouse, T. A. (1994). The aging of working memory. *Neuropsychology* 8, 535–543. doi: 10.1037/0894-4105.8.4.535

Salthouse, T. A., Mitchell, D. R. D., Skovronek, E., and Babcock, R. L. (1989). Effects of adult age and working memory on reasoning and spatial abilities. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 507–516. doi: 10.1037/0278-7393.15.3.507

Samanez-Larkin, G. R., Kuhnen, C. M., Yoo, D. J., and Knutson, B. (2010). Variability in nucleus accumbens activity mediates age-related suboptimal financial risk taking. *J. Neurosci.* 27, 1426–1434.

Samanez-Larkin, G. R., Levens, S. M., Perry, L. M., Dougherty, R. F., and Knutson, B. (2012). Frontostriatal white matter integrity mediates adult age differences in probabilistic reward learning. *J. Neurosci.* 32, 5333–5337. doi: 10.1523/JNEUROSCI.5756-11.2012

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of predicition and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593

Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning)*. Cambridge, MA: MIT Press.

Thorndike, E. L. (1911). *Animal Intelligence.* New York, NY: Macmillan.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208. doi: 10.1037/h0061626

Turner, M. L., and Engle, R. W. (1989). Is working memory capacity task dependent. *J. Mem. Lang.* 28, 127–154. doi: 10.1016/0749-596X(89)90040-5

Unsworth, N., Heitz, R. P., Dschrock, J. C., and Engle, R. W. (2005). An automated version of the operation span task. *Behav. Res. Methods* 37, 498–505. doi: 10.3758/BF03192720

Worthy, D. A., Gorlick, M. A., Pacheco, J. L., Schnyer, D. M., and Maddox, W. T. (2011). With age come wisdom: decision making in younger and older adults. *Psychol. Sci.* 22, 1375–1380. doi: 10.1177/0956797611420301

Worthy, D. A., and Maddox, W. T. (2012). Age-based differences in strategy use in choice tasks. *Front. Neurosci.* 5:145. doi: 10.3389/fnins.2011.00145

Wunderlich, K., Smittenaar, P., and Dolan, R. J. (2012). Dopamine enhances model-based over model-free behavior. *Neuron* 75, 418–424. doi: 10.1016/j.neuron.2012.03.042

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 10 July 2013; accepted: 06 December 2013; published online: 24 December 2013.*

*Citation: Eppinger B, Walter M, Heekeren HR and Li S-C (2013) Of goals and habits: age-related and individual differences in goal-directed decision-making. Front. Neurosci. 7:253. doi: 10.3389/fnins.2013.00253*

*This article was submitted to Decision Neuroscience, a section of the journal Frontiers in Neuroscience.*

frontiers in
NEUROSCIENCE

REVIEW ARTICLE
published: 19 April 2013
doi: 10.3389/fnins.2013.00062

# Prefrontal dopamine and behavioral flexibility: shifting from an "inverted-U" toward a family of functions

***Stan B. Floresco****

*Department of Psychology, Brain Research Centre, University of British Columbia, Vancouver, BC, Canada*

***Edited by:***
*Jeff A. Beeler, The University of Chicago, USA*

***Reviewed by:***
*Geoffrey Schoenbaum, University of Maryland School of Medicine, USA*
*Vincent Daniel Costa, National Institute of Mental Health, USA*

***Correspondence:**
*Stan B. Floresco, Department of Psychology, Brain Research Centre, University of British Columbia, 2136 West Mall, Vancouver, V6T 1Z4 BC, Canada.*
*e-mail: floresco@psych.ubc.ca*

Studies on prefrontal cortex (PFC) dopamine (DA) function have revealed its essential role in mediating a variety of cognitive and executive functions. A general principle that has emerged (primarily from studies on working memory) is that PFC DA, acting on $D_1$ receptors, regulates cognition in accordance to an "inverted-U" shaped function, so that too little or too much activity has detrimental effects on performance. However, contemporary studies have indicated that the receptor mechanisms through which mesocortical DA regulates different aspects of behavioral flexibility can vary considerably across different DA receptors and cognitive operations. This article will review psychopharmacological and neurochemical data comparing and contrasting the cognitive effects of antagonism and stimulation of different DA receptors in the medial PFC. Thus, set-shifting is dependent on a co-operative interaction between PFC $D_1$ and $D_2$ receptors, yet, supranormal stimulation of these receptors does not appear to have detrimental effects on this function. On the other hand, modification of cost/benefit decision biases in situations involving reward uncertainty is regulated in complex and sometimes opposing ways by PFC $D_1$ vs. $D_2$ receptors. When viewed collectively, these findings suggest that the "inverted-U" shaped dose-response curve underlying $D_1$ receptor modulation of working memory is not a one-size-fits-all function. Rather, it appears that mesocortical DA exerts its effects via a family of functions, wherein reduced or excessive DA activity can have a variety of effects across different cognitive domains.

**Keywords: prefrontal, dopamine, $D_1$, $D_2$, set-shifting, decision making, microdialysis, rats**

Brozoski et al. (1979) originally reported that depletion of dopamine (DA) in the prefrontal cortex (PFC) of monkeys impaired delayed responding in a manner comparable to complete removal of the frontal lobes. These seminal findings have since sparked a substantial amount of psychopharmacological, neurophysiological, and computational research on how mesocortical DA regulates complex forms of cognition. Much of the work stemming from these initial findings has focused on processes related to working memory, revealing that these functions are dependent primarily on PFC $D_1$ receptor activity. A particularly influential discovery from this line of research is that PFC $D_1$ receptor modulation of working memory takes the form of an "inverted-U" shaped curve (Arnsten, 1997; Zahrt et al., 1997; Williams and Castner, 2006), where suboptimal or excessive $D_1$ activity can have detrimental effects on cognition.

The notion that normal PFC functioning is dependent on an optimum range of DA activity, whereas "too little" or "too much" $D_1$ receptor stimulation has detrimental effects on working memory has become a cornerstone of our understanding of how mesocortical DA regulates cognition. However, the frontal lobes regulate a variety of other functions distinct from working memory, such as cognitive flexibility, cost/benefit decision making, and emotional processes. More contemporary studies have begun to elucidate how PFC DA may regulate these functions, and an emerging impression is that PFC DA regulation of these other functions differs considerably from mechanisms that facilitate working memory.

DA exerts its effects on PFC neural activity via multiple receptor subtypes. Both $D_1$-like and $D_2$-like ($D_2$, $D_4$) receptors are expressed within the PFC, although the subcellular localization of these receptors differs. Expression of $D_1$ receptors on principle pyramidal neurons appears to be substantially greater than $D_2$ receptors (Gaspar et al., 1995), whereas both types of receptors have been localized on GABAergic interneurons and may also reside on presynaptic excitatory glutamate terminals (Sesack et al., 1995; Mrzijak et al., 1996; Muly et al., 1998; Wedzony et al., 2001). Numerous studies have shown that activation of $D_1$, $D_2$, or $D_4$ receptors exerts complex and dissociable electrophysiological actions on the activity of different classes of PFC neurons that may either increase or decrease the excitability of these cells and differentially modulate PFC neural network activity, depending of a variety of factors (see Seamans and Yang, 2004 for a review). Moreover, recent studies have indicated that there may be separate population of PFC pyramidal neurons that preferentially express only $D_1$ or $D_2$ receptors (Gee et al., 2012; Seong and Carter, 2012). These anatomical and neurophysiological findings suggest that DA may exert differential effects on the activity of PFC neural networks which in turn may subserve a variety of distinct cognitive operations. Yet, despite these findings, the majority of studies on the role of PFC DA in functions such as working memory have

focused on the role of $D_1$ receptors, whereas until recently, the functional role of $D_2$ and $D_4$ receptors has been less clear. This review will highlight some recent advances in our understanding of how PFC DA regulates a variety of executive functions, focusing primarily on psychopharmacological and neurochemical data obtained from rodents, with an emphasis on the differences in the principles of operation through which medial PFC DA regulates different cognitive domains.

## DA, THE "INVERTED-U" AND WORKING MEMORY: IMPORTANT CAVEATS

One of the earliest and direct demonstrations that supranormal stimulation of PFC $D_1$ receptors can perturb working memory came from the seminal study by Zahrt et al. (1997). They showed that infusions of the full $D_1$ agonist SKF 81297 (0.01–0.1 μg) in the prelimbic region of the medial PFC of rats dose-dependently impaired delayed alternation on a T-maze task. An influential aspect of this paper was a summary figure, showing that treatment with a $D_1$ agonist or antagonist (SCH 23390) markedly reduced the proportion of correct responses when compared to control conditions or combined agonist/antagonist treatment. What was particularly striking about this synthesis was how actual empirical data were plotted to clearly demonstrate an "inverted-U" shaped function underlying dopaminergic modulation of working memory. However, an important point that is often overlooked is that impairments in delayed alternation induced by $D_1$ antagonism (the "too little" end of the curve) have been observed after *systemic* D1 receptor blockade. In contrast, a subsequent study using a near-identical task found that blockade of either $D_1$ or $D_2$ receptors in the medial PFC *did not impair* delayed alternation (Romanides et al., 1999). This discrepancy between the effects of systemic vs. local manipulations of DA activity indicates that caution is warranted when attributing the specific neural loci where systemic drug treatments may be acting to affect behavior and cognition. Note that in the aforementioned study, blockade of glutamate receptors did impair performance, indicating that working memory assessed in this manner is dependent on the integrity of excitatory transmission in the PFC. Yet, the fact that blockade of DA receptors in the rat medial PFC did not impair delayed alternation suggests that this form of delayed responding is not a particularly sensitive paradigm for assessing PFC DA regulation of working memory functions in rodents. Moreover, it suggests that certain aspects of working memory dependent on the PFC may nevertheless be relatively insensitive to reductions in mesocortical DA. This is in keeping with studies in primates showing that performance of a self-ordered sequencing task or a spatial delayed response task were both impaired by excitotoxic lesions of the PFC, yet PFC DA depletion only impaired delayed responding and left self-ordered working memory intact (Collins et al., 1998).

Another important principle underlying PFC DA modulation of working memory is the relative baseline levels of performance. Work by our group has used a delayed response variant of the radial-arm maze task (**Figure 1A**) utilizing a comparatively long delay (30 min) that, unlike delayed alternation, is sensitive to blockade of PFC $D_1$ (but not $D_2$) receptors (Seamans et al., 1998; **Figure 1B**, left; **Figure 5A**). We exploited this procedure to manipulate baseline performance by testing separate group of rats after either a typical, 30 min delay (when performance was good) or after an extended, 12-h delay (which degrades performance in control animals) (Floresco and Phillips, 2001). In keeping with previous findings, intra-PFC (prelimbic) infusion of the $D_1$ agonist SKF 81297 (0.05–0.2 μg) dose-dependently impaired working memory after the 30 min delay, compared to control rats that showed near-optimal performance (**Figure 1B**, right; **Figure 5A**). In contrast, control rats subjected to an extended 12 h delay made considerably more errors, presumably because the memory for the expected location of reward had degraded during this period. What was striking was that, under these conditions where performance was degraded, treatment with the same doses of the $D_1$ agonist had the diametrically opposite effect to that observed when performance was good, in that these treatments improved performance relative to controls. Similar results have been obtained with the same agonist using a within-subjects design in combination with a delayed-response task incorporating shorter delays (Chudasama and Robbins, 2004). Thus, pharmacological stimulation of PFC $D_1$ receptors does not always impair working memory, and can actually improve performance following degradation of the memory that the subject must "work" with (e.g., after longer delays). Note that degradations in performance induced by longer delays have been associated with reduced levels of mesocortical DA efflux compared to conditions where performance is good (Phillips et al., 2004; **Figure 1C**). Thus, differential effects of PFC $D_1$ stimulation on working memory may be mediated in part by the relative levels of mesocortical DA transmission, with good vs. poor performance linked to higher vs. lower levels of DA efflux. Under these conditions, exogenous stimulation of PFC $D_1$ receptors would either be expected to overstimulate these receptors (and impair good performance) or normalize levels of $D_1$ activity and improve performance, in keeping with the idea of the inverted-U shaped function.

Unlike PFC $D_1$ receptors, blockade of $D_2$ receptors has repeatedly been shown to not disrupt working memory in primates or rats (Sawaguchi and Goldman-Rakic, 1994; Seamans et al., 1998; Romanides et al., 1999), even though local application of $D_2$ agonists or antagonists augments or attenuates "response"-related firing of PFC neurons in monkeys performing an occulomotor delayed response task (Wang et al., 2004). Although the effects of PFC $D_2$ receptor stimulation on working memory performance have yet to be explored fully, one notable study revealed that prelimbic PFC infusions of a $D_2$ agonist disrupts delayed responding on a U-maze, whereas PFC $D_2$ antagonism reduced proactive interference (Druzin et al., 2000). Thus, under some conditions, PFC $D_2$ receptor modulation of working memory may take the form of a monotonic function (i.e., lower/higher levels of $D_2$ activation associated with better/poorer performance), in a manner that is distinct and antagonistic to the inverted-U shaped function underlying $D_1$ receptor modulation. However, as discussed below, the principles of operation through which different DA receptors interact to regulate other executive processes mediated by the frontal lobes can differ considerably from those underlying working memory.

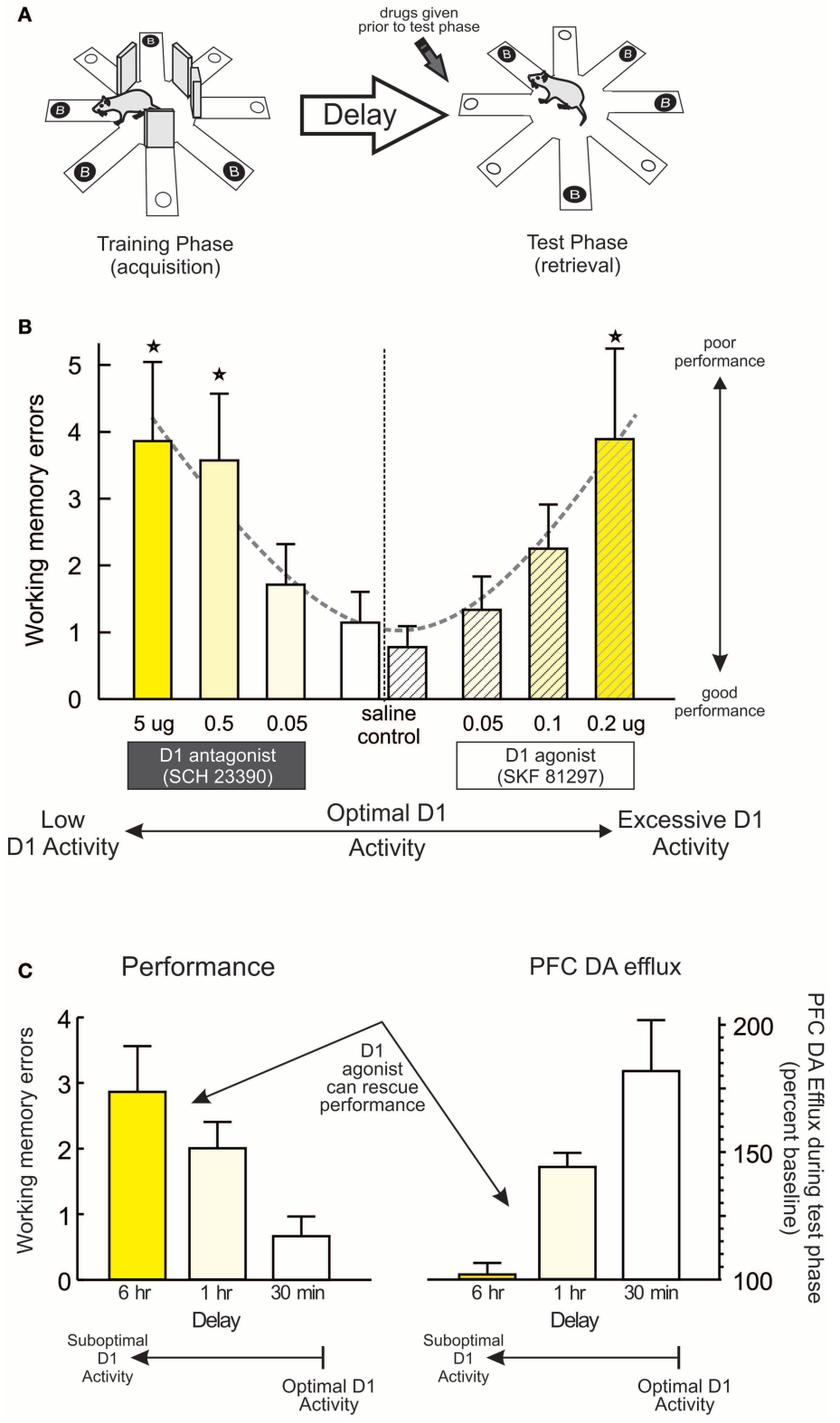



**FIGURE 1 | PFC D1 receptor modulation of working memory. (A)** A delayed response variant of the radial-arm maze task used to collect data presented in subsequent panels. The task consists of a training (acquisition) and a test (retrieval) phase. During the training phase, the rat must retrieve food from four randomly selected arms, with the other arms blocked. During a test phase occurring after a delay, arms that were blocked previously are now open and baited. DA drugs were administered prior to the test phase.

*(Continued)*

**FIGURE 1 | Continued**
**(B)** Infusions of a $D_1$ antagonist dose-dependently impaired working memory performance on this delayed-response task. Similarly, treatment with the D1 agonist SKF 81297 also impaired performance when infusions were made after a relatively short delay (30 min). These data have be re-plotted from those originally reported by Seamans et al. (1998) and Floresco and Phillips (2001) to highlight the effects of reduced vs. excessive D1 activation on performance. For this and all subsequent figures, dashed lines emphasize the dose-response function associated with reduced or excessive DA receptor stimulation. In the case of working memory, the effects present as a classic the U-shaped function, where reduced or excessive PFC D1 activity caused poorer performance relative to control conditions, numbers underneath each bar represents drug dose (in μg), and stars represent $p < 0.05$ vs. relative control treatments or groups. **(C)** Behavioral performance and peak increase in PFC DA efflux observed during the test phase of this task following a typical 30 min, or extended 1 or 6 h delays. Left panel shows that extending the delay period degrades performance and results in more working memory errors. Right panel shows that in these same animals, poorer performance was associated with reduced PFC DA efflux. Under these conditions, infusions of a D1 agonist can rescue performance. Adapted from Phillips et al. (2004).

## PREFRONTAL DA AND BEHAVIORAL FLEXIBILITY

Another key function of the mammalian PFC is to facilitate alterations in behavior in response to changing environmental demands (Dias et al., 1996; Brown and Bowman, 2002; Floresco et al., 2009). Behavioral flexibility is not a unitary phenomenon, but rather, may be viewed as a hierarchical process, ranging from simpler to more complex processes that are subserved by anatomically-distinct prefrontal and subcocortical regions. For example, extinction entails the suppression of a conditioned response elicited by a stimulus that no longer predicts reinforcement. Although the contribution of mesocortical DA to this form of flexibility remains to be characterized thoroughly, there have been reports that $D_2$ and $D_4$ receptors in the infralimbic medial PFC, may facilitate consolidation of fear extinction memories (Pfeiffer and Fendt, 2006; Mueller et al., 2010).

Reversal learning is a more complex form of flexibility engaged when an organism must discriminate between two or more stimuli, only one of which is associated with reinforcement. Reversal shifts require a switch between stimulus-reinforcement associations within a particular stimulus dimension (i.e., use the same basic strategy, but approach a different stimulus), a form of flexibility critically-dependent on the orbitofrontal PFC in both primates and rats (Dias et al., 1996, 1997; McAlonan and Brown, 2003). Unlike other forms of flexibility, reversal learning is generally unimpaired by global depletion of PFC DA (Roberts et al., 1994; Crofts et al., 2001). Rather, serotonin inputs to the orbital PFC appears to be the monoamine neurotransmitter that is of primary importance in modulating reversal learning (Clarke et al., 2004, 2005), although DA input to striatal regions also facilitates this form of flexibility (O'Neill and Brown, 2007; Clarke et al., 2011).

On the other hand, shifts between strategies, rules or attentional sets taps into higher-order cognitive functions, requiring attention be focused to multiple aspects of complex environmental stimuli. In humans, an inability to shift strategies is epitomized by impairments on the Wisconsin Card Sorting task. Patients with frontal lobe damage are initially able to sort cards by one dimension (e.g., color), but have great difficulty in altering their strategy when required to organize cards by another dimension, (number or shape), perseverating to the now incorrect strategy. Studies with laboratory animals have revealed that lesions/inactivation of the lateral PFC in primates or the medial PFC in rats do not affect initial discrimination learning, but profoundly impair the ability to inhibit an old strategy and utilize a new one (Dias et al., 1996, 1997; Ragozzino et al., 1999; Brown and Bowman, 2002; Floresco et al., 2008a), even though these manipulations do not affect reversal learning.

Much of the research on how mesocortical DA modulates behavioral flexibility has focused on attentional or strategy set-shifting. An initial report by Roberts et al. (1994) used an intradimensional/extradimensional (ID/ED) shifting task, wherein marmosets conducted a series of two-choice discriminations using complex stimuli (e.g., sets of lines overlaid onto different shapes). During the initial phases, subjects discriminated stimuli based on one stimulus dimension (e.g., lines), but during the critical ED phase, they had to shift their attention to the other stimulus dimension. Depletion of PFC DA actually improved ED set shifting, even though these manipulations disrupted working memory assessed with a spatial delayed-response task. The improvement in set shifting was later attributed to a disruption in attentional set formation, as a subsequent study showed that PFC DA depletion impaired repeated ID shifts within the same stimulus dimension (Crofts et al., 2001). However, this effect was only observed for one type of ED shift when animals were required to shift responding from a more difficult "lines" dimension to a "shapes" dimension. Nevertheless, these data indicate that mesocortical DA serves to stabilize representations, facilitating the ability to attend to relevant stimuli (Robbins, 2005; Robbins and Arnsten, 2009).

One way to assesses set-shifting ability in rodents that is amenable to psychopharmacological investigation is with a strategy-shifting task conducted either on a cross-maze or in an operant chamber. Rats initially learn to use either an egocentric response (e.g., always turn left) or visual-cue discrimination strategy (e.g., always approach the arm with a visual cue, located in the left, or right arm with equal frequency) to obtain reinforcement (see **Figure 2A**). During the shift, rats must cease using the previously-acquired strategy and learn the alternative discrimination. As has been observed with studies using ID/ED shifting tasks designed for rodents, strategy set-shifting is disrupted by inactivation of the medial, but not orbital PFC (Ragozzino et al., 1999; Birrell and Brown, 2000; McAlonan and Brown, 2003; Floresco et al., 2008a; Ghods-Sharifi et al., 2008). Another advantage of the strategy shifting task is that it permits a detailed analysis of the different types of errors committed during the shift, providing insight into whether impairments are due to enhanced perseverative responding or a deficit in acquiring or maintaining new strategies. Reversible inactivation of the medial PFC causes robust perseverative-type deficits when rats must shift from one strategy to another (Ragozzino et al., 1999; Floresco et al., 2008a).

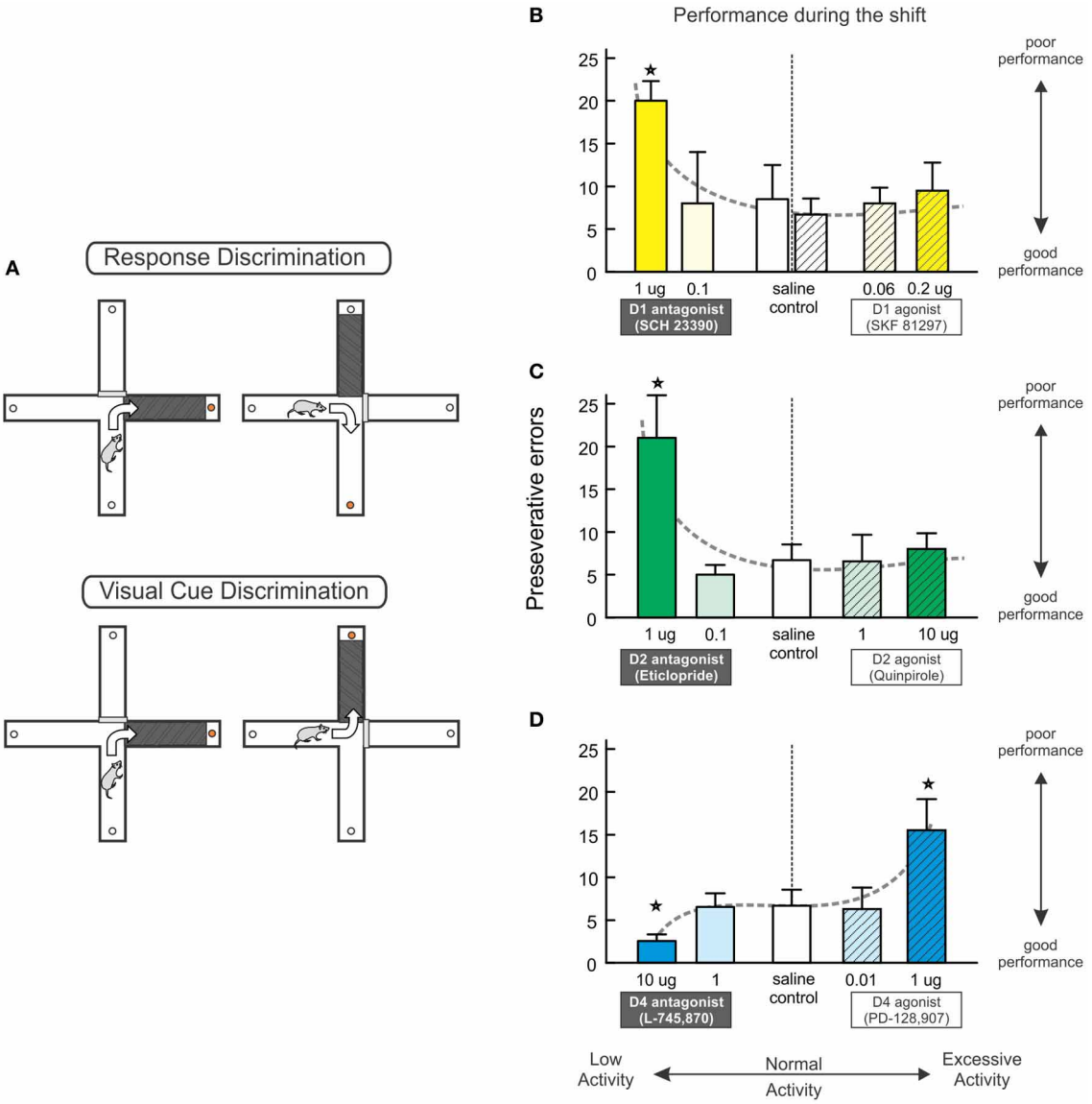



**FIGURE 2 | Multiple PFC DA receptors regulate set shifting. (A)** The set shifting task conducted on a cross-maze requires rats to initially learn one discrimination rule (e.g., always turn right) to receive food reinforcement (top). A visual-cue insert is randomly placed in one of the arms but does not reliably predict reward. During the set-shift (bottom), the rat is now required to use a visual-cue discrimination strategy, necessitating a shift from the old strategy and approach toward the previously-irrelevant cue. **(B–D)** The effects of blockade and stimulation of $D_1$, $D_2$, and D4 receptors on set shifting. These data have be re-plotted from those originally reported by Ragozzino (2002) and Floresco et al. (2006) to highlight the effects of reduced vs. excessive DA receptor activation on perseverative errors made during the shift. Blockade of PFC D1 **(B)** or D2 **(C)** receptors significantly impairs strategy set-shifting, whereas pharmacological stimulation of these receptors did not affect performance. **(D)** Blockade of $D_4$ receptors improves performance, whereas $D_4$ stimulation impaired set shifting. Stars represent $p < 0.05$.

Psychopharmacological studies have revealed some similarities, but also important differences in the receptor mechanisms through which PFC DA regulates set-shifting viz à viz working memory. Thus, akin to its important for working memory, medial prelimbic PFC $D_1$ receptor activity also facilitates strategy set-shifting, as blockade of these receptors with SCH 23390 induces severe perseverative deficits (Ragozzino, 2002, **Figure 2B**, left). However, a subsequent series of experiments by our group (Floresco et al., 2006) revealed blockade of PFC $D_2$ receptors also enhanced perseveration during set-shifting, indicating that, unlike working memory, this form of behavioral flexibility is critically-dependent on a cooperative interaction of both $D_1$ and $D_2$ receptors in the PFC (**Figure 2C**, left).

Further, dissociations in the DA receptor pharmacology underlying working memory and set-shifting were observed following administration of DA agonists into the prelimbic medial PFC. Infusions of the $D_1$ agonist, SKF 81297 (at doses known to affect working memory performance) neither impaired nor

improved set-shifting (**Figure 2B**, right). Note that the lack of effect of PFC $D_1$ stimulation is in keeping with findings from another study, where infusion of the agonist SKF 38393 did not alter set-shifting on an ID/ED task, although these manipulations did ameliorate impairments induced by repeated amphetamine (Fletcher et al., 2005). Similar to the lack of effect with a $D_1$ agonist, intra-PFC application of the $D_2$ agonist quinpirole also did not affect set-shifting (**Figure 2C**, right, and also see **Figure 5B**). Another interesting observation was that, unlike $D_1$ and $D_2$ receptors, PFC $D_4$ receptor modulation of set-shifting took the form of a negative sigmoidal function, as stimulation of these receptors impaired performance, and their blockade improved shifting relative to controls (**Figure 2D**). Collectively, these findings indicate that the construct of an "inverted-U" shaped function underlying $D_1$ (or $D_2$/$D_4$) receptor modulation of working memory does not appear to hold true for set-shifting functions mediated by the PFC. In this regard, it is plausible that combined stimulation of both of these receptors may have beneficial effects on set-shifting, as systemic treatment with the COMT inhibitor tolcapone has been reported to selectively increase PFC DA efflux and improve ED shifting (Tunbridge et al., 2004).

Additional insight into the contributions of PFC DA to set shifting comes from *in vivo* microdialysis studies conducted in freely-behaving rats performing a strategy set-shifting task similar to the one described above (Stefani and Moghaddam, 2006). These experiments also included two key control groups, the first being a yoked-reward group, where rats obtained reward on an intermittent schedule matched to rats performing the task but were not required to discriminate between arms or switch strategies. Thus, in this group, any choice led to either reward or no reward in accordance with a pattern of reinforcement experienced by a rat that actually performed the set-shifting task. However, in this instance, the specific response-reward contingencies were unpredictable from the perspective of the animals in this group. A second, reward-retrieval condition had rats obtain food on every trial, regardless of their choice.

The investigators observed that for rats trained on the set-shifting task, extracellular levels of PFC DA increased during learning of the initial rule, even though intact PFC DA does not appear to be necessary for learning simple discriminations (Ragozzino, 2002). When rats had to shift to a different rule in conflict with the first (a process that *is* dependent on PFC DA activity), PFC DA levels increased again, with a magnitude comparable to that observed during performance of a working memory task on a radial maze (∼80–100% above baseline; Phillips et al., 2004). Importantly, for rats trained on the set-shifting task, the relationship between PFC DA levels and performance during the shift did not reflect an "inverted-U" type function (i.e., moderate increases in DA associated with better performance compared to higher or lower levels). Instead, the relationship between the magnitude of DA efflux and behavioral performance was curvilinear, in that higher levels PFC DA efflux was associated with better performance during the shift (i.e., fewer trials required to achieve criterion performance). This finding is in keeping with the observation that pharmacological increases in PFC DA activity do not impair set-shifting, and may actually facilitate these functions in some situations. Interestingly, rats in the yoked-reward group (but not reward-retrieval group) displayed a profile of DA release similar to that observed in rats actually performing the set shift, despite the fact that the scheduling of reinforcement in this condition did not permit them to learn any reliable response-reward contingencies. This latter finding suggests that PFC DA transmission is particularly sensitive to situations where reward availability is unpredictable or uncertain. This increase in PFC DA transmission triggered by unexpected reward deliveries or omissions may serve as a signal that reward contingencies are changing and promote adaptations in behavior. Indeed, as will be discussed below, recent findings have shown decision making involving reward uncertainty is modulated in a particularly complex way by different DA receptors in the PFC.

## PREFRONTAL DA AND COST/BENEFIT DECISION MAKING

Since, the pioneering work of Damasio and colleagues showing that patients with damage to the ventromedial PFC were impaired on tasks designed to simulate real-life decisions in terms of uncertainty, reward and punishment (Bechara et al., 1994, 1999), there has been a growing interest in the neural circuitry underlying different forms of cost/benefit decision making. These types of situations require coordination of various cognitive and motivational processes to ensure that a decision maker adjusts choice biases in a flexible manner when cost/benefit contingencies change. A key component of decision making that can be assessed in rodents is the evaluation of costs associated with different actions relative to the rewards that may be obtained by those actions. In these studies, animals typically choose between smaller, readily-available rewards, or a larger/more palatable reward associated with some form of cost which can diminish the subjective value of objectively larger or more-preferred rewards. All things being equal, animals typically choose more (or "better") vs. less food, yet, imposition of certain costs lead to a "discounting" of preferred rewards. Costs that are effective in biasing choice behavior include (1) delays to reward delivery, (2) requiring animals to exert greater physical effort to obtain the reward, or (3) making reward delivery probabilistic (i.e., uncertain/risky).

Over the last 10 years, studies in rats have shown that different forms of cost/benefit decision making are regulated by anatomically-distinct regions of the frontal lobes, with the lateral orbital PFC playing a greater role in delay-related judgments, the dorsal anterior cingulate region of the medial PFC contributing to effort based decision making, and the prelimbic region of the medial PFC facilitating risk/reward judgments when reward probabilities are volatile (Walton et al., 2003; Winstanley et al., 2004; Rudebeck et al., 2006; St. Onge and Floresco, 2010; Zeeb et al., 2010). Although, each of these forms of decision making are sensitive systemic manipulations of DA transmission (Floresco et al., 2008b), there has been relatively little work on how mesocortical DA transmission regulates these decisions. Blockade of $D_1$ (but not $D_2$) receptors in the anterior cingulate reduces preference for larger rewards associated with a greater effort cost (Schweimer and Hauber, 2006). On the other hand, blockade of $D_1$ or $D_2$ receptors in the orbital PFC, or administration of $D_1$ receptor agonists or antagonists into the medial PFC increases delay discounting (Loos et al., 2010; Zeeb et al., 2010).

Work by our group has investigated the contribution of the rat prelimbic PFC to certain components risk-based decision making using a probabilistic discounting task, wherein rats choose between two options; a smaller, certain reward (1-pellet) or a larger uncertain (risky) option that may or may not yield 4-pellets (**Figure 3A**). The probability of obtaining the larger reward changes in a systematic manner over blocks of discrete, free-choice trials, ranging from 100 to 12.5%. Note that no explicit cues that signal changes in the odds of obtaining the larger reward are provided. Thus, in order to adjust their decision biases in an effective manner, rats must use internally-generated information to keep track of actions and outcomes (rewarded vs. non-rewarded choices) over multiple trials. This aspect of reward monitoring is dependent on the medial PFC, as inactivation of this region severely disrupts the ability to modify choice biases when reward probabilities change (St. Onge and Floresco, 2010). When the odds of obtaining the larger reward are initially good (100%) and gradually diminish over a session, PFC inactivation impairs shifting of decision biases toward the smaller/certain option in well-trained rats, which in this case results in an apparent increase in risky choice. Conversely, when the odds are initially poor (12.5%) and then increase, PFC inactivation retards shifts in bias toward the large/risky option, resulting in an overall decrease in risky choice. Thus, the medial PFC appears to play a critical role in detecting and tracking changes in action/outcome contingencies and reward availability, which in turn facilitates modifications in choice behavior when reward probabilities change.

We investigated the contribution of different DA receptors in the prelimbic medial PFC to this form of decision making, using doses of agonists and antagonists known to differentially affect working memory and set-shifting (St. Onge et al., 2011). Blockade of $D_2$ receptors with eticlopride induced an effect similar to PFC inactivation, impairing shifts in choice biases as reward probabilities decreased over time, which in this experiment manifested as an increase in risky choice (**Figure 3B**). Thus, $D_2$ receptor modulation of PFC neural activity facilitates modifications of decision biases in response to changes in risk/reward contingencies. In stark contrast, antagonism of PFC $D_1$ receptors with SCH 23390 induced the opposite effect of $D_2$ blockade (and PFC inactivation), causing a decrease in risky choice (**Figure 3C**). Thus, it appears that in some circumstances, $D_1$ and $D_2$ receptors regulate distinct and seemingly opposing functions related to risk-based decision making. Although, the mechanisms through which blockade of $D_1$ vs. $D_2$ receptors may induce opposing changes in behavior remains to be clarified, these effects may be related in part to actions of these receptors on separate populations of PFC pyramidal neurons (Gee et al., 2012; Seong and Carter, 2012), or their differential effects on the network activity of PFC neuronal populations (Durstewitz et al., 2000; Seamans and Yang, 2004).

Intra-PFC infusions of a $D_1$ agonist altered decision making in a manner symmetrical to $D_1$ blockade, inducing a moderate increase in risky choice that was not statistically-significant. Interestingly, these effects were numerically greater after treatment with the lower dose of SKF 81297 (0.1 μg) compared to the higher dose (0.4 μg; **Figure 3E**). A more pronounced disruption in decision making was induced by $D_2$ receptor stimulation with quinpirole. These treatments markedly flattened the discounting curve, as rats displayed no discernible discounting upon changes in reward probabilities (**Figure 3D**). Thus, excessive $D_2$ receptor activation severely interfered with the ability to adjust choice, causing rats to employ a simpler alternation strategy while maintaining a bias toward the large/risky option. This finding, in combination with the effects of eticlopride, suggests that the relative levels of both $D_1$ and $D_2$ receptor tone in the medial PFC has a critical impact on this aspect of decision making and either increasing or decreasing activity at either receptor interferes with performance.

Further difference in PFC $D_1/D_2$ modulation of different aspects of risk/reward decision making were unveiled upon examination of changes in reward and negative-feedback sensitivity induced by these treatments. Reward sensitivity was assessed by measuring the proportion of trials where subjects followed a risky "win" with another risky choice (a.ka., win-stay ratios), whereas, sensitivity to reward omissions was indexed by proportion of trials where rats shifted to the small/certain option after a non-rewarded risky choice (i.e., lose-shift ratios). Under control conditions, rats followed a risky win with another risky choice on 80–90% of these types of trials. Conversely, when rats played risky and were not rewarded, they chose the small/certain option on 25–30% of subsequent trials. Both of these processes were altered by PFC $D_1$ receptor manipulations in a particularly complex manner. Thus, reward sensitivity was not affected by reductions in $D_1$ tone but was increased by the lower dose of the $D_1$ agonist (**Figures 3F**, **5C**, left). Conversely, $D_1$ receptor blockade increased negative feedback sensitivity relative to control conditions, indicating that the decrease in risky choice induced by these treatments was primarily attributable to an increased sensitivity to reward omissions. This effect is similar to that observed after blockade of $D_1$ receptors in the nucleus accumbens (Stopper et al., 2013). On the flip side of the curve, $D_1$ stimulation had an opposite effect to $D_1$ antagonism, reducing lose-shift tendencies (**Figures 3G**, **5C**, right). With respect to $D_2$ receptors, blockade or stimulation increased or decreased reward sensitivity, respectively (**Figures 3H**, **5C**, left), whereas, either of these manipulations caused non-significant reductions in negative feedback sensitivity (**Figures 3I**, **5C**, right). Taken together, these data show how distinct aspects of risk/reward decision making can be affected by decreases or increases in mesocortical DA activity in manners that vary considerably across DA receptors. More generally, they further highlight that the specific functions describing how variations in PFC DA activity affect behavior are not uniform across cognitive domains.

One question that arose from the above-mentioned findings was how do fluctuations in mesocortical DA release relate to modifications in decision basis? To address this, we measured changes in PFC DA efflux with microdialysis in well-trained rats performing the same probabilistic discounting task (St. Onge et al., 2012). PFC DA levels corresponded to changes large/risky reward probabilities irrespective of whether the odds of obtaining the larger reward decreased or increased over a session (**Figure 4A**, yellow). Thus, when the odds were initially 100% and then decreased across blocks, there was a robust initial increase in PFC DA efflux

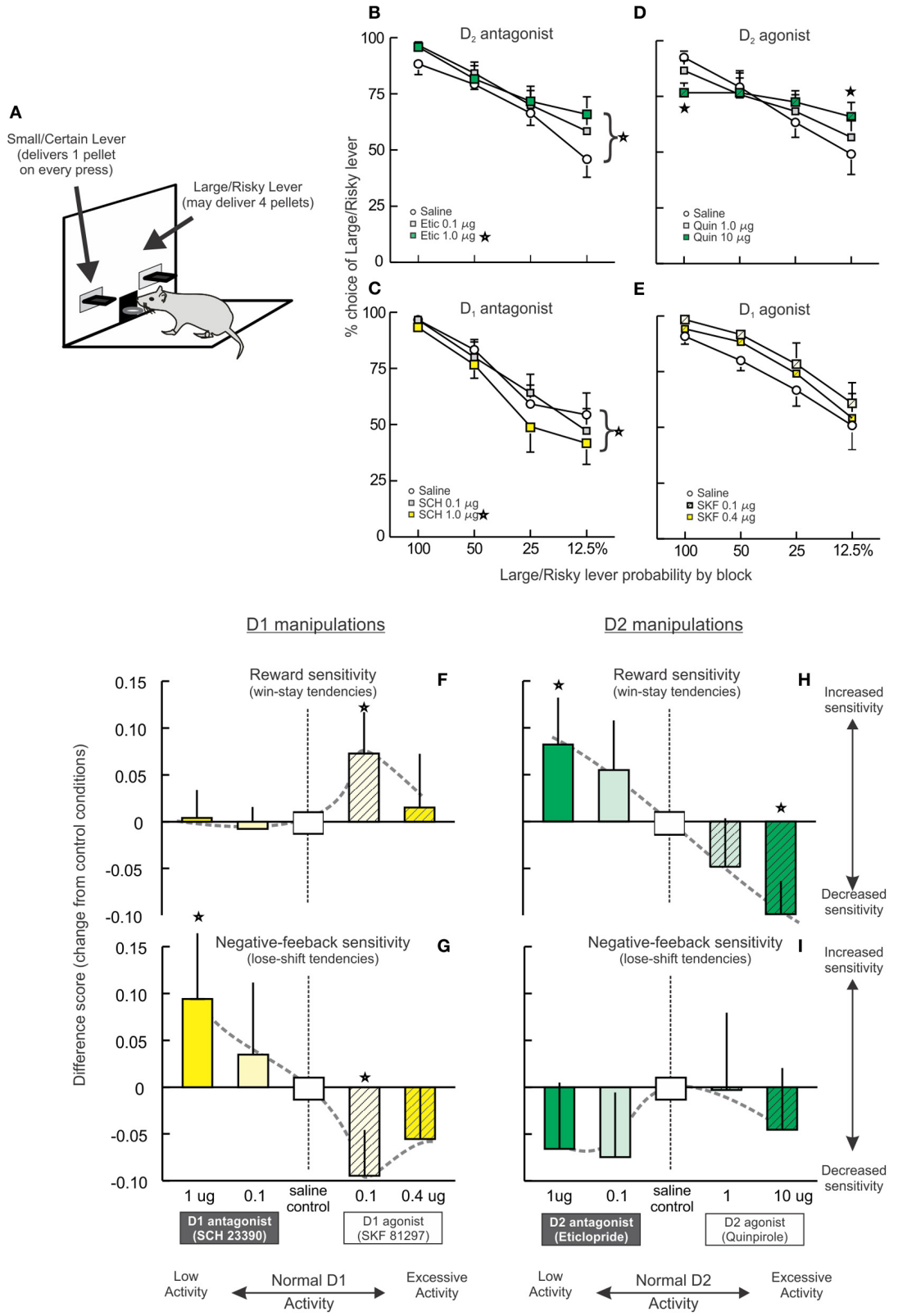



**FIGURE 3 | Opposing effects of PFC $D_1$ and $D_2$ receptor manipulations on risk-based decision making. (A)** The probabilistic discounting task required rats to choose between a small/certain reward option or a large/risky option. The probability of obtaining the larger reward changes in a systematic manner over blocks of free-choice trials. **(B–E)** Effects of PFC DA receptor

*(Continued)*

**FIGURE 3 | Continued**
manipulations on probabilistic discounting. Data are plotted in terms of percentage choice of the Large/Risky lever during free choice trials by probability block. **(B)** Blockade of D2 receptors retarded discounting and increased risky choice. **(C)** In contrast, blockade of PFC D1 receptors accelerated probabilistic discounting, reducing risky choice. **(D)** The $D_1$ agonist SKF 81297 induced a slight, non-significant increase in risky choice. **(E)** Infusions of the $D_2$ agonist quinpirole abolished discounting, decreasing risky choice during the initial block and increasing choice during the final block. **(F–I)** Effects on reward and negative-feedback sensitivity, indexed by win-stay and lose-shift ratios. For clarity and comparative purposes, the data are presented as difference scores between the ratios obtained on drug vs. control treatments (positive values indicate an increased ratio, negative values a decrease after drug treatment relative to control treatments). Adapted from St. Onge et al. (2011). Stars represent $p < 0.05$.

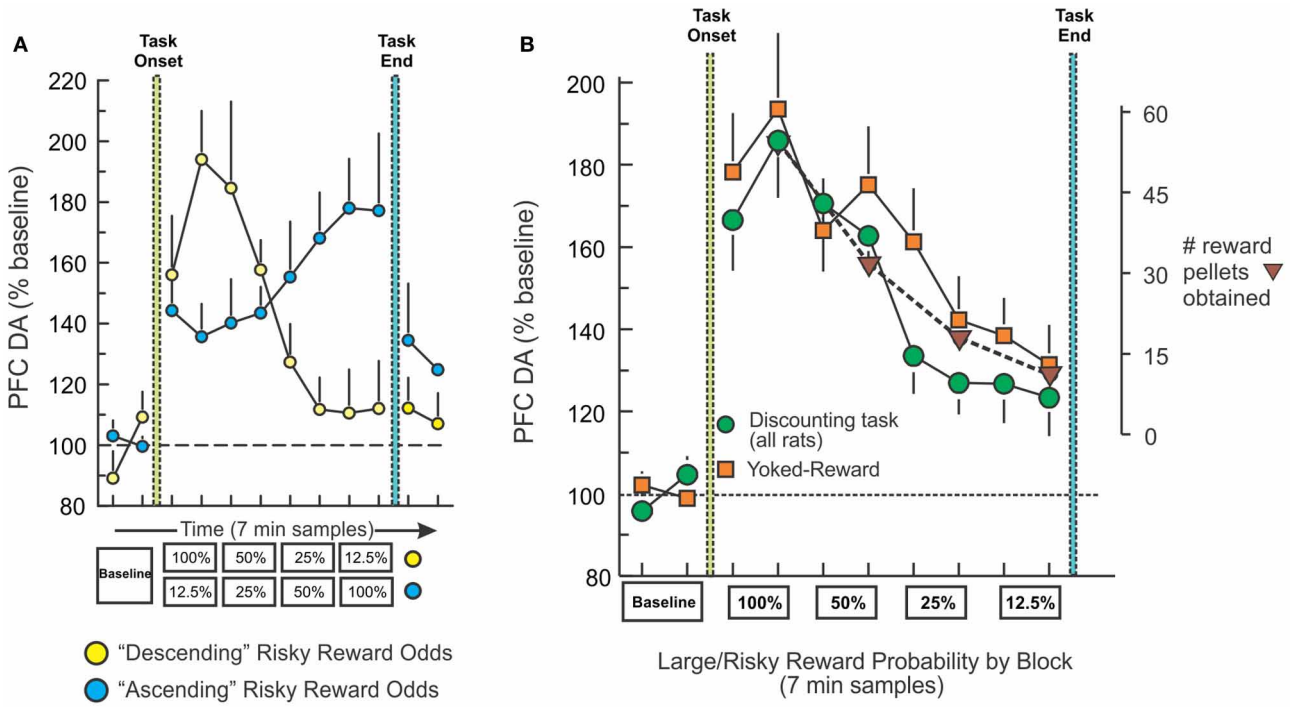



**FIGURE 4 | Fluctuations in PFC DA efflux during decision making track changes in reward rates over time. (A)** Percent change in basal PFC DA extracellular levels obtained from rats trained on the descending (yellow circles) and ascending (blue circles) variants of the probabilistic discounting task, plotted as a function of 7-min sample number. Rats tested on the descending version displayed an initial increase in DA that diminished as large/risky reward probabilities decreased, whereas those trained on the ascending version showed the opposite profile. **(B)** Change in PFC DA efflux for all rats trained on the both variants of the probabilistic discounting task (circles), plotted as a function of probability block. Combined data from rats in the yoked-reward experiment (squares) are also plotted. Triangles represent the number of reward pellets obtained by rats across task blocks. Changes in PFC DA efflux closely tracked changes in the relative amount of food obtain over the course of the session, irrespective of whether rats had to make decisions (task) or it the same amount of reward was delivered passively (yoked). Adapted from St. Onge et al. (2012).

(∼80–90% above baseline) that steadily declined over the session, whereas the opposite profile was observed when the odds were initially poor (12.5%) and subsequently increased (**Figure 4A**, blue). In this experiment, we included a key, yoked-reward control group consisting of rats that were not required to press any levers or make any decisions, but instead were accustomed to receiving food delivered passively on a schedule similar to rats performing the decision making task. Yoked rats displayed a profile of PFC DA efflux that was nearly identical to that observed during decision making, confirming that the fluctuations in PFC DA transmission during either condition corresponded primarily to changes in the relative rate of reward received (**Figure 4B**). These findings suggest that dopaminergic afferents to the frontal lobes convey information about changes in the relative amount of reward availability over time, irrespective of whether an organism actually has to do anything to retrieve that reward. However, these data suggest that in situations that require monitoring of changes in rates of reward delivery, dynamic fluctuations in tonic mesocortical DA levels may serve as a reward "running-rate meter," informing the PFC about changes in reward rates that can aid in adjusting choice accordingly (Niv et al., 2007).

The finding that PFC DA transmission is finely tuned to variations in reward availability provides additional insight into how pharmacological manipulations of DA activity might alter decision making. Thus, interfering with these dynamic signals via $D_1$ receptor blockade or stimulation would be expected to cause a discrepancy between the perceived vs. actual rates of reward obtained, leading to corresponding increases and decreases in risky choices. The fact that $D_2$ blockade altered decision making in a manner opposite to $D_1$ antagonism would suggest that $D_2$

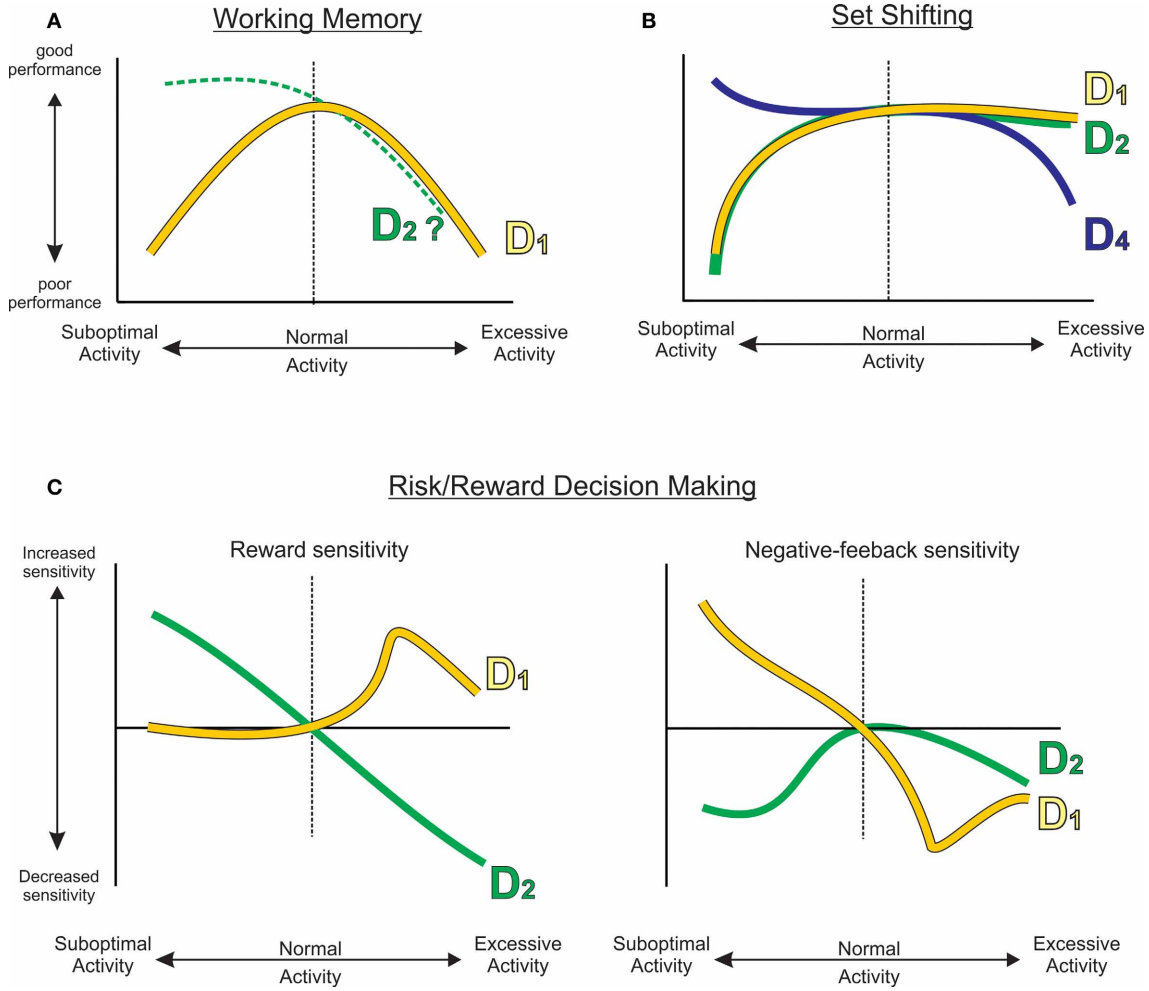



**FIGURE 5 | A family of functions describing how reduced or excessive PFC DA activity can affect different cognitive functions, including (A) working memory, (B) set shifting, or (C) different processes related to risk/reward decision making.** Curves have been extrapolated from empirical data (gray dashed curves) presented in **Figures 1**–**3**. The $D_2$ working memory curve was extrapolated from Druzin et al. (2000).

modulation of these functions may be less dependent on variations in extracellular PFC DA levels. However, the finding that $D_2$ receptor stimulation impaired probabilistic discounting implies that flooding $D_2$ receptors may disrupt the ability of a subgroup of PFC neurons to detect changes in PFC DA transmission over time, which may lead to more static patterns of choice.

## PFC DA AND COGNITION: A FAMILY OF FUNCTIONS

The findings reviewed here make it apparent that dopaminergic input to the frontal lobes is an essential component of the neural circuitry mediating a variety of cognitive and executive functions, including working memory, behavioral flexibility, and neuroeconomic processes related to cost/benefit decision making. Each of these requires distinct types of cognitive operations and functional neural circuits. Therefore, it is not surprising that the mechanisms by which DA exerts its effects are not unitary across these functions, but rather, each process relies on different patterns of activation of DA receptors. Thus, PFC $D_1$ receptor activity is of primary importance in mediating working memory, whereas, $D_1$ and $D_2$ receptors act either cooperatively or antagonistically to mediate functions related to behavioral flexibility and reward-related decision making. Moreover, although there is clear evidence that $D_1$ receptor modulation of working memory takes the form of an "inverted-U" shaped function, this profile is not necessarily shared by other receptors or across other PFC functions. A survey of the data reviewed here clearly demonstrates that, with respect to PFC DA, the "inverted-U" is not a one-size-fits-all function. Rather, it appears that mesocortical DA exerts its effects via a family of functions, wherein reduced vs. excessive DA activity may produce effects that are monotonic, sigmoidal, biphasic, exponential or polynomial across different cognitive domains (summarized in **Figure 5**).

The question remains: what are the potential mechanisms underlying these differential effects across cognitive domains? An answer may stem from contemporary theory on how these receptors differentially affect PFC neural network activity (Durstewitz et al., 2000; Seamans and Yang, 2004). $D_1$ receptors have been proposed to reduce the influence of weak inputs, stabilizing

network activity so that a subset of representation dominates PFC output. Conversely, $D_2$ activity attenuates inhibitory influences, allowing PFC neural ensembles to process multiple stimuli and/or representations, placing these networks in a more labile state that may permit changes in representations. With this conceptual framework in mind, it is likely that the cognitive operations underlying different functions would be mediated by distinct patterns of activity within PFC neural networks. Processes related to working memory require stable and persistent patterns of activity encoding information to be used across contexts or time (Goldman-Rakic, 1995; Lapish et al., 2008). The biophysical actions of $D_1$ receptors would be best suited for facilitating these patterns of activity. In comparison, shifting between different strategies has been linked to rapid reorganization of PFC neural ensemble activity that encodes different rules and action/outcome contingencies (Durstewitz et al., 2010). It is plausible that upon detection of rule changes, $D_2$ receptor activation destabilizes PFC network states, permitting the system to ascertain what the new course of action should be, and once a novel effective strategy has been recognized, stabilization of this new representation would be facilitated by $D_1$ receptor activity. Along similar lines, risk/reward decision making requires coordination between various cognitive processes, including those that facilitate flexible responding and action/outcome monitoring over time, which may be mediated by distinct populations of PFC neurons. By striking a fine balance between $D_1$ and $D_2$ receptor activity, mesocortical DA may help refine cost/benefit decisions between options of varying magnitude and uncertainty, with $D_1$ receptors promoting exploitation of current favorable circumstances and $D_2$ receptors facilitating exploration of more profitable ones when conditions change. Given these considerations, it is clear that a more comprehensive picture of how DA regulates frontal lobe functioning may be obtained not by painting every cognitive function with the same DA brush, but instead, taking into account the complex myriad of the neurophysiological actions of DA in combination with the neural network activity patterns underlying cognitive operations that subserve different PFC functions. Moreover, the advent of new technologies permitting manipulations of DA transmission in a more temporally and spatially specific manner will undoubtedly yield additional insight into how mesocortical DA regulates different forms of executive functioning. The picture that emerges from future studies of this kind will likely serve to both clarify and at the same time, further complicate our understanding of the functional contribution of PFC DA to cognition.

## ACKNOWLEDGMENTS

Much of the research described in this article was supported by grants from the Natural Sciences and Engineering Research Council of Canada and the Canadian Institutes of Health Research (MOP 89861) to Stan B. Floresco.

## REFERENCES

Arnsten, A. F. (1997). Catecholamine regulation of the prefrontal cortex. *J. Psychopharmacol.* 11, 151–162.

Bechara, A., Damasio, A. R., Damasio, H., and Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 50, 7–15.

Bechara, A., Damasio, H., Damasio, A. R., and Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *J. Neurosci.* 19, 5473–5481.

Birrell, J. M., and Brown, V. J. (2000). Medial frontal cortex mediates perceptual attentional set shifting in the rat. *J. Neurosci.* 20, 4320–4324.

Brown, V. J., and Bowman, E. M. (2002). Rodent models of prefrontal cortical function. *Trends Neurosci.* 25, 340–343.

Brozoski, T. S., Brown, R. M., Rosvold, H. E., and Goldman, P. S. (1979). Cognitive deficits caused by regional depletion of dopamine in prefrontal cortex of Rhesus monkey. *Science* 205, 929–932.

Chudasama, Y., and Robbins, T. W. (2004). Dopaminergic modulation of visual attention and working memory in the rodent prefrontal cortex. *Neuropsychopharmacology* 29, 1628–1636.

Clarke, H. F., Dalley, J. W., Crofts, H. S., Robbins, T. W., and Roberts, A. C. (2004). Cognitive inflexibility after prefrontal serotonin depletion. *Science* 304, 878–880.

Clarke, H. F., Hill, G. J., Robbins, T. W., and Roberts, A. C. (2011). Dopamine, but not serotonin, regulates reversal learning in the marmoset caudate nucleus. *J. Neurosci.* 31, 4290–4297.

Clarke, H. F., Walker, S. C., Crofts, H. S., Dalley, J. W., Robbins, T. W., and Roberts, A. C. (2005). Prefrontal serotonin depletion affects reversal learning but not attentional set shifting. *J. Neurosci.* 25, 532–538.

Collins, P., Roberts, A. C., Dias, R., Everitt, B. J., and Robbins, T. W. (1998). Perseveration and strategy in a novel spatial self-ordered sequencing task for nonhuman primates, effects of excitotoxic lesions and dopamine depletions of the prefrontal cortex. *J. Cogn. Neurosci.* 10, 332–354.

Crofts, H. S., Dalley, J. W., Collins, P., Van Denderen, J. C. M., Everitt, B. J., Robbins, T. W., et al. (2001). Differential effects of 6-OHDA lesions of the frontal cortex and caudate nucleus on the ability to acquire attentional set. *Cereb. Cortex* 11, 1015–1026.

Dias, R., Robbins, T. W., and Roberts, A. C. (1996). Primate analogue of the Wisconsin Card Sorting Test: effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behav. Neurosci.* 110, 872–886.

Dias, R., Robbins, T. W., and Roberts, A. C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test, restriction to novel situations and independence from "on-line" processing. *J. Neurosci.* 17, 9285–9297.

Druzin, M. Y., Kurzina, N. P., Malinina, E. P., and Kozlov, A. P. (2000). The effects of local application of D2 selective dopaminergic drugs into the medial prefrontal cortex of rats in a delayed spatial choice task. *Behav. Brain Res.* 109, 99–111.

Durstewitz, D., Seamans, J. K., and Sejnowski, T. J. (2000). Neurocomputational models of working memory. *Nat. Neurosci.* 3 (Suppl.), 1184–1191.

Durstewitz, D., Vittoz, N. M., Floresco, S. B., and Seamans, J. K. (2010). Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron* 66, 438–448.

Fletcher, P. J., Tenn, C. C., Rizos, Z., Lovic, V., and Kapur, S. (2005). Sensitization to amphetamine, but not PCP, impairs attentional set shifting: reversal by a D1 receptor agonist injected into the medial prefrontal cortex. *Psychopharmacology* 183, 190–200.

Floresco, S. B., Block, A. E., and Tse, M. T. (2008a). Inactivation of the medial prefrontal cortex of the rat impairs strategy set-shifting, but not reversal learning, using a novel, automated procedure. *Behav. Brain Res.* 190, 85–96.

Floresco, S. B., St. Onge, J. R., Ghods-Sharifi, S., and Winstanley, C. A. (2008b). Cortico-limbic-striatal circuits subserving different forms of cost-benefit decision making. *Cogn. Affect. Behav. Neurosci.* 8, 375–389.

Floresco, S. B., Magyar, O., Ghods-Sharifi, S., Vexelman, C., and Tse, M. T. (2006). Multiple dopamine receptor subtypes in the medial prefrontal cortex of the rat regulate set-shifting. *Neuropsychopharmacology* 31, 297–309.

Floresco, S. B., and Phillips, A. G. (2001). Delay-dependent modulation of memory retrieval by infusion of a dopamine D1 agonist into the rat medial prefrontal cortex. *Behav. Neurosci.* 115, 934–939.

Floresco, S. B., Zhang, Y., and Enomoto, T. (2009). Neural circuits subserving behavioral flexibility and their relevance to schizophrenia. *Behav. Brain Res.* 204, 396–409.

Gaspar, P., Bloch, B., and LeMoine, C. (1995). D1 and D2 receptor gene expression in rat frontal cortex: cellular localization in different classes of efferent neurons. *Eur. J. Neurosci.* 7, 1050–1063.

Gee, S., Ellwood, I., Patel, T., Luongo, F., Deisseroth, K., and Sohal, V. S. (2012). Synaptic activity unmasks dopamine D2 receptor modulation of a specific class of layer V pyramidal neurons in prefrontal cortex. *J. Neurosci.* 32, 4959–4971.

Ghods-Sharifi, S., Haluk, D. M., and Floresco, S. B. (2008). Differential effects of inactivation of the orbitofrontal cortex on strategy set shifting and reversal learning. *Neurobiol. Learn. Mem.* 89, 567–573.

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron* 14, 477–485.

Lapish, C. C., Durstewitz, D., Chandler, L. J., and Seamans, J. K. (2008). Successful choice behavior is associated with distinct and coherent network states in anterior cingulate cortex. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11963–11968.

Loos, M., Pattij, T., Janssen, M. C., Counotte, D. S., Schoffelmeer, A. N., Smit, A. B., et al. (2010). Dopamine receptor D1/D5 gene expression in the medial prefrontal cortex predicts impulsive choice in rats. *Cereb. Cortex* 20, 1064–1070.

McAlonan, K., and Brown, V. J. (2003). Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. *Behav. Neurosci.* 146, 97–103.

Mrzijak, L., Bergson, C., Pappy, M., Huff, R., Levenson, R., and Goldman-Rakic, P. S. (1996). Localization of dopamine D4 receptors in GABAergic neurons of the primate brain. *Nature* 381, 245–248.

Mueller, D., Bravo-Rivera, C., and Quirk, G. J. (2010). Infralimbic D2 receptors are necessary for fear extinction and extinction-related tone responses. *Biol. Psychiatry* 68, 1055–1060.

Muly, E. C., Szigeti, K., and Goldman-Rakic, P. S. (1998). D1 receptor in interneurons of macaque prefrontal cortex: distribution and subcellular localization. *J. Neurosci.* 18, 10553–10565.

Niv, Y., Daw, N. D., Joel, D., and Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology* 191, 507–520.

O'Neill, M., and Brown, V. J. (2007). The effect of striatal dopamine depletion and the adenosine A2A antagonist KW-6002 on reversal learning in rats. *Neurobiol. Learn. Mem.* 88, 75–81.

Pfeiffer, U. J., and Fendt, M. (2006). Prefrontal dopamine D4 receptors are involved in encoding fear extinction. *Neuroreport* 17, 847–850.

Phillips, A. G., Ahn, S., and Floresco, S. B. (2004). Magnitude of dopamine release in medial prefrontal cortex predicts accuracy of memory on a delayed response task. *J. Neurosci.* 24, 547–553.

Ragozzino, M. E. (2002). The effects of dopamine D1 receptor blockade on the prelimbic-infralimbic areas on behavioral flexibility. *Learn. Mem.* 9, 18–28.

Ragozzino, M. E., Detrich, S., and Kesner, R. P. (1999). Involvement of the prelimbic-infralimbic areas of the rodent prefrontal cortex in behavioral flexibility for place and response learning. *J. Neurosci.* 19, 4585–4594.

Robbins, T. W. (2005). Chemistry of the mind: neurochemical modulation of prefrontal cortical function. *J. Comp. Neurol.* 493, 140–146.

Robbins, T. W., and Arnsten, A. F. (2009). The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Annu. Rev. Neurosci.* 32, 267–287.

Roberts, A. C., De Salvia, M. A., Wilkinson, L. S., Collins, P., Muir, J. L., Everitt, B. J., et al. (1994). 6-Hydroxydopamine lesions of the prefrontal cortex in monkeys enhance performance on an analog of the Wisconsin Card Sort Test: possible interactions with subcortical dopamine. *J. Neurosci.* 14, 2531–2544.

Romanides, A. J., Duffy, P., and Kalivas, P. W. (1999). Glutamatergic and dopaminergic afferents to the prefrontal cortex regulate spatial working memory in rats. *Neuroscience* 92, 97–106.

Rudebeck, P. H., Walton, M. E., Smyth, A. N., Bannerman, D. M., and Rushworth, M. F. (2006). Separate neural pathways process different decision costs. *Nat. Neurosci.* 9, 1161–1168.

Sawaguchi, T., and Goldman-Rakic, P. S. (1994). The role of D1-dopamine receptor in working memory: local injections of dopamine antagonists into the prefrontal cortex of rhesus monkeys performing an oculomotor delayed-response task. *J. Neurosci.* 71, 515–528.

Schweimer, J., and Hauber, W. (2006). Dopamine D1 receptors in the anterior cingulate cortex regulate effort-based decision making. *Learn. Mem.* 13, 777–782.

Seamans, J. K., Floresco, S. B., and Phillips, A. G. (1998). D1 receptor modulation of hippocampal-prefrontal cortical circuits integrating spatial memory with executive functions in the rat. *J. Neurosci.* 18, 1613–1621.

Seamans, J. K., and Yang, C. R. (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog. Neurobiol.* 74, 1–57.

Seong, H. J., and Carter, A. G. (2012). D1 receptor modulation of action potential firing in a subpopulation of layer 5 pyramidal neurons in the prefrontal cortex. *J. Neurosci.* 32, 10516–10521.

Sesack, S. R., King, S. W., Bressler, C. N., Watson, S. J., and Lewis, D. A. (1995). Electron microscopic visualization of dopamine D2 receptors in the forebrain: cellular, regional, and species comparisons. *Soc. Neurosci. Abstr.* 21, 365.

Stefani, M. R., and Moghaddam, B. (2006). Rule learning and reward contingency are associated with dissociable patterns of dopamine activation in the rat prefrontal cortex, nucleus accumbens, and dorsal striatum. *J. Neurosci.* 26, 8810–8818.

St. Onge, J. R., Abhari, H., and Floresco, S. B. (2011). Dissociable contributions by prefrontal D1 and D2 receptors to risk-based decision making. *J. Neurosci.* 31, 8625–8633.

St. Onge, J. R., Ahn, S., Phillips, A. G., and Floresco, S. B. (2012). Dynamic fluctuations in dopamine efflux in the prefrontal cortex and nucleus accumbens during risk-based decision making. *J. Neurosci.* 32, 16880–16891.

St. Onge, J. R., and Floresco, S. B. (2010). Prefrontal cortical contribution to risk-based decision making. *Cereb. Cortex* 20, 1816–1828.

St. Onge, J. R., Stopper, C. M., Zahm, D. S., and Floresco, S. B. (2012). Separate prefrontal-subcortical circuits mediate different components of risk-based decision making. *J. Neurosci.* 32, 2886–2899.

Stopper, C. M., Khayambashi, S., and Floresco, S. B. (2013). Receptor-specific modulation of risk-based decision making by nucleus accumbens dopamine. *Neuropsychopharmacology* 38, 715–728.

Tunbridge, E. M., Bannerman, D. M., Sharp, T., and Harrison, P. J. (2004). Catechol-o-methyltransferase inhibition improves set-shifting performance and elevates stimulated dopamine release in the rat prefrontal cortex. *J. Neurosci.* 24, 5331–5335.

Walton, M. E., Bannerman, D. M., Alterescu, K., and Rushworth, M. F. (2003). Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *J. Neurosci.* 22, 10996–11003.

Wang, M., Vijayraghavan, S., and Goldman-Rakic, P. S. (2004). Selective $D_2$ receptor actions on the functional circuitry of working memory. *Science* 303, 853–856.

Wedzony, K., Chocyk, A., Mackowiak, M., Fijal, K., and Czyrak, A. (2001). Cortical localization of dopamine D4 receptors in the rat brain-immunocytochemical study. *J. Physiol. Pharmacol.* 51, 205–221.

Williams, G. V., and Castner, S. A. (2006). Under the curve: critical issues for elucidating D1 receptor function in working memory. *Neuroscience* 139, 263–276.

Winstanley, C. A., Theobald, D. E., Cardinal, R. N., and Robbins, T. W. (2004). Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *J. Neurosci.* 24, 4718–4722.

Zahrt, J., Taylor, J. R., Mathew, R. G., and Arnsten, A. F. (1997). Supranormal stimulation of D1 dopamine receptors in the rodent prefrontal cortex impairs working memory performance. *J. Neurosci.* 17, 8528–8535.

Zeeb, F. D., Floresco, S. B., and Winstanley, C. A. (2010). Contributions of the orbitofrontal cortex to impulsive choice: interactions with basal levels of impulsivity, dopamine signaling, and reward-related cues. *Psychopharmacology* 211, 87–98.

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 14 February 2013; paper pending published: 09 March 2013; accepted: 04 April 2013; published online: 19 April 2013.*

*Citation: Floresco SB (2013) Prefrontal dopamine and behavioral flexibility: shifting from an "inverted-U" toward a family of functions. Front. Neurosci. 7:62. doi: 10.3389/fnins.2013.00062*

*This article was submitted to Frontiers in Decision Neuroscience, a specialty of Frontiers in Neuroscience.*

frontiers in
NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 14 August 2013
doi: 10.3389/fnins.2013.00142

# Dopaminergic drug effects during reversal learning depend on anatomical connections between the orbitofrontal cortex and the amygdala

***Marieke E. van der Schaaf[1,2]*, Marcel P. Zwiers[1,2], Martine R. van Schouwenburg[1,2], Dirk E. M. Geurts[1,2], Arnt F. A. Schellekens[2], Jan K. Buitelaar[3], Robbert Jan Verkes[2] and Roshan Cools[1,2]***

[1] *Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands*
[2] *Department of Psychiatry, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands*
[3] *Department of Cognitive Neuroscience, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands*

***Edited by:***
*Jeff A. Beeler, The University of Chicago, USA*

***Reviewed by:***
*Kerstin Preuschoff, École Polytechnique Fédérale de Lausanne, Switzerland*
*Edythe D. London, University of California, Los Angeles, USA*

****Correspondence:***
*Marieke E. van der Schaaf, Donders Centre for Cognitive Neuroimaging, Radboud University Nijmegen, Kapittelweg 29, PO Box 9101, 6525 EN Nijmegen, Netherlands*
*e-mail: Marieke.vanderschaaf@donders.ru.nl*

Dopamine in the striatum is known to be important for reversal learning. However, the striatum does not act in isolation and reversal learning is also well-accepted to depend on the orbitofrontal cortex (OFC) and the amygdala. Here we assessed whether dopaminergic drug effects on human striatal BOLD signaling during reversal learning is associated with anatomical connectivity in an orbitofrontal-limbic-striatal network, as measured with diffusion tensor imaging (DTI). By using a fiber-based approach, we demonstrate that dopaminergic drug effects on striatal BOLD signal varied as a function of fractional anisotropy (FA) in a pathway connecting the OFC with the amygdala. Moreover, our experimental design allowed us to establish that these white-matter dependent drug effects were mediated via D2 receptors. Thus, white matter dependent effects of the D2 receptor agonist bromocriptine on striatal BOLD signal were abolished by co-administration with the D2 receptor antagonist sulpiride. These data provide fundamental insight into the mechanism of action of dopaminergic drug effects during reversal learning. In addition, they may have important clinical implications by suggesting that white matter integrity can help predict dopaminergic drug effects on brain function, ultimately contributing to individual tailoring of dopaminergic drug treatment strategies in psychiatry.

**Keywords: dopamine, striatum, amygdala, OFC, reversal learning, diffusion tensor imaging, bromocriptine, sulpiride**

## INTRODUCTION

Adequate dopamine neurotransmission is well-known to be important for reward and reversal learning and accumulating evidence indicates that the effects of dopamine on such learning implicate the striatum. Consistent with current theoretical modeling work (Badre and Frank, 2012), pharmacological functional magnetic resonance imaging (fMRI) studies have revealed dopaminergic drug effects on striatal BOLD signals during reversal learning (Cools et al., 2001, 2007; Dodds et al., 2008; Van Der Schaaf et al., 2012). In addition, positron emission tomography (PET) studies have demonstrated that reversal learning depends on striatal dopamine synthesis capacity (Cools et al., 2009) and D2 receptor availability in the striatum (Groman et al., 2011).

However, the striatum does not act in isolation and reversal learning is also well-accepted to depend on the interaction between striatum, orbitofrontal cortex (OFC) and amygdala (Iversen and Mishkin, 1970; Jones and Mishkin, 1972; Holland and Gallagher, 2004; Schoenbaum et al., 2009; Murray and Wise, 2010). This interaction is thought to be modulated by dopamine. Specifically, medium spiny neurons in the striatum and amygdala that receive glutamatergic projections from limbic and cortical regions also receive converging dopaminergic projections from the midbrain (Pennartz et al., 1994; Rosenkranz and Grace, 2002a; Sesack et al., 2003; Haber and Knutson, 2010). Animal studies have suggested that the effects of dopamine on these glutamatergic inputs are receptor specific, such that orbitofrontal inputs to the striatum are modulated by D2 receptor stimulation (Del Arco et al., 2007; Grace et al., 2007; Del Arco and Mora, 2009; Sesack and Grace, 2010) while orbitofrontal inputs to the amygdala and amygdala inputs to the striatum are modulated by D1 receptor stimulation (Rosenkranz and Grace, 2002b; Ambroggi et al., 2008; Sesack and Grace, 2010). These observations have led to the suggestion that dopamine regulates the degree to which the striatum, amygdala and OFC interact to integrate information about reward value, motivation and expectation and to ultimately facilitate adaptive and flexible behavior (Grace et al., 2007; Haber and Knutson, 2010; Pennartz et al., 2011).

Here we aim to provide evidence for such network effects of dopamine during human reversal learning by revisiting our recent pharmacological fMRI study that showed dopaminergic drug effects on striatal BOLD signal during reversal learning (Van Der Schaaf et al., 2012). Specifically, we ask whether these previously reported effects of dopamine on the striatum during reversal learning are associated with anatomical connectivity in an

orbitofrontal-limbic-striatal network, as measured with diffusion tensor imaging (DTI). A demonstration that drug effects are associated with individual differences in anatomical connectivity will not only address the question about whether dopamine's effects are associated with an orbitofrontal-limbic-striatal network of regions, but will also help elucidate individual trait factors that contribute to the known large variability in dopaminergic drug effects (Cools and D'esposito, 2011). Thus, individual differences in anatomical connections between the OFC, amygdala and striatum might predict the extent and direction of dopaminergic drug effects on reversal learning.

DTI is a non-invasive method to measure structural connectivity in humans. It measures the diffusion of water in tissue, which depends on the tight packing of cellular axons and myelin sheets that encapsulate the axon fibers. Two measures are generally obtained; fractional anisotropy (FA) and mean diffusivity (MD). FA is a measure of the directionality of water diffusion and has been associated with dense coherent bundling and myelination of axons. MD reflects the general presence of barriers to free diffusion and has been associated with overall cell density. Collectively, FA and MD provide information on the microstructural integrity and communicational efficacy of white matter fiber bundles (Beaulieu, 2002; Thomason and Thompson, 2011).

The hypothesis that individual differences in functional effects depend on anatomical connectivity as measured with DTI is grounded in prior work linking anatomical connectivity with individual differences in functional effects (Boorman et al., 2007; Cohen et al., 2008; Harsay et al., 2011; Samanez-Larkin et al., 2012). In addition, we have previously shown that dopaminergic drug effects on striatal BOLD signals during attention-shifting are associated with white matter integrity of dorsal fronto-striatal-thalamic pathways (Van Schouwenburg et al., 2013). These results concur with the known role of dorsal fronto-striatal-thalamic pathways in cognitive functions such as attention shifting (Dias et al., 1996). By contrast, reversal learning depends on a ventral orbitofronto-limbic-striatal network (Dias et al., 1996). In the present study we used a fiber based approach (Mandl et al., 2012) to substantiate the observation that dopaminergic drug effects can be predicted from anatomical connectivity, while also showing the neuroanatomical specificity of such findings. Based on the literature reviewed above, we predict that drug effects on striatal BOLD signal during reversal learning will depend on a ventral orbitofronto-limbic-striatal network and not on a dorsal fronto-thalamic-striatal network.

A final aim of this study was to assess the receptor specificity of the effects (Feldman et al., 1997). As described above, work with experimental animals has suggested that the dopaminergic modulation of interactions between the OFC, amygdala and striatum is dopamine receptor specific (Rosenkranz and Grace, 2002b; Del Arco et al., 2007; Grace et al., 2007; Ambroggi et al., 2008; Del Arco and Mora, 2009; Sesack and Grace, 2010). In addition, it has been demonstrated that reversal learning in monkeys specifically depends on D2 and not D1 receptor functioning (Lee et al., 2007). To address this issue in humans we employed a coadministration design. All subjects were scanned on four occasions: after administration of placebo; after administration of the dopamine D1/D2 receptor agonist bromocriptine; after administration of the dopamine D2 receptor antagonist sulpiride; and after combined administration of both sulpiride and bromocriptine. If drug effects are mediated by D2 receptors, then any significant effect of bromocriptine relative to placebo should be abolished by coadministration of sulpiride. If effects of bromocriptine are mediated by D1 receptors, then they should not be abolished by coadministration of sulpiride.

## METHODS

### SUBJECTS

The present study represents an extension of a previously published pharmacological fMRI study (Van Der Schaaf et al., 2012) with diffusion tensor images that were acquired from the same subjects during an intake session prior to the drug sessions. For this study, 28 healthy right handed volunteers with no relevant medical/psychiatric history 3 years prior to testing were tested after a medical screening [for screening procedure and exclusion criteria see Van Der Schaaf et al. (2012)]. They gave written informed consent approved by the local ethics committee (Commissie mensgebonden onderzoek, Arnhem-Nijmegen, number 2008/078, date 09-09-2008) and were compensated for participation. In total, 8 subjects were excluded from the fMRI analysis due to personal issues (1), technical issues (4), excessive head movement (2) and insufficient practice of the Dutch language (1) [see Van Der Schaaf et al. (2012) for further details on these exclusions]. Complete datasets including both DTI and all four fMRI sessions were available for twenty subjects (10 males, mean age: 22.7, range: 18.9–29.1).

### PROCEDURES AND PHARMACOLOGICAL DESIGN

Subjects were tested on four occasions, separated by at least 1 week. They were tested after oral intake of placebo, bromocriptine (Parlodel, Novartis®, 1.25 mg), sulpiride (Dogmatil, sanovaaventis®, 400 mg), and a combination of bromocriptine and sulpiride (sulpiride was administered 30 min prior to bromocriptine). Administration was randomized according to a counterbalanced, placebo controlled, double blind, double dummy design. The reversal learning task started 2¼ h after first drug intake with a duration of 60 min. Blood pressure, heart rate and subjective mood ratings [visual analog scales (Bond and Lader, 1974)] were taken 30 min before, 2 h after and 6 h after first drug intake. Blood samples were taken to determine the change in prolactin levels (Fitzgerald and Dinan, 2008) and were taken 30 min before and 2 h after first drug intake. Background neuropsychological tests (block completion, number cancellation, verbal fluency and digit span) were assessed 5 h after first drug intake. Drug effects on physiological measures were as expected with prolactin increases after intake of sulpiride and combined administration and prolactin and systolic blood pressure decreases after intake of bromocriptine. Analyses of the questionnaires and background neuropsychology are described in our previous report and revealed no significant drug effects on mood or general cognitive functioning. For further details on the screening and session procedures, physiology, mood and background neuropsychology see (Van Der Schaaf et al., 2012).

### REVERSAL LEARNING TASK

On each trial, subjects were presented with two simultaneously presented vertically adjacent stimuli, a face and a scene (location randomized). One of these stimuli was associated with reward and the other with punishment. One of the stimuli was selected by the computer (highlighted with a black border) and subjects were asked to predict the outcome associated with this preselected stimulus. After the prediction, indicated with a right index or middle finger button press (counterbalanced across subjects), the actual outcome was presented (100% deterministic). Note that these outcomes did not depend on subjects' responses but were directly coupled to the highlighted stimulus. The stimulus-outcome contingencies reversed after 4, 5, or 6 consecutive correct predictions. Such reversals were signaled by either an unexpected punishment (presented after a previously rewarded stimulus was highlighted) or an unexpected reward (presented after a previously punished stimulus was highlighted). On the trials directly following these unexpected outcomes (reversal trials), the same stimulus was highlighted again such that requirements for motor switching and prediction updating were matched between reward and punishment conditions. Accuracy on these reversal trials reflects how well-subjects updated stimulus-outcome associations after either unexpected rewards or unexpected punishments. The dependent variables used for the current report were striatal BOLD signaling during unexpected outcomes and the proportion of correct responses on reversal trials (see below).

### IMAGE ACQUISITION AND PREPROCESSING

Structural images were collected before the start of the experiment during screening using a 3-tesla Siemens MRI scanner with an 8 channel head coil. For each subject, a high resolution T1-weighted MP-RAGE anatomical scan (TE/TR = 3.03/2300 ms, flip angle = 8°, *FOV* = 256 mm × 256 mm × 192 mm, voxel size = 1 mm isotropic, GRAPPA acceleration factor 2) was obtained. Diffusion tensor images were acquired using a twice refocused spin-echo-planar imaging sequence to reduce spatial distortions caused by eddy currents (Reese et al., 2003). Sixteen subjects were scanned with the following protocol: 64 slices interleaved acquisition mode (TE/TR = 89/6700 ms, flip angle = 90°, *FOV* = 220 mm, voxel size = 2.2 mm isotropic). Acquisition consisted of 7 images without diffusion weighting ($b = 0$) and 61 images with diffusion weighting ($b = 1000\,\mathrm{s/mm^2}$) applied along the non-colinear directions. Four subjects were scanned with slightly modified protocol in which the TR was 8500 ms and images were acquired with partial instead of full Fourier with a slightly lower band width.

Raw diffusion weighted imaging (DWI) data were pre-processed using in-house software (Zwiers, 2010). The DTI images were realigned using rigid body transformations and mutual information as a cost function (SPM8). Susceptibility induced echo-planar imaging distortions were corrected by warping the images along the phase-encoding direction to the distortion-free T1 reference images (Studholme et al., 2000) using an in-house developed implementation (Visser et al., 2010). Diffusion tensors were then estimated using a robust artifact-insensitive compute algorithm (Zwiers, 2010). FA and MD measures were computed from the diffusion tensor eigenvalues. FA and MD maps were normalized to the T1 ICBM-template (MNI-space) using the unified segmentation parameters of the co-registered structural image. Images were then smoothed with a Gaussian kernel of 8 mm full width half maximum and masked with a full brain mask. Imaging parameters, pre-processing and analysis of the functional images, obtained during the drug sessions, are described elsewhere (Van Der Schaaf et al., 2012).

### GENERAL ANALYSIS STRATEGY

In our prior work we reported dopaminergic drug effects on striatal BOLD signal during reward and punishment reversal learning (Van Der Schaaf et al., 2012). These BOLD effects were centered on the ventral lateral putamen, a region that receives convergent inputs from both OFC and the amygdala (Draganski et al., 2008; Haber and Knutson, 2010). Here, we revisit our data and ask whether the observed dopaminergic drug effects on striatal BOLD signaling is associated with anatomical connections between the striatum, OFC and amygdala. Thus, we investigated individual differences in white matter integrity of anatomical pathways in an orbitofronto-limbic-striatal network, as indexed by diffusion tensor images that were acquired from the same subjects during an intake session prior to the drug sessions.

We used a fiber-based approach (Mandl et al., 2012) and focused on three anatomical white matter pathways of interest—(1) a pathway connecting the OFC with the striatum (Ongür and Price, 2000; Ogar and Gorno-Tempini, 2007; Haber and Knutson, 2010; Balleine et al., 2011), (2) a pathway connecting the amygdala with the striatum (Robbins et al., 1989; Everitt et al., 1991; Ambroggi et al., 2008), and (3) a pathway connecting the OFC with the amygdala (Baxter et al., 2000; Stalnaker et al., 2007)—, and one anatomical white matter pathway of no interest for control purposes [a pathway connecting the dorsal PFC (dPFC) with the striatum (Haber and Knutson, 2010)]. As described in the introduction, this additional pathway was included to demonstrate specificity of the effects to orbitofronto-limbic-striatal pathways, involved in reward processing and stimulus-outcome valuation. Thus, we anticipated that any effects would not extend to dorsal fronto-striatal pathways that have instead been associated with more cognitive processes and motor control (Alexander et al., 1990; Haber and Knutson, 2010). These study-specific anatomical volumes of interest were first created using probabilistic tractography (see probabilistic tractography section below) and average FA and MD values were extracted from each pathway. These FA and MD values were then used as independent predictor variables in multiple regression analyses with the drug-related change in striatal BOLD signal during reversal learning as the dependent variable (see statistical analysis section below).

### DEPENDENT VARIABLE I: SELECTION OF STRIATAL BOLD SIGNAL

Striatal BOLD signal was extracted for each drug session from the locus that exhibited the significant drug effect during reversal learning, as reported previously (Van Der Schaaf et al., 2012). This drug effect was centered on the left ventral putamen ($x$, $y$, $z = -22$, 18, 4, $p_{\mathrm{fwe}} = 0.03$) (**Figure 2A**) and reflected opposite modulation by the dopamine receptor antagonist sulpiride and the dopamine receptor agonist bromocriptine of BOLD signal change during unexpected relative to expected outcomes. Mean

beta estimates from this peak voxel were extracted with MarsBar software (Brett et al., 2002) for each drug session. The use of such a functionally defined timeseries is justified because the aim of our investigation was to account for variability in exactly this signal. Drug-related change in the extracted beta-values (representing signal during unexpected vs. expected outcomes) was then used as a dependent variable in linear regression analysis with the DTI-measurements as predictor variables (see below).

### DEPENDENT VARIABLE II: SELECTION OF THE BEHAVIORAL VALUES

The behavioral measures of interest were the valence-dependent and valence-independent reversal learning scores. These were calculated by computing, respectively, the difference between, and the average of the proportion of correct responses on reward and punishment reversal trials. The accuracy scores were arcsine transformed [$2 \times \arcsin(\sqrt{x})$] as is appropriate when the variance is proportional to the mean (Howell, 1997).

### PROBABILISTIC TRACTOGRAPHY: SELECTION OF FRONTAL-STRIATAL-LIMBIC PATHWAYS

Orbitofronto-limbic-striatal pathways are not yet included in white matter atlases. Accordingly, these study-specific anatomical pathways were created using probabilistic tractography as implemented in FMRIB's diffusion toolbox [See also (De Zeeuw et al., 2012; Mandl et al., 2012; Peper et al., 2013) for similar procedures]. In total four pathways were created; OFC—striatum, OFC—amygdala, amygdala—striatum and dorsal PFC—striatum. Masks used for tractography were defined in standard space using the AAL-template (Tzourio-Mazoyer et al., 2002) (**Figure1A**). Because the locus of the drug effects was centered on the left striatum we focussed our analysis on pathways in the left hemisphere. The left amygdala was defined as AAL-region 41, the left OFC as the gyrus rectus and orbito gyrus regions (Ogar and Gorno-Tempini, 2007) (AAL-regions 5, 9, 15, 25 and 2), the left dPFC as all left superior, middle and inferior frontal gyrus regions (AAL-regions 3, 7, 13, 23) and the left striatum as the left putamen and caudate nucleus (AAL-regions 71 and 73). Ventral and dorsal striatal subregions are not clearly separated by anatomical boundaries and best defined by its afferent projections from cortical areas (Haber and Knutson, 2010). Accordingly, we seeded our tractography from the OFC, dPFC and amygdala and used the whole striatum as waypoint mask.

For each pathway, waypoint (a.k.a. inclusion) and exclusion masks were defined as followed: (1) *OFC-Striatum*: Seed = OFC, waypoint = striatum, exclusion = dPFC, amygdala and planes excluding $x > 1$ and $y < -18$. (2) *Amygdala-Striatum*: Seed = amygdala, waypoint = striatum, exclusion = OFC, dPFC and planes excluding $x > 1$ and $y < -18$. (3) *OFC-Amygdala*: Seed = OFC, waypoint = amygdala, exclusion = dPFC, striatum, and planes excluding $x > 1$, and $y < -18$. (4) *dPFC-striatum*: Seed dPFC, waypoint: striatum, exclusion: OFC, amygdala and planes excluding $x > 1$ and $y < -18$. These masks were brought back into native space, using the inverse of the computed normalization parameters to create individual probabilistic diffusion pathways. Using FMRIB's diffusion toolbox [FMRIB's Software Library (FSL), bedpostx], fiber orientation probabilistic density functions were estimated at each voxel, allowing for multiple fiber directions (Behrens et al., 2007). Five thousand streamline samples per seed voxel were drawn through the probability density functions to form an estimate of the probability distribution of connections from each seeded voxel ("probtrackx" with a curvature threshold of 0.2). All pathways from the seed region that passed through the exclusion mask and all pathways that did not pass through the waypoint mask were discarded from the calculation of the connectivity distribution. The resulting connectivity distribution files are images in which the values at each voxel represent the number of samples between the seed and waypoint mask that passed through that voxel. These images were then brought back to standard space, using individual normalization parameters, thresholded to include voxels through which at least 1% of the samples passed, binarized and summed across subjects.

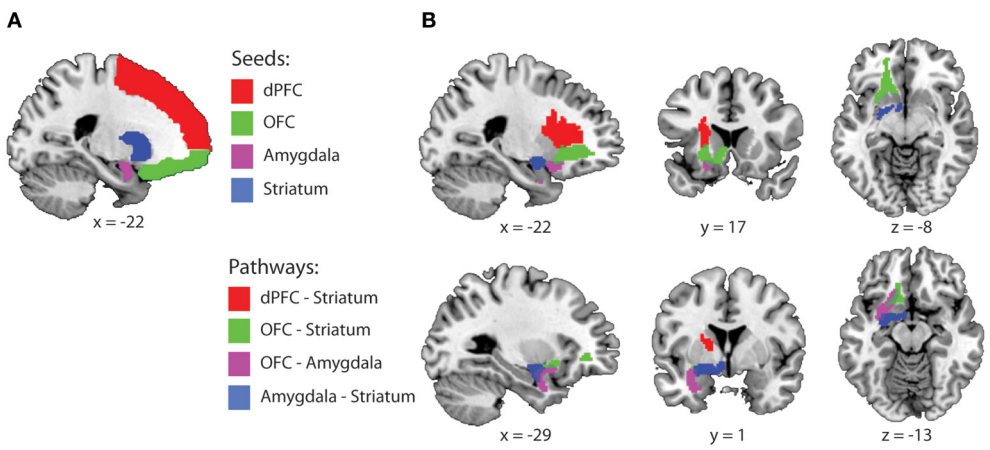



**FIGURE 1 | Fronto-striatal-limbic pathways. (A)** Seed and waypoint masks that were used for probabilistic tractography, displayed on a MNI-template. **(B)** The four study-specific anatomical ROIs used for FA and MD data extraction. Colored masks represent the binarized group masks for pathways that were present in at least 75% of all subjects. Abbreviations: dPFC, dorsal prefrontal cortex; OFC, Orbitofrontal cortex.

The 4 study-specific anatomical VOI's were created at the group level representing those pathways that were present in at least 75% of all subjects (**Figure 1B**). These are commonly used thresholds and are similar or more conservative compared with thresholds used in other fibre-based DTI tractography studies (Leh et al., 2007; Gutman et al., 2009; Mandl et al., 2012; Peper et al., 2013). Finally, the individual mean FA and MD values were extracted from each pathway.

### STATISTICAL ANALYSIS

Because we used different scanning protocols, the extracted FA and MD values of each pathway were first residualized with respect to protocol. Multiple linear regression analysis (SPSS, version 19.0.0) was done with the residualized FA values from the four pathways as predictor variables and the drug-related difference in beta values (BOLD) as dependent variable. A stepwise procedure was applied to include only those predictors that significantly contributed the model. The probability to enter or remove a predictor was set at 0.05 and 0.1, respectively (default). Consistent with our previous report, assessments of the different drug comparisons were done in a fixed a priori defined order. First, we investigated which of the pathways contributed to the effects of dopamine receptor stimulation (bromocriptine) relative to dopamine receptor blockade (sulpiride) on striatal BOLD. Next, for the pathways that were revealed in the first step, we assessed whether their contribution was driven by effects of bromocriptine relative to placebo or by effects of sulpiride relative to placebo. Finally, to establish the D2 receptor dependency of the observed effects, we assessed whether they were blocked by combined administration. The same procedures were used to assess associations between drug effects on behavior and FA values from these pathways. To further support the nature of our FA findings we also assessed the association between drug effect on BOLD and MD values in the pathways that yielded a significant relationship from the analysis described above. While FA values represent the orientation-dependence of water diffusion, which is directional in white matter fibers, MD values represent the overall magnitude of water diffusion. MD depends on fiber and membrane density and, in white matter, increases in MD have been associated with the degeneration of fiber bundles (Beaulieu, 2002; Thomason and Thompson, 2011). Accordingly, when, across subjects, higher FA values in white matter are accompanied by lower MD values, this likely reflects higher levels of fiber and membrane density within non-crossing fiber bundles. Conversely, when across subjects, higher FA values are accompanied by higher MD values, this possibly reflects selectively lower levels of fiber and membrane density within of one of the fiber bundles in a crossing fiber region.

Finally, for completion, main effects of drugs on behavior were assessed with repeated measures ANOVA with the within-subjects factors drug and valence (reward and punishment). The order of drug comparisons were assessed in the same a priori defined order as described above.

### SUPPLEMENTARY ANALYSIS

In addition to the volume of interest analyses, we conducted supplementary voxel-wise regression analysis at the whole-brain level, using random effects multiple regression procedures in SPM8 (http://www.fil.ion.ucl.ac.uk/spm). This allowed us to visualize the (physiological plausibility of) effects that were revealed to be statistically significant using the volume of interest analyses. To this end, individual FA-maps were submitted to a second-level one sample $T$-test and the drug-related changes in striatal BOLD signal were entered as a covariate of interest. Scan-protocol was entered as a covariate of no-interest. Voxels revealed by the covariate of interest represent white matter regions that exhibit a linear relationship between individual FA-values and drug effects on striatal BOLD. Effects are displayed for visualization purposes only at a threshold of $p < 0.001$ uncorrected for multiple comparisons (**Figure 3**). Next, probabilistic diffusion tractography (FMRIB's diffusion toolbox) was used to visualize the pathways connecting with the FA region revealed by the voxel-wise regression analysis as a seed. To this end, the FA seed region was defined as a 4 mm sphere around the peak voxel of the FA region ($x$, $y$, $z = -34$, 8, $-6$) revealed by the regression analysis. For each subject, this region was brought back into native space, using the inverse of the computed normalization parameters and used as a seed region for probabilistic tractography (same settings as above). The resulting connectivity distribution images were brought back to standard space, using individual normalization parameters, and tractography maps were thresholded to include only voxels through which at least 1% of all samples had passed. These individual maps were then binarized and summed across subjects to produce group probability maps.

## RESULTS

Linear regression analysis revealed a significant positive relationship between drug effects on striatal BOLD (bromocriptine—sulpiride) and FA values from the OFC-amygdala pathway [$F_{(1,\ 19)} = 8.33$, $R^2 = 0.32$, adjusted $R^2 = 0.28$, $\beta = 0.56$, $T = 2.89$, $p = 0.010$]. No significant contribution of the OFC-striatum ($\beta = 0.06$, $T = 0.22$, $p = 0.82$), dPFC-striatum ($\beta = 0.01$, $T = -0.03$, $p = 0.98$) or the amygdala-striatum pathways ($\beta = -0.04$, $T = -0.18$, $p = 0.86$) were revealed. FA values from the OFC-amygdala pathway were associated with the effects of bromocriptine relative to placebo on striatal BOLD [$F_{(1,\ 19)} = 5.63$, $R^2 = 0.24$, adjusted $R^2 = 0.19$, $\beta = 0.48$, $T = 2.37$, $p = 0.029$], but not of sulpiride relative to placebo (beta $= -0.14$, $p = 0.55$). Moreover, these white-matter dependent effects of bromocriptine on striatal BOLD were abolished by co-administration of both drugs; FA-values from the OFC—amygdala pathway correlated significantly with the effects of bromocriptine relative to combined administration on striatal BOLD [$F_{(1,\ 19)} = 10.73$, $R^2 = 0.37$, adjusted $R^2 = 0.34$, $\beta = 0.61$, $T = 3.28$, $p = 0.004$], but not with the effects of placebo relative to combined administration on striatal BOLD ($\beta = -0.23$, $p = 0.34$) (**Figure 2B**).

Subsequent correlation analyses with MD-values from the OFC–amygdala pathway revealed a negative relationship between the effect of bromocriptine relative to sulpiride on striatal BOLD and MD-values from the OFC-amygdala pathway ($\beta = -0.61$, $p = 0.004$). Thus, the found associations with FA were accompanied by associations with fiber density within the OFC-amygdala pathway. Taken together, these data show that bromocriptine

increased reversal-related striatal BOLD in subjects with high FA-values in the OFC-amygdala pathway, while it decreased striatal BOLD in subjects with low FA-values in this pathway. These effects were likely mediated via D2 receptor stimulation, as effects were abolished by co-administration with sulpiride.

There were no associations between FA-values and drug effects (reported here are effects of bromocriptine relative to placebo) on behavioral measures of valence-dependent reversal learning (OFC-amygdala: $\beta = -0.10$, $p = 0.68$; OFC-striatum: $\beta = -0.23$, $p = 0.34$; dPFC-striatum: $\beta = -0.03$, $p = 0.89$; Amygdala-striatum: $\beta = -0.27$, $p = 0.25$) or valence-independent reversal learning (OFC-amygdala: $\beta = -0.34$, $p = 0.14$; OFC-striatum: $\beta = -0.23$, $p = 0.33$; dPFC-striatum: $\beta = -0.25$, $p = 0.30$; amygdala-striatum: $\beta = -0.06$, $p = 0.81$).

For completeness, we also assessed drug effects on behavior irrespective of FA values. This revealed a trend toward opposite effects of bromocriptine and sulpiride on reward and punishment reversal learning [drug × valence: $F_{(1,\ 19)} = 4.2$, $P = 0.054$]. This was due to better punishment relative to reward learning after bromocriptine (raw accuracy scores ± standard error of the mean: reward: $0.90 \pm 0.02$; punishment: $0.92 \pm 0.01$), but better reward relative to punishment learning after sulpiride (reward: $0.93 \pm 0.02$; punishment: $0.90 \pm 0.02$). However, no drug by valence effects were seen when comparing bromocriptine with placebo (reward: $0.90 \pm 0.02$; punishment: $0.89 \pm 0.03$) [drug × valence: $F_{(1,\ 19)} = 1.8$, $P = 0.20$].

## SUPPLEMENTARY ANALYSES

Results from the brain-wide voxel wise regression analyses and subsequent tractography concurred with the results from the volumes of interest analyses reported above. Thus, brain-wide analysis revealed that FA in a region within the uncinate fasciculus, as identified with the JHU white matter tractography atlas, predicted drug effects on striatal BOLD signal (bromocriptine—sulpiride: $x, y, z = -34, 8, -6$, $T = 4.87$, $p_{\text{unc}} < 0.001$; bromocriptine—placebo: $x, y, z = -30, 10, -8$, $T = 3.84$, $p_{\text{unc}} < 0.001$; bromocriptine-combined: $x, y, z = -26, 4, -12$, $T = 6.51$, $p_{\text{unc}} < 0.001$) (**Table 1**, **Figure 3A**). Probabilistic tractography from this region revealed an extended network of pathways between the OFC, amygdala and striatum. Other pathways revealed by tractography included connections with the insular cortex and a pathway along the inferior longitudinal fasciculus along the hippocampus and toward the visual cortex. No pathways toward the thalamus or midbrain regions were seen. These tractography findings further support that our findings likely involve white matter integrity in the OFC-amygdala pathway, rather than direct fronto-striatal pathways, as the latter

**Table 1 | Whole brain results from the voxel-wise regression analysis of FA regions that showed a linear correlation with drug effects on striatal BOLD.**

| Side | *N* voxels | *T*-value | MNI | | |
|---|---|---|---|---|---|
| | | | *x* | *y* | *z* |
| **BROMOCRIPTINE—SULPIRIDE (POSITIVE)** | | | | | |
| L | 14 | 5.12 | −34 | 0 | 42 |
| L | **35** | **4.87** | **−34** | **8** | **−6** |
| R | 16 | 4.77 | 26 | 40 | 14 |
| L | 13 | 4.43 | −10 | 38 | 46 |
| L | 17 | 4.34 | −30 | −20 | −46 |
| **BROMOCRIPTINE—PLACEBO (POSITIVE)** | | | | | |
| R | 13 | 5.32 | 16 | −28 | 44 |
| L | **73** | **3.84** | **−30** | **10** | **−8** |
| **BROMOCRIPTINE—COMBINED(POSITIVE)** | | | | | |
| L | **196** | **6.51** | **−26** | **4** | **−12** |
| L | 14 | 4.48 | −30 | −14 | −46 |
| L | 11 | 3.93 | −18 | −54 | 12 |

*The regions that fell within our anatomically defined pathways are printed in bold. Data is presented with p < 0.001 uncorrected and extended threshold of > 10 voxels.*

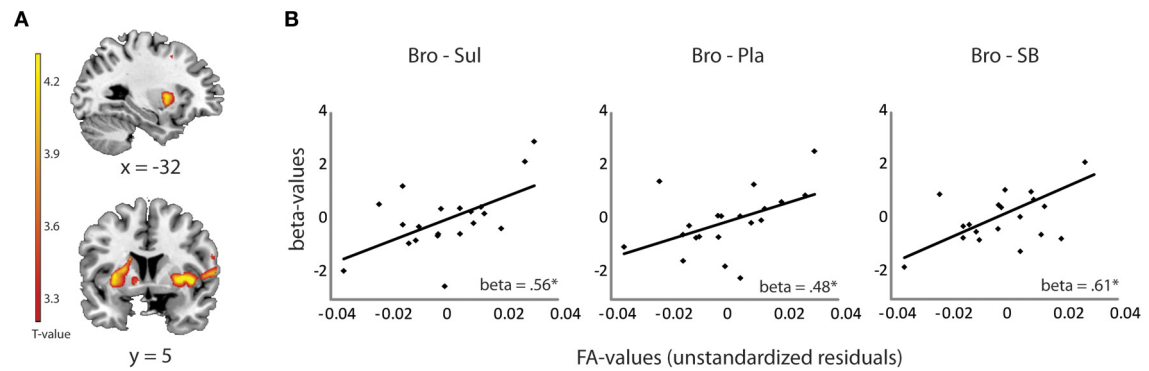



**FIGURE 2 | Relationship between drug effects on BOLD and FA-values, revealed by ROI analyses. (A)** Shown are effects of bromocriptine relative to sulpiride on striatal BOLD signal during unexpected outcomes in the reversal learning task [(unexpected—expected rewards) + (unexpected—expected punishments)] ($x, y, z = -22, 18, 4$, $p_{\text{fwe\_striatum}} = 0.03$). These effects were reported previously in Van Der Schaaf et al. (2012) and depended on working memory capacity. **(B)** Linear relationship between FA values in the OFC-amygdala pathway and the effects of bromocriptine relative to sulpiride (left), bromocriptine relative to placebo (middle) and bromocriptine relative to combined administration of both drugs (right) on striatal BOLD. Only significant effects are shown. $^*p < 0.05$ Abbreviations: Pla, placebo; Bro, bromocriptine; Sul, sulpiride; SB = combined administration of bromocriptine and sulpiride.

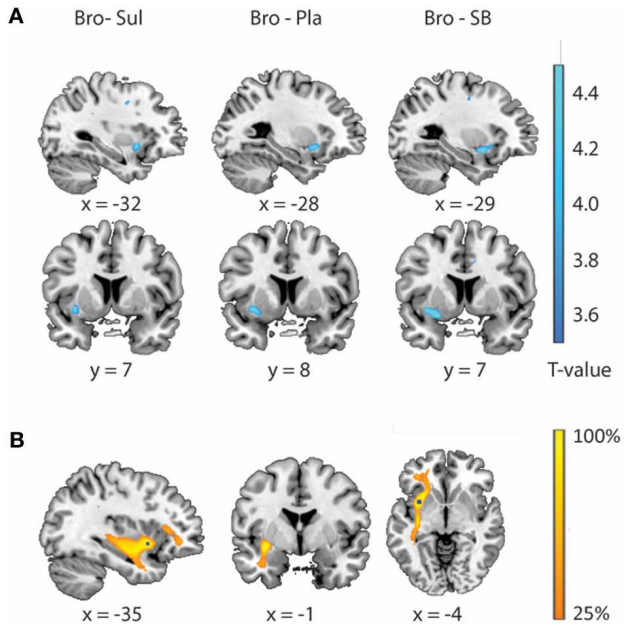



**FIGURE 3 | Relationship between drug effects on BOLD and FA-values, revealed by whole brain analyses. (A)** White matter regions showing a linear relationship between FA-values and the effects of bromocriptine relative to sulpiride (left), bromocriptine relative to placebo (middle) and bromocriptine relative to combined administration of both drugs (right) on striatal BOLD. **(B)** Summed tractography maps of the individual pathways that originate from the FA-region displayed in **Figure 3A** (left). Probabilistic tractography from this region revealed an extensive network of pathways between the OFC, amygdala and striatum. Image is thresholded to present those tracks that were present in at least 25% of the subjects.

typically also involve thalamic connections (Haber and Knutson, 2010) (**Figure 3B**).

## DISCUSSION

Dopaminergic drug effects have been shown to vary greatly between individuals (Cools and D'esposito, 2011). Here we provide evidence for an important link between dopaminergic drug effects during reversal learning and neuroanatomical integrity of connections between the OFC and amygdala. More specifically, we demonstrate that dopaminergic drug effects on striatal BOLD signal during reversal learning vary as a function of FA and MD in a pathway connecting the OFC with the amygdala. FA and MD rely on several microstructural properties, including the level of axon myelination, intact axonal membranes and fiber density (Beaulieu, 2002). Accordingly, our results support the hypothesis that dopaminergic drug effects on human striatal BOLD signal are associated with the neuronal communication efficiency of cortico-limbic projections. The implication of these findings is 2-fold. First, the data provide fundamental insight into the mechanism of action of dopaminergic drug effects on reward-related processing and reversal learning. Specifically, effects of D2 receptor stimulation during reversal learning involve an orbitofronto-limbic-striatal network. Second, they may have important clinical implications by suggesting that measures of white matter integrity can help predict dopaminergic drug effects on brain function, thus contributing ultimately to the individual tailoring of dopaminergic drug treatment strategies in psychiatry.

The drug effects on striatal BOLD signal were associated with white matter integrity of the pathway connecting the OFC with the amygdala (i.e., part of the uncinate fasciculus), and not by that of direct orbitofronto-striatal or amygdala-striatal projections. These findings extend previous non-pharmacological human DTI studies demonstrating that reward-related striatal BOLD responses (Camara et al., 2010) and associated functional connectivity (Cohen et al., 2008) are associated with white matter integrity of orbitofrontal-limbic-striatal pathways. Furthermore, we also showed that the drug effects during reversal learning were not associated with white matter integrity of dorsal fronto-striatal connections, which are suggested to be involved in more cognitive and motor processing (Alexander et al., 1990; Haber and Knutson, 2010). Indeed, our findings complement those from a recent study (Van Schouwenburg et al., 2013), in which we demonstrated that white matter integrity of a dorsal fronto-striatal-thalamic pathway was associated with drug effects on striatal BOLD signals during a form of attention-shifting that did not involve reward. Together, these data establish that associations between dopaminergic drug effects and white matter integrity are neuroanatomically specific and depend on task demands.

Our results are consistent with animal lesion work that has repeatedly demonstrated the crucial role of OFC-amygdala interactions in reversal learning (Baxter et al., 2000; Stalnaker et al., 2007; Schoenbaum et al., 2009). The OFC has originally been suggested to rapidly encode new associations and regulate reversal learning by directly driving areas such as the striatum (Thorpe et al., 1983). However, accumulating evidence indicates that the OFC instead contributes indirectly to the updating of stimulus-outcome associations by providing information about expected outcomes to other down-stream areas such as the amygdala (Stalnaker et al., 2007; Schoenbaum et al., 2009; Takahashi et al., 2009). Amygdala projections to the ventral striatum might then, in turn, mediate the effects of (updated) outcome-predictive stimuli on action selection. Indeed, electrophysiological responses in the ventral striatum (and associated behavioral responding) to relevant sensory stimuli critically depend on concomitant amygdala and dopamine inputs (Robbins et al., 1989; Everitt et al., 1991; Ambroggi et al., 2008; Pennartz et al., 2011). Accordingly, our results highlight the importance of indirect OFC-amygdala pathways in reversal learning by showing that dopaminergic modulation of striatal BOLD responses during reversal learning are not associated with white matter integrity of direct fronto-striatal pathways, but instead are associated with white matter integrity of the OFC-amygdala pathway. Together, these results provide fundamental insight into the mechanism by which dopamine changes brain function during reversal learning.

In addition, our experimental design allowed us to establish that these white-matter dependent drug effects were mediated by D2 receptors. Effects of the D2 receptor agonist bromocriptine on striatal BOLD signal were abolished by co-administration with

the D2 receptor antagonist sulpiride. This generally concurs with animal work demonstrating that reversal learning in monkeys is selectively mediated by D2 receptors but not D1 receptors (Lee et al., 2007). In addition, animal work has demonstrated that the effects of dopamine on the output of amygdala neurons are at least partially mediated by D2 receptors (Rosenkranz and Grace, 1999, 2002a; Grace and Rosenkranz, 2002; Bissière et al., 2003). While D2 receptor stimulation was found to potentiate sensory driven amygdala outputs to the striatum, D1 receptor stimulation was found to attenuate PFC inhibitory influences on amygdala output neurons (Rosenkranz and Grace, 1999, 2002a). Based on such experimental animal work, we speculate that bromocriptine potentiated sensory-driven amygdala output excitability to a greater extent in subjects with high communicational efficacy within the OFC-amygdala pathway than in those with low OFC-amygdala connectivity. It might be noted we cannot provide definitive evidence for this latter hypothesis, because DTI is inconclusive with regard to the direction in which information travels. Nevertheless, our results do converge with prior animal work and highlight the importance of D2 receptor stimulation for reversal learning.

One caveat of our study is that we did not find evidence for a direct relationship between white matter integrity of the OFC-amygdala pathway and drug effects on behavioral updating of stimulus-outcome associations. This is particularly surprising given that experimental animal work has demonstrated that the OFC and amygdala (Iversen and Mishkin, 1970; Jones and Mishkin, 1972) and their interaction (Baxter et al., 2000; Stalnaker et al., 2007) are crucial for behavioral performance on reversal learning tasks. Accordingly, we believe that our failure to observe correlations with drug effects on behavior might reflect a relative lack of sensitivity. Future work should reveal whether the present finding that white matter integrity of orbitofrontal-limbic-striatal pathways is associated with drug effects on brain function extends to behavior.

## ACKNOWLEDGMENTS

We thank Jan Leijtens for his assistance during medical screening and Joris van Elshout for his assistance during data collection. Roshan Cools was supported by a VIDI grant from the innovational Incentives Scheme of the Netherlands Organization for Scientific Research (NWO), a research grant to Kae Nakamura, Roshan Cools and Nataniel Daw from the Human Frontiers Science Program (RGP0036/2009-C) and a James McDonnell Scholar award. Roshan Cools has been a consultant to Abbott Laboratories, but she is not an employee or a stock shareholder. Jan K. Buitelaar has been in the past 3 years a consultant to / member of advisory board of/and/or speaker for Janssen Cilag BV, Eli Lilly, Bristol-Myer Squibb, Shering Plough, UCB, Shire, Novartis and Servier. He is not an employee of any of these companies, and not a stock shareholder of any of these companies. He has no other financial or material support, including expert testimony, patents, royalties. The authors declare no conflict of interest. Data will be made available upon request.

## REFERENCES

Alexander, G., Crutcher, M., and Delong, M. (1990). Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. *Prog. Brain Res.* 85, 119–165. doi: 10.1016/S0079-612362678-3

Ambroggi, F., Ishikawa, A., Fields, H., and Nicola, S. (2008). Basolateral amygdala neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. *Neuron* 59, 648–661. doi: 10.1016/j.neuron.2008.07.004

Badre, D., and Frank, M. (2012). Mechanisms of hierarchical reinforcement learning in cortico-striatal circuits 2: evidence from fMRI. *Cereb. Cortex* 22, 527–536.

Balleine, B., Leung, B., and Ostlund, S. (2011). The orbitofrontal cortex, predicted value, and choice. *Ann. N.Y. Acad. Sci.* 1239, 43–50. doi: 10.1111/j.1749-6632.2011.06270.x

Baxter, M., Parker, A., Lindner, C., Izquierdo, A., and Murray, E. (2000). Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *J. Neurosci.* 20, 4311–4319.

Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR Biomed.* 15, 435–455. doi: 10.1002/nbm.782

Behrens, T., Berg, H., Jbabdi, S., Rushworth, M., and Woolrich, M. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* 34, 144–155. doi: 10.1016/j.neuroimage.2006.09.018

Bissière, S., Humeau, Y., and Lüthi, A. (2003). Dopamine gates LTP induction in lateral amygdala by suppressing feedforward inhibition. *Nat. Neurosci.* 6, 587–592. doi: 10.1038/nn1058

Bond, A., and Lader, M. (1974). The use of analogue scales in rating subjective feelings. *Br. J. Med. Psychol.* 47, 211–429. doi: 10.1111/j.2044-8341.1974.tb02285.x

Boorman, E. D., O'shea, J., Sebastian, C., Rushworth, M. F., and Johansen-Berg, H. (2007). Individual differences in white-matter microstructure reflect variation in functional connectivity during choice. *Curr. Biol.* 17, 1426–1431. doi: 10.1016/j.cub.2007.07.040

Brett, M., Anton, J. L., Valabregue, R., and Poline, J. B. (2002). "Region of interest analysis using an SPM toolbox," in *International Conference on Functional Mapping of the Human Brain* (Sendai).

Camara, E., Rodriguez-Fornells, A., and Münte, T. (2010). Microstructural brain differences predict functional hemodynamic responses in a reward processing task. *J. Neurosci.* 30, 11398–11800. doi: 10.1523/JNEUROSCI.0111-10.2010

Cohen, M., Elger, C., and Weber, B. (2008). Amygdala tractography predicts functional connectivity and learning during feedback-guided decision-making. *Neuroimage* 39, 1396–1407. doi: 10.1016/j.neuroimage.2007.10.004

Cools, R., Barker, R., Sahakian, B., and Robbins, T. (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., and D'esposito, M. (2011). Inverted-u-shaped dopamine actions on human working memory and cognitive control. *Biol. Psychiatry* 69, e113–e125.

Cools, R., Frank, M. J., Gibbs, S. E., Miyakawa, A., Jagust, W., and D'esposito, M. (2009). Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. *J. Neurosci.* 29, 1538–1543. doi: 10.1523/JNEUROSCI.4467-08.2009

Cools, R., Lewis, S. J., Clark, L., Barker, R. A., and Robbins, T. W. (2007). L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. *Neuropsychopharmacology* 32, 180–189. doi: 10.1038/sj.npp.1301153

Del Arco, A., and Mora, F. (2009). Neurotransmitters and prefrontal cortex-limbic system interactions: implications for plasticity and psychiatric disorders. *J. Neural Transm.* 116, 941–952. doi: 10.1007/s00702-009-0243-8

Del Arco, A., Mora, F., Mohammed, A., and Fuxe, K. (2007). Stimulation of D2 receptors in the prefrontal cortex reduces PCP-induced hyperactivity, acetylcholine release and dopamine metabolism in the

nucleus accumbens. *J. Neural Transm.* 114, 185–193. doi: 10.1007/s00702-006-0533-3

De Zeeuw, P., Mandl, R., Hulshoff Pol, H., Van Engeland, H., and Durston, S. (2012). Decreased frontostriatal microstructural organization in attention deficit/hyperactivity disorder. *Hum. Brain Mapp*. 33, 1941–1951. doi: 10.1002/hbm.21335

Dias, R., Robbins, T., and Roberts, A. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* 380, 69–72. doi: 10.1038/380069a0

Dodds, C. M., Muller, U., Clark, L., Van Loon, A., Cools, R., and Robbins, T. W. (2008). Methylphenidate has differential effects on blood oxygenation level-dependent signal related to cognitive subprocesses of reversal learning. *J. Neurosci.* 28, 5976–5982. doi: 10.1523/JNEUROSCI.1153-08.2008

Draganski, B., Kherif, F., Klöppel, S., Cook, P., Alexander, D., Parker, G., et al. (2008). Evidence for segregated and integrative connectivity patterns in the human Basal Ganglia. *J. Neurosci.* 28, 7143–7152. doi: 10.1523/JNEUROSCI.1486-08.2008

Everitt, B., Morris, K., O'brien, A., and Robbins, T. (1991). The basolateral amygdala-ventral striatal system and conditioned place preference: further evidence of limbic-striatal interactions underlying reward-related processes. *Neuroscience* 42, 1–18. doi: 10.1016/0306-452290145-E

Feldman, R. S., Meyer, J. S., and Quenzer, L. F. (1997). *Principles of Neuropsychopharmacology*. Sunderland, MA: Sinauer Associates.

Fitzgerald, P., and Dinan, T. (2008). Prolactin and dopamine: what is the connection. A review article. *J. Psychopharmacol.* 22, 12–21. doi: 10.1177/0269216307087148

Grace, A., Floresco, S., Goto, Y., and Lodge, D. (2007). Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends Neurosci.* 30, 220–227. doi: 10.1016/j.tins.2007.03.003

Grace, A., and Rosenkranz, J. (2002). Regulation of conditioned responses of basolateral amygdala neurons. *Physiol. Behav.* 77, 489–493. doi: 10.1016/S0031-938400909-5

Groman, S., Lee, B., London, E., Mandelkern, M., James, A., Feiler, K., et al. (2011). Dorsal striatal D2-like receptor availability covaries with sensitivity to positive reinforcement during discrimination learning. *J. Neurosci.* 31, 7291–7299. doi: 10.1523/JNEUROSCI.0363-11.2011

Gutman, D., Holtzheimer, P., Behrens, T., Johansen-Berg, H., and Mayberg, H. (2009). A tractography analysis of two deep brain stimulation white matter targets for depression. *Biol. Psychiatry* 65, 276–282. doi: 10.1016/j.biopsych.2008.09.021

Haber, S., and Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 35, 4–30. doi: 10.1038/npp.2009.129

Harsay, H., Cohen, M., Reneman, L., and Ridderinkhof, K. (2011). How the aging brain translates motivational incentive into action: the role of individual differences in striatocortical white matter pathways. *Dev. Cogn. Neurosci.* 1, 530–539. doi: 10.1016/j.dcn.2011.06.005

Holland, P., and Gallagher, M. (2004). Amygdala-frontal interactions and reward expectancy. *Curr. Opin. Neurobiol.* 14, 148–203. doi: 10.1016/j.conb.2004.03.007

Howell, D. C. (1997). *Statistical Methods for Psychology*. Belmont, CA: Wadsworth Publishing Company.

Iversen, S., and Mishkin, M. (1970). Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Exp. Brain Res.* 11, 376–386.

Jones, B., and Mishkin, M. (1972). Limbic lesions and the problem of stimulus–reinforcement associations. *Exp. Neurol.* 36, 362–377. doi: 10.1016/0014-488690030-1

Lee, B., Groman, S., London, E. D., and Jentsch, J. D. (2007). Dopamine D2/D3 receptors play a specific role in the reversal of a learned visual discrimination in monkeys. *Neuropsychopharmacology* 32, 2125–2134. doi: 10.1038/sj.npp.1301337

Leh, S., Ptito, A., Chakravarty, M., and Strafella, A. (2007). Fronto-striatal connections in the human brain: a probabilistic diffusion tractography study. *Neurosci. Lett.* 419, 113–118. doi: 10.1016/j.neulet.2007.04.049

Mandl, R., Rais, M., Van Baal, G., Van Haren, N., Cahn, W., Kahn, R., et al. (2012). Altered white matter connectivity in never-medicated patients with schizophrenia. *Hum. Brain Mapp.* doi: 10.1002/hbm.22075. [Epub ahead of print].

Murray, E., and Wise, S. (2010). Interactions between orbital prefrontal cortex and amygdala: advanced cognition, learned responses and instinctive behaviors. *Curr. Opin. Neurobiol.* 20, 212–232. doi: 10.1016/j.conb.2010.02.001

Ogar, J., and Gorno-Tempini, M. (2007). "The orbitofrontal cortex and the insula," in *The Human Frontal Lobes; Functions and Disorders*, *2nd Edn.*, eds B. Miller and J. Cummings (New York, NY: The guilford press), 59–68.

Ongür, D., and Price, J. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb. Cortex* 10, 206–219.

Pennartz, C., Groenewegen, H., and Lopes Da Silva, F. (1994). The nucleus accumbens as a complex of functionally distinct neuronal ensembles: an integration of behavioural, electrophysiological and anatomical data. *Prog. Neurobiol.* 42, 719–780. doi: 10.1016/0301-008290025-6

Pennartz, C., Ito, R., Verschure, P., Battaglia, F., and Robbins, T. (2011). The hippocampal-striatal axis in learning, prediction and goal-directed behavior. *Trends Neurosci.* 34, 548–607. doi: 10.1016/j.tins.2011.08.001

Peper, J., Mandl, R., Braams, B., De Water, E., Heijboer, A., Koolschijn, P., et al. (2013). Delay discounting and frontostriatal fiber tracts: a combined DTI and MTR study on impulsive choices in healthy young adults. *Cereb. Cortex* 23, 1695–1702. doi: 10.1093/cercor/bhs163

Reese, T., Heid, O., Weisskoff, R., and Wedeen, V. (2003). Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. *Magn. Reson. Med.* 49, 177–182. doi: 10.1002/mrm.10308

Robbins, T., Cador, M., Taylor, J., and Everitt, B. (1989). Limbic-striatal interactions in reward-related processes. *Neurosci. Biobehav. Rev.* 13, 155–162. doi: 10.1016/S0149-763480025-9

Rosenkranz, J., and Grace, A. (1999). Modulation of basolateral amygdala neuronal firing and afferent drive by dopamine receptor activation *in vivo*. *J. Neurosci.* 19, 11027–11039.

Rosenkranz, J., and Grace, A. (2002a). Cellular mechanisms of infralimbic and prelimbic prefrontal cortical inhibition and dopaminergic modulation of basolateral amygdala neurons *in vivo*. *J. Neurosci.* 22, 324–361.

Rosenkranz, J. A., and Grace, A. A. (2002b). Cellular mechanisms of infralimbic and prelimbic prefrontal cortical inhibition and dopaminergic modulation of basolateral amygdala neurons *in vivo*. *J. Neurosci.* 22, 324–337.

Samanez-Larkin, G., Levens, S., Perry, L., Dougherty, R., and Knutson, B. (2012). Frontostriatal white matter integrity mediates adult age differences in probabilistic reward learning. *J. Neurosci.* 32, 5333–5340. doi: 10.1523/JNEUROSCI.5756-11.2012

Schoenbaum, G., Roesch, M., Stalnaker, T., and Takahashi, Y. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nat. Rev. Neurosci.* 10, 885–892.

Sesack, S., Carr, D., Omelchenko, N., and Pinto, A. (2003). Anatomical substrates for glutamate-dopamine interactions: evidence for specificity of connections and extrasynaptic actions. *Ann. N.Y. Acad. Sci.* 1003, 36–88. doi: 10.1196/annals.1300.066

Sesack, S., and Grace, A. (2010). Cortico-basal ganglia reward network: microcircuitry. *Neuropsychopharmacology* 35, 27–47. doi: 10.1038/npp.2009.93

Stalnaker, T., Franz, T., Singh, T., and Schoenbaum, G. (2007). Basolateral amygdala lesions abolish orbitofrontal-dependent reversal impairments. *Neuron* 54, 51–58. doi: 10.1016/j.neuron.2007.02.014

Studholme, C., Constable, R., and Duncan, J. (2000). Accurate alignment of functional EPI data to anatomical MRI using a physics-based distortion model. *IEEE Trans. Med. Imaging* 19, 1115–1142. doi: 10.1109/42.896788

Takahashi, Y., Roesch, M., Stalnaker, T., Haney, R., Calu, D., Taylor, A., et al. (2009). The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron* 62, 269–280. doi: 10.1016/j.neuron.2009.03.005

Thomason, M., and Thompson, P. (2011). Diffusion imaging, white matter, and psychopathology. *Annu. Rev. Clin. Psychol.* 7, 63–85. doi: 10.1146/annurev-clinpsy-032210-104507

Thorpe, S., Rolls, E., and Maddison, S. (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp. Brain Res.* 49, 93–115.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–362. doi: 10.1006/nimg.2001.0978

Van Der Schaaf, M., Van Schouwenburg, M., Geurts, D., Schellekens, A., Buitelaar, J., Verkes, R., et al. (2012). Establishing the dopamine dependency of human striatal signals during reward and punishment reversal learning. *Cereb. Cortex* doi: 10.1093/cercor/bhs344. [Epub ahead of print].

Van Schouwenburg, M., Zwiers, M., Van Der Schaaf, M., Geurts, D., Schellekens, A., Buitelaar, J., et al. (2013). Anatomical connection strength predicts dopaminergic drug effects on fronto-striatal function. *Psychopharmacology (Berl)* 227, 521–531. doi: 10.1007/s00213-013-3000-5

Visser, E., Quin, S., and Zwiers, M. (2010). "Echo-planar imaging distortion correction by contrained nonlinear coregistration improves group MRI," in *Presented at the Joint Annual Meeting ISMR-ESMRMB, poster 3459* (Stockholm).

Zwiers, M. (2010). Patching cardiac and head motion artefacts in diffusion-weighted images. *Neuroimage* 53, 565–640. doi: 10.1016/j.neuroimage.2010.06.014

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 29 April 2013; accepted: 21 July 2013; published online: 14 August 2013.*
*Citation: van Der Schaaf ME, Zwiers MP, van Schouwenburg MR, Geurts DEM, Schellekens AFA, Buitelaar JK, Verkes RJ and Cools R (2013) Dopaminergic drug effects during reversal learning depend on anatomical connections between the orbitofrontal cortex and the amygdala. Front. Neurosci. **7**:142. doi: 10.3389/fnins.2013.00142*

*This article was submitted to Frontiers in Decision Neuroscience, a specialty of Frontiers in Neuroscience.*

frontiers in
NEUROSCIENCE

REVIEW ARTICLE
published: 11 October 2013
doi: 10.3389/fnins.2013.00175

# Heads for learning, tails for memory: reward, reinforcement and a role of dopamine in determining behavioral relevance across multiple timescales

***Mathieu Baudonnat[1]*, Anna Huber[1,2], Vincent David[3] and Mark E. Walton[1]***

[1] *Department of Experimental Psychology, University of Oxford, Oxford, UK*
[2] *Department of Psychiatry, University of Oxford, Oxford, UK*
[3] *Centre National de la Recherche Scientifique, Unité Mixte de Recherche 5287, Institut de Neurosciences Cognitives et Intégratives d'Aquitaine, Université de Bordeaux, Talence, France*

***Edited by:***
*Jeff A. Beeler, The City University of New York, USA*

***Reviewed by:***
*Stan Floresco, University of British Columbia, Canada*
*Saleem Nicola, Albert Einstein College of Medicine, USA*

****Correspondence:***
*Mathieu Baudonnat and Mark E. Walton, Department of Experimental Psychology, University of Oxford, 9 South Parks Road, Oxford OX1 3UD, UK*
*e-mail: mathieu.baudonnat@psy.ox.ac.uk;*
*mark.walton@psy.ox.ac.uk*

Dopamine has long been tightly associated with aspects of reinforcement learning and motivation in simple situations where there are a limited number of stimuli to guide behavior and constrained range of outcomes. In naturalistic situations, however, there are many potential cues and foraging strategies that could be adopted, and it is critical that animals determine what might be behaviorally relevant in such complex environments. This requires not only detecting discrepancies with what they have recently experienced, but also identifying similarities with past experiences stored in memory. Here, we review what role dopamine might play in determining how and when to learn about the world, and how to develop choice policies appropriate to the situation faced. We discuss evidence that dopamine is shaped by motivation and memory and in turn shapes reward-based memory formation. In particular, we suggest that hippocampal-striatal-dopamine networks may interact to determine how surprising the world is and to either inhibit or promote actions at time of behavioral uncertainty.

**Keywords: dopamine, nucleus accumbens, hippocampus, reward, reinforcement learning, long term memory**

## INTRODUCTION

It is often assumed, when faced to an unfamiliar environment, that our main task is to learn about this new world via a process of exploration, gathering information through trial and error. As experimenters trying to study these processes, we present our subjects with novel stimuli, different ways of responding and different types of reinforcer in order to determine how they learn about these elements of their world. Sometimes then we might change associations or add new cues to try to understand how new information is represented or existing associations modified. Indeed, when using animals as our subjects, often a lot of effort is expended to ensure that the task environment shares few, if any, features that the animals might have encountered at a previous point in their lives to ensure that new learning can proceed uncontaminated by past experience. Several decades of work has helped to map out how learning and adaptive behavior in these types of environment might be represented in brain circuits, with differing angles of focus within widespread frontal-temporal-striatal networks and particularly their interactions with the neurotransmitter dopamine.

However, while these processes are inarguably critical for survival, this concentration of research into such constrained task set-ups has also indirectly diminished the amount of work on another critical factor in everyday learning and decision making. For what has sometimes been overlooked is the degree to which behavior in novel or changing environments relies not only on detecting discrepancies with what has recently been experienced, but also with identifying similarities with past situations stored in memory (Lisman and Grace, 2005; Gershman and Niv, 2010; Shohamy and Adcock, 2010). While the former process can be used to determine how the environment has changed, the latter in addition can provide a structure, based on past experience, to allow learning to proceed more rapidly. Ideally, these processes will interact dynamically to enable the rapid acquisition of beneficial behaviors in new situations by providing potential response strategies or by biasing attention to what are expected to be the relevant parts of the environment. However, the balance between relying on past strategies or adopting a new response pattern is delicate and in certain circumstances, particularly when the environment has fundamentally changed, a reliance on stored experience at the expense of new learning may also lead to inflexible and maladaptive responses.

A key question, therefore, is what role dopamine transmission might play in guiding *how* to learn and determining when to use or ignore choice policies implemented in previous comparable situations. While dopamine has tended to be tightly associated with aspects of reinforcement learning and motivation, there is increasing evidence that the pattern of dopamine activity can be shaped by both an animal's experience of the structure of an environment and even the long-term nutritional effects of a reinforcer. Moreover, as well as signaling reward predictions, dopamine activity and release can be influenced by novel cues and environments, and therefore might signal the potential importance of elements of the world in order to guide behavior toward the most valuable options (Ljungberg et al., 1992; Kakade and Dayan, 2002; Lisman and Grace, 2005). However, beyond its role

in guiding ongoing learning, there is another literature implicating dopamine in aspects of memory consolidation and retrieval at times distant from the original experience (Packard and White, 1989; Floel et al., 2008; Goto and Grace, 2008; Phillips et al., 2008; Shohamy and Adcock, 2010).

In this article, we will attempt to pull together these different strands—learning, familiarity, and memory—to provide a descriptive account of how dopamine transmission might facilitate adaptive behavior in complex, changing environments. Given the heterogeneity of dopamine responses in different terminal regions, for the sake of simplicity, we will focus most closely on phasic dopamine transmission in the ventral striatum/nucleus accumbens (NAc)—in other words, a transient change in dopamine levels lasting between a few hundred milliseconds and several seconds—and how this can have an impact on both short- and longer-term behavior. However, we acknowledge that there are likely different functions of dopamine activity measured across minutes, hours and even days, which might very well correlate with different phases of behavior (see, for example, Schultz, 2007). Moreover, dopamine release in other terminal regions, such as prefrontal cortex, may subserve similar but distinct roles in behavioral flexibility owing to the differences in receptor location and difference in clearance mechanisms and timing (Floresco, 2013).

Here, we will first review the evidence for the modulation of dopamine over different timescales and then will discuss what the behavioral consequences of such modulation would be in terms of patterns of dopamine release in terminal regions. Beyond a straightforward role in reinforcement learning, we will argue that phasic dopamine release here might act as a signal to motivate animals to engage with options at times of uncertainty in order to learn the best predictors of reward in an environment. In a final section, we will outline ideas, building on the work of several other groups (Lisman and Grace, 2005; Johnson et al., 2007; Shohamy and Wagner, 2008; Pennartz et al., 2011), about how hippocampal-striatal-midbrain networks might cooperate to allow such adaptive behaviors to emerge. Specifically, we postulate that this network is key to determine how surprising the world is and therefore how to use memory to shape learning.

## DOPAMINE ACROSS THE TIMESCALES

### DOPAMINE, REINFORCEMENT, AND ONLINE REWARD LEARNING

Reinforcers drive the everyday life of all individuals. Reinforcement can either involve punishment (e.g., pain) and induce avoidance behavior or be positive (e.g., reward) and motivate approach behavior. Unexpected delivery of a reward causes a brief increase in firing rate in a large population of putative midbrain dopamine containing cells as well as a phasic dopamine release in part of the ventral striatum such as the nucleus accumbens (Schultz, 1997; Day et al., 2007; Flagel et al., 2011). A prominent theory suggests that these signals do not directly encode the affective properties of the reward, but instead reflect the deviation at a particular moment in time between an animal's expectation of reward and new information about future rewards (Montague et al., 1996; Schultz, 1997; Aggarwal et al., 2012). This discrepancy—termed a reward prediction error (RPE)—can be used as a teaching signal by temporal difference learning models to enable learning about the long-term cached reward values associated with stimuli in the environment. In support of this idea, dopamine cell activity reflects whether stimuli provide new, useful information about the world (Waelti et al., 2001) and optogenetic driving of the dopamine system to artificially signal the presence of new information when stimuli are presented can cause reward associations to be formed (Steinberg et al., 2013). A similar quantitative and causal relationship has also been demonstrated for dopamine, RPE and action updating in instrumental tasks (Bayer and Glimcher, 2005; Adamantidis et al., 2011). Although fMRI can only provide an indirect measure of dopamine transmission via changes in blood oxygen level-dependent (BOLD) signals (Knutson and Gibbs, 2007), BOLD responses in human ventral tegmental area (VTA) and ventral striatum/NAc have also been shown to represent positive RPEs (D'Ardenne et al., 2008).

While most theoretical work has focused on this type of reward-driven dopamine activity, it is increasingly clear that there is a heterogeneity of response types that can be observed in rodents and primates between different putative dopamine neurons. For instance, while some dopamine containing neurons are modulated by the expected value of predictive stimuli, showing an increase in firing that scales both with anticipated future reward *and* a decrease in firing that scales with anticipate future punishment (e.g., Matsumoto and Hikosaka, 2009; Cohen et al., 2012), other neurons appear mainly to reflect anticipated future reward alone (Mirenowicz and Schultz, 1996; Joshua et al., 2008). Yet another population scales with the likelihood of any future relevant event, whether positive or negative (Matsumoto and Hikosaka, 2009), which has been suggested to be a signal encoding the motivational salience of a stimulus (Bromberg-Martin et al., 2010a). This latter response may relate to the well-known observation that dopamine neurons can briefly respond to unexpected novel stimuli, which have no direct association with any reinforcer (Ljungberg et al., 1992; Horvitz et al., 1997). Alternatively, these responses to novel stimuli could reflect a signal to promote exploration to gain new information (Kakade and Dayan, 2002; Bromberg-Martin and Hikosaka, 2009). A similar range of responses to rewards and punishment can also be seen at the time of reinforcer delivery (Brischoux et al., 2009; Matsumoto and Hikosaka, 2009; Fiorillo et al., 2013a).

Moreover, recent studies have shown that the unfolding activity patterns of an individual dopamine neuron may also encode multiple signals across different timescales. Some dopamine neurons have been shown to exhibit a short-latency, brief phasic response that correlates with reward value, followed by a slower change in activity that scales with reward uncertainty (Fiorillo et al., 2003). In other studies, dopamine cell firing may initially code for the initial surprise and/or intensity of a stimulus before evolving to signal the motivational value of an upcoming outcome (Nomoto et al., 2010; Fiorillo et al., 2013b). The same population of dopamine neurons may also come to reflect different parameters of the local reward environment at different points in a trial (Bromberg-Martin et al., 2010c).

In spite of the fact that dopamine transmission has often been viewed as a regionally−homogenous reinforcement signal broadcast to all terminal regions (Schultz, 2002), a range of

dopamine release patterns has been observed in different striatal sub-regions. For instance, changes in extracellular dopamine levels measured with microdialysis in response to cues predicting either reward or punishment, or to the receipt of reward or punishment itself, can be different in the NAc core or shell regions (Ito et al., 2000; Pezze et al., 2001; Bassareo et al., 2002). The same is also true for brief phasic changes in dopamine measured at sub-second time resolution with fast-scan cyclic voltammetry, with variations observed across different parts of the NAc (core vs. shell) and dorsal striatum (Brown et al., 2011; Badrinarayan et al., 2012). Whether these differences relate to the anatomical diversity that was described above (Brischoux et al., 2009; Lammel et al., 2011), patterns of afferent input to the dopamine cells in the midbrain (Besson et al., 2012; Tolu et al., 2013), local influences on dopamine release from afferent input in the terminal region (Floresco et al., 1998; Threlfell et al., 2012), differences in the temporal resolution of electrophysiology compared to electrochemistry and microdialysis (Schultz, 2007), or combinations of all is an area of active research.

Taken together, all the above evidence demonstrates that the dopamine system is able to reactively signal events of potential significance, to reflect the discrepancy between current predictions and the discounted sum of prospective rewards, and to update these predictions as new information is acquired, which are prerequisites of trial-and-error, model-free, associative learning. Even though the precise degree of heterogeneity of dopamine cell responses remains contentious (Fiorillo et al., 2013b; Schultz, 2013), there are clearly diverse dopamine release patterns across striatum. As we will go on to discuss, there is also increasing evidence that the dopamine systems may not simply encode such short-term information, but may as well interact with other structures to allow stored information about task structure and motivational parameters to influence dopamine release. This, we will argue, may enable dopamine to provide a signal that influences what to learn and how to behave in particular contexts.

## DOPAMINE RESPONSES SCULPTED BY MEMORY AND MOTIVATION

While it is well known that, in agreement with RPE theories, dopamine cell activity and dopamine release in the NAc adapts during associative learning to reflect the earliest consistent predictor of future reward and experienced history of reinforcement (Schultz, 1997; Nakahara et al., 2004; Flagel et al., 2011), it is also apparent that dopamine in response to cues and outcomes can be shaped by task structure and memory. For instance, dopamine cell activity at the time of reward delivery scales to the potential range of available outcomes signaled by a stimulus, with a similar increase in activity across a 10-fold range of reward sizes (Tobler et al., 2005). Moreover, after extensive training on tasks where the consequences of one trial have direct impact on the likelihood of reward in a subsequent trial (e.g., a deterministic reversal learning task where one option is always rewarded and the other not or a sequential response task where one target is always rewarded within a block of 4 trials), some dopamine neurons come to represent values partially inferred from the overall reward structure rather than just from direct recent experience (Nakahara et al., 2004; Bromberg-Martin et al., 2010b). In human fMRI studies, ventral striatal/NAc BOLD signals have been shown to incorporate information about task structure in the RPE signals (Daw et al., 2011) and, intriguingly, it has recently been shown that systemic L-DOPA enhances the use of such model-based information (Wunderlich et al., 2012).

Studies of dopamine release in the striatum have also found that training can selectively modulate transmission in different regions. As discussed in an earlier section, unpredicted reward given to naïve animals only caused increases in dopamine in the NAc core region and not in the NAc shell, or the dorsomedial or dorsolateral striatum. However, after training on a simple cued instrumental task, now unpredicted reward *did* evoke dopamine release in the dorsomedial striatum as well as the NAc core (Brown et al., 2011). While this might just reflect an increase in the number of dopamine neurons recruited following any reinforced training, the fact that this occurred outside the task structure suggests this change could instead relate to the fact that reward has become a relevant event for guiding responding.

Decision parameters that drive dopamine transmission can also be dissociably influenced by the amount of experience of the task. In one recent study, rats were trained on a two-option decision making paradigm, where one option required a particular amount of work to gain a particular size of reward and either the work or reward of the alternative was systematically varied. This meant that animals' choices were either guided by a difference in the effort required to gain the reward (different cost, same benefit) or by a difference in the eventual payoff for taking a particular option (same cost, different benefit). In the "benefit" conditions, dopamine release elicited when either of the two options was presented consistently reflected the anticipated future reward associated with that option. This cue-elicited effect was unchanged over a range of testing sessions (**Figure 1A**). However, a different pattern was observed in the "cost" conditions. Now, differential dopamine release to the low as compared to the standard cost option was only recorded in early training, when the change in effort was unexpected, but after several sessions of experience this difference disappeared (Gan et al., 2010) (**Figure 1B**). This was not caused by any detectable differences in behavior as the animals showed equal preference for the low effort and the high reward options and continued to rapidly update their responses when the cost-benefit contingencies changed across sessions.

Exactly why and when dopamine adapts in response to the task structure and what role dopamine plays in shaping learning of task structure is currently a matter of some debate (Gershman and Niv, 2010; McDannald et al., 2012; Nakahara and Hikosaka, 2012). In the cost-benefit study described above, one idea is that NAc dopamine release is reflecting uncertainty over the temporal statistics of reward delivery. In the cost conditions, even though the cue-to-reward period varies according to the animals' choice, the overall average trial-to-trial reward rate is relatively static as the inter-trial intervals (ITIs) were adjusted as a function of the animals' choices. Such a proposition is supported by recent evidence from a Pavlovian task, which showed that NAc dopamine adapts to the temporal variability of cue-reward pairings over training. More specifically, cue-elicited dopamine decreased after extensive presentation of the cue-reward associations, although was then restored if the cue was unexpectedly presented at a

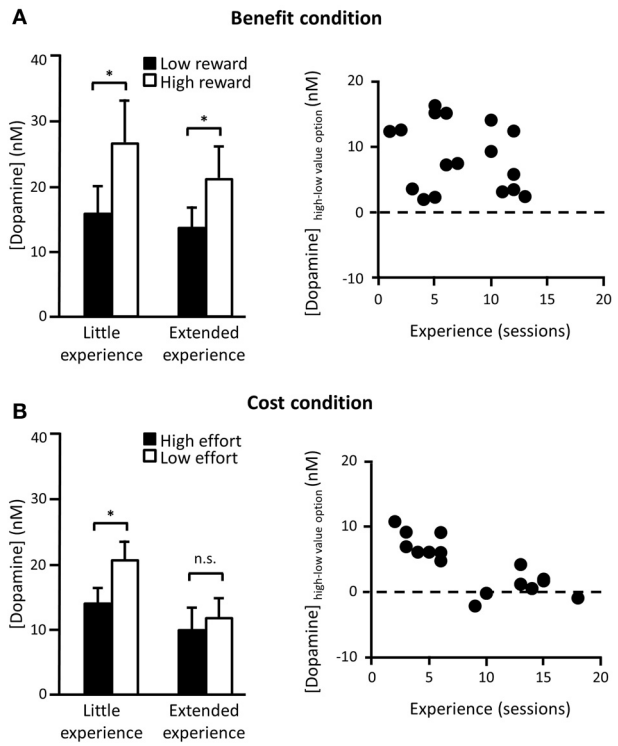



**FIGURE 1 | Cue-elicited dopamine release in NAc core during a two-option decision making task. (A)** Left panel: Average (mean + s.e.m.) phasic dopamine release in NAc core elicited by cues signaling availability of either a known low (2 pellets) or a high reward (4 pellets) option for a fixed cost (16 lever presses) in rats who have either had little (<9 sessions) or extended (≥9 sessions) experience with these contingencies. Data taken from forced trials (where only one option was available) after animals were choosing the high reward option on ≥75% of choice trials. Contingencies changed every session so rats had to flexibly update associations in each session. Right panel: Difference in cue-evoked peak dopamine release between cues signaling high and low reward options as a function of number of sessions in which they had experienced these contingencies. There was no correlation between experience of dopamine-based benefit encoding. **(B)** Left panel: Same as **(A)** except now the benefit was fixed (2 pellets) and the cost varied across options (low effort = 2 lever presses; high effort = 16 lever presses). Right panel: Same as **(A)** for high and low effort costs. Now there was a significant reduction in dopamine cost encoding as a function of experience. *n.s.* differences not significant; *significant at $p < 0.05$. [adapted from Gan et al. (2010)].

shorter ITI (Clark et al., 2013). A second, related idea is that, while dopamine may preferentially encode the anticipated benefits of a course of action, there may also be an initial boost in dopamine release to any unpredicted, uncertain event to motivate exploration and investigation of that option (Kakade and Dayan, 2002; Phillips et al., 2007; Walton et al., 2011). Once a settled pattern of responding has been established in a stable environment, however, the NAc dopamine signal may not be required to sustain performance.

All of the above studies have looked at how dopamine is shaped by learned or inferred predictions of proximal rewards. However, it is important to remember that dopamine transmission is also strongly affected by motivational state. While food reward consistently increases dopamine efflux in hungry animals, the effect is much reduced in sated rats (Wilson et al., 1995; Bassareo and Di Chiara, 1999). This effect appears specific to sensory properties of a particular food as while there is no increase in dopamine levels in the NAc measured with microdialysis when presented a food type that had previously been consumed, dopamine efflux in this situation still occurs when given a novel foodstuff (Ahn and Phillips, 1999). Such effects may reflect an important modulatory role of peptides such as insulin, leptin, and ghrelin, which act on neurons in the VTA and regions that target the VTA such as the lateral hypothalamus and can therefore affect dopamine transmission (Abizaid, 2009; Domingos et al., 2011; Mebel et al., 2012).

Dopamine transmission is also sensitive to distal post-ingestive effects, such as the nutritional, calorific and metabolic consequences of particular rewards (de Araujo et al., 2008). Therefore, while unexpected receipt, or cues signaling the impending delivery, of sucrose pellets consistently evokes dopamine release, this is attenuated when a sweet but calorie-free saccharine reward is used instead (Beeler et al., 2012; McCutcheon et al., 2012). Again, it is likely that these effects may be heterogeneous across the striatum depending on task and on the nutritional content. In recent studies looking at the effects of direct, intra-gastric infusions of fat (i.e., bypassing the taste receptors entirely), dopamine levels increased in dorsal striatum whereas NAc core dopamine decreased as a function of fat density (Ferreira et al., 2012). However, it remains to be determined whether dopamine levels update to reflect an inference about how behavior should be prioritized given an animal's current motivational state or instead are only altered after experiencing a particular food in a particular motivational state.

These motivational influences on dopamine are intriguing as they demonstrate that dopamine transmission can be modulated by distal, as well as proximal, consequences of reinforcers, which poses an extreme credit assignment problem given the time that must elapse between the predictive cues, ingestion and the metabolic effects of these rewards. More importantly, they also act as a reminder that theories of dopamine-mediated reinforcement learning and behavior should incorporate motivational parameters. An interesting possibility is that such a mechanism could underpin the seemingly paradoxical decisions that have been observed in the foraging literature where animals' choices depend on memory for context-dependent utility (Pompilio and Kacelnik, 2010).

## DOPAMINE, SYNAPTIC PLASTICITY, AND LONG-TERM MEMORY

In the previous section, we detailed evidence that dopamine transmission is influenced by memory of reward structure and motivational context. However, it is also important to remember that dopamine has also been implicated directly in long-term memory processes themselves. Long-term potentiation (LTP) and long-term depression (LTD) are thought to be critical at the cellular level to underlie memory formation and long-lasting changes in synaptic plasticity (Kandel, 2001). Dopamine has been identified as a strong modulator of these cellular adaptations (Wickens, 2009; Lovinger, 2010). For instance, D1 antagonists block the induction of LTP in striatum (Kerr and Wickens, 2001) and both D1 and D2 receptors appear necessary for striatal LTD (Calabresi

et al., 1992). Dopamine is also required to enable spike-timing dependent LTP or LTD in dorsal striatum (Pawlak and Kerr, 2008). Indeed, it is likely that the precise timing of dopamine transmission at striatal synapses has a significant effect on the direction of plasticity (Wickens, 2009).

At the molecular level, strengthening synapse communication is critical in setting up a network supporting both the acquisition and the recall of a particular learning experience. This relies on the co-release of glutamate and dopamine at target synapses. D1-like and D2-like receptor activation leads to the activation of two competing molecular pathways. While D1 receptors are coupled with a Gαs protein which positively modulates adenylate cyclases, D2 receptors are coupled with a Gαi protein which inhibits adenylate cyclases (Siegel et al., 1999; Hyman et al., 2006). Adenylate cyclases are responsible for the activation of various protein kinases and molecular cascades, which lead to activation of transcription factors (e.g., the phosphorylated form of the cAMP response element binding protein, pCREB) which in turn induce the transcription of immediate early genes (e.g., *c-fos*). The resulting proteins underlie the systemic consolidation necessary for long-term memory storage and recall (Huang and Kandel, 1995; Frankland and Bontempi, 2005).

Much of the work looking at the role of dopamine in memory formation, consolidation and recall at a behavioral level has concentrated on the effects of direct hippocampal dopamine interference (Bethus et al., 2010), which largely goes beyond the scope of this review. However, there are some findings that also imply a role for striatal dopamine itself in the encoding and consolidation of memories. It has been known for a while that dopamine-dependent potentiation of corticostriatal synaptic efficacy correlates with the speed of acquisition of intracranial self-stimulation, which in essence provides a cellular correlate of the standard, short-term reinforcement learning described in earlier sections (Reynolds et al., 2001). However, there are also some more unexpected reported effects of post-training dopamine manipulations on reinforcement that occur long after reward receipt. For instance, Dalley and colleagues reported that infusion of either a D1 or NMDA antagonist into the NAc given *after* a Pavlovian conditioning experiment blocked acquisition of autoshaping responses (Dalley et al., 2005). Similarly, in an inhibitory avoidance task, post-training injection of dopamine in the NAc shell, but not the NAc core, enhanced the retention of the conditioning (LaLumiere et al., 2005).

These studies indicate that dopamine's reinforcing effect can be temporally dissociated from the receipt of reward. Moreover, dopamine may even play some role in consolidation of memories for unreinforced items, similar to the way in which dopamine is activated in response to the presentation of novel stimuli. Dopamine lesions to the NAc core, but not to the shell, impaired a familiarity discrimination test with objects 24 h after the initial presentation, and NAc shell (and core to some extent) dopamine lesions affected location familiarity responses (Nelson et al., 2010).

#### INTERIM SUMMARY: DOPAMINE ACROSS THE TIMESCALES

In the above sections, we have described how dopamine acts at both short- and long-term timescales. Its most well defined function is that it allows the detection of discrepancies between predictions and outcomes at the time of an event. At a cellular level, dopamine also plays an important role in the storage of past experiences into memory. However, it is also becoming increasingly clear that all of these effects can be shaped by memory, motivation and internal state. Dopamine cannot be described as providing a homogeneous reinforcement signal as dopamine's role in these processes are clearly both site- and task-specific, with effects in a particular striatal region dependent on the state of the environment and of the animal.

In the next sections, we will build on these points to outline a possible framework that might help explain how phasic dopamine can function at different timescales. In particular, we will focus on two main aspects: (1) the behavioral consequences of dopamine release in striatal regions, and (2) the function of the anatomical networks in which these striatal regions are embedded and how dopamine might facilitate selection of one system over another.

## DOPAMINE TRANSMISSION, LEARNING, AND STRATEGY SELECTION

### DOPAMINE TRANSMISSION ACROSS THE STRIATUM

A critical issue when considering the role of dopamine concerns the question of what behavioral effect heterogeneous dopamine transmission has in different terminal regions. As has been discussed above, dopamine cell activity in many circumstances correlates highly with RPE signals. However, dopamine *release* in terminal regions in these situations suggests a role beyond a passive process of learning.

Across a range of studies, dopamine in the NAc, particularly in response to cues—and particularly in the core region—is required to activate animals to engage in a behavioral response. For instance, dopamine is only required to learn about cue-reward relationships in situations where cues acquire Pavlovian incentive values, which thus promote approach behavior, rather than simply being predictors of reward (Di Ciano et al., 2001; Dalley et al., 2002; Flagel et al., 2011). Similarly, dopamine transmission in cued decision making tasks seems strongly tied to the advantageous response elicited by a cue, whether to gain reward or avoid punishment, rather than just the predictive cue itself, and in some situations, can be elicited by an internal drive to respond in the absence of any external stimulus (Roitman et al., 2004; Yun et al., 2004; Oleson et al., 2012; Wassum et al., 2012). At least for the NAc core, this may be in the form of a general motivational drive rather than a representation of the particular sensory properties of the outcome. Lesioning the NAc core or blocking D1 receptors in this region disrupts general motivational arousal associated with cues during Pavlovian-instrumental transfer, but the former manipulation has no effect on outcome-specific versions of this task (Lex and Hauber, 2008; Corbit and Balleine, 2011).

In both Pavlovian and cued instrumental situations, dopamine may only be critical when there is some uncertainty or novelty about the environment, whether in terms of the consequences associated with a choice or the particular actions required to obtain a reward (cf. Nicola, 2010). This is not to say that NAc dopamine plays any direct role at the time of a choice in guiding the selection of one alternative over another. Several different

studies have shown now that phasic dopamine reflects the anticipated future benefit of whatever option will be chosen, even in cases where this is not the most valuable available option (Morris et al., 2006; Walton et al., 2011, but see Roesch et al., 2007).

Instead, phasic release in NAc may play two complementary roles: first, to act to energize animals to engage with a response based on its anticipated cached value, especially in situations when the environment changes and reward associations need to be updated; and, second, to enable them to learn about *behaviorally-relevant* consequences associated with cues in the environment (Phillips et al., 2007; Nicola, 2010; Walton et al., 2011). Although NAc dopamine is required to promote reward seeking even in the most simple Pavlovian situations, little is currently known about how this occurs in more complex, naturalistic environments where there are multiple potential cues and an unknown range of potential outcomes and it is necessary to determine which parameters are useful to guide behavior. Although we know of no studies to date that have directly investigated this issue with respect to NAc dopamine signaling, there is some recent evidence from fMRI that the ventral striatum/NAc might play an particular role in extracting information relevant for learning about reward (Klein-Flugge et al., 2011). In this study, participants underwent Pavlovian conditioning using stimuli that varied trial-by-trial in both their associated reward magnitude and the delay-to-reward. Interspersed were instrumental timing estimation trials where they had to predict when (but not what size) reward would appear in order to accumulate points that determined how much money they would receive for participating. While both precise reward magnitude and timing prediction errors were observable in the midbrain, as predicted by temporal difference learning models, ventral striatum/NAc BOLD signals *only* reflected the timing RPE signals required to guide subsequent choices and not the task-irrelevant reward RPE signals. Future studies will be needed to determine if these BOLD signal changes are driven by dopamine transmission. However, it may be that NAc core dopamine only signals a subset of relevant events, playing a particular role in motivating animals to learn strategies to improve their current state. As we will discuss in a later section, phasic NAc dopamine does not seem to be required when simply switching behavior to maintain a previous state (i.e., in most reversal tasks) (Haluk and Floresco, 2009).

Inspite of there being a number of examples showing that the patterns of rapid dopamine release are frequently divergent in the NAc core and shell, the function of dopamine in the latter structure is not yet clear. There is some evidence for a potential role of NAc shell in spatial processing, with infusions of a dopamine antagonist decreasing place conditioning but not cue conditioning (Ito and Hayen, 2011) and with dopamine efflux here being influenced by projections from ventral hippocampus (Legault et al., 2000). However, this seems unlikely to define its primary function given that there are many events distinct from spatial context that elicit NAc shell dopamine transmission and, in fact, the ventral hippocampus is arguably also more concerned with emotional responses to uncertainty, conflict detection and response inhibition than with spatial processing (Bannerman et al., 2004; Abela et al., 2013). Instead, it seems possible that NAc shell dopamine is important for signaling the occurrence of novel and potentially salient events, particularly in the case when there is ambiguity over the cause of that event. Tuning down NAc shell dopamine when uncertainty is resolved might facilitate an appropriate allocation of attention to the environment only to behaviorally relevant events. In partial support of this idea, it is notable that, while the NAc core (and NAc core dopamine), but not the NAc shell, has been implicated as being critical for beneficial choice behavior where there is guaranteed reward for any response (for example, effort- or delay-based decision making), the shell region appears to play a more critical role than the core region in biasing decisions when there is uncertainty about reward (Sokolowski and Salamone, 1998; Ghods-Sharifi and Floresco, 2010; Stopper and Floresco, 2011). Intriguingly, there is some evidence suggesting that NAc core and shell dopamine might play complementary and possibly antagonistic functions in some circumstances (Ito and Hayen, 2011), which may reflect the degree to which the overall statistics of the environment are known and how responses are being allocated.

Compared to the mesolimbic pathways to NAc, the nigrostriatal dopamine projections to dorsal striatum have been tied more closely to action selection and action reinforcement. Nonetheless, there is likely no simple, neat divide between the motivational and motor components of dopamine-dependent behavior, not least as the activity of many putative dopamine cells in both substantia nigra pars compacta and VTA correlates with RPE signals (Everitt and Robbins, 2005; Wise, 2009). It is known that dopamine in dorsomedial striatum is necessary to detect the contingency between actions and their consequences (Lex and Hauber, 2010) and, like in the NAc, dopamine levels also track the acquisition of a reinforced instrumental action and are sensitive to satiety (Ostlund et al., 2011). However, compared to NAc, where DA release appears to energize a decision policy selected elsewhere, there is also some evidence that DA in dorsal striatum may directly bias the choices to be made, particularly when there is evidence of a requirement to change behavior. Selective stimulation of the D1- or D2-receptor expressing striatal neurons during a probabilistic decision making task in this region can increase the incidence of a contralateral or ipsilateral action respectively following an unrewarded action (Tai et al., 2012).

Therefore, while there may be a common role of dopamine release across the striatum in helping reduce ambiguity through motivating cue-driven behavior or detecting the consequences of a novel event or response (Costa, 2007; Redgrave et al., 2008), the effect of dopamine transmission will be shaped by the properties of the terminal region and the networks in which this region is embedded. To explore this further, in the next sections, we will consider some examples of how and why this might occur, with a particular focus on NAc dopamine.

## DOPAMINE, STRATEGY SELECTION, AND BEHAVIORAL RELEVANCE IN COMPLEX CHANGEABLE ENVIRONMENTS

In naturalistic situations, the environment consists of multiple cues and there are multiple potential responses that could be made at any one time, the relevance of which change constantly over time. Consequently, a foraging animal will need to rely on different cues to locate food depending on its availability at a particular time, its current motivational requirements and the

environment in which the animal is operating, and to update its search strategy accordingly. These various cues and their relations to outcomes can be learned through experience and then retrieved from memory to guide behavior efficiently when the animal is again faced with comparable situations in the future. As we discussed in the previous sections, as well as being important for learning cue- and response-outcome associations through trial-and-error, dopamine transmission can be sculpted by experience and therefore may be critical to guide when and how stored memories are used and updated when the environment changes. To enable appropriate learning decision and decision making to take place, general rules can be established to facilitate trial-and-error learning to which exceptions are then added. We would contend that dopamine is involved in these learning and memory processes by playing a key role in guiding the initial search strategy when environmental contingencies are uncertain, putting the neural network in a state to learn about the consequences of choices, and also in determining when to reactivate parts of this network when encountering novel situations for which previously acquired strategies may be useful.

One prominent brain structure candidate for supporting the switch between updating current estimates of the world, using a previously stored memory or starting new learning is the NAc. Haluk and Floresco (2009) have argued that dopamine in the NAc is not required simply to update behavior when contingencies reverse, but is instead key to allow the shifting from a cue-driven response strategy, where spatial location is irrelevant, to a spatially-guided response strategy, where the cue location now should be ignored. They showed that either pharmacological blockade of NAc D1 receptors or stimulation of D2 receptors impairs this type of strategy shift. Interestingly, the D1 receptor antagonist had no comparable effect when the response strategy was simply reversed, demonstrating that there is not some general role of NAc phasic dopamine in altering choice strategies when there is no change to the overall reward statistics of the environment. In a separate study, it was demonstrated that NAc tonic dopamine levels also markedly increase when rats switch between response strategies, much more so than when initially acquiring the task, as well as in control conditions where the reward contingencies are deliberately kept uncertain (Stefani and Moghaddam, 2006).

Although Haluk and Floresco describe one of their conditions as requiring a "spatial response strategy," the navigational component in an operant chamber is necessarily sparse and in other settings, such as a water maze (McDonald and White, 1994; Porte et al., 2008), the radial arm maze or a Y-maze, an egocentric response strategy can be dissociable from an allocentric spatial one. In many naturalistic situations, spatial-, response-, and cue-learning will not necessarily proceed independently and the predictive value of each will have to be weighed up against one another, with the optimal choice strategy obviously dependent on the particular task environment. As well as competing for control, these systems may also cooperate during the initial stages of learning (White and McDonald, 2002). Several lines of evidence suggest that partially separate frontal-striatal-temporal networks underpin these different forms of learning and also guide which should be used to guide behavior (White and McDonald, 2002; Porte et al., 2008). Therefore, a key open question in the framework of the current review is to try to pinpoint how dopamine in these interconnected networks may guide attention to the appropriate parts of the environment in order to make advantageous foraging choices.

Other than Haluk and Floresco (2009), few studies to date have directly tried to address how dopamine might help arbitrate between spatial- and cue-guided behavior. The first clue came from the work of Packard and White (1991) in which they manipulated dopamine after training their animals in either a "win-shift" or "win-stay" task, which they had previously shown to depend on the hippocampus or caudate nucleus, respectively. They found that immediate post-training injection of dopamine agonists in the dorsal hippocampus selectively improved win-shift retention whereas injections into the posterior ventrolateral caudate nucleus improved the acquisition of the win-stay task. They argued that dopamine acts to modulate the functioning of the structure into which it is infused and potentially acts to reduce the interference of one strategy over another in the early stage of learning that can slow down the acquisition of the task (Packard and White, 1991).

Recently, Baudonnat et al. (2011) also investigated how different types of reward influence the selection and acquisition of learning strategies. Initially, they showed that mice were able to learn whether to use spatial location or intramaze visual cues to guide decision-making in a Y-maze when correct choices were reinforced with natural (food) reward (**Figures 2A,B**). As discussed in a previous section, it is well known that unpredicted reward drives phasic dopamine and that dopamine transmission at target synapses can induce the activation of molecular pathways leading to CREB phosphorylation and modification of synaptic strength (Dudman et al., 2003). Therefore, in order to investigate the cellular mechanism involved in the different types of learning, they measured pCREB levels in different candidate brain regions after the last behavioral testing session. Similar to the study by Packard and White (1991), pCREB was found to be specifically increased in the hippocampus after acquiring the spatial task whereas the increase was mainly present in the dorsal striatum after the cued task. By contrast, pCREB was highly expressed in the NAc independently of the task type (**Figure 2C**).

The above example demonstrates the cellular effects in hippocampus—dorsal striatum—NAc regions during appropriate reinforcement learning. However, drugs of abuse can also pharmacologically hijack the dopamine system and result in maladaptive patterns of behavior. To determine how an excess of dopamine might affect learning strategies, Baudonnat and colleagues carried out the same Y-maze experiment except that instead of receiving food reward for correct responses, the mice received intra-VTA injections of morphine, which has been shown to disinhibit dopamine neurons and induce dopamine release in target structures (Matthews and German, 1984; Johnson and North, 1992; Nugent et al., 2007; Baudonnat et al., 2011). While the animals learned the cued version of the task at a comparable rate with either natural or pharmacological reward, mice reinforced with morphine were unable to acquire the spatial strategy (**Figure 2B**). pCREB staining demonstrated that morphine given for correct responses caused increased

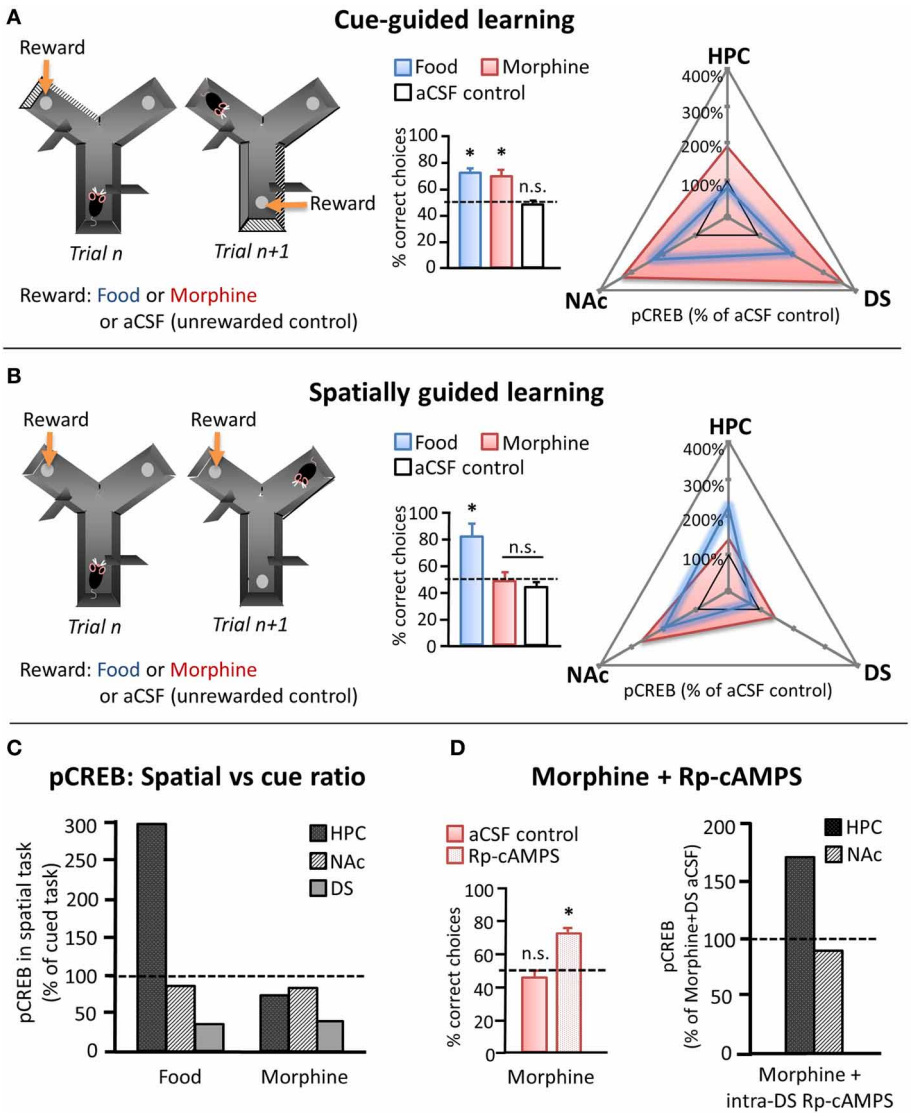



**FIGURE 2 | Effects of natural and pharmacological reinforcers on the acquisition of cue- and spatially-guided learning strategies. (A, B)** Left panels depict schematics of a cue-guided **(A)** and a spatially guided **(B)** version of a Y-maze decision making task. Correct responses were guided either by intramaze visual cues or by spatial location, respectively, and were reinforced in separate groups of mice by either food reward or intra-VTA morphine infusions. Control mice received intra-VTA aCSF infusions and no food reward at the "correct" location. Middle panels depict choice performance on the 10th day of training on the respective task (chance performance = 50%, marked with dashed line). Right panels depict pCREB levels measured in NAc, dorsal striatum (DS), and the hippocampus (HPC) of food- and morphine-reinforced mice after 10 days of training on the cued (**A**: upper) or the spatial (**B**: middle) version of the Y-maze task, normalized to pCREB levels observed in aCSF controls. **(C)** Relative changes in pCREB levels in HPC, NAc, and DS after training on the spatial as compared to the cued task (cued task pCREB = 100%, dashed line). **(D)** Left: Effect of daily intra-DS injection of either the PKA inhibitor Rp-cAMPS or aCSF on the spatial version of the Y maze in morphine reinforced mice. Left: Choice performance on the 10th training day. Right: pCREB levels measured in NAc and HPC after 10 days of training on the spatial task in morphine-reinforced mice that received intra-DS injections of Rp-cAMPS, normalized to morphine-reinforced aCSF controls (dashed line). *n.s.* differences not significant, *significant at $p < 0.05$. [adapted from Baudonnat et al. (2011)].

dopamine-related plasticity in both the NAc and the dorsal striatum in both tasks but was correlated with a marked decrease in pCREB expression in dorsal hippocampus in the spatial task compared to when reinforced with food.

One possibility was that the increase in dorsal and ventral striatal dopamine release originating from the intra-VTA morphine injection, even to predicted rewards, would bias choices to be driven by striatal networks at the expense of hippocampal ones. In support of this hypothesis, inhibition of the protein kinase A pathway in the dorsal striatum, thus potentially down-regulating the consequences of drug-induced dopamine transmission, enabled mice to learn the spatial task even with intra-VTA morphine as the reward and restored pCREB expression in the hippocampus (**Figure 2D**).

This study leaves open several interesting questions concerning how exactly dopamine transmission might dynamically regulate how learning and attention are flexibly allocated and what the longer-term behavioral consequences of this type of dopamine-driven synaptic plasticity might be when there are multiple potential strategies to guide foraging choices. Nonetheless, it clearly demonstrates that different reinforcers, and by extension dopamine transmission, can have a strong influence on the acquisition and consolidation of an appropriate choice strategy, and which brain regions/systems are prioritized during decision making. Similarly, it also shows that the parameters of normal reward-guided dopamine release and the resulting dopamine-mediated synaptic plasticity are critical to allow the correct information to be identified and retained. Thus, modulation of the magnitude and timing of dopamine release may be key to allowing animals to determine when to use, update or discard past experience when encountering novel situations.

## TEMPORAL LOBE INFLUENCES ON DOPAMINE TRANSMISSION DURING LEARNING AND STRATEGY SELECTION

In the previous sections, we have tried to highlight the fact that short- and more long-term dopamine-dependent processes are not mutually exclusive but instead are mutually interacting in order to produce adaptive behavior. Moreover, dopamine transmission in target structures can be locally modulated and may therefore act locally to facilitate acquisition and selection of appropriate foraging strategies. What is not yet clear is: (1) how dopamine transmission affects the way in which potentially competing networks for valuation and behavior are selected; and (2) how dopamine release in terminal regions and afferent modulators of dopamine transmission interact to signal the current state of the world. In the remainder of this review, we will briefly outline some ideas about these processes. Specifically, we will focus on how dopamine transmission may interact with hippocampal circuits through the midbrain and NAc in a bi-directional manner to allow us to learn and behave appropriately in uncertain and changing environments.

As briefly discussed in a previous section, there are strong inputs from the hippocampus into the NAc in both rats and primates (Brog et al., 1993; French and Totterdell, 2002; Friedman et al., 2002), particularly—though not exclusively—from ventral hippocampus to the medial NAc in primates/NAc shell in rodents (NB. the core and shell are difficult to characterize based on cytoarchitecture in primates). In rodents, some of these synapses may converge with amygdala and medial frontal cortex inputs (French and Totterdell, 2002, 2003) and may predominantly target NAc medium spiny neurons associated with the direct pathway (MacAskill et al., 2012). These circuits are also involved in the regulation of VTA DA neuron firing and excitability (Floresco et al., 2001; Lodge and Grace, 2006). Grace et al. (2007), in particular, have highlighted the potential importance of the hippocampus—NAc—ventral pallidum—VTA circuit for altering the activity states of midbrain dopamine neurons, which can therefore act to gate glutamate-driven burst firing. By extension, the modification of the basal activity of DA neurons makes them more likely to produce phasic burst firing when a novel event is detected or when a reward-predictive cue is presented (Grace et al., 2007; Aggarwal et al., 2012). However, this loop is not the only one involved in regulating VTA DA neuronal activity. Several hippocampal—VTA networks, coexisting via different relays in the brain (such as the lateral septum), modulate dopamine cell activity, along with multiple other pathways from cortical regions such as the orbitofrontal and medial frontal cortex (Lodge, 2011; Luo et al., 2011). The precise contribution of the inputs to midbrain dopamine cells or the afferents to striatal regions targeted by dopamine fibers to patterns of local dopamine *release* remains to be determined.

How might these circuits facilitate learning and adaptive foraging through NAc dopamine? As we described in the previous sections, phasic dopamine release in the NAc appears not only to correlate with predictions of future rewards and deviations from these predictions, but can also be driven by novel stimuli, shaped by uncertainty about the structure of the environment and about what information should be used to guide choices, and modified by motivational state and the post-ingestive consequences of reward. At a behavioral level, dopamine release, particularly in the NAc core, can promote approach behavior to cues in the environment, though is unlikely to play a leading role in setting a behavioral policy. Moreover, though dopamine release in NAc in response to both reward-predicting cues and unexpected reward may often be formally consistent with an RPE (i.e., both signal a deviation from a past prediction of future rewards based on newly received information), the two signals may in fact be regulated via dissociable processes (Wanat et al., 2013). This potentially allows for separable influences of afferent structures on cue- and outcome-driven dopamine transmission.

While the main function of the hippocampus is often described in terms of spatial memory, there is also an extensive literature demonstrating that this brain structure also plays a key role in encoding the predictability and regularity of events as well as signaling mismatches or conflicts in the incoming information (Honey et al., 1998; Gray and McNaughton, 2000; Strange et al., 2005; Kumaran and Maguire, 2007; Sanderson and Bannerman, 2012; Schapiro et al., 2012). Given the anatomical and functional connections between the hippocampus, dopaminergic midbrain and the NAc, there is the possibility that these circuits interact dynamically not only to detect novelty, but also to determine the *behavioral relevance* of a new ambiguous cue or environment. One way this might occur is through hippocampal modulations of outcome-driven dopamine release (how surprising is a cue given the current state of the environment and how unexpected is the outcome given past expectations in this state), which in turn can influence synaptic plasticity and the efficiency with which particular inputs to the NAc can affect activity in this region (see Floresco, 2007). This may well be a bi-directional mechanism, with deviations in stimulus-surprise (associative or contextual novelty) being directed from hippocampus to influence mesolimbic dopamine release and the magnitude of outcome-surprise signaled by the extent of NAc dopamine release, which is then directly or indirectly communicated to temporal lobe structures. For instance, as suggested by Gershman, Niv and colleagues in a series of papers (Gershman et al., 2010; Gershman and Niv, 2010), the magnitude of an RPE signaled by dopamine release might be used to indicate whether or not an environment has

fundamentally changed and therefore whether previous choice strategies should be updated, consolidated or discarded in favor of a new state. Therefore, if there is an abrupt change in reward contingencies (for instance, going from a seldom-reinforced to a fully-reinforced situation), the consequent sudden change in outcome-driven dopamine could provide a signal in terminal regions that the animal is in a new context (or new task state, in reinforcement learning terms) and should therefore not integrate the new evidence with past events but should instead treat them as separate situations and start learning anew.

These anatomical loops can therefore potentially account for the effect of dopamine in both short-term encoding of reward associations and longer-term memory recall and updating. The presentation of a familiar situation can cause the re-activation of the memory related to this state. At this time point, the pattern of the cue-outcome contingencies within a behavioral strategy will be critical as to whether the memory is used, updated or discarded because of the labile nature of a re-activated memory (Kuhl et al., 2010). Therefore, by acting in concert, hippocampus – NAc – dopamine circuits could allow organisms not only to work out what cues are relevant in the environment, but also, when faced with a seemingly familiar situation, to determine whether or not to generalize based on stored experience (Shohamy and Wagner, 2008; Wimmer et al., 2012). Specifically, through signaling how similar the current state of the environment is to stored associations—the cues and context via the hippocampus and accompanying reward contingencies via dopamine—these circuits may signal when to integrate separate events if there are statistical regularities between them or, by contrast, when to separate memories and promote new reinforcement learning if there are notable discrepancies.

However, as well as shaping learning and strategy selection, another potential key role of the hippocampus may be the modulation of dopamine-dependent Pavlovian approach behavior by suppressing inappropriate choices or facilitating exploratory responses. Several lines of evidence suggest that one output of hippocampal computations may be to inhibit ongoing behavior in order to prevent impulsive, disadvantageous behavior (Gray and McNaughton, 2000; Mariano et al., 2009; Bannerman et al., 2012; Abela et al., 2013). Therefore, in situations, for instance at times of uncertainty, where there is conflict over which cues in the environment are most relevant to guide behavior or where a superficially tempting option should be resisted in order to obtain a larger benefit in the future, the hippocampus may act as a regulator of NAc dopamine transmission, reducing the likelihood of a prepotent cue-driven response being elicited before the potential future consequences are considered. Consistent with this, Floresco and colleagues found that preventing hippocampal afferents from interacting with NAc dopamine transmission using an asymmetric disconnection procedure caused rats to inappropriately return to previously sampled arms in a foraging task, suggesting that these circuits are required to suppress previously reinforced spatial behaviors (see Floresco, 2007). Conversely, novelty-induced dopamine release to unexpected cues with no current reward associations could provide the motivational drive for animals to approach such a cue in order to gain information about its significance.

In summary, by influencing dopamine transmission in the NAc, the hippocampus can help the dopamine system to: (1) shape appropriate behavior toward the behaviorally- and motivationally-relevant elements of the environment, (2) code the degree of uncertainty between cue-outcome relationships, and (3) elicit molecular cascades strengthening or weakening the re-activated network. Consequently, any novel cue-outcome association can either be integrated in an existing memory or new memories can be laid down instead.

## CONCLUSION

In this review, we have tried to illustrate: first, how dopamine release is not only critical for learning but also to motivate animals to learn about the world at times of uncertainty; second, how past experience can shape dopamine-dependent learning; and, third, how dopamine might play a role not only in the initial learning of cue-reward associations but also in determining when to use stored experience and also when to consolidate associations between stimuli and outcome into memory. Furthermore, we have gone on to suggest ways in which the hippocampus might interact with NAc dopamine to facilitate these processes and to enable animals to react to what is behaviorally relevant in the given environment.

As is common with most comparable reviews, we gladly acknowledge that there are still many details that remain to be fleshed out and many complexities that have been glossed over for the sake of coherence. For instance, throughout, we have concentrated mainly on phasic dopamine release at the expense of slower tonic changes (although the two are likely related). As was discussed in an earlier section, even at the "slow, phasic" timescale (~0.5–10 s post-event) dopamine activity can evolve over time to represent several different parameters. Similarly, quite how different regions of the hippocampus interact with dopamine transmission across different parts of the striatum during learning and behavior remains to be determined. We would contend that a general computation might be shared across structures (for instance, determining statistical regularities of events and inhibiting ongoing behavior when there is conflict, for the hippocampus), even if the specific information provided by, for instance, ventral hippocampus to the NAc shell region may be different to dorsal hippocampus and NAc core.

Finally, while we have focused on hippocampus—VTA—NAc loops for simplicity, it is improbable that other temporal and frontal lobe regions are not also required to enable these processes to work optimally. For instance, orbitofrontal cortex, which receives hippocampal input and projects to the VTA, has been shown to provide input to allow dopamine cells to disambiguate similar states (for instance, being in a reward port following a choice), particularly when there is a delay between a choice and its consequences (Takahashi et al., 2011). Basolateral amygdala can also attenuate NAc cue-driven dopamine (Jones et al., 2010). How different nodes in this network interact to generate appropriate learning and behavior will be key questions to be addressed over the coming years in order to enable us to understand these processes in the complex, changeable and uncertain environments

within which we live. Our hope is that, by investigating these networks more deeply and their interactions with dopamine release at different timescales, we may gain new insights to understand pathologies related to dopamine dysfunction such as schizophrenia where learning and behavior can become unconstrained by the parameters of the local environment.

## ACKNOWLEDGMENTS

This work was supported by a Wellcome Trust Research Career Development Fellowship to Mark E. Walton. The authors would like to thank David Bannerman for valuable input about hippocampal function and David Dupret for discussions on neuroanatomy.

## REFERENCES

Abela, A. R., Dougherty, S. D., Fagen, E. D., Hill, C. J., and Chudasama, Y. (2013). Inhibitory control deficits in rats with ventral hippocampal lesions. *Cereb. Cortex* 23, 1396–1409. doi: 10.1093/cercor/bhs121

Abizaid, A. (2009). Ghrelin and dopamine: new insights on the peripheral regulation of appetite. *J. Neuroendocrinol.* 21, 787–793. doi: 10.1111/j.1365-2826.2009.01896.x

Adamantidis, A. R., Tsai, H. C., Boutrel, B., Zhang, F., Stuber, G. D., Budygin, E. A., et al. (2011). Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior. *J. Neurosci.* 31, 10829–10835. doi: 10.1523/JNEUROSCI.2246-11.2011

Aggarwal, M., Hyland, B. I., and Wickens, J. R. (2012). Neural control of dopamine neurotransmission: implications for reinforcement learning. *Eur. J. Neurosci.* 35, 1115–1123. doi: 10.1111/j.1460-9568.2012.08055.x

Ahn, S., and Phillips, A. G. (1999). Dopaminergic correlates of sensory-specific satiety in the medial prefrontal cortex and nucleus accumbens of the rat. *J. Neurosci.* 19, RC29.

Badrinarayan, A., Wescott, S. A., Vander Weele, C. M., Saunders, B. T., Couturier, B. E., Maren, S., et al. (2012). Aversive stimuli differentially modulate real-time dopamine transmission dynamics within the nucleus accumbens core and shell. *J. Neurosci.* 32, 15779–15790. doi: 10.1523/JNEUROSCI.3557-12.2012

Bannerman, D. M., Bus, T., Taylor, A., Sanderson, D. J., Schwarz, I., Jensen, V., et al. (2012). Dissecting spatial knowledge from spatial choice by hippocampal NMDA receptor deletion. *Nat. Neurosci.* 15, 1153–1159. doi: 10.1038/nn.3166

Bannerman, D. M., Rawlins, J. N., McHugh, S. B., Deacon, R. M., Yee, B. K., Bast, T., et al. (2004). Regional dissociations within the hippocampus–memory and anxiety. *Neurosci. Biobehav. Rev.* 28, 273–283. doi: 10.1016/j.neubiorev.2004.03.004

Bassareo, V., De Luca, M. A., and Di Chiara, G. (2002). Differential expression of motivational stimulus properties by dopamine in nucleus accumbens shell versus core and prefrontal cortex. *J. Neurosci.* 22, 4709–4719.

Bassareo, V., and Di Chiara, G. (1999). Modulation of feeding-induced activation of mesolimbic dopamine transmission by appetitive stimuli and its relation to motivational state. *Eur. J. Neurosci.* 11, 4389–4397. doi: 10.1046/j.1460-9568.1999.00843.x

Baudonnat, M., Guillou, J. L., Husson, M., Vandesquille, M., Corio, M., Decorte, L., et al. (2011). Disrupting effect of drug-induced reward on spatial but not cue-guided learning: implication of the striatal protein kinase A/cAMP response element-binding protein pathway. *J. Neurosci.* 31, 16517–16528. doi: 10.1523/JNEUROSCI.1787-11.2011

Bayer, H. M., and Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141. doi: 10.1016/j.neuron.2005.05.020

Beeler, J. A., McCutcheon, J. E., Cao, Z. F., Murakami, M., Alexander, E., Roitman, M. F., et al. (2012). Taste uncoupled from nutrition fails to sustain the reinforcing properties of food. *Eur. J. Neurosci.* 36, 2533–2546. doi: 10.1111/j.1460-9568.2012.08167.x

Besson, M., David, V., Baudonnat, M., Cazala, P., Guilloux, J. P., Reperant, C., et al. (2012). Alpha7-nicotinic receptors modulate nicotine-induced reinforcement and extracellular dopamine outflow in the mesolimbic system in mice. *Psychopharmacology (Berl.)* 220, 1–14. doi: 10.1007/s00213-011-2422-1

Bethus, I., Tse, D., and Morris, R. G. (2010). Dopamine and memory: modulation of the persistence of memory for novel hippocampal NMDA receptor-dependent paired associates. *J. Neurosci.* 30, 1610–1618. doi: 10.1523/JNEUROSCI.2721-09.2010

Brayda-Bruno, L., Mons, N., Yee, B. K., Micheau, J., Abrous, D. N., Nogues, X., et al. (2013). Partial loss in septo-hippocampal cholinergic neurons alters memory-dependent measures of brain connectivity without overt memory deficits. *Neurobiol. Dis.* 54, 372–381. doi: 10.1016/j.nbd.2013.01.010

Brischoux, F., Chakraborty, S., Brierley, D. I., and Ungless, M. A. (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4894–4899. doi: 10.1073/pnas.0811507106

Brog, J. S., Salyapongse, A., Deutch, A. Y., and Zahm, D. S. (1993). The patterns of afferent innervation of the core and shell in the "accumbens" part of the rat ventral striatum: immunohistochemical detection of retrogradely transported fluorogold. *J. Comp. Neurol.* 338, 255–278. doi: 10.1002/cne.903380209

Bromberg-Martin, E. S., and Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron* 63, 119–126. doi: 10.1016/j.neuron.2009.06.009

Bromberg-Martin, E. S., Matsumoto, M., and Hikosaka, O. (2010a). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68, 815–834. doi: 10.1016/j.neuron.2010.11.022

Bromberg-Martin, E. S., Matsumoto, M., Hong, S., and Hikosaka, O. (2010b). A pallidus-habenula-dopamine pathway signals inferred stimulus values. *J. Neurophysiol.* 104, 1068–1076. doi: 10.1152/jn.00158.2010

Bromberg-Martin, E. S., Matsumoto, M., Nakahara, H., and Hikosaka, O. (2010c). Multiple timescales of memory in lateral habenula and dopamine neurons. *Neuron* 67, 499–510. doi: 10.1016/j.neuron.2010.06.031

Brown, H. D., McCutcheon, J. E., Cone, J. J., Ragozzino, M. E., and Roitman, M. F. (2011). Primary food reward and reward-predictive stimuli evoke different patterns of phasic dopamine signaling throughout the striatum. *Eur. J. Neurosci.* 34, 1997–2006. doi: 10.1111/j.1460-9568.2011.07914.x

Calabresi, P., Maj, R., Mercuri, N. B., and Bernardi, G. (1992). Coactivation of D1 and D2 dopamine receptors is required for long-term synaptic depression in the striatum. *Neurosci. Lett.* 142, 95–99. doi: 10.1016/0304-3940(92)90628-K

Clark, J. J., Collins, A. L., Sanford, C. A., and Phillips, P. E. (2013). Dopamine encoding of Pavlovian incentive stimuli diminishes with extended training. *J. Neurosci.* 33, 3526–3532. doi: 10.1523/JNEUROSCI.5119-12.2013

Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B., and Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* 482, 85–88. doi: 10.1038/nature10754

Corbit, L. H., and Balleine, B. W. (2011). The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. *J. Neurosci.* 31, 11786–11794. doi: 10.1523/JNEUROSCI.2711-11.2011

Costa, R. M. (2007). Plastic corticostriatal circuits for action learning: what's dopamine got to do with it. *Ann. N.Y. Acad. Sci. U.S.A.* 1104, 172–191. doi: 10.1196/annals.1390.015

Dalley, J. W., Chudasama, Y., Theobald, D. E., Pettifer, C. L., Fletcher, C. M., and Robbins, T. W. (2002). Nucleus accumbens dopamine and discriminated approach learning: interactive effects of 6-hydroxydopamine lesions and systemic apomorphine administration. *Psychopharmacology (Berl.)* 161, 425–433. doi: 10.1007/s00213-002-1078-2

Dalley, J. W., Laane, K., Theobald, D. E., Armstrong, H. C., Corlett, P. R., Chudasama, Y., et al. (2005). Time-limited modulation of appetitive Pavlovian memory by D1 and NMDA receptors in the nucleus accumbens. *Proc. Natl. Acad. Sci. U.S.A.* 102, 6189–6194. doi: 10.1073/pnas.0502080102

D'Ardenne, K., McClure, S. M., Nystrom, L. E., and Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human

ventral tegmental area. *Science* 319, 1264–1267. doi: 10.1126/science.1150605

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., and Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69, 1204–1215. doi: 10.1016/j.neuron.2011.02.027

Day, J. J., Roitman, M. F., Wightman, R. M., and Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.* 10, 1020–1028. doi: 10.1038/nn1923

de Araujo, I. E., Oliveira-Maia, A. J., Sotnikova, T. D., Gainetdinov, R. R., Caron, M. G., Nicolelis, M. A., et al. (2008). Food reward in the absence of taste receptor signaling. *Neuron* 57, 930–941. doi: 10.1016/j.neuron.2008.01.032

Di Ciano, P., Cardinal, R. N., Cowell, R. A., Little, S. J., and Everitt, B. J. (2001). Differential involvement of NMDA, AMPA/kainate, and dopamine receptors in the nucleus accumbens core in the acquisition and performance of pavlovian approach behavior. *J. Neurosci.* 21, 9471–9477.

Domingos, A. I., Vaynshteyn, J., Voss, H. U., Ren, X., Gradinaru, V., Zang, F., et al. (2011). Leptin regulates the reward value of nutrient. *Nat. Neurosci.* 14, 1562–1568. doi: 10.1038/nn.2977

Dudman, J. T., Eaton, M. E., Rajadhyaksha, A., Macias, W., Taher, M., Barczak, A., et al. (2003). Dopamine D1 receptors mediate CREB phosphorylation via phosphorylation of the NMDA receptor at Ser897-NR1. *J. Neurochem.* 87, 922–934. doi: 10.1046/j.1471-4159.2003.02067.x

Everitt, B. J., and Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* 8, 1481–1489. doi: 10.1038/nn1579

Ferreira, J. G., Tellez, L. A., Ren, X., Yeckel, C. W., and de Araujo, I. E. (2012). Regulation of fat intake in the absence of flavour signalling. *J. Physiol.* 590, 953–972. doi: 10.1113/jphysiol.2011.218289

Fiorillo, C. D., Song, M. R., and Yun, S. R. (2013a). Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli. *J. Neurosci.* 33, 4710–4725. doi: 10.1523/JNEUROSCI.3883-12.2013

Fiorillo, C. D., Yun, S. R., and Song, M. R. (2013b). Diversity and homogeneity in responses of midbrain dopamine neurons. *J. Neurosci.* 33, 4693–4709. doi: 10.1523/JNEUROSCI.3886-12.2013

Fiorillo, C. D., Tobler, P. N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902. doi: 10.1126/science.1077349

Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., et al. (2011). A selective role for dopamine in stimulus-reward learning. *Nature* 469, 53–57. doi: 10.1038/nature09588

Floel, A., Garraux, G., Xu, B., Breitenstein, C., Knecht, S., Herscovitch, P., et al. (2008). Levodopa increases memory encoding and dopamine release in the striatum in the elderly. *Neurobiol. Aging* 29, 267–279. doi: 10.1016/j.neurobiolaging.2006.10.009

Floresco, S. B. (2007). Dopaminergic regulation of limbic-striatal interplay. *J. Psychiatry Neurosci.* 32, 400–411.

Floresco, S. B. (2013). Prefrontal dopamine and behavioral flexibility: shifting from an "inverted-U" toward a family of functions. *Front. Neurosci.* 7:62. doi: 10.3389/fnins.2013.00062

Floresco, S. B., Todd, C. L., and Grace, A. A. (2001). Glutamatergic afferents from the hippocampus to the nucleus accumbens regulate activity of ventral tegmental area dopamine neurons. *J. Neurosci.* 21, 4915–4922.

Floresco, S. B., Yang, C. R., Phillips, A. G., and Blaha, C. D. (1998). Basolateral amygdala stimulation evokes glutamate receptor-dependent dopamine efflux in the nucleus accumbens of the anaesthetized rat. *Eur. J. Neurosci.* 10, 1241–1251. doi: 10.1046/j.1460-9568.1998.00133.x

Frankland, P. W., and Bontempi, B. (2005). The organization of recent and remote memories. *Nat. Rev. Neurosci.* 6, 119–130. doi: 10.1038/nrn1607

French, S. J., and Totterdell, S. (2002). Hippocampal and prefrontal cortical inputs monosynaptically converge with individual projection neurons of the nucleus accumbens. *J. Comp. Neurol.* 446, 151–165. doi: 10.1002/cne.10191

French, S. J., and Totterdell, S. (2003). Individual nucleus accumbens-projection neurons receive both basolateral amygdala and ventral subicular afferents in rats. *Neuroscience* 119, 19–31. doi: 10.1016/S0306-4522(03)00150-7

Friedman, D. P., Aggleton, J. P., and Saunders, R. C. (2002). Comparison of hippocampal, amygdala, and perirhinal projections to the nucleus accumbens: combined anterograde and retrograde tracing study in the Macaque brain. *J. Comp. Neurol.* 450, 345–365. doi: 10.1002/cne.10336

Gan, J. O., Walton, M. E., and Phillips, P. E. (2010). Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nat. Neurosci.* 13, 25–27. doi: 10.1038/nn.2460

Gershman, S. J., Blei, D. M., and Niv, Y. (2010). Context, learning, and extinction. *Psychol. Rev.* 117, 197–209. doi: 10.1037/a0017808

Gershman, S. J., and Niv, Y. (2010). Learning latent structure: carving nature at its joints. *Curr. Opin. Neurobiol.* 20, 251–256. doi: 10.1016/j.conb.2010.02.008

Ghods-Sharifi, S., and Floresco, S. B. (2010). Differential effects on effort discounting induced by inactivations of the nucleus accumbens core or shell. *Behav. Neurosci.* 124, 179–191. doi: 10.1037/a0018932

Goto, Y., and Grace, A. A. (2008). Dopamine modulation of hippocampal-prefrontal cortical interaction drives memory-guided behavior. *Cereb. Cortex* 18, 1407–1414. doi: 10.1093/cercor/bhm172

Grace, A. A., Floresco, S. B., Goto, Y., and Lodge, D. J. (2007). Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends Neurosci.* 30, 220–227. doi: 10.1016/j.tins.2007.03.003

Gray, J. A., and McNaughton, N. (eds.), (2000). *The Neuropsychology of Anxiety.* Oxford: Oxford Univ. Press.

Haluk, D. M., and Floresco, S. B. (2009). Ventral striatal dopamine modulation of different forms of behavioral flexibility. *Neuropsychopharmacology* 34, 2041–2052. doi: 10.1038/npp.2009.21

Honey, R. C., Watt, A., and Good, M. (1998). Hippocampal lesions disrupt an associative mismatch process. *J. Neurosci.* 18, 2226–2230.

Horvitz, J. C., Stewart, T., and Jacobs, B. L. (1997). Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. *Brain Res.* 759, 251–258. doi: 10.1016/S0006-8993(97)00265-5

Huang, Y. Y., and Kandel, E. R. (1995). D1/D5 receptor agonists induce a protein synthesis-dependent late potentiation in the CA1 region of the hippocampus. *Proc. Natl. Acad. Sci. U.S.A.* 92, 2446–2450. doi: 10.1073/pnas.92.7.2446

Hyman, S. E., Malenka, R. C., and Nestler, E. J. (2006). Neural mechanisms of addiction: the role of reward-related learning and memory. *Annu. Rev. Neurosci.* 29, 565–598. doi: 10.1146/annurev.neuro.29.051605.113009

Ito, R., Dalley, J. W., Howes, S. R., Robbins, T. W., and Everitt, B. J. (2000). Dissociation in conditioned dopamine release in the nucleus accumbens core and shell in response to cocaine cues and during cocaine-seeking behavior in rats. *J. Neurosci.* 20, 7489–7495.

Ito, R., and Hayen, A. (2011). Opposing roles of nucleus accumbens core and shell dopamine in the modulation of limbic information processing. *J. Neurosci.* 31, 6001–6007. doi: 10.1523/JNEUROSCI.6588-10.2011

Johnson, A., van der Meer, M. A., and Redish, A. D. (2007). Integrating hippocampus and striatum in decision-making. *Curr. Opin. Neurobiol.* 17, 692–697. doi: 10.1016/j.conb.2008.01.003

Johnson, S. W., and North, R. A. (1992). Opioids excite dopamine neurons by hyperpolarization of local interneurons. *J. Neurosci.* 12, 483–488.

Jones, J. L., Day, J. J., Aragona, B. J., Wheeler, R. A., Wightman, R. M., and Carelli, R. M. (2010). Basolateral amygdala modulates terminal dopamine release in the nucleus accumbens and conditioned responding. *Biol. Psychiatry* 67, 737–744. doi: 10.1016/j.biopsych.2009.11.006

Joshua, M., Adler, A., Mitelman, R., Vaadia, E., and 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, 11673–11684. doi: 10.1523/JNEUROSCI.3839-08.2008

Kakade, S., and Dayan, P. (2002). Dopamine: generalization and bonuses. *Neural Netw.* 15, 549–559. doi: 10.1016/S0893-6080(02)00048-5

Kandel, E. R. (2001). The molecular biology of memory storage: a dialogue between genes and synapses. *Science* 294, 1030–1038. doi: 10.1126/science.1067020

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

Klein-Flugge, M. C., Hunt, L. T., Bach, D. R., Dolan, R. J., and

Behrens, T. E. (2011). Dissociable reward and timing signals in human midbrain and ventral striatum. *Neuron* 72, 654–664. doi: 10.1016/j.neuron.2011.08.024

Knutson, B., and Gibbs, S. E. (2007). Linking nucleus accumbens dopamine and blood oxygenation. *Psychopharmacology (Berl.)* 191, 813–822. doi: 10.1007/s00213-006-0686-7

Kuhl, B. A., Shah, A. T., DuBrow, S., and Wagner, A. D. (2010). Resistance to forgetting associated with hippocampus-mediated reactivation during new learning. *Nat Neurosci* 13, 501–506. doi: 10.1038/nn.2498

Kumaran, D., and Maguire, E. A. (2007). Match mismatch processes underlie human hippocampal responses to associative novelty. *J. Neurosci.* 27, 8517–8524. doi: 10.1523/JNEUROSCI.1677-07.2007

LaLumiere, R. T., Nawar, E. M., and McGaugh, J. L. (2005). Modulation of memory consolidation by the basolateral amygdala or nucleus accumbens shell requires concurrent dopamine receptor activation in both brain regions. *Learn. Mem.* 12, 296–301. doi: 10.1101/lm.93205

Lammel, S., Ion, D. I., Roeper, J., and Malenka, R. C. (2011). Projection-specific modulation of dopamine neuron synapses by aversive and rewarding stimuli. *Neuron* 70, 855–862. doi: 10.1016/j.neuron.2011.03.025

Legault, M., Rompre, P. P., and Wise, R. A. (2000). Chemical stimulation of the ventral hippocampus elevates nucleus accumbens dopamine by activating dopaminergic neurons of the ventral tegmental area. *J. Neurosci.* 20, 1635–1642.

Lex, A., and Hauber, W. (2008). Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. *Learn. Mem.* 15, 483–491. doi: 10.1101/lm.978708

Lex, B., and Hauber, W. (2010). The role of dopamine in the prelimbic cortex and the dorsomedial striatum in instrumental conditioning. *Cereb. Cortex* 20, 873–883. doi: 10.1093/cercor/bhp151

Lisman, J. E., and Grace, A. A. (2005). The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* 46, 703–713. doi: 10.1016/j.neuron.2005.05.002

Ljungberg, T., Apicella, P., and Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* 67, 145–163.

Lodge, D. J. (2011). The medial prefrontal and orbitofrontal cortices differentially regulate dopamine system function. *Neuropsychopharmacology* 36, 1227–1236. doi: 10.1038/npp.2011.7

Lodge, D. J., and Grace, A. A. (2006). The hippocampus modulates dopamine neuron responsivity by regulating the intensity of phasic neuron activation. *Neuropsychopharmacology* 31, 1356–1361. doi: 10.1038/sj.npp.1300963

Lovinger, D. M. (2010). Neurotransmitter roles in synaptic modulation, plasticity and learning in the dorsal striatum. *Neuropharmacology* 58, 951–961. doi: 10.1016/j.neuropharm.2010.01.008

Luo, A. H., Tahsili-Fahadan, P., Wise, R. A., Lupica, C. R., and Aston-Jones, G. (2011). Linking context with reward: a functional circuit from hippocampal CA3 to ventral tegmental area. *Science* 333, 353–357. doi: 10.1126/science.1204622

MacAskill, A. F., Little, J. P., Cassel, J. M., and Carter, A. G. (2012). Subcellular connectivity underlies pathway-specific signaling in the nucleus accumbens. *Nat. Neurosci.* 15, 1624–1626. doi: 10.1038/nn.3254

Mariano, T. Y., Bannerman, D. M., McHugh, S. B., Preston, T. J., Rudebeck, P. H., Rudebeck, S. R., et al. (2009). Impulsive choice in hippocampal but not orbitofrontal cortex-lesioned rats on a nonspatial decision-making maze task. *Eur. J. Neurosci.* 30, 472–484. doi: 10.1111/j.1460-9568.2009.06837.x

Matsumoto, M., and Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459, 837–841. doi: 10.1038/nature08028

Matthews, R. T., and German, D. C. (1984). Electrophysiological evidence for excitation of rat ventral tegmental area dopamine neurons by morphine. *Neuroscience* 11, 617–625. doi: 10.1016/0306-4522(84)90048-4

McCutcheon, J. E., Beeler, J. A., and Roitman, M. F. (2012). Sucrose-predictive cues evoke greater phasic dopamine release than saccharin-predictive cues. *Synapse* 66, 346–351. doi: 10.1002/syn.21519

McDannald, M. A., Takahashi, Y. K., Lopatina, N., Pietras, B. W., Jones, J. L., and Schoenbaum, G. (2012). Model-based learning and the contribution of the orbitofrontal cortex to the model-free world. *Eur. J. Neurosci.* 35, 991–996. doi: 10.1111/j.1460-9568.2011.07982.x

McDonald, R. J., and White, N. M. (1994). Parallel information processing in the water maze: evidence for independent memory systems involving dorsal striatum and hippocampus. *Behav. Neural Biol.* 61, 260–270. doi: 10.1016/S0163-1047(05)80009-3

Mebel, D. M., Wong, J. C., Dong, Y. J., and Borgland, S. L. (2012). Insulin in the ventral tegmental area reduces hedonic feeding and suppresses dopamine concentration via increased reuptake. *Eur. J. Neurosci.* 36, 2336–2346. doi: 10.1111/j.1460-9568.2012.08168.x

Mirenowicz, J., and Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379, 449–451. doi: 10.1038/379449a0

Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.

Morris, G., Nevet, A., Arkadir, D., Vaadia, E., and Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* 9, 1057–1063. doi: 10.1038/nn1743

Nakahara, H., and Hikosaka, O. (2012). Learning to represent reward structure: a key to adapting to complex environments. *Neurosci. Res.* 74, 177–183. doi: 10.1016/j.neures.2012.09.007

Nakahara, H., Itoh, H., Kawagoe, R., Takikawa, Y., and Hikosaka, O. (2004). Dopamine neurons can represent context-dependent prediction error. *Neuron* 41, 269–280. doi: 10.1016/S0896-6273(03)00869-9

Nelson, A. J., Thur, K. E., Marsden, C. A., and Cassaday, H. J. (2010). Dissociable roles of dopamine within the core and medial shell of the nucleus accumbens in memory for objects and place. *Behav. Neurosci.* 124, 789–799. doi: 10.1037/a0021114

Nicola, S. M. (2010). The flexible approach hypothesis: unification of effort and cue-responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. *J. Neurosci.* 30, 16585–16600. doi: 10.1523/JNEUROSCI.3958-10.2010

Nomoto, K., Schultz, W., Watanabe, T., and Sakagami, M. (2010). Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli. *J. Neurosci.* 30, 10692–10702. doi: 10.1523/JNEUROSCI.4828-09.2010

Nugent, F. S., Penick, E. C., and Kauer, J. A. (2007). Opioids block long-term potentiation of inhibitory synapses. *Nature* 446, 1086–1090. doi: 10.1038/nature05726

Oleson, E. B., Gentry, R. N., Chioma, V. C., and Cheer, J. F. (2012). Subsecond dopamine release in the nucleus accumbens predicts conditioned punishment and its successful avoidance. *J. Neurosci.* 32, 14804–14808. doi: 10.1523/JNEUROSCI.3087-12.2012

Ostlund, S. B., Wassum, K. M., Murphy, N. P., Balleine, B. W., and Maidment, N. T. (2011). Extracellular dopamine levels in striatal subregions track shifts in motivation and response cost during instrumental conditioning. *J. Neurosci.* 31, 200–207. doi: 10.1523/JNEUROSCI.4759-10.2011

Packard, M. G., and White, N. M. (1989). Memory facilitation produced by dopamine agonists: role of receptor subtype and mnemonic requirements. *Pharmacol. Biochem. Behav.* 33, 511–518. doi: 10.1016/0091-3057(89)90378-X

Packard, M. G., and White, N. M. (1991). Dissociation of hippocampus and caudate nucleus memory systems by posttraining intracerebral injection of dopamine agonists. *Behav. Neurosci.* 105, 295–306. doi: 10.1037/0735-7044.105.2.295

Pawlak, V., and Kerr, J. N. (2008). Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. *J. Neurosci.* 28, 2435–2446. doi: 10.1523/JNEUROSCI.4402-07.2008

Pennartz, C. M., Ito, R., Verschure, P. F., Battaglia, F. P., and Robbins, T. W. (2011). The hippocampal-striatal axis in learning, prediction and goal-directed behavior. *Trends Neurosci.* 34, 548–559. doi: 10.1016/j.tins.2011.08.001

Pezze, M. A., Heidbreder, C. A., Feldon, J., and Murphy, C. A. (2001). Selective responding of nucleus accumbens core and shell dopamine to aversively conditioned contextual and discrete stimuli. *Neuroscience* 108, 91–102. doi: 10.1016/S0306-4522(01)00403-1

Phillips, A. G., Vacca, G., and Ahn, S. (2008). A top-down perspective on dopamine, motivation and memory. *Pharmacol.*

*Biochem. Behav.* 90, 236–249. doi: 10.1016/j.pbb.2007.10.014

Phillips, P. E., Walton, M. E., and Jhou, T. C. (2007). Calculating utility: preclinical evidence for cost-benefit analysis by mesolimbic dopamine. *Psychopharmacology (Berl.)* 191, 483–495. doi: 10.1007/s00213-006-0626-6

Pompilio, L., and Kacelnik, A. (2010). Context-dependent utility overrides absolute memory as a determinant of choice. *Proc. Natl. Acad. Sci. U.S.A.* 107, 508–512. doi: 10.1073/pnas.0907250107

Porte, Y., Buhot, M. C., and Mons, N. E. (2008). Spatial memory in the Morris water maze and activation of cyclic AMP response element-binding (CREB) protein within the mouse hippocampus. *Learn. Mem.* 15, 885–894. doi: 10.1101/lm.1094208

Redgrave, P., Gurney, K., and Reynolds, J. (2008). What is reinforced by phasic dopamine signals. *Brain Res. Rev.* 58, 322–339. doi: 10.1016/j.brainresrev.2007.10.007

Reynolds, J. N., Hyland, B. I., and Wickens, J. R. (2001). A cellular mechanism of reward-related learning. *Nature* 413, 67–70. doi: 10.1038/35092560

Roesch, M. R., Calu, D. J., and Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* 10, 1615–1624. doi: 10.1038/nn2013

Roitman, M. F., Stuber, G. D., Phillips, P. E., Wightman, R. M., and Carelli, R. M. (2004). Dopamine operates as a subsecond modulator of food seeking. *J. Neurosci.* 24, 1265–1271. doi: 10.1523/JNEUROSCI.3823-03.2004

Sanderson, D. J., and Bannerman, D. M. (2012). The role of habituation in hippocampus-dependent spatial working memory tasks: evidence from GluA1 AMPA receptor subunit knockout mice. *Hippocampus* 22, 981–994. doi: 10.1002/hipo.20896

Schapiro, A. C., Kustner, L. V., and Turk-Browne, N. B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Curr. Biol.* 22, 1622–1627. doi: 10.1016/j.cub.2012.06.056

Schultz, W. (1997). Dopamine neurons and their role in reward mechanisms. *Curr. Opin. Neurobiol.* 7, 191–197. doi: 10.1016/S0959-4388(97)80007-4

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron* 36, 241–263. doi: 10.1016/S0896-6273(02)00967-4

Schultz, W. (2007). Behavioral dopamine signals. *Trends Neurosci.* 30, 203–210. doi: 10.1016/j.tins.2007.03.007

Schultz, W. (2013). Updating dopamine reward signals. *Curr. Opin. Neurobiol.* 23, 229–238. doi: 10.1016/j.conb.2012.11.012

Shohamy, D., and Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends Cogn. Sci.* 14, 464–472. doi: 10.1016/j.tics.2010.08.002

Shohamy, D., and Wagner, A. D. (2008). Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron* 60, 378–389. doi: 10.1016/j.neuron.2008.09.023

Siegel, G. J., Agranoff, B. W., Albers, R. W., Fisher, S. K., and Uhler, M. D. (eds.), (1999). *Basic Neurochemistry, 6th Edn*. Philadelphia: Lippincott Williams and Wilkins.

Sokolowski, J. D., and Salamone, J. D. (1998). The role of accumbens dopamine in lever pressing and response allocation: effects of 6-OHDA injected into core and dorsomedial shell. *Pharmacol. Biochem. Behav.* 59, 557–566. doi: 10.1016/S0091-3057(97)00544-3

Stefani, M. R., and Moghaddam, B. (2006). Rule learning and reward contingency are associated with dissociable patterns of dopamine activation in the rat prefrontal cortex, nucleus accumbens, and dorsal striatum. *J. Neurosci.* 26, 8810–8818. doi: 10.1523/JNEUROSCI.1656-06.2006

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., and Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nat. Neurosci.* 16, 966–973. doi: 10.1038/nn.3413

Stopper, C. M., and Floresco, S. B. (2011). Contributions of the nucleus accumbens and its subregions to different aspects of risk-based decision making. *Cogn. Affect. Behav. Neurosci.* 11, 97–112. doi: 10.3758/s13415-010-0015-9

Strange, B. A., Duggins, A., Penny, W., Dolan, R. J., and Friston, K. J. (2005). Information theory, novelty and hippocampal responses: unpredicted or unpredictable. *Neural Netw.* 18, 225–230. doi: 10.1016/j.neunet.2004.12.004

Tai, L. H., Lee, A. M., Benavidez, N., Bonci, A., and Wilbrecht, L. (2012). Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat. Neurosci.* 15, 1281–1289. doi: 10.1038/nn.3188

Takahashi, Y. K., Roesch, M. R., Wilson, R. C., Toreson, K., O'Donnell, P., Niv, Y., et al. (2011). Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nat. Neurosci.* 14, 1590–1597. doi: 10.1038/nn.2957

Threlfell, S., Lalic, T., Platt, N. J., Jennings, K. A., Deisseroth, K., and Cragg, S. J. (2012). Striatal dopamine release is triggered by synchronized activity in cholinergic interneurons. *Neuron* 75, 58–64. doi: 10.1016/j.neuron.2012.04.038

Tobler, P. N., Fiorillo, C. D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642–1645. doi: 10.1126/science.1105370

Tolu, S., Eddine, R., Marti, F., David, V., Graupner, M., Pons, S., et al. (2013). Co-activation of VTA DA and GABA neurons mediates nicotine reinforcement. *Mol. Psychiatry* 18, 382–393. doi: 10.1038/mp.2012.83

Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412, 43–48. doi: 10.1038/35083500

Walton, M. E., Gan, J. O., and Phillips, P. E. M. (2011). "The influence of dopamine in generating action from motivation," in *Neural Basis of Motivational and Cognitive Control*, ed R. B. Mars (Cambridge, MA: MIT Press), 163–187.

Wanat, M. J., Bonci, A., and Phillips, P. E. (2013). CRF acts in the midbrain to attenuate accumbens dopamine release to rewards but not their predictors. *Nat. Neurosci.* 16, 383–385. doi: 10.1038/nn.3335

Wassum, K. M., Ostlund, S. B., and Maidment, N. T. (2012). Phasic mesolimbic dopamine signaling precedes and predicts performance of a self-initiated action sequence task. *Biol. Psychiatry* 71, 846–854. doi: 10.1016/j.biopsych.2011.12.019

White, N. M., and McDonald, R. J. (2002). Multiple parallel memory systems in the brain of the rat. *Neurobiol. Learn. Mem.* 77, 125–184. doi: 10.1006/nlme.2001.4008

Wickens, J. R. (2009). Synaptic plasticity in the basal ganglia. *Behav. Brain Res.* 199, 119–128. doi: 10.1016/j.bbr.2008.10.030

Wilson, C., Nomikos, G. G., Collu, M., and Fibiger, H. C. (1995). Dopaminergic correlates of motivated behavior: importance of drive. *J. Neurosci.* 15, 5169–5178.

Wimmer, G. E., Daw, N. D., and Shohamy, D. (2012). Generalization of value in reinforcement learning by humans. *Eur. J. Neurosci.* 35, 1092–1104. doi: 10.1111/j.1460-9568.2012.08017.x

Wise, R. A. (2009). Roles for nigrostriatal–not just mesocorticolimbic–dopamine in reward and addiction. *Trends Neurosci.* 32, 517–524. doi: 10.1016/j.tins.2009.06.004

Wunderlich, K., Smittenaar, P., and Dolan, R. J. (2012). Dopamine enhances model-based over model-free choice behavior. *Neuron* 75, 418–424. doi: 10.1016/j.neuron.2012.03.042

Yun, I. A., Wakabayashi, K. T., Fields, H. L., and Nicola, S. M. (2004). The ventral tegmental area is required for the behavioral and nucleus accumbens neuronal firing responses to incentive cues. *J. Neurosci.* 24, 2923–2933. doi: 10.1523/JNEUROSCI.5282-03.2004

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 14 July 2013; accepted: 09 September 2013; published online: 11 October 2013.*

*Citation: Baudonnat M, Huber A, David V and Walton ME (2013) Heads for learning, tails for memory: reward, reinforcement and a role of dopamine in determining behavioral relevance across multiple timescales. Front. Neurosci. 7:175. doi: 10.3389/fnins.2013.00175*

*This article was submitted to Decision Neuroscience, a section of the journal Frontiers in Neuroscience.*

frontiers in
NEUROSCIENCE

REVIEW ARTICLE
published: 07 June 2013
doi: 10.3389/fnins.2013.00096

# On the role of subsecond dopamine release in conditioned avoidance

***Erik B. Oleson and Joseph F. Cheer****

*Department of Anatomy and Neurobiology, School of Medicine, University of Maryland, Baltimore, MD, USA*

***Edited by:***
*Sean B. Ostlund, University of California at Los Angeles, USA*

***Reviewed by:***
*Christopher Cain, Nathan S. Kline Institute for Pscyhiatric Research, USA*
*Niall P. Murphy, University of California at Los Angeles, USA*

***Correspondence:**
*Joseph F. Cheer, Department of Anatomy and Neurobiology, School of Medicine, University of Maryland, 20 Penn Street, Baltimore, MD 21201, USA*
*e-mail: jchee001@umaryland.edu*

Using shock avoidance procedures to study conditioned behavioral responses has a rich history within the field of experimental psychology. Such experiments led to the formulation of the general concept of negative reinforcement and specific theories attempting to explain escape and avoidance behavior, or why animals choose to either terminate or prevent the presentation of an aversive event. For example, the two-factor theory of avoidance holds that cues preceding an aversive event begin to evoke conditioned fear responses, and these conditioned fear responses reinforce the instrumental avoidance response. Current neuroscientific advances are providing new perspectives into this historical literature. Due to its well-established role in reinforcement processes and behavioral control, the mesolimbic dopamine system presented itself as a logical starting point in the search for neural correlates of avoidance and escape behavior. We recently demonstrated that phasic dopamine release events are inhibited by stimuli associated with aversive events but increased by stimuli preceding the successful avoidance of the aversive event. The latter observation is inconsistent with the second component of the two-factor theory of avoidance and; therefore, led us propose a new theoretical explanation of conditioned avoidance: (1) fear is initially conditioned to the warning signal and dopamine computes this fear association as a decrease in release, (2) the warning signal, now capable of producing a negative emotional state, suppresses dopamine release and behavior, (3) over repeated trials the warning signal becomes associated with safety rather than fear; dopaminergic neurons already compute safety as an increase in release and begin to encode the warning signal as the earliest predictor of safety (4) the warning signal now promotes conditioned avoidance *via* dopaminergic modulation of the brain's incentive-motivational circuitry.

**Keywords: dopamine, voltammetry, conditioned avoidance, nucleus accumbens, fear conditioning**

## INTRODUCTION TO CONDITIONED AVOIDANCE

Conditioned avoidance is an acquired behavioral response that results in the prevention of an aversive event. Conditioned avoidance was first described by one of Ivan Pavlov's chief scientific rivals, Bechterev (1913) before being introduced to American psychology by Watson (1916). Ironically, Watson adopted Bekhterev's experimental approach of investigating "associated" motoric avoidance responses in an attempt to validate Pavlov's work on classical conditioning (Bolles, 1972). While it is well known that Pavlov clearly demonstrated that dogs exhibit a strong salivary reflex to stimuli previously associated with food (Pavlov, 2003), Watson found odor-evoked conditional reflexes of the human parotid gland to be elusive (Lashley, 1916; Watson, 1916). Thus, in an attempt to observe a conditioned reflex in human subjects, Watson turned to Bekhterev's experimental design (**Figure 1**), in which: electrodes capable of delivery faradaic stimulation are placed under the palm and finger of a human subject, the hand is exposed to a mild electrical shock that is preceded by a bell (2 s prior to shock), finger movement eliminates electric shock by breaking the circuit between the two electrodes, motoric finger responses are measured by a lever that supports a writing lever (Bechterev, 1913; Watson, 1916). Under these conditions, finger withdrawal initially occurred in response to the electric shock, but within a few trials finger withdrawal began to occur to the bell—thereby leading to the complete avoidance of electric shock (Watson, 1916). This conditioned behavioral response to a shock-predictive cue proved to be highly replicable across subjects, ages and species (Watson, 1916). Although Watson interpreted the aforementioned response as a conditioned reflex, today we recognize this behavioral action as a conditioned avoidance response that is energized by the incentive-motivational circuitry of the brain. One of the major theories involved in integrating motivational theory with conditioned avoidance is the two-process theory of avoidance (Miller, 1948; Mowrer and Aiken, 1954). In general, this theory holds that conditioned fear responses resulting from Pavlovian learning motivate avoidance behavior through fear reduction. The first factor of this theory describes the Pavlovian associations that are established between the aversive stimulus (shock) and the preceding cue (the bell in Watson's experiment). The second factor of this theory states that the fear evoked by the preceding cue functions to reinforce the avoidance response. Over the course

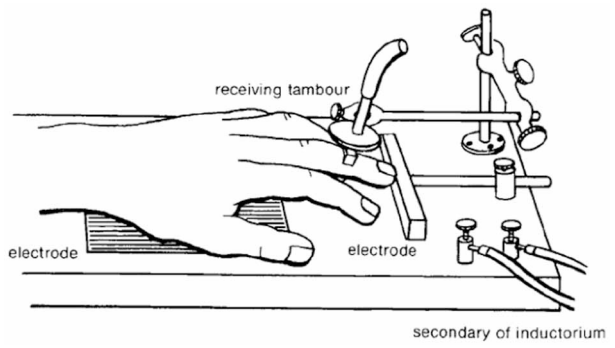



**FIGURE 1 | Illustration of the first conditioned avoidance (initially described as a conditioned reflex) experiment conducted in America by John B Watson.** Electrodes were placed under the hand and finger of a human subject. An auditory stimulus was presented prior to the delivery of electrical shock. A recording device allowed for the detection of finger movements evoked by the shock and the preceding auditory stimulus. Within a few trials, finger withdrawal began to occur to the auditory stimulus. The conditioned finger withdrawal broke the circuit between the two electrodes, which was necessary for the delivery of electric shock. Originally published in Watson (1916).

of the century, investigators developed various methodological adaptations to study conditioned avoidance using experimental animals, most prominently shuttle boxes (Warner, 1932) and operant chambers (Skinner, 1938). It is not the intention of this review to focus on theoretical intricacies of avoidance learning. Instead, we would like to refer the reader to recent reviews focusing on the learning mechanisms that might contribute to the development of avoidance behavior (Depue and Collins, 1999; Moutoussis et al., 2008; Maia, 2010). The present review will focus primarily on the role of the mesolimbic dopamine system during behavior maintained in a signaled operant avoidance procedure. In particular, we will describe how subsecond dopamine release relates to discrete cues during conditioned avoidance and escape responses. Here, it is critical to understand the distinction between avoidance and escape responses. Specifically, an escape response is defined as an action resulting in the cessation of an ongoing aversive stimulus; whereas, an avoidance response is defined as an action preventing the presentation of the aversive stimulus. Two discrete cues will be discussed. A warning signal (a cue light in our case) predicts the potential occurrence of an aversive event; a safety signal (a tone in our case) indicates that the aversive event was successfully avoided or terminated.

## DOPAMINE, INCENTIVE MOTIVATION AND CONDITIONED AVOIDANCE

When experimental psychologists began considering the phenomenon of conditioned avoidance in the middle of the twentieth century, they were relatively unsatisfied with Watson's interpretation that the avoidance response is simply a conditioned reflex resulting from classical conditioning (Bolles, 1972). Alternative explanations began to emerge, many of which described conditioned avoidance as a reinforcement process influenced by the experimental subject's motivation to avoid or terminate the aversive stimulus (e.g., Miller, 1948; Mowrer and Aiken, 1954). The purely psychological view that incentive-motivation (defined as the energizing effects of an encounter with an otherwise neutral stimulus that has acquired motivational importance through prior association, Wise, 2004) might influence the maintenance of conditioned avoidance is supported by modern neuroscientific research.

Before we discuss a role for subsecond dopamine release in conditioned avoidance, it is important to first briefly overview the neural circuitry involved in centrally representing incentive salience. One of the most studied components of the motivational circuitry of the brain is the nucleus accumbens. This brain region has been referred to as a limbic-motor (Mogenson et al., 1980) and Pavlovian-instrumental (Cardinal et al., 2002) interface—both of which appropriately represent the importance of the nucleus accumbens during an avoidance task in which a subject's behavior is effected by their motivational state and conditioned predictors of aversive stimuli. Of note, the nucleus accumbens integrates input from amygdalar and prefrontal cortical regions that carry information regarding the motivational value of stimuli maintaining reinforcement processing before energizing ongoing behavior (Cardinal et al., 2002). The mesolimbic dopamine pathway is theorized to modulate the integration of these motivational circuits by stamping-in stimulus-reinforcement associations, thereby strengthening the incentive value ascribed to previously neutral stimuli (e.g., warning signal) and motivating the conditioned behavioral response (Wise, 2004), or in this case conditioned avoidance.

The mesolimbic dopamine system is a neural pathway that originates from A10 dopamine neurons in the ventral tegmental area of the midbrain and projects to the brain's motivational circuitry, most prominently the nucleus accumbens, amygdala and prefrontal cortices (Swanson, 1982; Spanagel and Weiss, 1999). During ongoing behavior, two distinct patterns of dopamine release occur. Midbrain dopamine neurons typically fire at low frequencies of 1–5 Hz, which is thought to produce a tone on high affinity dopamine D2 receptors in the nucleus accumbens (Grace, 1991; Dreyer et al., 2010). Experimentally, one can detect tonic dopamine levels using techniques like *in vivo* microdialysis, which allow for neurochemical detection on a timescale of minutes. In contrast, when animals are presented with motivationally salient stimuli, A10 dopamine neurons fire in high frequency bursts ($\geq$20 Hz). These high frequency bursts of dopaminergic neural activity produce transient increases in dopamine concentration in terminal fields (e.g., nucleus accumbens). Dopamine concentration transients are detectable at the neurochemical level within terminal fields of the mesolimbic dopamine system using fast-scan cyclic voltammetry, an electrochemical technique that allows for the detection of dopamine on the millisecond timescale. Importantly, only neurochemical techniques like fast-scan cyclic voltammetry provide the temporal resolution necessary to measure dopamine release events evoked by a warning signal in a standard conditioned avoidance procedure.

Pharmacological, lesion, genetic and microdialysis studies have been conducted over the last few decades to demonstrate a general role for dopamine in conditioned avoidance. Animals fail to acquire avoidance following 6-hydroxydopamine lesions

of midbrain dopamine neurons, a deficit that is reversed by the restoration of dopamine levels using L-dopa treatment (Cooper et al., 1973; Zis et al., 1974). Intriguingly, only deficits in avoidance responses are observed, as opposed to responses motivated by the termination of ongoing shock (i.e., escape responses) (Fibiger et al., 1975). Similar observations are reported during the maintenance of conditioned avoidance. Lesions of dopamine terminals in the striatum in general (Amalric and Koob, 1987) and ventral striatum (i.e., nucleus accumbens) in particular (McCullough et al., 1993) are sufficient to impair conditioned avoidance. Systemic administration of dopamine receptor antagonists reliably disrupts avoidance responding without significantly impairing escape behavior (Arnt, 1982). Likewise, locally infusing a dopamine receptor antagonist into the nucleus accumbens alone is sufficient to impair the maintenance of conditioned avoidance (Wadenberg et al., 1990). Using recently developed genetic technology (Darvas et al., 2011) restored dopamine in specific brain regions that were otherwise dopamine-deficient. They found that while the entire striatum and amygdala are necessary for the acquisition of conditioned avoidance, only the striatum is required for the maintenance of conditioned avoidance (Darvas et al., 2011). These findings are in agreement with previous work demonstrating that the amygdala, while important for aversively motivated learning (Ledoux and Muller, 1997; LeDoux, 2003), plays a more specific role in the acquisition rather than the maintenance of instrumental avoidance behavior (Poremba and Gabriel, 1999). In addition to the amygdala and nucleus accumbens, it is important to note that the Gabriel lab has discovered that cingular-thalamic circuitry is also necessary for avoidance learning (Gabriel, 1993). For example, lesions of the anterior cingulate cortex or the limbic thalamus impair acquisition of conditioned avoidance (Gabriel et al., 1989, 1991). Microdialysis studies have demonstrated that dopamine levels are generally increased in the prefrontal cortex and striatum during the acquisition (Dombrowski et al., 2012) and maintenance (McCullough et al., 1993; Feenstra et al., 2001) of conditioned avoidance. Together these studies demonstrated that dopamine plays a general role in the maintenance of conditioned avoidance.

Recently, Kapur (2003), Kapur et al. (2005) generated an incentive-motivation based theory that offers a specific role for dopamine in conditioned avoidance as they attempted to explain why antipsychotics are efficacious in modulating conditioned avoidance. Their theory is based on the observation that all effective antipsychotics antagonize dopamine D2 receptors and disrupt conditioned avoidance. In fact, conditioned avoidance is a classic animal model used to screen for the efficacy of antipsychotic drugs and their dopamine antagonizing properties (Kapur et al., 2005; Smith et al., 2005). This observation led these investigators to speculate that the development of a hyperdopaminergic state in schizophrenia leads to an aberrant assignment of incentive salience to environmental stimuli, thereby promoting psychosis (Kapur, 2003), and the effectiveness of antipsychotics to disrupt conditioned avoidance is due to their ability to block subsecond dopaminergic encoding of the warning signal after it has acquired incentive value (Kapur et al., 2005; Smith et al., 2005). If this theory is correct, discrete dopamine release events time-locked to the warning signal should be detected during the maintenance of conditioned avoidance.

## SUBSECOND DOPAMINE RELEASE DURING WARNING SIGNAL PRESENTATION

To investigate whether subsecond dopamine release is altered by the presentation of a warning signal, we used fast-scan cyclic voltammetry to assess subsecond dopaminergic release events in the nucleus accumbens core during behavior maintained in an operant signaled shock avoidance procedure (**Figure 2**). In this task, a stimulus light was presented as a warning signal for 2 s prior to the delivery of recurring foot shocks. During this 2 s warning signal, a response lever was extended into an operant chamber which, if depressed, resulted in the immediate retraction of the lever and a 20 s safety period signaled by a tone (i.e., safety signal). Animals could initiate an avoidance response by pressing the lever during the 2 s warning signal, entirely preventing shock. Alternatively, once shocks commenced, animals could initiate an escape response by pressing the lever during this punishment period, terminating shock. This experimental design allowed us to assess dopamine signaling during warning signal presentation, safety periods and during two distinct behavioral responses—avoidance and escape. It is important to note that, regardless of the methodology used (i.e., operant or shuttle box), avoidance and escape responses are distinct. This distinction was originally noted in one of the first conditioned avoidance experiments using a shuttle box with a hurdle that separated a shock-free side from a shock side (Bolles, 1972). In this early study, Warner reported that animals would scramble under the hurdle during escape responses, but jump over the hurdle during avoidance responses (Warner, 1932). He further went on to study the unique behavioral responses produced independently by either the shock or the warning signal and found that the shock produced scampering reactions whereas the warning signal produced more calculated, coordinated reactions (Warner, 1932). In the operant signaled shock avoidance task used in our study, we also observed distinct escape and avoidance reactions. Early in training, during which only operant escape responses occur, we observed several unique behavior reactions to the shock: jumping up the wall, attacking the lever and freezing. Interestingly, an unintentional (i.e., not experimenter intended outcome) avoidance response sometimes emerged early in training as well. In certain instances animals attempted to avoid shock by grounding themselves. As in Watson's early finger avoidance study (1916), electrical continuity is only maintained if the rat is in contact between two electrodes or, in our case, two electrified bars comprising the grid floor of the operant chamber. Occasionally, animals balanced their hind paws on a single bar while propping their front paws on a side of the operant chamber, thereby breaking the continuity of the electrical circuit and avoiding footshock. As the contingencies of reinforcement were learned, however, these unintended behaviors begin to dissipate until consistently maintained avoidance and escape behaviors emerged. In our first study on this subject (Oleson et al., 2012), we only recorded dopamine from animals in our operant avoidance task after they began avoiding footshock in ∼50% of trials. At this point in training, we visually observed one of two distinct

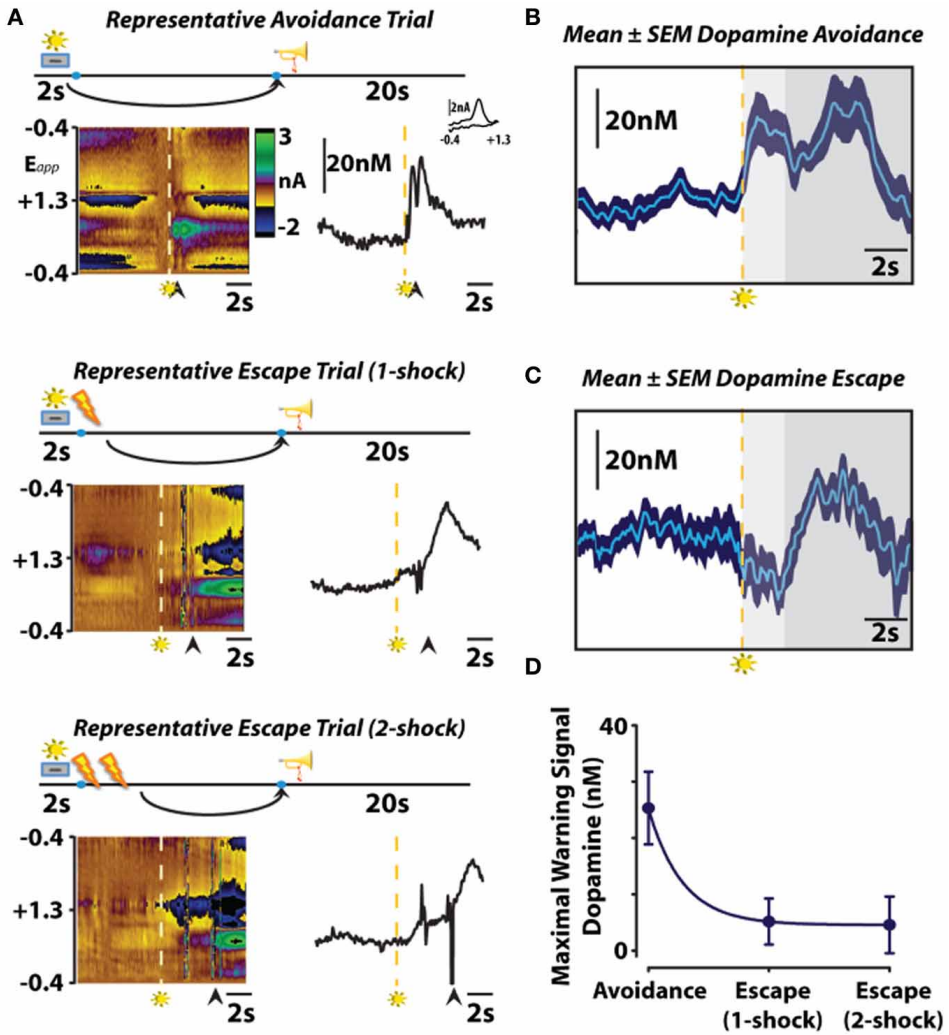



**FIGURE 2 | The role of subsecond dopamine release during conditioned avoidance. (A)** Changes in subsecond dopamine release observed in different response types observed in a single session. Representative color plots (left) and dopamine concentration traces (right) show avoidance (top), one-footshock escape (middle), and two-footshock escape (bottom) responses. Left, the *y*-axis represents the scan potential (E*pp*, V) applied to the electrode, the *x*-axis represents time, and the *z*-axis represents current. Inspection of the color plot allows for the identification of dopamine over time. Dopamine can be identified in the color plot by assessing for changes in current at the oxidation (+0.6V) and reduction (−0.2V) potentials for dopamine. Right, representative dopamine concentration traces plotted as a function of time with the inset showing the cyclic voltammograms for dopamine. Arrows indicate lever responses, lightning bolts indicate footshocks, trumpets indicate safety periods, levers + lights indicate warning signals. **(B,C)** Mean ± SEM dopamine concentration traces from all avoidance and escape responses. Maximal warning signal duration is representative by the light gray fill, subsequent safety periods are represented by the dark gray fill. **(D)** Maximal dopamine concentration evoked by warning signal presentation predicts conditioned avoidance. Originally published in Oleson et al. (2012).

behavioral reactions in response to warning signal presentation. When the animal successfully avoided footshock, an uninhibited motor sequence directed at the lever was observed upon presentation of the warning signal. When the animal escaped footshock, a hesitation—presumably a fear-induced freezing response—was observed upon presentation of the warning signal. While it is well established that amygdalar modulation of prefrontal cortical activity is critically important in the expression of conditioned fear (Davis, 1992; Morgan and LeDoux, 1995; Garcia et al., 1999), dopaminergic modulation of striatal input may be involved in the expression of the freezing response. The canonical view of the basal ganglia holds that the striatum outputs two parallel projections, the direct and indirect pathways, which either excite or inhibit behavioral activity, respectively. According to this canonical view, dopamine release events are theorized to promote behavioral activation by increasing activity along the direct pathway by acting on $G_s$ coupled dopamine D1 receptors, whereas decreases in dopamine release may inhibit behavioral activation by increasing activity along the indirect pathway by acting on $G_{i/o}$ coupled dopamine D2 receptors (DeLong and

Wichmann, 2007). A recent optogenetic study supported this conceptualization by demonstrating that selective activation of striatal dopamine D1 receptor expressing neurons of the direct pathway promotes behavioral activation, while selective activation of striatal dopamine D2 receptor expressing neurons of the indirect pathway promotes freezing behavior (Kravitz et al., 2010). Thus, it is possible that dopamine may contribute to the expression of a freezing response, although additional optogentic studies should be conducted to directly assess for this possibility within the context of conditioned fear. It is also important to note that, rather than solely causing avoidance or freezing responses by activating dopamine D1 or D2 receptors, dopamine concentration changes within the striatum are thought to modulate converging amygdalar, hippocampal and prefrontal input (Floresco et al., 2001; Brady and O'Donnell, 2004) to control behavioral activation.

As animals displayed either directed avoidance or inhibited freezing responses to warning signal presentation, it might be expected, therefore that distinct dopaminergic responses accompany these divergent behavioral reactions. In accordance with our behavioral observation, dichotomous dopaminergic responses occurred at the warning signal during avoidance and escape behavior. When animals successfully avoided footshock, dopamine release increased during warning signal presentation as would be predicted if dopamine was motivating the avoidance response. Importantly, the warning signal evoked increase in dopamine concentration reliably predicted when an animal would successfully avoid foot shock. Trial-by-trial analysis revealed that the maximal dopamine concentration time-locked to warning signal presentation sharply decreased during trials in which animals failed to avoid and was significantly lower during escape responses irrespective of the number of footshocks received. Averaging dopamine concentrations during escape trials revealed that dopamine levels not only failed to increase during presentation of the warning signal presentation, dopamine release events actually ceased at warning signal onset when the animals failed to avoid. This latter finding is somewhat reminiscent of the previously described classical psychological theory called the two-process theory of avoidance (Mowrer, 1951). The first factor of this theory posits that fear becomes conditioned to the warning signal; the second factor suggests that the conditioned fear that is evoked by the warning signal is what reinforces the instrumental avoidance response *via* fear reduction. To further test whether our dopamine data align with the first-factor of this theory, we measured whether dopamine release in the nucleus accumbens core is also suppressed during classical fear associations by employing a standard fear-conditioning model. In this fear-conditioned model, animals were conditioned to an auditory stimulus predicting inescapable footshock before we measured dopamine release 24 h later during repeated presentations of the cue alone (**Figure 3**). As was observed at the warning-signal during escape responses, the fear-associated auditory stimulus produced a decrease in dopamine concentration transients (Oleson et al., 2012), a phenomenon that appears to be exclusive to the core, as opposed to the shell, subregion of the nucleus accumbens (Badrinarayan et al., 2012). This finding supports the first factor of the two-process theory of avoidance that the warning signal can evoke conditioned fear responses, and reveals that dopamine neurons compute this conditioned fear response as a decrease in the frequency of dopamine release events. These data fail to align with the second factor of two-process theory, however, as dopamine

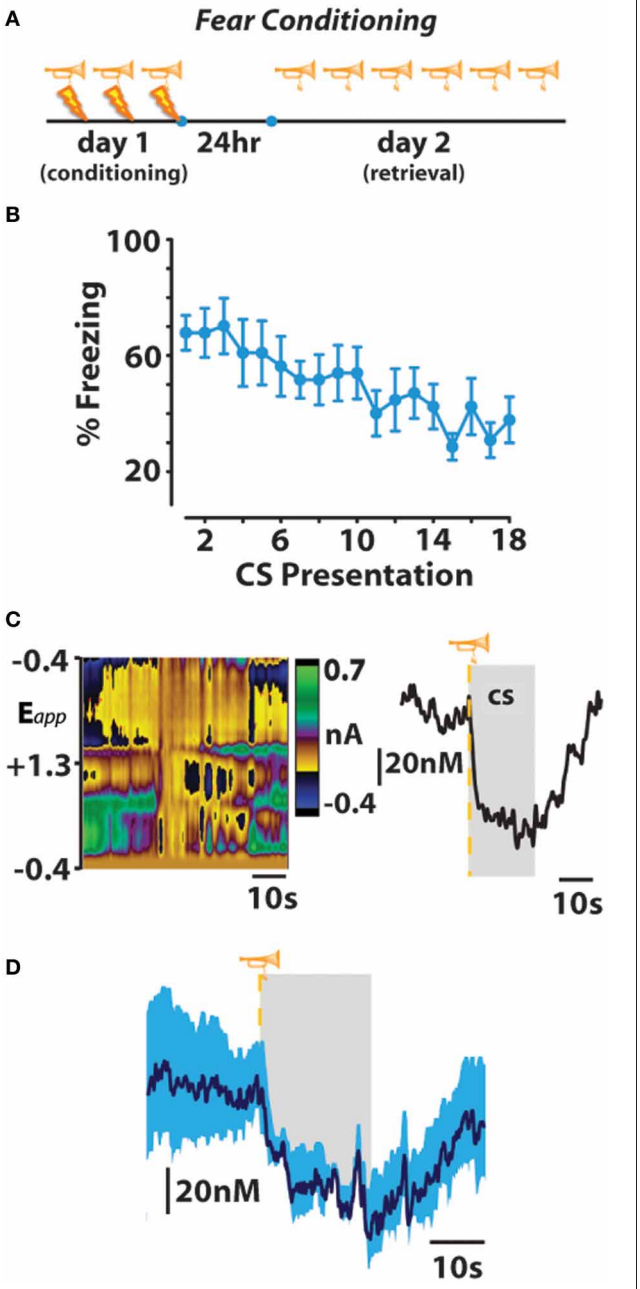



**FIGURE 3 | Fear-conditioned stimuli freeze behavior and subsecond dopamine release events. (A,B)** An otherwise neutral stimulus (trumpet) previously conditioned to inescapable footshock (lightning bolt) produces freezing behavior that extinguishes across repeated trials of conditioned stimulus (CS) presentation on fear-memory retrieval day. **(C)** Representative color plot (left) and corresponding dopamine concentration trace (right) show a CS-induced decrease in dopamine release. Gray represents CS duration. **(D)** Mean ± SEM dopamine concentration trace during presentations of the fear-conditioned CS. Originally published in Oleson et al. (2012).

release accompanies the presentation of the warning signal when animals successfully avoid foot shock. Rather, fear may become irrelevant during conditioned avoidance in a well-trained animal. The warning signal no longer evokes fear, and fear reduction is no longer the primary motivator of behavior. Instead of evoking a fear response, the warning signal becomes associated exclusively with a positive outcome—avoidance. At this point the warning signal motivates behavior similarly to a reward-predictive cue, by stimulating the incentive-motivational circuitry of the brain.

The observation that dopamine begins to increase to the warning signal during avoidance trials suggests that the fear response originally elicited by the warning signal can dissipate over time, as the prediction of a positive outcome (i.e., successful avoidance) becomes more prominent. These findings support other recent work demonstrating that the representation of a conditioned cue can switch between appetitive and aversive stimuli over repeated pairings (Nasser and McNally, 2012) and was predicted by early experimental psychologists. In fact, it has long been reported that animals become less fearful during conditioned avoidance. In one of Richard Solomon's early experiments studying the extinction of the avoidance response, he noted that the animals "learn to relax" in the presence of the warning signal (Solomon et al., 1953). The possibility that the fear response evoked by the warning signal begins to dissipate over time was objectively tested in a subsequent study (Kamin et al., 1963), in which: rats were trained to respond for food in an operant chamber, then trained to avoid shock by responding to an auditory warning signal in a shuttle-box for either 1, 3, 9, or 27 trials, then retested in the operant chamber while periodic presentations of the warning signal occurred during food maintained responding. It was found that the warning signal was less effective at suppressing food maintained responding after 27 trials of conditioned avoidance in comparison to animals with less extensive behavioral histories (Kamin et al., 1963). Importantly, in each of these examples, the fear response evoked by the conditioned stimulus begins to dissipate while the avoidance response remains strong—so strong it is incredibly difficult to extinguish (Solomon et al., 1953). Thus, fear is unlikely to motivate effective avoidance responses in the well-trained rat. Instead, we propose that the strength of the avoidance response is bolstered by increases in dopamine release evoked by the warning signal through higher order reinforcement processes, and these warning signal evoked dopamine release events are capable of motivating avoidance behavior by modulating the incentive-motivational circuitry of the brain. It is also possible that these warning signal evoked dopamine release events might contribute to stimulus-response, or habit, learning. Habit learning reflects the formation of higher order stimulus-response associations (e.g., warning signal-avoidance) that are capable of reinforcing behavioral action but do not become encoded as a goal themselves; thus, rendering the behavior resistant to extinction despite primary reinforcer devaluation (Everitt and Robbins, 2005). Under these circumstances, dopaminergic encoding of the warning signal likely remains critical for the maintenance of conditioned avoidance, although a hierarchical shift of warning-signal evoked dopamine release toward brain regions more implicated in habitual behavior (e.g., dorsal striatum) may contribute (Willuhn et al., 2012). However, a transition to habit formation in this particular behavior may be critically linked to the animal's training history and may also be influenced by individual differences. For example, an animal whose responding is completely dominated by avoidance behavior may always respond to the warning signal even if the shock is removed. On the other hand, an animal that primarily shows escape behavior will extinguish responding when the shock is discontinued because the unconditioned stimulus is the primary driver of the action.

## SUBSECOND DOPAMINE RELEASE DURING SAFETY SIGNAL PRESENTATION

As occurs following the presentation of rewarding stimuli (Schultz et al., 1997), we observed an increase in dopamine release during the safety signal that was indistinguishable between avoidance and escape responses (**Figure 2**). Thus, the elimination of aversive stimuli is processed by dopamine neurons similarly to the receipt of reward, regardless of the representation of the preceding warning signal or whether or not foot shock actually occurred. These data are in agreement with recently published work showing that the relief of pain increases dopamine release in the nucleus accumbens (Navratilova et al., 2012), and further support the notion that avoidance or removal of negative stimuli produces negative reinforcement via mesolimbic dopamine release. This finding supports the theory that the safety signal acquires positive reinforcement value that is capable of promoting avoidance behavior by functioning as a positive conditioned reinforcer (Dinsmoor, 1954, 2001). Several previous studies directly assessed the positive reinforcing effectiveness of the safety signal. Early reports demonstrated that a tone, previously associated with a safety period, is capable of increasing rates of responding to a frequency required to produce the tone alone (Weisman and Litner, 1969). Dinsmoor and colleagues extended upon this finding by demonstrating that presentation of a conditioned safety signal increased rates of responding in a shock avoidance task in which the reinforcing operandum remained available between aversive events (Dinsmoor and Sears, 1973). Rescorla (1969) further proved the reinforcing strength the safety signal holds over avoidance behavior by showing that animals choose a shock-terminating operandum that produces a safety signal over one that simply stops shock. Together, these studies suggest that the safety signal acquires positive reinforcing value capable of promoting avoidance, and dopamine release encodes safety as an increase in release. However, it should also be noted that the warning signal and its dopaminergic correlate is a stronger determinant of the behavioral action than the safety signal and its dopaminergic correlate. That is, only the warning signal evoked dopamine concentration predicts an animal's behavioral response, as dopamine increased during the safety signal regardless of whether safety was reached by escape or avoidance of footshock.

## TONIC vs. PHASIC DOPAMINE

All neurochemical data introduced within the subsequent two sections describe subsecond dopamine release events resulting from the phasic activation of A10 dopamine neurons.

It is important to note that these phasic dopamine data are distinct from previous accounts of tonic dopamine release obtained using microdialysis. For example, microdialysis studies report that tonic brain dopamine levels are generally increased during both conditioned avoidance (McCullough et al., 1993; Feenstra et al., 2001) and fear conditioning (Young et al., 1993; Wilkinson et al., 1998). As previously suggested (McGinty et al., 2011; Oleson et al., 2012), we believe these seemingly contradictory results can be explained by the possibility that aversive stimuli selectively suppress phasic dopamine release while concurrently enhancing tonic dopamine release. In this sense, tonic patterns of dopamine release may serve as an opponent-process (Solomon and Corbit, 1974) to phasic dopamine release evoked by aversive stimuli. It has also been suggested that phasic and tonic dopaminergic encoding of aversive stimuli might vary between subregions of the nucleus accumbens (Badrinarayan et al., 2012). Advances in microdialysis technology offering greater temporal and spatial resolution (Perry et al., 2009) will allow for the clarification of whether these relationships between phasic and tonic dopamine release exist.

## SYNTHESIZING OUR NEUROCHEMICAL OBSERVATIONS WITH THE HISTORICAL PSYCHOLOGICAL LITERATURE LED US TO FORMULATE THE FOLLOWING 4-FACTOR DOPAMINERGIC THEORY OF SIGNALED OPERANT AVOIDANCE

(1) As in the original two-process theory of avoidance, fear is initially conditioned to the warning signal and dopamine computes this fear association as a decrease in release.
(2) The conditioned fear evoked by the warning signal elicits a freezing response, which actually inhibits operant avoidance.
(3) Over repeated trials the warning signal becomes associated with safety rather than fear. Dopaminergic neurons already compute safety as an increase in release. Similarly to the temporal difference model of reinforcement learning (Schultz et al., 1997), dopamine release begins to encode the warning signal as the earliest predictor of safety through a positive prediction error, as the animal's expectation of a negative outcome (being shocked) is violated when avoidance takes place.
(4) The warning signal now promotes conditioned avoidance *via* dopaminergic modulation of brain's incentive-motivational circuitry.

This new model, inspired by recent neurochemical findings, is based upon our conceptualization of the associative structure of the avoidance memory. Specifically, we speculate that early in training the safety signal is associated with the alleviation or avoidance of shock and the warning signal is associated with fear (as in the two-process theory of avoidance); later in training, the safety signal remains associated with the alleviation/avoidance of shock while the warning signal becomes associated with the successful avoidance of foot shock through a reinforcement learning mechanism. The role that temporal difference reinforcement learning may play in transition of cue-evoked dopamine from the safety signal to the warning signal during conditioned avoidance has been previously discussed in detail (Hollon et al., 2013). Briefly, temporal difference reinforcement learning is driven by the error between temporally successive predictions (Sutton, 1988) and midbrain dopamine neurons acquire reward-predicting responses to conditioned cues (Schultz et al., 1997). As detailed by Hollon et al. (2013), our data suggest that midbrain dopamine neurons can acquire predictive responses to negative reinforcers (e.g., warning signal predicts safety) and this learning mechanism might contribute to the development of conditioned avoidance. A longitudinal study assessing for changes in dopamine release to the warning and safety signals over training, would provide additional support for the role of temporal difference reinforcement learning in the acquisition of conditioned avoidance and offer clarification regarding the nature of the safety signal. As it stands, it is possible that the safety signal is more akin to a confirmation of shock avoidance/termination rather than a true signal of safety. Dopaminergic models of temporal difference reinforcement learning predict that dopamine neurons would stop encoding the safety signal as they begin to encode the warning signal. If the safety signal were a confirmatory signal, dopaminergic encoding of the safety signal should persist irrespective of training history. It is also important to note that we do not believe that such computational learning theories are at odds with psychological theories involving the role of dopamine in motivation. On the contrary, as previously described in detail (McClure et al., 2003) many commonalities between the reinforcement learning and motivation literatures exist.

Our conditioned avoidance model predicts that the warning signal is ultimately more important than the safety signal in promoting successful avoidance, as only the warning signal evoked-dopamine response predicts the behavioral outcome (i.e., avoidance vs. escape). It should be noted that this model is only intended to apply to operant signaled shock avoidance tasks. We still believe the mesolimbic dopamine system may function in Sidman operant avoidance tasks, where operant avoidance is maintained without an exteroceptive warning signal (Sidman, 1953), as an anticipatory timing signal (Bromberg-Martin et al., 2010)—although additional experiments are required to test this hypothesis. Also, certain factors of our theory (e.g., factor 2) might be more difficult to detect using a shuttle box because a directed instrumental response is not required for avoidance. Finally, we would like to add that the fourth factor of our model that the warning-signal evoked dopamine release actually promotes successful avoidance, is currently being experimentally assessed using optogenetic technology. These studies will directly test whether the role of dopamine in conditioned avoidance is causal or merely an epiphenomenon, and further discern if the role of dopamine in conditioned avoidance is related to reinforcement learning, motivational processes or, as we predict, both.

## ACKNOWLEDGMENTS

We would like to thank Dr. Lindsey Hamilton for helpful comments in the preparation of this manuscript.

## REFERENCES


Amalric, M., and Koob, G. F. (1987). Depletion of dopamine in the caudate nucleus but not in nucleus accumbens impairs reaction-time performance in rats. *J. Neurosci.* 7, 2129–2134.

Arnt, J. (1982). Pharmacological specificity of conditioned avoidance response inhibition in rats: inhibition by neuroleptics and correlation to dopamine receptor blockade. *Acta Pharmacol. Toxicol.* 51, 321–329.

Badrinarayan, A., Wescott, S. A., Vander Weele, C. M., Saunders, B. T., Couturier, B. E., Maren, S., et al. (2012). Aversive stimuli differentially modulate real-time dopamine transmission dynamics within the nucleus accumbens core and shell. *J. Neurosci.* 32, 15779–15790. doi: 10.1523/JNEUROSCI.3557-12.2012

Bechterev, V. M. (1913). *The Psychologie Objective*. Paris: Alcan.

Bolles, R. C. (1972). The avoidance learning problem. *Psychol. Learn. Motiv. Adv. Res. Theory* 6, 97.

Brady, A. M., and O'Donnell, P. (2004). Dopaminergic modulation of prefrontal cortical input to nucleus accumbens neurons *in vivo*. *J. Neurosci.* 24, 1040–1049. doi: 10.1523/JNEUROSCI.4178-03.2004

Bromberg-Martin, E. S., Matsumoto, M., and Hikosaka, O. (2010). Distinct tonic and phasic anticipatory activity in lateral habenula and dopamine neurons. *Neuron* 67, 144–155. doi: 10.1016/j.neuron.2010.06.016

Cardinal, R. N., Parkinson, J. A., Hall, J., and Everitt, B. J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci. Biobehav. Rev.* 26, 321–352. doi: 10.1016/S0149-7634(02)00007-6

Cooper, B. R., Breese, G. R., Grant, L. D., and Howard, J. L. (1973). Effects of 6-hydroxydopamine treatments on active avoidance responding: evidence for involvement of brain dopamine. *J. Pharmacol. Exp. Therapeut.* 185, 358–370.

Darvas, M., Fadok, J. P., and Palmiter, R. D. (2011). Requirement of dopamine signaling in the amygdala and striatum for learning and maintenance of a conditioned avoidance response. *Learn. Mem.* 18, 136–143. doi: 10.1101/lm.2041211

Davis, M. (1992). "The role of the amygdala in conditioned fear," in The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction, ed J. P. Aggleton (New York, NY: Wiley-Liss), 255–306.

DeLong, M. R., and Wichmann, T. (2007). Circuits and circuit disorders of the basal ganglia. *Arch. Neurol.* 64, 20. doi: 10.1001/archneur.64.1.20

Depue, R. A., and Collins, P. F. (1999). Neurobiology of the structure of personality: dopamine, facilitation of incentive motivation, and extraversion. *Behav. Brain Sci.* 22, 491–517. doi: 10.1017/S0140525X99002046

Dinsmoor, J. A. (1954). Punishment: I. The avoidance hypothesis. *Psychol. Rev.* 61, 34. doi: 10.1037/h0062725

Dinsmoor, J. A. (2001). Stimuli inevitably generated by behavior that avoids electric shock are inherently reinforcing. *J. Exp. Anal. Behav.* 75, 311. doi: 10.1901/jeab.2001.75-311

Dinsmoor, J. A., and Sears, G. W. (1973). Control of avoidance by a response-produced stimulus. *Learn. Motiv.* 4, 284–293. doi: 10.1016/0023-9690(73)90018-0

Dombrowski, P. A., Maia, T. V., Boschen, S. L., Bortolanza, M., Wendler, E., Schwarting, R. K., et al. (2012). Evidence that conditioned avoidance responses are reinforced by positive prediction errors signaled by tonic striatal dopamine. *Behav. Brain Res.* 241, 112–119. doi: 10.1016/j.bbr.2012.06.031

Dreyer, J. K., Herrik, K. F., Berg, R. W., and Hounsgaard, J. D. (2010). Influence of phasic and tonic dopamine release on receptor activation. *J. Neurosci.* 30, 14273–14283. doi: 10.1523/JNEUROSCI.1894-10.2010

Everitt, B. J., and Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* 8, 1481–1489. doi: 10.1038/nn1579

Feenstra, M. G., Vogel, M., Botterblom, M. H., Joosten, R. N., de Bruin, J. P. (2001). Dopamine and noradrenaline efflux in the rat prefrontal cortex after classical aversive conditioning to an auditory cue. *Eur. J. Neurosci.* 13, 1051–1054. doi: 10.1046/j.0953-816x.2001.01471.x

Fibiger, H. C., Zis, A. P., and Phillips, A. G. (1975). Haloperidol-induced disruption of conditioned avoidance responding: attenuation by prior training or by anticholinergic drugs. *Eur. J. Pharmacol.* 30, 309–314. doi: 10.1016/0014-2999(75)90114-4

Floresco, S. B., Blaha, C. D., Yang, C. R., and Phillips, A. G. (2001). Modulation of hippocampal and amygdalar-evoked activity of nucleus accumbens neurons by dopamine: cellular mechanisms of input selection. *J. Neurosci.* 21, 2851–2860.

Gabriel, M. (1993). "Discriminative avoidance learning: a model system," in Neurobiology of Cingulate Cortex and Limbic Thalamus: A Comprehensive Handbook, eds B. A. Vogt and G. Michael (Cambridge, MA: Birkhäuser), 478–523.

Gabriel, M., Kubota, Y., Sparenborg, S., Straube, K., and Vogt, B. A. (1991). Effects of cingulate cortical lesions on avoidance learning and training-induced unit activity in rabbits. *Exp. Brain Res.* 86, 585–600. doi: 10.1007/BF00230532

Gabriel, M., Sparenborg, S., and Kubota, Y. (1989). Anterior and medial thalamic lesions, discriminative avoidance learning, and cingulate cortical neuronal activity in rabbits. *Exp. Brain Res.* 76, 441–457. doi: 10.1007/BF00247901

Garcia, R., Vouimba, R. M., Baudry, M., and Thompson, R. F. (1999). The amygdala modulates prefrontal cortex activity relative to conditioned fear. *Nature* 402, 294–296. doi: 10.1038/46286

Grace, A. A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience* 41, 1–24. doi: 10.1016/0306-4522(91)90196-U

Hollon, N. G., Soden, M. E., and Wanat, M. J. (2013). Dopaminergic prediction errors persevere in the nucleus accumbens core during negative reinforcement. *J. Neurosci.* 33, 3253–3255. doi: 10.1523/JNEUROSCI.5762-12.2013

Kamin, L. J., Brimer, C. J., and Black, A. H. (1963). Conditioned suppression as a monitor of fear of the CS in the course of avoidance training. *J. Comp. Physiol. Psychol.* 56, 497. doi: 10.1037/h0047966

Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am. J. Psychiatry* 160, 13–23. doi: 10.1176/appi.ajp.160.1.13

Kapur, S., Mizrahi, R., and Li, M. (2005). From dopamine to salience to psychosis—linking biology, pharmacology and phenomenology of psychosis. *Schizophr. Res.* 79, 59–68. doi: 10.1016/j.schres.2005.01.003

Kravitz, A. V., Freeze, B. S., Parker, P. R., Kay, K., Thwin, M. T., Deisseroth, K., et al. (2010). Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature* 466, 622–626. doi: 10.1038/nature09159

Lashley, K. S. (1916). Reflex secretion of the human parotid gland. *J. Exp. Psychol.* 1, 461. doi: 10.1037/h0073282

LeDoux, J. (2003). The emotional brain, fear, and the amygdala. *Cell. Mol. Neurobiol.* 23, 727–738.

Ledoux, J. E., and Muller, J. (1997). Emotional memory and psychopathology. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 352, 1719–1726. doi: 10.1098/rstb.1997.0154

Maia, T. V. (2010). Two-factor theory, the actor-critic model, and conditioned avoidance. *Learn. Behav.* 38, 50–67. doi: 10.3758/LB.38.1.50

McClure, S. M., Daw, N. D., and Read Montague, P. (2003). A computational substrate for incentive salience. *Trends Neurosci.* 26, 423–428. doi: 10.1016/S0166-2236(03)00177-2

McCullough, L. D., Sokolowski, J. D., and Salamone, J. D. (1993). A neurochemical and behavioral investigation of the involvement of nucleus accumbens dopamine in instrumental avoidance. *Neuroscience* 52, 919–925. doi: 10.1016/0306-4522(93)90538-Q

McGinty, V. B., Hayden, B. Y., Heilbronner, S. R., Dumont, E. C., Graves, S. M., Mirrione, M. M., et al. (2011). Emerging, reemerging, and forgotten brain areas of the reward circuit: notes from the 2010 motivational neural networks conference. *Behav. Brain Res.* 225, 348–357. doi: 10.1016/j.bbr.2011.07.036

Miller, N. E. (1948). Studies of fear as an acquirable drive: I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *J. Exp. Psychol.* 38, 89.

Mogenson, G. J., Jones, D. L., and Yim, C. Y. (1980). From motivation to action: functional interface between the limbic system and the motor system. *Prog. Neurobiol.* 14, 69–97. doi: 10.1016/0301-0082(80)90018-0

Morgan, M. A., and LeDoux, J. E. (1995). Differential contribution of dorsal and ventral medial prefrontal cortex to the acquisition and extinction of conditioned fear in rats. *Behav. Neurosci.* 109, 681.

Moutoussis, M., Bentall, R. P., Williams, J., and Dayan, P. (2008). A temporal difference account of avoidance learning. *Network* 19, 137–160. doi: 10.1080/09548980802192784

Mowrer, O. H. (1951). Two-factor learning theory: summary and comment. *Psychol. Rev.* 58, 350.

Mowrer, O. H., and Aiken, E. G. (1954). Contiguity vs. drive-reduction in

conditioned fear: temporal variations in conditioned and unconditioned stimulus. *Am. J. Psychol.* 67, 26–38.

Nasser, H. M., and McNally, G. P. (2012). Appetitive–aversive interactions in Pavlovian fear conditioning. *Behav. Neurosci.* 126, 404. doi: 10.1037/a0028341

Navratilova, E., Xie, J. Y., Okun, A., Qu, C., Eyde, N., Ci, S., et al. (2012). Pain relief produces negative reinforcement through activation of mesolimbic reward–valuation circuitry. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20709–20713. doi: 10.1073/pnas.1214605109

Oleson, E. B., Gentry, R. N., Chioma, V. C., and Cheer, J. F. (2012). Subsecond dopamine release in the nucleus accumbens predicts conditioned punishment and its successful avoidance. *J. Neurosci.* 32, 14804–14808. doi: 10.1523/JNEUROSCI.3087-12.2012

Pavlov, I. P. (2003). *Conditioned Reflexes*, Mineola: Dover Publications.

Perry, M., Li, Q., and Kennedy, R. T. (2009). Review of recent advances in analytical techniques for the determination of neurotransmitters. *Anal. Chim. Acta* 653, 1–22. doi: 10.1016/j.aca.2009.08.038

Poremba, A., and Gabriel, M. (1999). Amygdala neurons mediate acquisition but not maintenance of instrumental avoidance behavior in rabbits. *J. Neurosci.* 19, 9635–9641.

Rescorla, R. A. (1969). Establishment of a positive reinforcer through contrast with shock. *J. Comp. Physiol. Psychol.* 67, 260.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593

Sidman, M. (1953). Avoidance conditioning with brief shock and no exteroceptive warning signal. *Science* 118, 157–158. doi: 10.1126/science.118.3058.157

Skinner, B. F. (1938). *The Behavior of Organisms: An Experimental Analysis*, New York, NY: Appleton-Century.

Smith, A. J., Becker, S., and Kapur, S. (2005). A computational model of the functional role of the ventral-striatal D2 receptor in the expression of previously acquired behaviors. *Neural Comput.* 17, 361–395. doi: 10.1162/0899766053011546

Solomon, R. L., and Corbit, J. D. (1974). An opponent-process theory of motivation: I. Temporal dynamics of affect. *Psychol. Rev.* 81, 119–145. doi: 10.1037/h0036128

Solomon, R. L., Kamin, L. J., and Wynne, L. C. (1953). Traumatic avoidance learning: the outcomes of several extinction procedures with dogs. *J. Abnorm. Psychol.* 48, 291. doi: 10.1037/h0058943

Spanagel, R., and Weiss, F. (1999). The dopamine hypothesis of reward: past and current status. *Trends Neurosci.* 22, 521–527. doi: 10.1016/S0166-2236(99)01447-2

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Mach. Learn.* 3, 9–44. doi: 10.1007/BF00115009

Swanson, L. W. (1982). The projections of the ventral tegmental area and adjacent regions: a combined fluorescent retrograde tracer and immunofluorescence study in the rat. *Brain Res. Bull.* 9, 321–353. doi: 10.1016/0361-9230(82)90145-9

Wadenberg, M. L., Ericson, E., Magnusson, O., and Ahlenius, S. (1990). Suppression of conditioned avoidance behavior by the local application of (-) sulpiride into the ventral, but not the dorsal, striatum of the rat. *Biol. Psychiatry* 28, 297–307. doi: 10.1016/0006-3223(90)90657-N

Warner, L. H. (1932). The association span of the white rat. *J. Genet. Psychol.* 41, 57–90. doi: 10.1080/08856559.1932.9944143

Watson, J. B. (1916). The place of the conditioned-reflex in psychology. *Psychol. Rev.* 23, 89. doi: 10.1037/h0070003

Weisman, R. G., and Litner, J. S. (1969). Positive conditioned reinforcement of Sidman avoidance behavior in rats. *J. Comp. Physiol. Psychol.* 68, 597.

Wilkinson, L. S., Humby, T., Killcross, A. S., Torres, E. M., Everitt, B. J., and Robbins, T. W. (1998). Dissociations in dopamine release in medial prefrontal cortex and ventral striatum during the acquisition and extinction of classical aversive conditioning in the rat. *Eur. J. Neurosci.* 10, 1019–1026. doi: 10.1046/j.1460-9568.1998.00119.x

Willuhn, I., Burgeno, L. M., Everitt, B. J., and Phillips, P. E. (2012). Hierarchical recruitment of phasic dopamine signaling in the striatum during the progression of cocaine use. *Proc. Natl. Acad. Sci.U.S.A.* 109, 20703–20708. doi: 10.1073/pnas.1213460109

Wise, R. A. (2004). Dopamine, learning and motivation. *Nat. Rev. Neurosci.* 5, 483–494. doi: 10.1038/nrn1406

Young, A. M., Joseph, M. H., and Gray, J. A. (1993). Latent inhibition of conditioned dopamine release in rat nucleus accumbens. *Neuroscience* 54, 5–9. doi: 10.1016/0306-4522(93)90378-S

Zis, A. P., Fibiger, H. C., and Phillips, A. G. (1974). Reversal by L-dopa of impaired learning due to destruction of the dopaminergic nigro-neostriatal projection. *Science* 185, 960–962. doi: 10.1126/science.185.4155.960

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 28 March 2013; accepted: 20 May 2013; published online: 07 June 2013.*

*Citation: Oleson EB and Cheer JF (2013) On the role of subsecond dopamine release in conditioned avoidance. Front. Neurosci. 7:96. doi: 10.3389/fnins.2013.00096*

*This article was submitted to Frontiers in Decision Neuroscience, a specialty of Frontiers in Neuroscience.*

frontiers in
NEUROSCIENCE

REVIEW ARTICLE
published: 04 July 2013
doi: 10.3389/fnins.2013.00114

# Dopamine imbalance in Huntington's disease: a mechanism for the lack of behavioral flexibility

***Jane Y. Chen, Elizabeth A. Wang, Carlos Cepeda and Michael S. Levine****

*Intellectual and Developmental Disabilities Research Center, Semel Institute for Neuroscience and Human Behavior and the Brain Research Institute, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA*

***Edited by:***
*Giselle Petzinger, University of Southern California, USA*

***Reviewed by:***
*Krishna P. Miyapuram, University of Trento, Italy*
*John P. Walsh, University of Southern California, USA*

****Correspondence:***
*Michael S. Levine, Intellectual and Developmental Disabilities Research Center, Semel Institute for Neuroscience and Human Behavior, School of Medicine, University of California Los Angeles, Room 58-258, 760 Westwood Plaza, Los Angeles, CA 90095, USA*
*e-mail: mlevine@mednet.ucla.edu*

Dopamine (DA) plays an essential role in the control of coordinated movements. Alterations in DA balance in the striatum lead to pathological conditions such as Parkinson's and Huntington's diseases (HD). HD is a progressive, invariably fatal neurodegenerative disease caused by a genetic mutation producing an expansion of glutamine repeats and is characterized by abnormal dance-like movements (chorea). The principal pathology is the loss of striatal and cortical projection neurons. Changes in brain DA content and receptor number contribute to abnormal movements and cognitive deficits in HD. In particular, during the early hyperkinetic stage of HD, DA levels are increased whereas expression of DA receptors is reduced. In contrast, in the late akinetic stage, DA levels are significantly decreased and resemble those of a Parkinsonian state. Time-dependent changes in DA transmission parallel biphasic changes in glutamate synaptic transmission and may enhance alterations in glutamate receptor-mediated synaptic activity. In this review, we focus on neuronal electrophysiological mechanisms that may lead to some of the motor and cognitive symptoms of HD and how they relate to dysfunction in DA neurotransmission. Based on clinical and experimental findings, we propose that some of the behavioral alterations in HD, including reduced behavioral flexibility, may be caused by altered DA modulatory function. Thus, restoring DA balance alone or in conjunction with glutamate receptor antagonists could be a viable therapeutic approach.

**Keywords: Huntington's disease, behavioral inflexibility, dopamine, glutamate, electrophysiology**

## INTRODUCTION

Huntington's disease (HD) is an inherited, autosomal dominant, and progressive neurodegenerative disorder caused by a mutation in the huntingtin gene (*HTT*) resulting in an abnormally long polyglutamine (CAG >40) repeat (The Huntington's Disease Collaborative Research Group, 1993). It is characterized by involuntary dance-like movements (chorea) in the early stages, then akinesia and dystonia in the late stages. Other symptoms include psychiatric alterations and cognitive deterioration (Bonelli and Hofmann, 2007). Cognitive disturbances affecting learning, memory processes, as well as attention and executive function emerge early in the course of the disease and become prominent in the advanced stages (Brandt and Butters, 1986; Peinemann et al., 2005; Wang et al., 2012). A juvenile form of HD also occurs, generally when the length of CAG repeats is >60. These patients develop epileptic seizures and intellectual decline associated with a more rapidly progressing course of the disease (Andrew et al., 1993; Seneca et al., 2004).

In HD, the most striking neuropathology is massive loss of medium-sized spiny neurons (MSNs) in the striatum (Vonsattel and Difiglia, 1998), as well as laminar thinning and white matter loss in the cerebral cortex (Rosas et al., 2006). Other structures such as the globus pallidus, thalamus, hypothalamus, subthalamic nucleus (STN), and substantia nigra also are affected, particularly in the later stages (Kremer et al., 1990; Heinsen et al., 1996; Petersen et al., 2005). Although the symptomatology of HD is classically attributed to striatal and cortical neuronal loss, studies have demonstrated that neuronal dysfunction precedes cell death (Tobin and Signer, 2000; Levine et al., 2004). For example, psychiatric, cognitive, and motor symptoms can and often appear alongside cellular and synaptic alterations in the absence of neuronal loss (Vonsattel and Difiglia, 1998).

This review examines the role of striatal dopamine (DA) in HD. We focus on neuronal electrophysiological mechanisms that may lead to some of the motor and cognitive symptoms of HD and how they relate to dysfunction in DA neurotransmission. Data from human and animal studies are reviewed with particular emphasis on alterations of the DA system and how they relate to behavioral inflexibility. The central thesis is that the major symptoms of HD can be associated with biphasic changes in DA transmission and its modulatory role on glutamate (GLU) receptor function. Thus, treatments of HD symptoms should take into account and be tailored according to the temporal progression of neurotransmitter and receptor changes. Before elaborating on these changes, we first need to understand the role of the DA system and its interactions in normal neuronal function, particularly in the striatum.

## STRIATAL ORGANIZATION

GABAergic projection MSNs comprise 90–95% of striatal neurons (Kita and Kitai, 1988) and receive glutamatergic inputs primarily from the cortex as well as specific thalamic nuclei (Kemp

and Powell, 1971; Smith et al., 2004). There are two striatal projection pathways (**Figure 1**), each with distinct MSN populations expressing different DA receptors and neuropeptides (Graybiel, 2000). The direct pathway consists of MSNs expressing DA D1 receptors, substance P, and dynorphin (Vincent et al., 1982; Haber and Nauta, 1983; Gerfen et al., 1990). It projects monosynaptically to the substantia nigra pars reticulata and the internal segment of the globus pallidus (Albin et al., 1989; Gerfen et al., 1990). The indirect pathway is composed of MSNs that express D2 receptors, adenosine $A_{2A}$ receptors, and enkephalin (Gerfen et al., 1990; Schiffmann and Vanderhaeghen, 1993; Steiner and Gerfen, 1999), and projects to the external segment of the globus pallidus (Gerfen, 1992; Bolam et al., 2000). The external segment of the globus pallidus, in turn, projects to the STN (Albin et al., 1989). Electrophysiological studies using mice expressing enhanced green fluorescent protein (EGFP) in MSNs enriched with D1 or D2 DA receptors demonstrated that, although direct and indirect pathway neurons display similar basic membrane properties, indirect pathway MSNs are more excitable and thus may be more susceptible to abnormal GLU release or receptor dysfunction (Kreitzer and Malenka, 2007; Cepeda et al., 2008). This is partially due to a difference in dendritic surface area, where indirect pathway MSNs have fewer primary dendrites than direct pathway MSNs, suggesting that the increased excitability of indirect pathway MSNs partially results from a higher membrane input resistance due to their more compact morphology (Gertler et al., 2008; Flores-Barrera et al., 2010). The remaining 5–10% of striatal neurons are interneurons, which are divided into two main groups: GABAergic interneurons, which provide feedforward inhibition to MSNs (Tepper et al., 2008); and cholinergic interneurons, which are responsible for acetylcholine levels in the striatum (Bolam et al., 1984; Zhou et al., 2002).

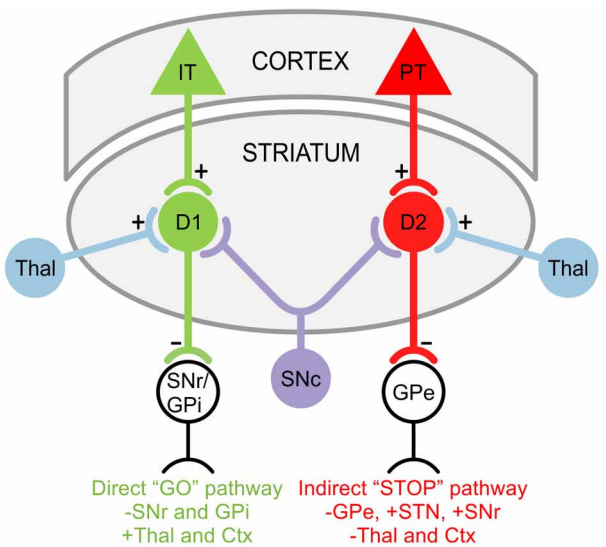



**FIGURE 1 | Striatal projection pathways.** In the direct "GO" pathway, MSNs expressing DA D1 receptors receive inputs from intratelencephalically projecting (IT) neurons in the cortex (Ctx) and project to the substantia nigra pars reticulata (SNr) as well as the internal segment of the globus pallidus (GPi). In the indirect "STOP" pathway, MSNs expressing DA D2 receptors receive inputs from pyramidal tract (PT) neurons in the Ctx and project to the external segment of the globus pallidus (GPe). The GPe, in turn, projects to the STN and SNr. Both D1 and D2 MSNs also receive afferents from the substantia nigra pars compacta (SNc) and thalamus (Thal).

The striatum can also be described as a mosaic of two functionally distinct compartments. The striosome compartment is enriched with μ-opioid receptors while the surrounding extrastriosomal matrix contains neurons that express acetylcholinesterase, somatostatin, and calbindin (Gerfen, 1984). GABAergic striosomal neurons innervate DA neurons in the substantia nigra pars compacta and reticulata, essentially forming a third striatal output pathway (Gerfen, 1984; Jimenez-Castellanos and Graybiel, 1989). Since interactions between striosomes and the extrastriosomal matrix are involved in drug-induced stereotypies (Saka et al., 2004; Canales, 2005), it has been proposed that the striosomal system may change the set point of DA neurons (Canales and Graybiel, 2000). This, in turn, could modulate DA neurotransmission in the basal ganglia and alter the occurrence of stereotypic behaviors (Graybiel, 2000). As discussed later, pathological changes in the striosome compartment could underlie dysregulation of DA release in the early stages of HD.

## MODULATORY ROLE OF DA IN THE BRAIN

The modulatory effects of DA are better understood if considered as a representation of an inverted "U" shaped function. This concept suggests that too much or too little DA perturbs cognitive function (Williams and Castner, 2006; Vijayraghavan et al., 2007). Furthermore, maximum efficiency in behavioral and cognitive performance is a result of maintaining an optimal DA level, where imbalances cause decreased efficiency (Dickinson and Elvevag, 2009). As an extension, we can say that in the dorsal striatum, increases or decreases in DA alter motor behavior.

One of the main functions of DA in the brain is to enhance the signal-to-noise ratio. This can be achieved by at least 3 different mechanisms: (1) DA can modulate neuronal firing in a selective manner. For example, studies in awake rats show that iontophoresis of DA induces excitation of motor-related, and inhibition of non-motor-related neurons (Pierce and Rebec, 1995). Also, the effect of D1 agonists on neuronal firing can be excitatory or inhibitory depending on the membrane potential of the cell. At hyperpolarized potentials, D1 receptor activation is inhibitory, whereas at depolarized potentials, it is excitatory (Hernandez-Lopez et al., 1997). (2) DA affects responses evoked by GLU in a differential manner. Responses evoked by activation of α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionic acid (AMPA) receptors are reduced by DA, whereas responses evoked by activation of N-methyl-D-aspartate (NMDA) receptors are increased by DA (Cepeda et al., 1993; Levine et al., 1996; Cepeda and Levine, 1998). In general, activation of D1 receptors enhances GLU responses whereas activation of D2 receptors decreases these responses (Cepeda et al., 1993). (3) DA also can select excitatory inputs to the striatum (Flores-Hernandez et al., 1997) and thus act as a filter for less active inputs (Bamford et al., 2004). These effects are probably mediated by presynaptic D2 receptors located on corticostriatal GLU terminals (Cepeda et al., 2001). DA modulation of neurotransmitter release also is influenced by

endocannabinoid production and retrograde activation of presynaptic corticostriatal CB1 receptors (Maejima et al., 2001; Patel et al., 2003; Kreitzer and Malenka, 2005).

## DA AND BEHAVIORAL INFLEXIBILITY

Behavioral inflexibility is defined as a failure to shift between behaviors and the inability to adapt behavior to changes in environmental stimuli. Lack of behavioral flexibility depends on the inability to stop ongoing behaviors and is mediated by a discrete cortico-basal ganglia circuit (Aron and Poldrack, 2006; Aron et al., 2007). Although behavioral routines are often stereotyped through learning and result in habit formation, extremely repetitive behaviors (stereotypies) appear to be prominent symptoms in various neuropsychiatric disorders and addiction. These range from impaired behavioral inhibition in attention deficit/hyperactivity disorder and inability to suppress emotions in autism spectrum disorders, to repetitive twitches or vocalizations in Tourette's syndrome, movement fixation in obsessive-compulsive disorder, punding due to over-medication of Parkinson's disease patients, and may include some of the involuntary movements in HD. Despite the wide range of behavioral phenotypes in these disorders, central features of these behaviors are DA-dependent and related to striatal dysfunction (Frank et al., 2004; Beste et al., 2010).

Changes in the DA system have long been implicated in human cognitive inflexibility. However, patients with DA impairments do not show deficits on all tasks that assess cognitive flexibility. Specifically, DA function in the striatum involves set-shifting between object features but is not involved in shifting between abstract rules (Cools et al., 2006; Dang et al., 2012). Patients with disorders of the basal ganglia, such as in Parkinson's Disease or HD, routinely show cognitive inflexibility as demonstrated by impaired performance on the Wisconsin Card Sorting Test and attentional set-shifting tests (Owen et al., 1993; Lawrence et al., 1996).

In the striatum, different subregions are involved in specific behavioral strategies and learning. Rats with lesions of the lateral striatum have deficits in motor skill learning and arbitrary stimulus-response associations (Reading et al., 1991; Devan et al., 1999), whereas those with lesions of the medial striatum have impairments in spatial and reversal learning (Whishaw et al., 1987; Pisa and Cyr, 1990). Furthermore, the medial striatum plays a role in switching between navigational strategies in response to changes in the environment (Mizumori et al., 2000). The dorsomedial striatum also is necessary for maintaining and executing a new strategy. Failure to maintain a proper response pattern by shifting strategies results in behavioral inflexibility (Ragozzino, 2007). Additionally, reversal learning and trait impulsivity in mice is associated with DA receptor density in the midbrain (Dalley et al., 2007; Lee et al., 2009). Taken together, these studies indicate that the striatum and DA neurotransmission play a crucial role in determining behavioral flexibility.

Stereotypies in rodents are an extreme form of behavioral inflexibility that manifest as rigid, repetitive movements. These include excessive grooming, sniffing, rearing, as well as locomotion, and may be more manifest during social isolation and stress (Ridley, 1994). Stereotypies present as behavioral abnormalities with little flexibility and high repetition, often similar to addictive states. Drugs that act on the DA system can produce stereotyped behaviors in a dose-dependent manner. For example, low doses of amphetamine and cocaine induce repetitive locomotion while high doses cause more focal stereotypy, such as sniffing and grooming (Cooper and Dourish, 1990). Striatal cocaine administration also results in impaired reversal learning (Stalnaker et al., 2009), further indicating that aberrant DA transmission results in behavioral inflexibility. The intensity of drug-induced stereotypies is determined by striatal DA, where rats with high extracellular DA levels demonstrate complex stereotypic behavior, including syntactic grooming (Berridge et al., 2005). In fact, robust stereotypies in rats similar to those induced by amphetamine and cocaine can be induced by striatal infusions of D1 and D2 receptor agonists (Waszczak et al., 2002).

It would be misleading, however, to think that only DA alterations are involved in behavioral inflexibility. In fact, the capacity for attentional shifts and inhibition of ongoing motor activity by salient stimuli seems to depend on thalamostriatal inputs onto cholinergic interneurons (Ding et al., 2010). These aspiny interneurons have rich terminal connections and are implicated in stereotypic behavior as well as associative learning (Aosaki et al., 1994, 2010). For example, striatal application of the muscarinic receptor antagonist pirenzepine impairs reversal learning, indicating that these cholinergic receptors play a role in the shifting of response patterns (Tzavos et al., 2004). Thus, cholinergic interneurons may also play an important role in the loss of cognitive and behavioral flexibility in pathological conditions including HD.

## DA ALTERATIONS IN HUNTINGTON'S DISEASE

Alterations in DA function play a significant role in the motor and cognitive symptoms of HD. Here we will discuss changes in DA transmission that may underlie the neuropathological changes in HD. There is evidence from studies in HD patients that increased DA release induces chorea while a reduction in DA leads to akinesia (Bird, 1980; Spokes, 1980), thus giving rise to the biphasic movement symptoms of early and late HD. The idea that aberrant DA signaling underlies behavioral abnormalities was first proposed as a predictive test when asymptomatic offspring of individuals with HD developed dyskinesias in response to levodopa (L-DOPA) administration (Klawans et al., 1970). The hypothesis was that stimulation of DA receptors was involved in the production of dyskinesias as a basic mechanism of chorea. Early studies indicating an involvement of the DA nigrostriatal pathway in HD demonstrated increased levels of DA in postmortem brains of HD patients and showed that DA-depleting agents and DA receptor agonists can be used with therapeutic benefit (Bird, 1980; Spokes, 1980). Later, neurochemical studies of HD patients suggested that increased DA occurs in the early stages of the disease (Garrett and Soares-Da-Silva, 1992) while postmortem studies of late-stage HD patients showed reduced levels of caudate DA and homovanillic acid, the principal DA metabolite (Bernheimer et al., 1973; Kish et al., 1987). Thus, it was thought that DA levels in HD may show biphasic, time-dependent changes, with early increases followed by late decreases (**Table 1**).

**Table 1 | DA in human HD and animal models.**

| | Early stage | Late stage |
|---|---|---|
| **HUMAN HD** | | |
| DA levels in striatum | Increased<br>Garrett and Soares-Da-Silva, 1992 | Decreased<br>Bernheimer et al., 1973; Kish et al., 1987 |
| DA receptor density | Decreased<br>Joyce et al., 1988; Richfield et al., 1991; Van Oostrom et al., 2009 | Decreased<br>Antonini et al., 1996; Weeks et al., 1996 |
| DAT | Not determined | Decreased<br>Backman et al., 1997; Ginovart et al., 1997; Suzuki et al., 2001 |
| **ANIMAL MODELS** | | |
| DA levels | Increased*<br>*tgHD rat model<br>Jahanshahi et al., 2010 | Decreased<br>Hickey et al., 2002; Johnson et al., 2006; Callahan and Abercrombie, 2011 |
| DA receptors | Decreased<br>Cha et al., 1998; Bibb et al., 2000; Ariano et al., 2002; Petersen et al., 2002b | Decreased<br>Pouladi et al., 2012 |
| DAT | Not determined | Not determined |

During the early phase of HD, neuropathological studies have shown that discrete islands of neuronal loss and astrocytosis appear in the striosomes almost exclusively, whereas in the late phase, cell loss increasingly occurs in the matrix compartment (Hedreen and Folstein, 1995). As MSNs from the striosomes project to the substantia nigra pars compacta, it may be that early degeneration of these inhibitory neurons produces hyperactivity of the DA pathway, contributing to chorea and other early clinical manifestations of HD. Studies using positron emission tomography, autoradiography, and markers for pre- and postsynaptic neurons have observed reduced striatal D1 and D2 DA receptor density, even in asymptomatic HD patients, further indicating that DA signaling is disrupted early in HD (Joyce et al., 1988; Richfield et al., 1991; Van Oostrom et al., 2009). These observations have been confirmed by imaging studies, which reported reduced striatal D1 and D2 receptors in both HD patients and asymptomatic HD mutation carriers (Antonini et al., 1996; Weeks et al., 1996). There also is a progressive reduction of D1 and D2 receptor binding in the temporal and frontal cortices (Ginovart et al., 1997; Pavese et al., 2003). Striatal and cortical loss of DA receptors in presymptomatic and early stage HD patients have been correlated with early cognitive decline, which may reflect altered synaptic plasticity and lead to deficits in cognitive processes such as attention, executive function, learning, and memory (Backman and Farde, 2001).

Studies also have examined DA transporter (DAT) density as both an index of DA neurotransmission and a correlate of clinical status (Hwang and Yao, 2011). DAT is a key regulator of DA receptor stimulation and, in turn, affects locomotion and cognitive function. DA transmission is initiated by DA release from the presynaptic terminal and is terminated by its reuptake through DAT. In fact, postmortem analyses of brains from HD patients have shown reduced striatal DAT binding and reduced levels of vesicular monoamine transporter type-2, which is used to estimate the extent of DA innervation (Backman et al., 1997; Ginovart et al., 1997; Suzuki et al., 2001). This indicates that the reduction in DAT binding likely results from a loss of DA nigrostriatal terminals, consistent with the view that the dystonic late-stage symptoms of HD may arise in part from critical reductions in DA input.

## DA IN ANIMAL MODELS OF HD

Animal models of HD have been available for more than 30 years, beginning with the first neurotoxin-based models in which chemically-induced striatal lesions reproduced HD neuropathology, providing insights into the mechanisms underlying striatal cell death (Difiglia, 1990; Brouillet et al., 1999). After the discovery of the HD gene, transgenic and knock-in rodent models were generated. These better replicated the processes and mechanisms underlying the slow development of the human disease far beyond endpoint analyses. We have previously reviewed the phenotypic properties of a number of these models (Cepeda et al., 2010; Raymond et al., 2011). Here, we will briefly describe those that have been used for electrophysiological studies examining DA neurotransmission.

The most widely used mouse model for electrophysiology is the R6/2 line, a transgenic fragment model expressing exon 1 of *HTT* with ∼150 CAG repeats (Mangiarini et al., 1996). R6/2 mice display a very rapidly progressing phenotype, similar to the juvenile form of HD in humans. In these mice, overt symptoms begin to appear at 5–7 weeks of age and become fully manifest after 8 weeks. The R6/1 transgenic mouse model, with ∼110 CAG repeats and less mutant *HTT* expression than the R6/2, displays similar phenotypic alterations but in a more protracted form (Mangiarini et al., 1996). HD mouse models with full-length mutant *HTT* include the yeast artificial chromosome model with 128 CAG repeats (YAC128) and the bacterial artificial chromosome model with 97 CAG/CAA repeats (BACHD) (Slow et al., 2003; Gray et al., 2008). These models show a longer development of the HD phenotype and thus are generally studied at the early (1.5–2 months of age) and late stages (12 months of age), corresponding roughly to periods of hyperkinesia and hypokinesia, respectively. In contrast to transgenic mice where the mutant *HTT* is randomly inserted into the mouse genome, knock-in mouse models have the CAG expansion inserted into the mouse huntingtin gene, which allows gene expression in its appropriate genomic and protein context (Menalled, 2005). The transgenic rat model of HD (tgHD) carries a truncated huntingtin cDNA fragment with 51 CAG repeats (Von Horsten et al., 2003). The tgHD model and most

knock-in mouse models also manifest a slow progression of the HD phenotype.

There is evidence that DA release is reduced in transgenic mouse models in the late stages of the disease, consistent with what is proposed to occur in human HD. There is a progressive reduction in striatal DA levels in both R6/2 and YAC128 mice concomitant with motor abnormalities (Hickey et al., 2002; Johnson et al., 2006; Callahan and Abercrombie, 2011). Furthermore, motorically asymptomatic R6/2 mice show a significant reduction in DA metabolites by 4 weeks of age (Mochel et al., 2011). Deficits in DA levels and/or release have been attributed to either impaired vesicle loading or a reduction in DA reserve pool vesicles available for mobilization (Suzuki et al., 2001; Ortiz et al., 2010). The tgHD rat model displays an increase in striatal DA levels and DA neurons at the early symptomatic stage in two main sources of striatal DA input, the substantia nigra pars compacta and the ventral tegmental area (Jahanshahi et al., 2010). However, these rats also show impaired DA release dynamics, as demonstrated by a reduction in evoked release of DA (Ortiz et al., 2012). Since these results from animal models are not entirely consistent, future studies on DA release dynamics in HD will be needed to parse out changes in DA levels that occur in the early and late disease stages (**Table 1**).

In agreement with analyses of HD patients, striatal D1 and D2 receptors also are compromised in HD mouse models. Striatal D1 and D2 receptor binding is reduced early, with deficiencies in DA signaling seen in R6/2 and R6/1 mice (Cha et al., 1998; Bibb et al., 2000; Ariano et al., 2002; Petersen et al., 2002a). Significant reductions also are seen in mRNA levels of striatal D1 and D2 receptors in late stage YAC128 mice, but not in BACHD mice (Pouladi et al., 2012). It is unclear why these differences occur between the two full-length models.

The traditional view of behavioral abnormalities in HD proposes that hyperkinetic choreic movements in the early stages result from initial dysfunction of D2-enriched indirect pathway MSNs, while hypokinesia during the late stages is a consequence of further defects in D1-enriched direct pathway MSNs (Spektor et al., 2002). This view has been challenged by recent data obtained in experimental mouse models of HD (YAC128 and BACHD) crossed with mice expressing EGFP in direct and indirect pathway neurons. In the early hyperkinetic stage (1.5 months of age), direct pathway MSNs receive more excitatory inputs than control animals, whereas indirect pathway MSNs are not as affected. In contrast, in the late hypokinetic stage (12 months of age) both pathways receive less excitatory inputs compared to controls (André et al., 2011b; Galvan et al., 2012).

DAT dysregulation also may mediate key alterations in DA neurotransmission and behavior in HD mouse models. A marked reduction of DAT immunoreactivity is observed in the striatum of R6/2 mice (Stack et al., 2007). DAT knock-out mice present not only neuropathological but also behavioral hallmarks of HD, i.e., elevated striatal extracellular DA levels, selective MSN degeneration, and locomotor hyperactivity (Giros et al., 1996; Jones et al., 1998; Cyr et al., 2006; Crook and Housman, 2012). Additionally, studies of DAT knock-out mice crossed with a knock-in mouse model of HD demonstrate an increase in stereotypic behavior that emerges at 6 months of age before returning to baseline by 12 months. Wild-type mice crossed with these knock-in HD mice merely demonstrate a similar but less pronounced biphasic pattern of locomotor alteration (Cyr et al., 2006). Thus, it can be concluded that enhanced DA transmission in HD mice exacerbates the behavioral phenotype of the disease.

## DA AND SYNAPTIC PLASTICITY IN HD

Striatal long-term depression (LTD), a long-lasting decrease in the efficacy of GLU synapses, can be induced through high frequency afferent stimulation or sustained postsynaptic membrane depolarization paired with activation of presynaptic metabotropic GLU receptors (Calabresi et al., 1994; Kreitzer and Malenka, 2005). Additionally, acetylcholine and activation of DA D2 and endocannabinoid CB1 receptors is necessary for LTD induction (Wang et al., 2006; Singla et al., 2007). Induction of striatal long-term potentiation (LTP), a long-lasting increase in the efficacy of GLU synapses, requires activation of DA D1, NMDA, and muscarinic acetylcholine receptors (Calabresi et al., 1999; Kerr and Wickens, 2001). LTD is more easily induced in the dorsolateral and caudal striatum while LTP is more prevalent in the dorsomedial and rostral striatum (Partridge et al., 2000; Spencer and Murphy, 2000; Smith et al., 2001).

The 3-nitropropionic acid (3-NP) toxin model shows an increase in NMDA receptor-dependent LTP at cortico-striatal synapses (Akopian et al., 2008). This form of LTP is mediated by D1 receptors and can be reversed by exogenous addition of DA or a D2 receptor agonist. In genetic HD mouse models, DA levels and receptor numbers are altered, resulting in impaired synaptic plasticity. Furthermore, R6/2 mice display a significant reduction in D1-receptor mediated LTP in the striatum (Kung et al., 2007). Impaired LTP in the medial prefrontal cortex of presymptomatic R6/1 mice can be reversed by D1 receptor agonists (Dallerac et al., 2011). Additionally, layer II/III cells in the perirhinal cortex of symptomatic R6/1 mice are unable to support LTD, which may be a result of reductions in D2 receptor activation (Cummings et al., 2006). Paired-pulse profiles, which are measures of short-term plasticity, are aberrant in cortical slices from R6/1 mice. Instead of exhibiting paired-pulse depression seen in control mice, mutants show a more facilitatory profile. Quinpirole, a D2 receptor agonist, produces a profile that resembles age-matched controls and restores LTD (Cummings et al., 2006). Evidence that D1 receptor agonists rescue impaired LTP while D2 receptor agonists rescue impaired LTD show that there is much promise in therapeutics targeting DA modulation of synaptic plasticity. These are functional consequences that hold important implications for ameliorating the cognitive deficits in HD.

As cholinergic transmission and DA are involved in both LTD and LTP, disturbances of the DA-acetylcholine balance in synaptic plasticity could lead to behavioral deficits. In several HD rodent models, LTP does not occur in cholinergic interneurons. As a consequence, MSNs do not display depotentiation, a process induced by low frequency stimulation that leads to reversion of LTP and requires activation of muscarinic receptors (Picconi et al., 2006). This lack of depotentiation may represent a synaptic mechanism for early behavioral abnormalities observed in HD (Picconi et al., 2006).

## DA AND GLU RECEPTOR INTERACTIONS IN HD

Although it is unknown why MSNs preferentially degenerate in HD, one major hypothesis has been that MSNs are more susceptible to excitotoxicity. This theory posits that an excess of excitatory neurotransmitters such as GLU and/or overactivation of GLU receptors, particularly the NMDA receptor, mediate MSN neurodegeneration. Overactivity of NMDA receptors can induce cell death through sustained neuronal membrane depolarization, unchecked $Ca^{2+}$ influx, and/or mitochondrial dysfunction (Difiglia, 1990; Coyle and Puttfarcken, 1993). In addition, although DA exists in high concentrations in the striatum, studies also suggest a toxic role for DA in which cell death is accelerated through increases in free radical production (Hastings et al., 1996; Jakel and Maragos, 2000; Wersinger et al., 2004; Hastings, 2009). In striatal cultures derived from R6/2 mice, MSNs undergo DA-mediated oxidative stress and apoptosis (Petersen et al., 2001). Further, DAT knock-out mice are hypersensitive to 3-NP striatal damage (Fernagut et al., 2002).

DA and GLU neurotransmission are intimately intertwined. Understanding this interplay could help elucidate the cause of biphasic DA changes in human HD. In animal models of HD, biphasic changes in corticostriatal GLU transmission are characterized by initial increases in GLU synaptic activity followed by later decreases (Klapstein et al., 2001; Cepeda et al., 2003; Joshi et al., 2009; André et al., 2011a). Early increases in GLU are associated with cortical hyperexcitability (Cepeda et al., 2003; Spampanato et al., 2008; Cummings et al., 2009) and loss of D2 receptors contributes to increased synaptic activity. Stimulation of corticostriatal neurons has been shown to activate DA release in the striatum (Nieoullon et al., 1978). In addition, DA neurons that modulate GLU release in the corticostriatal pathway are subject to afferent GLU regulation, which is suggested by the presence of GLU receptors on DA neurons (Meltzer et al., 1997). There is substantial evidence for a direct cortico-nigral projection (Afifi et al., 1974; Kornhuber et al., 1984; Naito and Kita, 1994) and work in rodents demonstrates that this pathway both directly and indirectly regulates the firing pattern of DA neurons (Maurice et al., 1999; Sesack and Carr, 2002). Other studies indicate that stimulation of GLU receptors on DA neurons increases DA release in both the substantia nigra and in DA innervated areas (Mintz et al., 1986; Kalivas et al., 1989; Murase et al., 1993). Thus, if DA neuron firing is regulated by frontal cortical neurons, the activity of which is upregulated in early HD, the biphasic trends of DA levels in early and late human HD may be correlated with the biphasic changes of GLU release by cortical afferents. This indicates that biphasic changes in DA levels during early and late HD parallel changes occurring in GLU transmission.

In forebrain neurons, which receive both DA and GLU input, a diminished signal-to-noise ratio can impair both motor and cognitive functions (Kiyatkin and Rebec, 1996). Furthermore, a reduction in DA diminishes the strength of the GLU signal above background activity (Kiyatkin and Rebec, 1996). Recently, Hong and Rebec (2012) developed a theoretical framework suggesting that inflexibility rather than inconsistency is the more relevant problem to explain changes during aging and neurodegeneration. Dysfunction in the DA and GLU systems restricts their ability to modulate neural noise. With aging and neurodegeneration, the range over which DA and GLU can be modulated is decreased, leading to dysfunctional neuronal communication, increased neural noise, and inflexibility in brain activity (Hong and Rebec, 2012). Increased neural noise is evident in HD, appearing as a decrease in burst activity and a loss of correlation in the firing patterns of pairs of neurons in the striatum of HD mice (Miller et al., 2008). As a consequence, behavioral adaptations in response to environmental challenges are reduced.

DA and GLU signaling pathways can synergistically enhance MSN sensitivity to huntingtin toxicity. Studies demonstrate that this deleterious process occurs through D1 but not D2 receptor activation (Tang et al., 2007; Paoletti et al., 2008) and are in agreement with previous studies demonstrating that DA and D1 receptor agonists enhance excitotoxicity (Cepeda and Levine, 1998; McLaughlin et al., 1998). D1 receptor-mediated potentiation of NMDA responses, which holds key functional consequences in HD, has been verified in the cortex and striatum (Cepeda et al., 1993; Wang and O'donnell, 2001; Flores-Hernandez et al., 2002). For example, D1 receptor-induced cell death in MSNs of knock-in HD mice is increased with pretreatment with NMDA when compared with cells from wild-type mice (Paoletti et al., 2008). In neurons from YAC128 mice or Q111 knock-in mice, the convergence of DA and GLU signaling pathways leads to $Ca^{2+}$ overload, resulting in excitotoxic processes such as induction of mitochondrial depolarization and caspase activation (Cepeda et al., 2001; Zeron et al., 2002, 2004; Tang et al., 2007; Paoletti et al., 2008).

While D1-NMDA receptor activation is thought to be neurotoxic, activation of D2 receptors reduces NMDA receptor responses and thus may be neuroprotective (Lee et al., 2002; Bozzi and Borrelli, 2006; Blanke and Vandongen, 2009). For example, activation of D2 receptors by quinpirole reduces the toxicity of both NMDA and kainic acid in rat striatal neurons (Cepeda and Levine, 1998), as well as in mesencephalic and cortical neurons (Sawada et al., 1998; Kihara et al., 2002). However, an exclusive role for D1 receptor activation in mediating MSN degeneration is contradicted by evidence that blocking D2 receptor stimulation significantly reverses DA potentiation of mutant huntingtin-induced MSN cell death (Charvin et al., 2005). As cultured striatal neurons can be protected by antagonism of D1 and D2 receptors, it is possible that both D1 and D2 receptor activation might contribute to neurotoxicity (Davis et al., 2002; Bozzi and Borrelli, 2006). Thus, the exact nature of DA and NMDA interactions are dynamic and complex, indicating a need for further investigation into the differential effects of D1 and D2 activation on GLU signaling in the HD striatum.

## DA AGONISTS AND ANTAGONISTS AS TREATMENTS FOR HD

Since the abnormalities in the DA system appear to underlie many of the behavioral symptoms of HD, DA agonists, antagonists, and/or stabilizers may provide potential treatment options (**Table 2**). Conceptually, DA stabilizers (or partial agonists) increase or decrease DA receptor activity depending on the level of DA tone. HD patients treated with aripiprazole, a partial D2 receptor agonist, demonstrate improvements in chorea, but not cognitive function (Brusa et al., 2009). A recent phase 3 clinical trial of the DA stabilizer pridopidine demonstrated

**Table 2 | Available and potential treatments.**

| | |
|---|---|
| **HUMAN HD** | |
| Tetrabenazine | Well-supported antichoreatic effects but frequent adverse reactions limit its usefulness (Huntington Study Group, 2006). |
| D2 antagonists | *Haloperidol*: a traditional D2 antagonist; improves chorea, but does not increase functional capacity (Bonelli and Wenning, 2006).<br>*Olanzapine and risperidone*: atypical antipsychotic drugs with D2 antagonist properties; improve chorea and behavioral disturbances (Squitieri et al., 2001; Duff et al., 2008). |
| D2 agonists | *Bromocriptine*: effects are both positive and negative (Frattola et al., 1977; Caraceni et al., 1980).<br>*Lisuride*: limited positive effects (Caraceni et al., 1980; Frattola et al., 1983).<br>*Aripiprazole*: a partial D2 agonist; improves chorea but not cognitive function (Brusa et al., 2009). |
| Other DA drugs | *Pridopidine*: a DA stabilizer; produces slight improvements in motor dysfunction (De Yebenes et al., 2011).<br>*L-DOPA*: possibly useful for treatment of rigidity (Racette and Perlmutter, 1998). |
| **ANIMAL MODELS** | |
| Tetrabenazine | Alleviates motor alterations and reduces striatal loss in both early and late stages (Tang et al., 2007; Wang and Morris, 2010; André et al., 2011a). |
| D1 antagonist | *SCH23390*: rescues electrophysiological changes in excitatory and inhibitory synaptic transmission in direct pathway MSNs (André et al., 2011a). |
| D1 agonist | *SKF38393*: reverses impaired LTP in the medial prefrontal cortex of presymptomatic R6/1 mice (Dallerac et al., 2011). |
| D2 antagonist | *Haloperidol*: early and chronic treatment significantly reduces striatal toxicity in the tgHD rat model (Charvin et al., 2008). |
| D2 agonist | *Quinpirole*: restores the ability of transgenic cortical slices to support LTD (Cummings et al., 2006). |

improvements in hand movements, gait, and balance of HD patients as defined by the unified HD rating scale (De Yebenes et al., 2011). Although these changes fell short of the primary efficacy threshold, the slight improvements in motor dysfunction without any deleterious side effects suggest that treatments targeted toward DA imbalance may have therapeutic benefits.

Current treatment options for HD are limited and confined to antidopaminergic agents for motor symptoms while there are virtually no therapeutics for cognitive deterioration (Venuto et al., 2012). Additionally, clinical results of these treatments seem contradictory, possibly reflecting the dynamic and time-dependent changes that occur in the DA system as the disease progresses (Mochel et al., 2011). For example, both D2 agonists and antagonists have demonstrated clinical benefits for improvement of HD motor symptoms (Tedroff et al., 1999; Haskins and Harrison, 2000; Brusa et al., 2009). Conventional antipsychotic drugs, such as the D2 antagonist haloperidol, are used in clinical practice, but they do not improve functional capacity (Bonelli and Wenning, 2006). Atypical antipsychotic drugs with D2 antagonist properties such as olanzapine, risperidone, quetiapine, and ziprasidone, can improve chorea and impact a larger range of behavioral disturbances with a reduced risk of side effects (Squitieri et al., 2001; Bonelli et al., 2003; Alpay and Koroshetz, 2006; Duff et al., 2008). D2 agonists like bromocriptine and lisuride have also demonstrated therapeutic potential in HD (Frattola et al., 1977, 1983; Caraceni et al., 1980).

As the early stages of HD may reflect a hyperdopaminergic stage, drugs that reduce DA tone can be beneficial during the choreic movement phase (Mochel et al., 2011). DA-depleting agents such as tetrabenazine (TBZ), which inhibits vesicular monoamine transporter type-2 and decreases DA content in presynaptic vesicles, have been shown to reduce chorea (Huntington Study Group, 2006). Currently, TBZ is the only drug formally approved for treatment of Huntington's chorea by a regulatory agency (Mestre and Ferreira, 2012).

*In vivo* and *in vitro* studies of animal models support a role for DA inhibitors in protecting HD MSNs from cell death. The rationale follows and agrees with experimental and clinical findings suggesting that DA tone is elevated during the early stages of the disease. In YAC128 mice, TBZ alleviates motor deficits and reduces striatal loss in both early and late stages (Tang et al., 2007; Wang et al., 2010). TBZ also rescues the increased stereotypies in 1–2 month old YAC128 and BACHD mice (André et al., 2011a). D1 receptor antagonists rescue the changes in excitatory synaptic transmission of direct pathway MSNs that occur in the early symptomatic phase of YAC128 and BACHD mice, suggesting that tonic activation of D1 receptors may underlie early dysfunction of D1 MSNs (André et al., 2011a). Similarly, D1 receptor antagonists prevent DA/GLU-induced MSN death in YAC128 mice (Tang et al., 2007). In a lentivirus-based rat model, striatal toxicity is reduced by early and chronic treatment with haloperidol (Charvin et al., 2008). However, this evidence is complicated by the fact that haloperidol, a putative D2 receptor antagonist, also modulates NMDA receptor function (Fletcher and Macdonald, 1993; Ilyin et al., 1996; Arvanov et al., 1997). Predictably, DA antagonists may be more beneficial when administered with other neuroprotective drugs such as memantine, a NMDA receptor antagonist, as a combination therapy (Wu et al., 2006).

HD mouse models have demonstrated the therapeutic potential of not only DA antagonists, but also DA agonists. For example, in fully symptomatic R6/2 mice, replacement of reduced DA levels by chronic treatment with L-DOPA yields short-term improvements in the HD behavioral phenotype whereas long-term treatment impairs survival and rotarod performance (Hickey et al., 2002). Additionally, D1 receptor agonists rescue cortical LTP impairment and deficits in synaptic plasticity of R6/1 mice (Dallerac et al., 2011), suggesting that increasing DA levels could improve cognitive dysfunction. Since some treatments may only be suitable early or late in disease progression, effective therapies need to be temporally oriented to accommodate differential changes in DA levels throughout the course of the disease.

## CONCLUSIONS AND FUTURE DIRECTIONS

While the role of DA in Parkinson's disease is well-established, its role in HD is less well-understood. Although an association between chorea and excess DA levels had long been suspected, a causal link was not demonstrated until TBZ was shown to alleviate abnormal movements in HD. Other less known alterations in early symptomatic patients, such as cognitive changes, impulsivity, gambling, and hypersexuality, could also associate with perturbations of the DA system (Fedoroff et al., 1994; Stout et al., 2001; Rosenblatt, 2007; Beglinger et al., 2008; Jhanjee et al., 2011). TBZ can treat chorea and other early symptoms by reducing DA, but it can also have deleterious effects on cognitive function. Understanding time- and region-dependent alterations in DA function throughout the course of the disease will help in discovering better therapeutic strategies. Selective manipulation of DA-producing neurons, such as using optogenetics in animal models and potentially in human patients, may open new and exciting alternatives.

While much knowledge on the role of DA in HD has been gathered in the past few years, many questions remain unanswered and should be the focus of future endeavors. The traditional view that D2 MSNs are more vulnerable in HD is beginning to change due to emerging data from experimental animal models. Based on evidence reviewed here, one may think that, in fact, D1 MSNs should be more vulnerable to the HD mutation, i.e., they become dysfunctional in the early stage of HD and D1-NMDA receptor interactions enhance neurotoxicity. Therefore, the standing question should be reformulated to ask why D1 MSNs are less susceptible in HD. Do they have a neuroprotective mechanism that D2 MSNs lack? Recent studies using mice expressing EGFP in D1 or D2 cells point in that direction. For example, fluorescence-activated cell sorting array analyses showed that the transcription factor Zfp521, which is enriched in D1 MSNs, is anti-apoptotic (Lobo et al., 2008). Specifically, Zfp521 promotes proliferation, delays differentiation, and reduces apoptosis (Shen et al., 2011).

Another important question is: what causes early perturbations in DA release? Is it the loss of striosome MSN projections to the substantia nigra pars compacta, increased activity along the cortico-nigral projection, or dysregulation of DA release due to loss of D2 auto-receptors? On a similar note, since there are at least two splice variants for D2 receptors, a short D2S (mostly presynaptic) and a long D2L (mostly postsynaptic) form, which one is reduced in early HD? In the striatum, DA D2 auto-receptor function is mediated by synapsin III, a phosphoprotein that is specifically involved in regulating vesicular reserve pools and DA release in the striatum (Feng et al., 2002; Kile et al., 2010). In brains of R6/2 mice and HD patients, there is a progressive loss of complexins, synaptic proteins similar to syntaxin III that are involved in synaptogenesis and modulate neurotransmitter release (Freeman and Morton, 2004). If a similar reduction in synapsin III occurs, this could explain increased DA transmission in early HD and a consequent loss of behavioral flexibility. In agreement, reversal learning can be improved by increasing levels of synapsin III (Laughlin et al., 2011). Thus far, it is unknown whether or not presynaptic D2 auto- or hetero-receptors are lost before postsynaptic receptors (Sandstrom et al., 2010). However, selective agonists of D2 auto-receptors produce long-lasting suppression of extracellular brain DA levels *in vivo* and could provide promising therapeutic benefits for HD (Pifl et al., 1988).

As shown in this review, our knowledge of changes in DA function in HD has made substantial strides, particularly after the introduction of genetic rodent models. However, many more questions remain. Answering these questions is within reach and use of these animal models should help understand the early mechanisms of striatal DA dysfunction and its role in behavioral alterations.

## ACKNOWLEDGMENTS

Supported by USPHS NS41574, a contract from CHDI Inc., and the Hereditary Disease Foundation.

## REFERENCES

Afifi, A. K., Bahuth, N. B., Kaelber, W. W., Mikhael, E., and Nassar, S. (1974). The cortico-nigral fibre tract. An experimental Fink-Heimer study in cats. *J. Anat.* 118, 469–476.

Akopian, G., Crawford, C., Beal, M. F., Cappelletti, M., Jakowec, M. W., Petzinger, G. M., et al. (2008). Decreased striatal dopamine release underlies increased expression of long-term synaptic potentiation at corticostriatal synapses 24 h after 3-nitropropionic-acid-induced chemical hypoxia. *J. Neurosci.* 28, 9585–9597. doi: 10.1523/JNEUROSCI.5698-07.2008

Albin, R. L., Young, A. B., and Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends Neurosci.* 12, 366–375. doi: 10.1016/0166-2236(89)90074-X

Alpay, M., and Koroshetz, W. J. (2006). Quetiapine in the treatment of behavioral disturbances in patients with Huntington's disease. *Psychosomatics* 47, 70–72. doi: 10.1176/appi.psy.47.1.70

André, V. M., Cepeda, C., Fisher, Y. E., Huynh, M., Bardakjian, N., Singh, S., et al. (2011a). Differential electrophysiological changes in striatal output neurons in Huntington's disease. *J. Neurosci.* 31, 1170–1182. doi: 10.1523/JNEUROSCI.3539-10.2011

André, V. M., Fisher, Y. E., and Levine, M. S. (2011b). Altered balance of activity in the striatal direct and indirect pathways in mouse models of Huntington's disease. *Front. Syst. Neurosci.* 5:46. doi: 10.3389/fnsys.2011.00046

Andrew, S. E., Goldberg, Y. P., Kremer, B., Telenius, H., Theilmann, J., Adam, S., et al. (1993). The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nat. Genet.* 4, 398–403. doi: 10.1038/ng0893-398

Antonini, A., Leenders, K. L., Spiegel, R., Meier, D., Vontobel, P., Weigell-Weber, M., et al. (1996). Striatal glucose metabolism and dopamine D2 receptor binding in asymptomatic gene carriers and patients with Huntington's disease. *Brain* 119(Pt 6), 2085–2095. doi: 10.1093/brain/119.6.2085

Aosaki, T., Graybiel, A. M., and Kimura, M. (1994). Effect of the nigrostriatal dopamine system on acquired neural responses in the striatum of behaving monkeys. *Science* 265, 412–415. doi: 10.1126/science.8023166

Aosaki, T., Miura, M., Suzuki, T., Nishimura, K., and Masuda, M. (2010). Acetylcholine-dopamine balance hypothesis in the striatum: an update. *Geriatr. Gerontol. Int.* 10(Suppl. 1), S148–S157. doi: 10.1111/j.1447-0594.2010.00588.x

Ariano, M. A., Aronin, N., Difiglia, M., Tagle, D. A., Sibley, D. R., Leavitt, B. R., et al. (2002). Striatal neurochemical changes in transgenic models of Huntington's disease. *J. Neurosci. Res.* 68, 716–729. doi: 10.1002/jnr.10272

Aron, A. R., Durston, S., Eagle, D. M., Logan, G. D., Stinear, C. M., and Stuphorn, V. (2007). Converging evidence for a fronto-basal-ganglia network for inhibitory control of action and cognition. *J. Neurosci.* 27, 11860–11864. doi: 10.1523/JNEUROSCI.3644-07.2007

Aron, A. R., and Poldrack, R. A. (2006). Cortical and subcortical contributions to Stop signal response inhibition: role of the subthalamic nucleus. *J. Neurosci.* 26, 2424–2433. doi: 10.1523/JNEUROSCI.4682-05.2006

Arvanov, V. L., Liang, X., Schwartz, J., Grossman, S., and Wang, R. Y. (1997). Clozapine and haloperidol modulate N-methyl-D-aspartate- and non-N-methyl-D-aspartate receptor-mediated neurotransmission in rat prefrontal cortical neurons *in vitro*. *J. Pharmacol. Exp. Ther.* 283, 226–234.

Backman, L., and Farde, L. (2001). Dopamine and cognitive functioning: brain imaging findings in Huntington's disease and normal aging. *Scand. J. Psychol.* 42, 287–296. doi: 10.1111/1467-9450.00238

Backman, L., Robins-Wahlin, T. B., Lundin, A., Ginovart, N., and Farde, L. (1997). Cognitive deficits in Huntington's disease are predicted by dopaminergic PET markers and brain volumes. *Brain* 120(Pt 12), 2207–2217. doi: 10.1093/brain/120.12.2207

Bamford, N. S., Zhang, H., Schmitz, Y., Wu, N. P., Cepeda, C., Levine, M. S., et al. (2004). Heterosynaptic dopamine neurotransmission selects sets of corticostriatal terminals. *Neuron* 42, 653–663. doi: 10.1016/S0896-6273(04)00265-X

Beglinger, L. J., Paulsen, J. S., Watson, D. B., Wang, C., Duff, K., Langbehn, D. R., et al. (2008). Obsessive and compulsive symptoms in prediagnosed Huntington's disease. *J. Clin. Psychiatry* 69, 1758–1765. doi: 10.4088/JCP.v69n1111

Bernheimer, H., Birkmayer, W., Hornykiewicz, O., Jellinger, K., and Seitelberger, F. (1973). Brain dopamine and the syndromes of Parkinson and Huntington. Clinical, morphological and neurochemical correlations. *J. Neurol. Sci.* 20, 415–455. doi: 10.1016/0022-510X(73)90175-5

Berridge, K. C., Aldridge, J. W., Houchard, K. R., and Zhuang, X. (2005). Sequential super-stereotypy of an instinctive fixed action pattern in hyper-dopaminergic mutant mice: a model of obsessive compulsive disorder and Tourette's. *BMC Biol.* 3:4. doi: 10.1186/1741-7007-3-4

Beste, C., Willemssen, R., Saft, C., and Falkenstein, M. (2010). Response inhibition subprocesses and dopaminergic pathways: basal ganglia disease effects. *Neuropsychologia* 48, 366–373. doi: 10.1016/j.neuropsychologia.2009.09.023

Bibb, J. A., Yan, Z., Svenningsson, P., Snyder, G. L., Pieribone, V. A., Horiuchi, A., et al. (2000). Severe deficiences in dopamine signaling in presymptomatic Huntington's disease mice. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6809–6814. doi: 10.1073/pnas.120166397

Bird, E. D. (1980). Chemical pathology of Huntington's disease. *Annu. Rev. Pharmacol. Toxicol.* 20, 533–551. doi: 10.1146/annurev.pa.20.040180.002533

Blanke, M. L., and Vandongen, A. M. J. (2009). "Activation mechanisms of the NMDA receptor," in *Biology of the NMDA Receptor*, ed A. M. J. Vandongen (Boca Raton, FL: CRC Press), Chapter 13.

Bolam, J. P., Hanley, J. J., Booth, P. A., and Bevan, M. D. (2000). Synaptic organisation of the basal ganglia. *J. Anat.* 196(Pt 4), 527–542. doi: 10.1046/j.1469-7580.2000.19640527.x

Bolam, J. P., Wainer, B. H., and Smith, A. D. (1984). Characterization of cholinergic neurons in the rat neostriatum. A combination of choline acetyltransferase immunocytochemistry, Golgi-impregnation and electron microscopy. *Neuroscience* 12, 711–718. doi: 10.1016/0306-4522(84)90165-9

Bonelli, R. M., and Hofmann, P. (2007). A systematic review of the treatment studies in Huntington's disease since 1990. *Expert Opin. Pharmacother.* 8, 141–153. doi: 10.1517/14656566.8.2.141

Bonelli, R. M., and Wenning, G. K. (2006). Pharmacological management of Huntington's disease: an evidence-based review. *Curr. Pharm. Des.* 12, 2701–2720. doi: 10.2174/138161206777698693

Bonelli, R. M., Mayr, B. M., Niederwieser, G., Reisecker, F., and Kapfhammer, H. P. (2003). Ziprasidone in Huntington's disease: the first case reports. *J. Psychopharmacol.* 17, 459–460. doi: 10.1177/0269881103174009

Bozzi, Y., and Borrelli, E. (2006). Dopamine in neurotoxicity and neuroprotection: what do D2 receptors have to do with it. *Trends Neurosci.* 29, 167–174. doi: 10.1016/j.tins.2006.01.002

Brandt, J., and Butters, N. (1986). The neuropsychology of Huntington's disease. *Trends Neurosci.* 9, 118–120. doi: 10.1016/0166-2236(86)90039-1

Brouillet, E., Conde, F., Beal, M. F., and Hantraye, P. (1999). Replicating Huntington's disease phenotype in experimental animals. *Prog. Neurobiol.* 59, 427–468. doi: 10.1016/S0301-0082(99)00005-2

Brusa, L., Orlacchio, A., Moschella, V., Iani, C., Bernardi, G., and Mercuri, N. B. (2009). Treatment of the symptoms of Huntington's disease: preliminary results comparing aripiprazole and tetrabenazine. *Mov. Disord.* 24, 126–129. doi: 10.1002/mds.22376

Calabresi, P., Centonze, D., Gubellini, P., and Bernardi, G. (1999). Activation of M1-like muscarinic receptors is required for the induction of corticostriatal LTP. *Neuropharmacology* 38, 323–326. doi: 10.1016/S0028-3908(98)00199-3

Calabresi, P., Pisani, A., Mercuri, N. B., and Bernardi, G. (1994). Post-receptor mechanisms underlying striatal long-term depression. *J. Neurosci.* 14, 4871–4881.

Callahan, J. W., and Abercrombie, E. D. (2011). *In vivo* dopamine efflux is decreased in striatum of both fragment (R6/2) and full-length (YAC128) transgenic mouse models of Huntington's disease. *Front. Syst. Neurosci.* 5:61. doi: 10.3389/fnsys.2011.00061

Canales, J. J. (2005). Stimulant-induced adaptations in neostriatal matrix and striosome systems: transiting from instrumental responding to habitual behavior in drug addiction. *Neurobiol. Learn. Mem.* 83, 93–103. doi: 10.1016/j.nlm.2004.10.006

Canales, J. J., and Graybiel, A. M. (2000). A measure of striatal function predicts motor stereotypy. *Nat. Neurosci.* 3, 377–383. doi: 10.1038/73949

Caraceni, T. A., Girotti, F., Giovannini, P., Pederzoli, M., and Parati, E. A. (1980). Effects of DA agonist in Huntington disease hyperkinesia. *Ital. J. Neurol. Sci.* 1, 155–161. doi: 10.1007/BF02335845

Cepeda, C., Andre, V. M., Yamazaki, I., Wu, N., Kleiman-Weiner, M., and Levine, M. S. (2008). Differential electrophysiological properties of dopamine D1 and D2 receptor-containing striatal medium-sized spiny neurons. *Eur. J. Neurosci.* 27, 671–682. doi: 10.1111/j.1460-9568.2008.06038.x

Cepeda, C., Buchwald, N. A., and Levine, M. S. (1993). Neuromodulatory actions of dopamine in the neostriatum are dependent upon the excitatory amino acid receptor subtypes activated. *Proc. Natl. Acad. Sci. U.S.A.* 90, 9576–9580. doi: 10.1073/pnas.90.20.9576

Cepeda, C., Cummings, D. M., Andre, V. M., Holley, S. M., and Levine, M. S. (2010). Genetic mouse models of Huntington's disease: focus on electrophysiological mechanisms. *ASN Neuro* 2, e00033. doi: 10.1042/AN20090058

Cepeda, C., Hurst, R. S., Altemus, K. L., Flores-Hernandez, J., Calvert, C. R., Jokel, E. S., et al. (2001). Facilitated glutamatergic transmission in the striatum of D2 dopamine receptor-deficient mice. *J. Neurophysiol.* 85, 659–670.

Cepeda, C., Hurst, R. S., Calvert, C. R., Hernandez-Echeagaray, E., Nguyen, O. K., Jocoy, E., et al. (2003). Transient and progressive electrophysiological alterations in the corticostriatal pathway in a mouse model of Huntington's disease. *J. Neurosci.* 23, 961–969.

Cepeda, C., and Levine, M. S. (1998). Dopamine and N-methyl-D-aspartate receptor interactions in the neostriatum. *Dev. Neurosci.* 20, 1–18. doi: 10.1159/000017294

Cha, J. H., Kosinski, C. M., Kerner, J. A., Alsdorf, S. A., Mangiarini, L., Davies, S. W., et al. (1998). Altered brain neurotransmitter receptors in transgenic mice expressing a portion of an abnormal human huntington disease gene. *Proc. Natl. Acad. Sci. U.S.A.* 95, 6480–6485. doi: 10.1073/pnas.95.11.6480

Charvin, D., Roze, E., Perrin, V., Deyts, C., Betuing, S., Pages, C., et al. (2008). Haloperidol protects striatal neurons from dysfunction induced by mutated huntingtin *in vivo*. *Neurobiol. Dis.* 29, 22–29. doi: 10.1016/j.nbd.2007.07.028

Charvin, D., Vanhoutte, P., Pages, C., Borrelli, E., and Caboche, J. (2005). Unraveling a role for dopamine in Huntington's disease: the dual role of reactive oxygen species and D2 receptor stimulation. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12218–12223. doi: 10.1073/pnas.0502698102

Cools, R., Ivry, R. B., and D'esposito, M. (2006). The human striatum is necessary for responding to changes in stimulus relevance. *J. Cogn. Neurosci.* 18, 1973–1983. doi: 10.1162/jocn.2006.18.12.1973

Cooper, S., and Dourish, C. (1990). *Neurobiology of Stereotyped Behavior*. Oxford: Clarendon Press.

Coyle, J. T., and Puttfarcken, P. (1993). Oxidative stress, glutamate, and neurodegenerative disorders. *Science* 262, 689–695. doi: 10.1126/science.7901908

Crook, Z. R., and Housman, D. E. (2012). Dysregulation of dopamine receptor D2 as a sensitive measure for Huntington disease pathology in model mice. *Proc. Natl. Acad. Sci. U.S.A.* 109, 7487–7492. doi: 10.1073/pnas.1204542109

Cummings, D. M., Andre, V. M., Uzgil, B. O., Gee, S. M., Fisher, Y. E., Cepeda, C., et al. (2009). Alterations in cortical excitation and inhibition in genetic mouse models of Huntington's disease. *J. Neurosci.* 29, 10371–10386. doi: 10.1523/JNEUROSCI.1592-09.2009

Cummings, D. M., Milnerwood, A. J., Dallerac, G. M., Waights, V., Brown, J. Y., Vatsavayai, S. C., et al. (2006). Aberrant cortical synaptic plasticity and dopaminergic dysfunction in a mouse model of huntington's disease. *Hum. Mol. Genet.* 15, 2856–2868. doi: 10.1093/hmg/ddl224

Cyr, M., Sotnikova, T. D., Gainetdinov, R. R., and Caron, M. G. (2006). Dopamine enhances motor and neuropathological consequences of polyglutamine expanded huntingtin. *FASEB J.* 20, 2541–2543. doi: 10.1096/fj.06-6533fje

Dallerac, G. M., Vatsavayai, S. C., Cummings, D. M., Milnerwood, A. J., Peddie, C. J., Evans, K. A., et al. (2011). Impaired long-term potentiation in the prefrontal cortex of Huntington's disease mouse models: rescue by D1 dopamine receptor activation. *Neurodegener. Dis.* 8, 230–239. doi: 10.1159/000322540

Dalley, J. W., Fryer, T. D., Brichard, L., Robinson, E. S., Theobald, D. E., Laane, K., et al. (2007). Nucleus accumbens D2/3 receptors predict trait impulsivity and cocaine reinforcement. *Science* 315, 1267–1270. doi: 10.1126/science.1137073

Dang, L. C., Donde, A., Madison, C., O'neil, J. P., and Jagust, W. J. (2012). Striatal dopamine influences the default mode network to affect shifting between object features. *J. Cogn. Neurosci.* 24, 1960–1970. doi: 10.1162/jocn_a_00252

Davis, S., Brotchie, J., and Davies, I. (2002). Protection of striatal neurons by joint blockade of D1 and D2 receptor subtypes in an *in vitro* model of cerebral hypoxia. *Exp. Neurol.* 176, 229–236. doi: 10.1006/exnr.2002.7926

De Yebenes, J. G., Landwehrmeyer, B., Squitieri, F., Reilmann, R., Rosser, A., Barker, R. A., et al. (2011). Pridopidine for the treatment of motor function in patients with Huntington's disease (MermaiHD): a phase 3, randomised, double-blind, placebo-controlled trial. *Lancet Neurol.* 10, 1049–1057. doi: 10.1016/S1474-4422(11)70233-2

Devan, B. D., McDonald, R. J., and White, N. M. (1999). Effects of medial and lateral caudate-putamen lesions on place- and cue-guided behaviors in the water maze: relation to thigmotaxis. *Behav. Brain Res.* 100, 5–14. doi: 10.1016/S0166-4328(98)00107-7

Dickinson, D., and Elvevag, B. (2009). Genes, cognition and brain through a COMT lens. *Neuroscience* 164, 72–87. doi: 10.1016/j.neuroscience.2009.05.014

Difiglia, M. (1990). Excitotoxic injury of the neostriatum: a model for Huntington's disease. *Trends Neurosci.* 13, 286–289. doi: 10.1016/0166-2236(90)90111-M

Ding, J. B., Guzman, J. N., Peterson, J. D., Goldberg, J. A., and Surmeier, D. J. (2010). Thalamic gating of corticostriatal signaling by cholinergic interneurons. *Neuron* 67, 294–307. doi: 10.1016/j.neuron.2010.06.017

Duff, K., Beglinger, L. J., O'rourke, M. E., Nopoulos, P., Paulson, H. L., and Paulsen, J. S. (2008). Risperidone and the treatment of psychiatric, motor, and cognitive symptoms in Huntington's disease. *Ann. Clin. Psychiatry* 20, 1–3. doi: 10.1080/10401230701844802

Fedoroff, J. P., Peyser, C., Franz, M. L., and Folstein, S. E. (1994). Sexual disorders in Huntington's disease. *J. Neuropsychiatry Clin. Neurosci.* 6, 147–153.

Feng, J., Chi, P., Blanpied, T. A., Xu, Y., Magarinos, A. M., Ferreira, A., et al. (2002). Regulation of neurotransmitter release by synapsin III. *J. Neurosci.* 22, 4372–4380.

Fernagut, P. O., Diguet, E., Jaber, M., Bioulac, B., and Tison, F. (2002). Dopamine transporter knock-out mice are hypersensitive to 3-nitropropionic acid-induced striatal damage. *Eur. J. Neurosci.* 15, 2053–2056. doi: 10.1046/j.1460-9568.2002.02047.x

Fletcher, E. J., and Macdonald, J. F. (1993). Haloperidol interacts with the strychnine-insensitive glycine site at the NMDA receptor in cultured mouse hippocampal neurones. *Eur. J. Pharmacol.* 235, 291–295. doi: 10.1016/0014-2999(93)90148-B

Flores-Barrera, E., Vizcarra-Chacon, B. J., Tapia, D., Bargas, J., and Galarraga, E. (2010). Different corticostriatal integration in spiny projection neurons from direct and indirect pathways. *Front. Syst. Neurosci.* 4:15. doi: 10.3389/fnsys.2010.00015

Flores-Hernandez, J., Cepeda, C., Hernandez-Echeagaray, E., Calvert, C. R., Jokel, E. S., Fienberg, A. A., et al. (2002). Dopamine enhancement of NMDA currents in dissociated medium-sized striatal neurons: role of D1 receptors and DARPP-32. *J. Neurophysiol.* 88, 3010–3020. doi: 10.1152/jn.00361.2002

Flores-Hernandez, J., Galarraga, E., and Bargas, J. (1997). Dopamine selects glutamatergic inputs to neostriatal neurons. *Synapse* 25, 185–195.

Frank, M. J., Seeberger, L. C., and O'reilly R, C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306, 1940–1943. doi: 10.1126/science.1102941

Frattola, L., Albizzati, M. G., Alemani, A., Bassi, S., Ferrarese, C., and Trabucchi, M. (1983). Acute treatment of Huntington's chorea with lisuride. *J. Neurol. Sci.* 59, 247–253. doi: 10.1016/0022-510X(83)90042-4

Frattola, L., Albiazzati, M. G., Spano, P. F., and Trabucchi, M. (1977). Treatment of Huntington's chorea with bromocriptine. *Acta Neurol. Scand.* 56, 37–45. doi: 10.1111/j.1600-0404.1977.tb01407.x

Freeman, W., and Morton, A. J. (2004). Regional and progressive changes in brain expression of complexin II in a mouse transgenic for the Huntington's disease mutation. *Brain Res. Bull.* 63, 45–55. doi: 10.1016/j.brainresbull.2003.12.004

Galvan, L., André, V. M., Wang, E. A., Cepeda, C., and Levine, M. S. (2012). Functional differences between direct and indirect striatal output pathways in Huntington's disease. *J. Huntingt. Dis.* 1, 17–25.

Garrett, M. C., and Soares-Da-Silva, P. (1992). Increased cerebrospinal fluid dopamine and 3, 4-dihydroxyphenylacetic acid levels in Huntington's disease: evidence for an overactive dopaminergic brain transmission. *J. Neurochem.* 58, 101–106. doi: 10.1111/j.1471-4159.1992.tb09283.x

Gerfen, C. R. (1984). The neostriatal mosaic: compartmentalization of corticostriatal input and striatonigral output systems. *Nature* 311, 461–464. doi: 10.1038/311461a0

Gerfen, C. R. (1992). The neostriatal mosaic: multiple levels of compartmental organization. *Trends Neurosci.* 15, 133–139. doi: 10.1016/0166-2236(92)90355-C

Gerfen, C. R., Engber, T. M., Mahan, L. C., Susel, Z., Chase, T. N., Monsma, F. J., et al. (1990). D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science* 250, 1429–1432. doi: 10.1126/science.2147780

Gertler, T. S., Chan, C. S., and Surmeier, D. J. (2008). Dichotomous anatomical properties of adult striatal medium spiny neurons. *J. Neurosci.* 28, 10814–10824. doi: 10.1523/JNEUROSCI.2660-08.2008

Ginovart, N., Lundin, A., Farde, L., Halldin, C., Backman, L., Swahn, C. G., et al. (1997). PET study of the pre- and post-synaptic dopaminergic markers for the neurodegenerative process in Huntington's disease. *Brain* 120(Pt 3), 503–514. doi: 10.1093/brain/120.3.503

Giros, B., Jaber, M., Jones, S. R., Wightman, R. M., and Caron, M. G. (1996). Hyperlocomotion and indifference to cocaine and amphetamine in mice lacking the dopamine transporter. *Nature* 379, 606–612. doi: 10.1038/379606a0

Gray, M., Shirasaki, D. I., Cepeda, C., Andre, V. M., Wilburn, B., Lu, X. H., et al. (2008). Full-length human mutant huntingtin with a stable polyglutamine repeat can elicit progressive and selective neuropathogenesis in BACHD mice. *J. Neurosci.* 28, 6182–6195. doi: 10.1523/JNEUROSCI.0857-08.2008

Graybiel, A. M. (2000). The basal ganglia. *Curr. Biol.* 10, R509–R511. doi: 10.1016/S0960-9822(00)00593-5

Haber, S. N., and Nauta, W. J. (1983). Ramifications of the globus pallidus in the rat as indicated by patterns of immunohistochemistry. *Neuroscience* 9, 245–260. doi: 10.1016/0306-4522(83)90291-9

Haskins, B. A., and Harrison, M. B. (2000). Huntington's disease. *Curr. Treat. Options Neurol.* 2, 243–262. doi: 10.1007/s11940-000-0007-0

Hastings, T. G. (2009). The role of dopamine oxidation in mitochondrial dysfunction: implications for Parkinson's disease. *J. Bioenerg. Biomembr.* 41, 469–472. doi: 10.1007/s10863-009-9257-z

Hastings, T. G., Lewis, D. A., and Zigmond, M. J. (1996). Role of oxidation in the neurotoxic effects of intrastriatal dopamine injections. *Proc. Natl. Acad. Sci. U.S.A.* 93, 1956–1961. doi: 10.1073/pnas.93.5.1956

Hedreen, J. C., and Folstein, S. E. (1995). Early loss of neostriatal striosome neurons in Huntington's disease. *J. Neuropathol. Exp. Neurol.* 54, 105–120. doi: 10.1097/00005072-199501000-00013

Heinsen, H., Rub, U., Gangnus, D., Jungkunz, G., Bauer, M., Ulmar, G., et al. (1996). Nerve cell loss in the thalamic centromedian-parafascicular complex in patients with Huntington's disease. *Acta. Neuropathol.* 91, 161–168. doi: 10.1007/s004010050408

Hernandez-Lopez, S., Bargas, J., Surmeier, D. J., Reyes, A., and Galarraga, E. (1997). D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca2+ conductance. *J. Neurosci.* 17, 3334–3342.

Hickey, M. A., Reynolds, G. P., and Morton, A. J. (2002). The role of dopamine in motor symptoms in the R6/2 transgenic mouse model of Huntington's disease. *J. Neurochem.* 81, 46–59. doi: 10.1046/j.1471-4159.2002.00804.x

Hong, S. L., and Rebec, G. V. (2012). Biological sources of inflexibility in brain and behavior with aging and neurodegenerative diseases. *Front. Syst. Neurosci.* 6:77. doi: 10.3389/fnsys.2012.00077

Huntington Study Group. (2006). Tetrabenazine as antichorea therapy in Huntington disease: a randomized controlled trial. *Neurology* 66, 366–372. doi: 10.1212/01.wnl.0000198586.85250.13

Hwang, W. J., and Yao, W. J. (2011). SPECT study of the nigrostriatal dopaminergic system in Huntington's disease. *J. Neuroimaging* 2, 192–196. doi: 10.1111/j.1552-6569.2011.00671.x

Ilyin, V. I., Whittemore, E. R., Guastella, J., Weber, E., and Woodward, R. M. (1996). Subtype-selective inhibition of N-methyl-D-aspartate receptors by haloperidol. *Mol. Pharmacol.* 50, 1541–1550.

Jahanshahi, A., Vlamings, R., Kaya, A. H., Lim, L. W., Janssen, M. L., Tan, S., et al. (2010). Hyperdopaminergic status in experimental Huntington disease. *J. Neuropathol. Exp. Neurol.* 69, 910–917. doi: 10.1097/NEN.0b013e3181ee005d

Jakel, R. J., and Maragos, W. F. (2000). Neuronal cell death in Huntington's disease: a potential role for dopamine. *Trends Neurosci.* 23, 239–245. doi: 10.1016/S0166-2236(00)01568-X

Jhanjee, A., Anand, K. S., and Bajaj, B. K. (2011). Hypersexual features in Huntington's disease. *Singapore Med. J.* 52, e131–e133.

Jimenez-Castellanos, J., and Graybiel, A. M. (1989). Compartmental origins of striatal efferent projections in the cat. *Neuroscience* 32, 297–321. doi: 10.1016/0306-4522(89)90080-8

Johnson, M. A., Rajan, V., Miller, C. E., and Wightman, R. M. (2006). Dopamine release is severely compromised in the R6/2 mouse model of Huntington's disease. *J. Neurochem.* 97, 737–746. doi: 10.1111/j.1471-4159.2006.03762.x

Jones, S. R., Gainetdinov, R. R., Jaber, M., Giros, B., Wightman, R. M., and Caron, M. G. (1998). Profound neuronal plasticity in response to inactivation of the dopamine transporter. *Proc. Natl. Acad. Sci. U.S.A.* 95, 4029–4034. doi: 10.1073/pnas.95.7.4029

Joshi, P. R., Wu, N. P., Andre, V. M., Cummings, D. M., Cepeda, C., Joyce, J. A., et al. (2009). Age-dependent alterations of corticostriatal activity in the YAC128 mouse model of Huntington disease. *J. Neurosci.* 29, 2414–2427. doi: 10.1523/JNEUROSCI.5687-08.2009

Joyce, J. N., Lexow, N., Bird, E., and Winokur, A. (1988). Organization of dopamine D1 and D2 receptors in human striatum: receptor autoradiographic studies in Huntington's disease and schizophrenia. *Synapse* 2, 546–557. doi: 10.1002/syn.890020511

Kalivas, P. W., Duffy, P., and Barrow, J. (1989). Regulation of the mesocorticolimbic dopamine system by glutamic acid receptor subtypes. *J. Pharmacol. Exp. Ther.* 251, 378–387.

Kemp, J. M., and Powell, T. P. (1971). The termination of fibres from the cerebral cortex and thalamus upon dendritic spines in the caudate nucleus: a study with the Golgi method. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 262, 429–439. doi: 10.1098/rstb.1971.0105

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

Kihara, T., Shimohama, S., Sawada, H., Honda, K., Nakamizo, T., Kanki, R., et al. (2002). Protective effect of dopamine D2 agonists in cortical neurons via the phosphatidylinositol 3 kinase cascade. *J. Neurosci. Res.* 70, 274–282. doi: 10.1002/jnr.10426

Kile, B. M., Guillot, T. S., Venton, B. J., Wetsel, W. C., Augustine, G. J., and Wightman, R. M. (2010). Synapsins differentially control dopamine and serotonin release. *J. Neurosci.* 30, 9762–9770. doi: 10.1523/JNEUROSCI.2071-09.2010

Kish, S. J., Shannak, K., and Hornykiewicz, O. (1987). Elevated serotonin and reduced dopamine in subregionally divided Huntington's disease striatum. *Ann. Neurol.* 22, 386–389. doi: 10.1002/ana.410220318

Kita, H., and Kitai, S. T. (1988). Glutamate decarboxylase immunoreactive neurons in rat neostriatum: their morphological types and populations. *Brain Res.* 447, 346–352. doi: 10.1016/0006-8993(88)91138-9

Kiyatkin, E. A., and Rebec, G. V. (1996). Dopaminergic modulation of glutamate-induced excitations of neurons in the neostriatum and nucleus accumbens of awake, unrestrained rats. *J. Neurophysiol.* 75, 142–153.

Klapstein, G. J., Fisher, R. S., Zanjani, H., Cepeda, C., Jokel, E. S., Chesselet, M. F., et al. (2001). Electrophysiological and morphological changes in striatal spiny neurons in R6/2 Huntington's disease transgenic mice. *J. Neurophysiol.* 86, 2667–2677.

Klawans, H. C., Paulson, G. W., and Barbeau, A. (1970). Predictive test for Huntington's chorea. *Lancet* 2, 1185–1186. doi: 10.1016/S0140-6736(70)90367-3

Kornhuber, J., Kim, J. S., Kornhuber, M. E., and Kornhuber, H. H. (1984). The cortico-nigral projection: reduced glutamate content in the substantia nigra following frontal cortex ablation in the rat. *Brain Res.* 322, 124–126. doi: 10.1016/0006-8993(84)91189-2

Kreitzer, A. C., and Malenka, R. C. (2005). Dopamine modulation of state-dependent endocannabinoid release and long-term depression in the striatum. *J. Neurosci.* 25, 10537–10545. doi: 10.1523/JNEUROSCI.2959-05.2005

Kreitzer, A. C., and Malenka, R. C. (2007). Endocannabinoid-mediated rescue of striatal LTD and motor deficits in Parkinson's disease models. *Nature* 445, 643–647. doi: 10.1038/nature05506

Kremer, H. P., Roos, R. A., Dingjan, G., Marani, E., and Bots, G. T. (1990). Atrophy of the hypothalamic lateral tuberal nucleus in Huntington's disease. *J. Neuropathol. Exp. Neurol.* 49, 371–382. doi: 10.1097/00005072-199007000-00002

Kung, V. W., Hassam, R., Morton, A. J., and Jones, S. (2007). Dopamine-dependent long term potentiation in the dorsal striatum is reduced in the R6/2 mouse model of Huntington's disease. *Neuroscience* 146, 1571–1580. doi: 10.1016/j.neuroscience.2007.03.036

Laughlin, R. E., Grant, T. L., Williams, R. W., and Jentsch, J. D. (2011). Genetic dissection of behavioral flexibility: reversal learning in mice. *Biol. Psychiatry* 69, 1109–1116. doi: 10.1016/j.biopsych.2011.01.014

Lawrence, A. D., Sahakian, B. J., Hodges, J. R., Rosser, A. E., Lange, K. W., and Robbins, T. W. (1996). Executive and mnemonic functions in early Huntington's disease. *Brain* 119(Pt 5), 1633–1645. doi: 10.1093/brain/119.5.1633

Lee, B., London, E. D., Poldrack, R. A., Farahi, J., Nacca, A., Monterosso, J. R., et al. (2009). Striatal dopamine d2/d3 receptor availability is reduced in methamphetamine dependence and is linked to impulsivity. *J. Neurosci.* 29, 14734–14740. doi: 10.1523/JNEUROSCI.3765-09.2009

Lee, F. J., Xue, S., Pei, L., Vukusic, B., Chery, N., Wang, Y., et al. (2002). Dual regulation of NMDA receptor functions by direct protein-protein interactions with the dopamine D1 receptor. *Cell* 111, 219–230. doi: 10.1016/S0092-8674(02)00962-5

Levine, M. S., Cepeda, C., Hickey, M. A., Fleming, S. M., and Chesselet, M. F. (2004). Genetic mouse models of Huntington's and Parkinson's diseases: illuminating but imperfect. *Trends Neurosci.* 27, 691–697. doi: 10.1016/j.tins.2004.08.008

Levine, M. S., Li, Z., Cepeda, C., Cromwell, H. C., and Altemus, K. L. (1996). Neuromodulatory actions of dopamine on synaptically-evoked neostriatal responses in slices. *Synapse* 24, 65–78. doi: 10.1002/syn.890240102

Lobo, M. K., Yeh, C., and Yang, X. W. (2008). Pivotal role of early

B-cell factor 1 in development of striatonigral medium spiny neurons in the matrix compartment. *J. Neurosci. Res.* 86, 2134–2146. doi: 10.1002/jnr.21666

Maejima, T., Ohno-Shosaku, T., and Kano, M. (2001). Endogenous cannabinoid as a retrograde messenger from depolarized postsynaptic neurons to presynaptic terminals. *Neurosci. Res.* 40, 205–210. doi: 10.1016/S0168-0102(01)00241-3

Mangiarini, L., Sathasivam, K., Seller, M., Cozens, B., Harper, A., Hetherington, C., et al. (1996). Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell* 87, 493–506. doi: 10.1016/S0092-8674(00)81369-0

Maurice, N., Deniau, J. M., Glowinski, J., and Thierry, A. M. (1999). Relationships between the prefrontal cortex and the basal ganglia in the rat: physiology of the cortico-nigral circuits. *J. Neurosci.* 19, 4674–4681.

McLaughlin, B. A., Nelson, D., Erecinska, M., and Chesselet, M. F. (1998). Toxicity of dopamine to striatal neurons *in vitro* and potentiation of cell death by a mitochondrial inhibitor. *J. Neurochem.* 70, 2406–2415. doi: 10.1046/j.1471-4159.1998.70062406.x

Meltzer, L. T., Christoffersen, C. L., and Serpa, K. A. (1997). Modulation of dopamine neuronal activity by glutamate receptor subtypes. *Neurosci. Biobehav. Rev.* 21, 511–518. doi: 10.1016/S0149-7634(96)00030-9

Menalled, L. B. (2005). Knock-in mouse models of Huntington's disease. *NeuroRx* 2, 465–470. doi: 10.1602/neurorx.2.3.465

Mestre, T. A., and Ferreira, J. J. (2012). An evidence-based approach in the treatment of Huntington's disease. *Parkinsonism Relat. Disord.* 18, 316–320. doi: 10.1016/j.parkreldis.2011.10.021

Miller, B. R., Dorner, J. L., Shou, M., Sari, Y., Barton, S. J., Sengelaub, D. R., et al. (2008). Up-regulation of GLT1 expression increases glutamate uptake and attenuates the Huntington's disease phenotype in the R6/2 mouse. *Neuroscience* 153, 329–337. doi: 10.1016/j.neuroscience.2008.02.004

Mintz, I., Hammond, C., Guibert, B., and Leviel, V. (1986). Stimulation of the subthalamic nucleus enhances the release of dopamine in the rat substantia nigra. *Brain Res.* 376, 406–408. doi: 10.1016/0006-8993(86)90209-X

Mizumori, S. J. Y., Ragozzino, K. E., and Cooper, B. G. (2000). Location and head direction representation in the dorsal striatum of rats. *Psychobiology* 28, 441–462.

Mochel, F., Durant, B., Durr, A., and Schiffmann, R. (2011). Altered dopamine and serotonin metabolism in motorically asymptomatic R6/2 mice. *PLoS ONE* 6:e18336. doi: 10.1371/journal.pone.0018336

Murase, S., Grenhoff, J., Chouvet, G., Gonon, F. G., and Svensson, T. H. (1993). Prefrontal cortex regulates burst firing and transmitter release in rat mesolimbic dopamine neurons studied *in vivo*. *Neurosci. Lett.* 157, 53–56. doi: 10.1016/0304-3940(93)90641-W

Naito, A., and Kita, H. (1994). The cortico-nigral projection in the rat: an anterograde tracing study with biotinylated dextran amine. *Brain Res.* 637, 317–322. doi: 10.1016/0006-8993(94)91252-1

Nieoullon, A., Cheramy, A., and Glowinski, J. (1978). Release of dopamine evoked by electrical stimulation of the motor and visual areas of the cerebral cortex in both caudate nuclei and in the substantia nigra in the cat. *Brain Res.* 145, 69–83. doi: 10.1016/0006-8993(78)90797-7

Ortiz, A. N., Kurth, B. J., Osterhaus, G. L., and Johnson, M. A. (2010). Dysregulation of intracellular dopamine stores revealed in the R6/2 mouse striatum. *J. Neurochem.* 112, 755–761. doi: 10.1111/j.1471-4159.2009.06501.x

Ortiz, A. N., Osterhaus, G. L., Lauderdale, K., Mahoney, L., Fowler, S. C., Von Horsten, S., et al. (2012). Motor function and dopamine release measurements in transgenic Huntington's disease model rats. *Brain Res.* 1450, 148–156. doi: 10.1016/j.brainres.2012.02.042

Owen, A. M., Roberts, A. C., Hodges, J. R., Summers, B. A., Polkey, C. E., and Robbins, T. W. (1993). Contrasting mechanisms of impaired attentional set-shifting in patients with frontal lobe damage or Parkinson's disease. *Brain* 116(Pt 5), 1159–1175. doi: 10.1093/brain/116.5.1159

Paoletti, P., Vila, I., Rife, M., Lizcano, J. M., Alberch, J., and Gines, S. (2008). Dopaminergic and glutamatergic signaling crosstalk in Huntington's disease neurodegeneration: the role of p25/cyclin-dependent kinase 5. *J. Neurosci.* 28, 10090–10101. doi: 10.1523/JNEUROSCI.3237-08.2008

Partridge, J. G., Tang, K. C., and Lovinger, D. M. (2000). Regional and postnatal heterogeneity of activity-dependent long-term changes in synaptic efficacy in the dorsal striatum. *J. Neurophysiol.* 84, 1422–1429.

Patel, S., Rademacher, D. J., and Hillard, C. J. (2003). Differential regulation of the endocannabinoids anandamide and 2-arachidonylglycerol within the limbic forebrain by dopamine receptor activity. *J. Pharmacol. Exp. Ther.* 306, 880–888. doi: 10.1124/jpet.103.054270

Pavese, N., Andrews, T. C., Brooks, D. J., Ho, A. K., Rosser, A. E., Barker, R. A., et al. (2003). Progressive striatal and cortical dopamine receptor dysfunction in Huntington's disease: a PET study. *Brain* 126, 1127–1135. doi: 10.1093/brain/awg119

Peinemann, A., Schuller, S., Pohl, C., Jahn, T., Weindl, A., and Kassubek, J. (2005). Executive dysfunction in early stages of Huntington's disease is associated with striatal and insular atrophy: a neuropsychological and voxel-based morphometric study. *J. Neurol. Sci.* 239, 11–19. doi: 10.1016/j.jns.2005.07.007

Petersen, A. A., Larsen, K. E., Behr, G. G., Romero, N., Przedborski, S., Brundin, P., et al. (2001). Brain-derived neurotrophic factor inhibits apoptosis and dopamine-induced free radical production in striatal neurons but does not prevent cell death. *Brain Res. Bull.* 56, 331–335. doi: 10.1016/S0361-9230(01)00580-9

Petersen, A., Chase, K., Puschban, Z., Difiglia, M., Brundin, P., and Aronin, N. (2002a). Maintenance of susceptibility to neurodegeneration following intrastriatal injections of quinolinic acid in a new transgenic mouse model of Huntington's disease. *Exp. Neurol.* 175, 297–300. doi: 10.1006/exnr.2002.7885

Petersen, A., Puschban, Z., Lotharius, J., Nicniocaill, B., Wiekop, P., O'connor, W. T., et al. (2002b). Evidence for dysfunction of the nigrostriatal pathway in the R6/1 line of transgenic Huntington's disease mice. *Neurobiol. Dis.* 11, 134–146. doi: 10.1006/nbdi.2002.0534

Petersen, A., Gil, J., Maat-Schieman, M. L., Bjorkqvist, M., Tanila, H., Araujo, I. M., et al. (2005). Orexin loss in Huntington's disease. *Hum. Mol. Genet.* 14, 39–47. doi: 10.1093/hmg/ddi004

Picconi, B., Passino, E., Sgobio, C., Bonsi, P., Barone, I., Ghiglieri, V., et al. (2006). Plastic and behavioral abnormalities in experimental Huntington's disease: a crucial role for cholinergic interneurons. *Neurobiol. Dis.* 22, 143–152. doi: 10.1016/j.nbd.2005.10.009

Pierce, R. C., and Rebec, G. V. (1995). Iontophoresis in the neostriatum of awake, unrestrained rats: differential effects of dopamine, glutamate and ascorbate on motor- and nonmotor-related neurons. *Neuroscience* 67, 313–324. doi: 10.1016/0306-4522(95)00012-8

Pifl, C., Pichler, L., Kobinger, W., and Hornykiewicz, O. (1988). The dopamine autoreceptor agonist, B-HT 920, preferentially reduces brain dopamine release *in vivo*: biochemical indices of brain dopamine, noradrenaline and serotonin in ventriculocisternal perfusates in the cat. *Eur. J. Pharmacol.* 153, 33–44. doi: 10.1016/0014-2999(88)90585-7

Pisa, M., and Cyr, J. (1990). Regionally selective roles of the rat's striatum in modality-specific discrimination learning and forelimb reaching. *Behav. Brain Res.* 37, 281–292. doi: 10.1016/0166-4328(90)90140-A

Pouladi, M. A., Stanek, L. M., Xie, Y., Franciosi, S., Southwell, A. L., Deng, Y., et al. (2012). Marked differences in neurochemistry and aggregates despite similar behavioural and neuropathological features of Huntington disease in the full-length BACHD and YAC128 mice. *Hum. Mol. Genet.* 21, 2219–2232. doi: 10.1093/hmg/dds037

Racette, B. A., and Perlmutter, J. S. (1998). Levodopa responsive parkinsonism in an adult with Huntington's disease. *J. Neurol. Neurosurg. Psychiatr.* 65, 577–579. doi: 10.1136/jnnp.65.4.577

Ragozzino, M. E. (2007). The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. *Ann. N.Y. Acad. Sci.* 1121, 355–375. doi: 10.1196/annals.1401.013

Raymond, L. A., André, V. M., Cepeda, C., Gladding, C. M., Milnerwood, A. J., and Levine, M. S. (2011). Pathophysiology of Huntington's disease: time-dependent alterations in synaptic and receptor function. *Neuroscience* 15, 252–273. doi: 10.1016/j.neuroscience.2011.08.052

Reading, P. J., Dunnett, S. B., and Robbins, T. W. (1991). Dissociable roles of the ventral, medial and lateral striatum on the acquisition and performance of a complex visual stimulus-response habit. *Behav. Brain Res.* 45, 147–161. doi: 10.1016/S0166-4328(05)80080-4

Richfield, E. K., O'brien, C. F., Eskin, T., and Shoulson, I. (1991). Heterogeneous dopamine receptor changes in early and late Huntington's disease. *Neurosci. Lett.* 132, 121–126. doi: 10.1016/0304-3940(91)90448-3

Ridley, R. M. (1994). The psychology of perserverative and stereotyped behaviour. *Prog. Neurobiol.* 44, 221–231. doi: 10.1016/0301-0082(94)90039-6

Rosas, H. D., Tuch, D. S., Hevelone, N. D., Zaleta, A. K., Vangel, M., Hersch, S. M., et al. (2006). Diffusion tensor imaging in presymptomatic and early Huntington's disease: selective white matter pathology and its relationship to clinical measures. *Mov. Disord.* 21, 1317–1325. doi: 10.1002/mds.20979

Rosenblatt, A. (2007). Neuropsychiatry of Huntington's disease. *Dialogues Clin. Neurosci.* 9, 191–197.

Saka, E., Goodrich, C., Harlan, P., Madras, B. K., and Graybiel, A. M. (2004). Repetitive behaviors in monkeys are linked to specific striatal activation patterns. *J. Neurosci.* 24, 7557–7565. doi: 10.1523/JNEUROSCI.1072-04.2004

Sandstrom, M. I., Steffes-Loydahl, S., Jayaprakash, N., Wolfram-Aduan, A., and Dunbar, G. L. (2010). "Early dysfunction of neural transmission and cognitive processing in Huntington's disease," in *Huntington's Disease - Core Concepts and Current Advances*, ed N. Tunali (New York, NY: InTech), 201–232.

Sawada, H., Ibi, M., Kihara, T., Urushitani, M., Akaike, A., Kimura, J., et al. (1998). Dopamine D2-type agonists protect mesencephalic neurons from glutamate neurotoxicity: mechanisms of neuroprotective treatment against oxidative stress. *Ann. Neurol.* 44, 110–119. doi: 10.1002/ana.410440117

Schiffmann, S. N., and Vanderhaeghen, J. J. (1993). Adenosine A2 receptors regulate the gene expression of striatopallidal and striatonigral neurons. *J. Neurosci.* 13, 1080–1087.

Seneca, S., Fagnart, D., Keymolen, K., Lissens, W., Hasaerts, D., Debulpaep, S., et al. (2004). Early onset Huntington disease: a neuronal degeneration syndrome. *Eur. J. Pediatr.* 163, 717–721. doi: 10.1007/s00431-004-1537-3

Sesack, S. R., and Carr, D. B. (2002). Selective prefrontal cortex inputs to dopamine cells: implications for schizophrenia. *Physiol. Behav.* 77, 513–517. doi: 10.1016/S0031-9384(02)00931-9

Shen, S., Pu, J., Lang, B., and McCaig, C. D. (2011). A zinc finger protein Zfp521 directs neural differentiation and beyond. *Stem Cell Res. Ther.* 2, 20. doi: 10.1186/scrt61

Singla, S., Kreitzer, A. C., and Malenka, R. C. (2007). Mechanisms for synapse specificity during striatal long-term depression. *J. Neurosci.* 27, 5260–5264. doi: 10.1523/JNEUROSCI.0018-07.2007

Slow, E. J., Van Raamsdonk, J., Rogers, D., Coleman, S. H., Graham, R. K., Deng, Y., et al. (2003). Selective striatal neuronal loss in a YAC128 mouse model of Huntington disease. *Hum. Mol. Genet.* 12, 1555–1567. doi: 10.1093/hmg/ddg169

Smith, R., Musleh, W., Akopian, G., Buckwalter, G., and Walsh, J. P. (2001). Regional differences in the expression of corticostriatal synaptic plasticity. *Neuroscience* 106, 95–101. doi: 10.1016/S0306-4522(01)00260-3

Smith, Y., Raju, D. V., Pare, J. F., and Sidibe, M. (2004). The thalamostriatal system: a highly specific network of the basal ganglia circuitry. *Trends Neurosci.* 27, 520–527. doi: 10.1016/j.tins.2004.07.004

Spampanato, J., Gu, X., Yang, X. W., and Mody, I. (2008). Progressive synaptic pathology of motor cortical neurons in a BAC transgenic mouse model of Huntington's disease. *Neuroscience* 157, 606–620. doi: 10.1016/j.neuroscience.2008.09.020

Spektor, B. S., Miller, D. W., Hollingsworth, Z. R., Kaneko, Y. A., Solano, S. M., Johnson, J. M., et al. (2002). Differential D1 and D2 receptor-mediated effects on immediate early gene induction in a transgenic mouse model of Huntington's disease. *Brain Res. Mol. Brain Res.* 102, 118–128. doi: 10.1016/S0169-328X(02)00216-4

Spencer, J. P., and Murphy, K. P. (2000). Bi-directional changes in synaptic plasticity induced at corticostriatal synapses *in vitro*. *Exp. Brain Res.* 135, 497–503. doi: 10.1007/s002210000523

Spokes, E. G. (1980). Neurochemical alterations in Huntington's chorea: a study of post-mortem brain tissue. *Brain* 103, 179–210. doi: 10.1093/brain/103.1.179

Squitieri, F., Cannella, M., Piorcellini, A., Brusa, L., Simonelli, M., and Ruggieri, S. (2001). Short-term effects of olanzapine in Huntington disease. *Neuropsychiatry Neuropsychol. Behav. Neurol.* 14, 69–72.

Stack, E. C., Dedeoglu, A., Smith, K. M., Cormier, K., Kubilus, J. K., Bogdanov, M., et al. (2007). Neuroprotective effects of synaptic modulation in Huntington's disease R6/2 mice. *J. Neurosci.* 27, 12908–12915. doi: 10.1523/JNEUROSCI.4318-07.2007

Stalnaker, T. A., Takahashi, Y., Roesch, M. R., and Schoenbaum, G. (2009). Neural substrates of cognitive inflexibility after chronic cocaine exposure. *Neuropharmacology* 56(Suppl. 1), 63–72. doi: 10.1016/j.neuropharm.2008.07.019

Steiner, H., and Gerfen, C. R. (1999). Enkephalin regulates acute D2 dopamine receptor antagonist-induced immediate-early gene expression in striatal neurons. *Neuroscience* 88, 795–810. doi: 10.1016/S0306-4522(98)00241-3

Stout, J. C., Rodawalt, W. C., and Siemers, E. R. (2001). Risky decision making in Huntington's disease. *J. Int. Neuropsychol. Soc.* 7, 92–101. doi: 10.1017/S1355617701711095

Suzuki, M., Desmond, T. J., Albin, R. L., and Frey, K. A. (2001). Vesicular neurotransmitter transporters in Huntington's disease: initial observations and comparison with traditional synaptic markers. *Synapse* 41, 329–336. doi: 10.1002/syn.1089

Tang, T. S., Chen, X., Liu, J., and Bezprozvanny, I. (2007). Dopaminergic signaling and striatal neurodegeneration in Huntington's disease. *J. Neurosci.* 27, 7899–7910. doi: 10.1523/JNEUROSCI.1396-07.2007

Tedroff, J., Ekesbo, A., Sonesson, C., Waters, N., and Carlsson, A. (1999). Long-lasting improvement following (-)-OSU6162 in a patient with Huntington's disease. *Neurology* 53, 1605–1606. doi: 10.1212/WNL.53.7.1605

Tepper, J. M., Wilson, C. J., and Koos, T. (2008). Feedforward and feedback inhibition in neostriatal GABAergic spiny neurons. *Brain Res. Rev.* 58, 272–281. doi: 10.1016/j.brainresrev.2007.10.008

The Huntington's Disease Collaborative Research Group. (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 72, 971–983. doi: 10.1016/0092-8674(93)90585-E

Tobin, A. J., and Signer, E. R. (2000). Huntington's disease: the challenge for cell biologists. *Trends Cell Biol.* 10, 531–536. doi: 10.1016/S0962-8924(00)01853-5

Tzavos, A., Jih, J., and Ragozzino, M. E. (2004). Differential effects of M1 muscarinic receptor blockade and nicotinic receptor blockade in the dorsomedial striatum on response reversal learning. *Behav. Brain Res.* 154, 245–253. doi: 10.1016/j.bbr.2004.02.011

Van Oostrom, J. C., Dekker, M., Willemsen, A. T., De Jong, B. M., Roos, R. A., and Leenders, K. L. (2009). Changes in striatal dopamine D2 receptor binding in pre-clinical Huntington's disease. *Eur. J. Neurol.* 16, 226–231. doi: 10.1111/j.1468-1331.2008.02390.x

Venuto, C. S., McGarry, A., Ma, Q., and Kieburtz, K. (2012). Pharmacologic approaches to the treatment of Huntington's disease. *Mov. Disord.* 27, 31–41. doi: 10.1002/mds.23953

Vijayraghavan, S., Wang, M., Birnbaum, S. G., Williams, G. V., and Arnsten, A. F. (2007). Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nat. Neurosci.* 10, 376–384. doi: 10.1038/nn1846

Vincent, S., Hokfelt, T., Christensson, I., and Terenius, L. (1982). Immunohistochemical evidence for a dynorphin immunoreactive striato-nigral pathway. *Eur. J. Pharmacol.* 85, 251–252. doi: 10.1016/0014-2999(82)90477-0

Von Horsten, S., Schmitt, I., Nguyen, H. P., Holzmann, C., Schmidt, T., Walther, T., et al. (2003). Transgenic rat model of Huntington's disease. *Hum. Mol. Genet.* 12, 617–624. doi: 10.1093/hmg/ddg075

Vonsattel, J. P., and Difiglia, M. (1998). Huntington disease. *J. Neuropathol. Exp. Neurol.* 57, 369–384. doi: 10.1097/00005072-199805000-00001

Wang, E., Cepeda, C., and Levine, M. S. (2012). Cognitive deficits in Huntington's disease: insights from animal models. *Curr. Trans. Geriatr. Gerontol. Rep.* 1, 29–38. doi: 10.1007/s13670-011-0005-y

Wang, H., Chen, X., Li, Y., Tang, T. S., and Bezprozvanny, I. (2010). Tetrabenazine is neuroprotective in Huntington's disease mice. *Mol. Neurodegener.* 5, 18. doi: 10.1186/1750-1326-5-18

Wang, J., and O'donnell, P. (2001). D(1) dopamine receptors potentiate nmda-mediated excitability increase in layer V prefrontal cortical pyramidal neurons. *Cereb. Cortex* 11, 452–462. doi: 10.1093/cercor/11.5.452

Wang, S. H., and Morris, R. G. (2010). Hippocampal-neocortical interactions in memory formation, consolidation, and reconsolidation. *Annu. Rev. Psychol.* 61, 49–79, C41–C44.

Wang, Z., Kai, L., Day, M., Ronesi, J., Yin, H. H., Ding, J., et al. (2006).

Dopaminergic control of corticostriatal long-term synaptic depression in medium spiny neurons is mediated by cholinergic interneurons. *Neuron* 50, 443–452. doi: 10.1016/j.neuron.2006.04.010

Waszczak, B. L., Martin, L. P., Finlay, H. E., Zahr, N., and Stellar, J. R. (2002). Effects of individual and concurrent stimulation of striatal D1 and D2 dopamine receptors on electrophysiological and behavioral output from rat basal ganglia. *J. Pharmacol. Exp. Ther.* 300, 850–861. doi: 10.1124/jpet.300.3.850

Weeks, R. A., Piccini, P., Harding, A. E., and Brooks, D. J. (1996). Striatal D1 and D2 dopamine receptor loss in asymptomatic mutation carriers of Huntington's disease. *Ann. Neurol.* 40, 49–54. doi: 10.1002/ana.410400110

Wersinger, C., Chen, J., and Sidhu, A. (2004). Bimodal induction of dopamine-mediated striatal neurotoxicity is mediated through both activation of D1 dopamine receptors and autoxidation. *Mol. Cell. Neurosci.* 25, 124–137. doi: 10.1016/j.mcn.2003.10.002

Whishaw, I. Q., Mittleman, G., Bunch, S. T., and Dunnett, S. B. (1987). Impairments in the acquisition, retention and selection of spatial navigation strategies after medial caudate-putamen lesions in rats. *Behav. Brain Res.* 24, 125–138. doi: 10.1016/0166-4328(87)90250-6

Williams, G. V., and Castner, S. A. (2006). Under the curve: critical issues for elucidating D1 receptor function in working memory. *Neuroscience* 139, 263–276. doi: 10.1016/j.neuroscience.2005.09.028

Wu, J., Tang, T., and Bezprozvanny, I. (2006). Evaluation of clinically relevant glutamate pathway inhibitors in *in vitro* model of Huntington's disease. *Neurosci. Lett.* 407, 219–223. doi: 10.1016/j.neulet.2006.08.036

Zeron, M. M., Fernandes, H. B., Krebs, C., Shehadeh, J., Wellington, C. L., Leavitt, B. R., et al. (2004). Potentiation of NMDA receptor-mediated excitotoxicity linked with intrinsic apoptotic pathway in YAC transgenic mouse model of Huntington's disease. *Mol. Cell. Neurosci.* 25, 469–479. doi: 10.1016/j.mcn.2003.11.014

Zeron, M. M., Hansson, O., Chen, N., Wellington, C. L., Leavitt, B. R., Brundin, P., et al. (2002). Increased sensitivity to N-methyl-D-aspartate receptor-mediated excitotoxicity in a mouse model of Huntington's disease. *Neuron* 33, 849–860. doi: 10.1016/S0896-6273(02)00615-3

Zhou, F. M., Wilson, C. J., and Dani, J. A. (2002). Cholinergic interneuron characteristics and nicotinic properties in the striatum. *J. Neurobiol.* 53, 590–605. doi: 10.1002/neu.10150

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 08 March 2013; accepted: 13 June 2013; published online: 04 July 2013.*

*Citation: Chen JY, Wang EA, Cepeda C and Levine MS (2013) Dopamine imbalance in Huntington's disease: a mechanism for the lack of behavioral flexibility. Front. Neurosci. 7:114. doi: 10.3389/fnins.2013.00114*

*This article was submitted to Frontiers in Decision Neuroscience, a specialty of Frontiers in Neuroscience.*