



ARTICLE

Activation of D1 receptors affects human reactivity and flexibility to valued cues

Alexander Soutschek^{1,2}, Rouba Kozak³, Nicholas de Martinis⁴, William Howe⁵, Christopher J. Burke², Ernst Fehr^{6,6}, Alexander Jetter⁷ and Philippe N. Tobler^{2,6}

Reward-predicting cues motivate goal-directed behavior, but in unstable environments humans must also be able to flexibly update cue-reward associations. While the capacity of reward cues to trigger motivation ('reactivity') as well as flexibility in cue-reward associations have been linked to the neurotransmitter dopamine in humans, the specific contribution of the dopamine D1 receptor family to these behaviors remained elusive. To fill this gap, we conducted a randomized, placebo-controlled, double-blind pharmacological study testing the impact of three different doses of a novel D1 agonist (relative to placebo) on reactivity to reward-predicting cues (Pavlovian-to-instrumental transfer) and flexibility of cue-outcome associations (reversal learning). We observed that the impact of the D1 agonist crucially depended on baseline working memory functioning, which has been identified as a proxy for baseline dopamine synthesis capacity. Specifically, increasing D1 receptor stimulation strengthened Pavlovian-to-instrumental transfer in individuals with high baseline working memory capacity. In contrast, higher doses of the D1 agonist improved reversal learning only in individuals with low baseline working memory functioning. Our findings suggest a crucial and baseline-dependent role of D1 receptor activation in controlling both cue reactivity and the flexibility of cue-reward associations.

Neuropsychopharmacology (2020) 45:780–785; <https://doi.org/10.1038/s41386-020-0617-z>

INTRODUCTION

Learned associations between environmental stimuli and rewards are crucial determinants of human behavior. For example, seeing the sign of our favorite coffee chain may trigger the desire to go in and enjoy a cup. Environmental cues inform us about which behaviors will lead to desired outcomes and can thus motivate continued performance of previously rewarded behavior ('reactivity'). In unstable environments, however, cue-reward associations are subject to change and perseveration of previously rewarded behaviors can become dysfunctional. Humans thus also need the capacity to flexibly update stimulus-reward associations and learn new ones that are more likely to yield rewards ('flexibility'). Dysfunctions in both reactivity and flexibility to cue-reward associations belong to the core symptoms of several psychiatric disorders, including addiction or schizophrenia [1–3]. It is thus crucial to obtain a better understanding of the neural mechanisms regulating the impact of cue-reward associations on human behavior.

The neurotransmitter dopamine has been proposed to play a central role in mediating both reactivity to reward cues and the capacity to flexibly update the underlying associative framework by controlling activity in a broad cortico-striatal network [4, 5]. Indeed, both cortical and striatal dopamine may mediate the impact of value-related cues on behavior. Dopamine D1 receptors (D1R) dominate in the direct striatal "go" pathway and in prefrontal cortex [6]. Prefrontal D1R availability was related to

reward learning [7], and prefrontal D1R activation may enhance cue-triggered goal-directed behavior by strengthening prefrontal goal representations [8]. Likewise, according to a recent theoretical model on striatal dopamine [9], striatal D1R activity both facilitates the updating of cue-outcome associations after unexpected positive feedback and increases instrumental behavior in response to reward-predicting cues. These assumptions are supported by evidence showing that modulating dopaminergic activity (targeting either unspecifically D1R/D2Rs or specifically D2Rs) changes reactivity to reward-predicting cues [10–12] as well as re-learning of stimulus-reward associations [13, 14]. In contrast, the specific contribution of D1R activation on behaviors supported by cue-reward associations in humans has remained unknown. Findings from animal studies suggest that D1R activation facilitates instrumental responses to reward-predicting cues [15, 16]. In the domain of reversal learning, prefrontal D1R stimulation might improve updating of maladaptive cue-outcome relationships [17, 18], but studies on striatal D1Rs observed no such effect [19, 20]. In humans, the assumption that the impact of cue-outcome associations on behavior is mediated by D1Rs is supported by both genetic and positron emission tomography (PET) evidence [21–23]. However, due to the correlative nature of these studies, the causal involvement of D1R activation remains elusive.

In humans, the causal contribution of D1Rs has not been as thoroughly investigated so far due to the relative paucity of

¹Department of Psychology, Ludwig Maximilian University Munich, Munich, Germany; ²Zurich Center for Neuroeconomics, Department of Economics, University of Zurich, Zurich, Switzerland; ³Takeda Pharmaceuticals International, Cambridge, MA, USA; ⁴Praxis Precision Medicines, Cambridge, MA, USA; ⁵School of Neuroscience, Virginia Tech, Blacksburg, VA, USA; ⁶Neuroscience Center Zurich, University of Zurich, Swiss Federal Institute of Technology Zurich, Zurich, Switzerland and ⁷Department of Clinical Pharmacology and Toxicology, University Hospital Zurich, University of Zurich, Zurich, Switzerland

Correspondence: Alexander Soutschek (Alexander.Soutschek@psy.lmu.de)

These authors contributed equally: Alexander Jetter, Philippe N. Tobler

Received: 13 September 2019 Revised: 2 December 2019 Accepted: 9 January 2020

Published online: 21 January 2020

selective D1 agonists. Using the novel D1 agonist PF-06412562 with a non-catechol structure [24, 25], the current study aims to address this gap and to provide the first test in humans of the capacity of D1Rs to mediate cue reactivity, as well as the ability to update cue-outcome contingencies in dynamic contexts. It is important to note that the relationship between dopaminergic activity and various aspects of behavior follows a non-linear, e.g., inverted u-shaped, relationship [26–29]. The impact of dopaminergic manipulations on these behaviors is predicted to crucially depend on baseline dopamine synthesis capacity [14], and previous studies have shown that both striatal and prefrontal dopaminergic activity are related to working memory function [26, 30–33]. Furthermore, working memory capacity (WMC) represents a proxy for dopamine synthesis capacity [34, 35]. Thus in the present study, we modeled the impact of the D1R stimulation on behavior as function of individual WMC measured before drug intake.

MATERIALS AND METHODS

Participants

The study protocol was approved by the Research Ethics Committee (2016-01693) of the canton of Zurich as well as by the Swiss agency for therapeutic products (Swissmedic, 2017DR1021). The study was also registered on ClinicalTrials.gov (NCT03181841). All participants gave written informed consent before the start of screening. From the screened volunteers, 120 participants (59 females, mean age = 22.57 years, range 18–28) fulfilled all inclusion criteria and were invited to the main experimental session. Participants received 480 Swiss francs for their participation and a monetary bonus depending on their choices (see below).

Study design and procedures

The study was a monocentric, randomized, double-blind, and placebo-controlled clinical phase 1 trial. The 120 participants were randomly assigned to one of four experimental arms, one receiving placebo (lactose) and the others a single dose of the D1 agonist PF-06412562 (either 6, 15, or 30 mg). PF-06412562 is selective for D1Rs compared with D2Rs (Supplementary methods). The D1 agonist was well tolerated by participants (Table S1), with only mild to moderate side effects including tiredness (39 participants in total, including placebo group), headache [34], nausea [17], vomiting [4], and dizziness [4]. Participants were also unable to distinguish between drug and placebo, $\chi^2(1) < 1, p = 0.52$.

The study was conducted in the SNS lab at the University hospital Zurich and entailed three sessions: in session 1 (duration = 1 h), participants were screened for exclusion criteria (for details, see [36]) and filled in questionnaires for reward sensitivity (BIS/BAS [37]), verbal intelligence (MWT-B [38]), and impulsivity (BIS-11 [39]). We also assessed participants’ baseline time and risk preferences (details are reported in [36]) as well as working memory performance using the digit span backward test. For none of these baseline measures we found significant differences between the experimental groups (Table 1).

At the start of session 2 (7–21 days after session 1; duration = 9 h), participants again performed the digit span task backward to measure working memory performance before drug intake [35]. After drug administration, participants stayed in the lab and were monitored for potential side effects. Pharmacokinetic samples of PF-06412562 and its metabolite PF-06663872 were measured exactly 4 h and 8 h after drug intake. Five hours after drug intake (i.e., around the first absorption peak), participants performed the following computerized tasks in balanced, pseudo-random order: time preference task, risk preference task, effort preference task, reversal learning task, Pavlovian-to-instrumental transfer (PIT) task, and exploration/exploitation task. In this article, we report the

Table 1. Demographics and mean baseline measures (in session 1) as function of administered dose. Standard deviations are in parentheses. Results of group comparisons using Pearson chi-square (for sex) and ANOVAs (all other variables) are listed under “p value”.

	Placebo (n = 30)	6 mg (n = 30)	15 mg (n = 30)	30 mg (n = 30)	p value
Age	23.0 (2.1)	22.9 (2.5)	22.1 (2.1)	22.3 (2.1)	0.30
Sex	16 female	18 female	12 female	15 female	0.48
Weight	71 (12)	75 (17)	67 (11)	68 (11)	0.11
BMI	23.0 (3.2)	23.9 (3.7)	22.7 (2.3)	22.6 (2.5)	0.28
BIS/BAS	5.8 (0.6)	5.7 (0.6)	5.9 (0.6)	6.0 (0.6)	0.25
BIS-11	1.9 (0.3)	1.8 (0.2)	1.9 (0.2)	1.9 (0.2)	0.21
MWTB	29.7 (3.2)	28.6 (4.4)	27.7 (3.6)	29.8 (2.6)	0.07
Digit span	5.6 (2.2)	6.3 (2.0)	6.3 (1.2)	5.9 (2.1)	0.41

results only of the PIT and the reversal learning task as they measure the impact of cue predicting motivationally relevant outcomes on behavior. In contrast, the time, risk, and effort preference tasks are reported in a separate article [36] because they measure economic preferences rather than reactivity or flexibility to reward-predicting cues. The exploration/exploitation task suffered from a problem in the script in 25 participants, thus precluding interpretable findings. Finally, in session 3 (6–8 days after session 2; duration = 0.5 h), participants were again screened for side effects and performed time and risk preference tasks as well the digit span backward task.

Behavioral assessments

PIT task. The PIT task followed the standard three-phase PIT design, where an instrumental conditioning phase (Fig. 1a) is followed by a Pavlovian phase (Fig. 1b) and, finally, a transfer-test phase (Fig. 1c) [40, 41]. Participants performed a computerized version of the PIT task [40, 41] that was programmed using the Cogent toolbox in Matlab. Procedures for the instrumental conditioning and the Pavlovian phase are described in the Supplementary methods. In the critical transfer-test phase, each cue appeared ten times. Importantly, the four different CS types allowed us to differentiate between specific and general PIT [42]. Specific PIT refers to the behavioral impact of cues associated with a particular reward. In the current design, we measured specific PIT with the number of key presses for a reward (e.g., smarties) during the presentation of the CS paired with this reward (CS_{reward}; e.g., the CS for smarties) relative to the key presses during the CS associated with the other reward (CS_{other reward}; e.g., the CS for popcorn). In contrast, general PIT is the behavioral impact of conditioned cues associated with unrelated appetitive outcomes. As measure of general PIT, we calculated the number of key presses (which were linked to popcorn or smarties) during presentation of the CS for cashews (CS_{reward}) relative to the CS associated with no outcomes (CS_{no reward}). In addition to the transfer-test phase involving the Pavlovian cues, participants also performed ten trials where only a white square was presented without Pavlovian cues, similar to the instrumental phase. This allowed us to test for a potential impact of the D1 agonist on noncued behavioral responses. This phase was performed either before or after the PIT phase (counterbalanced across participants).

Reversal learning task. We adopted and recoded a task design that allows distinguishing between reward and punishment reversal learning [13, 14, 43]. On each trial, participants viewed two visual stimuli, a face and a landscape picture, on the left and right screen sides. One of these stimuli was associated with reward

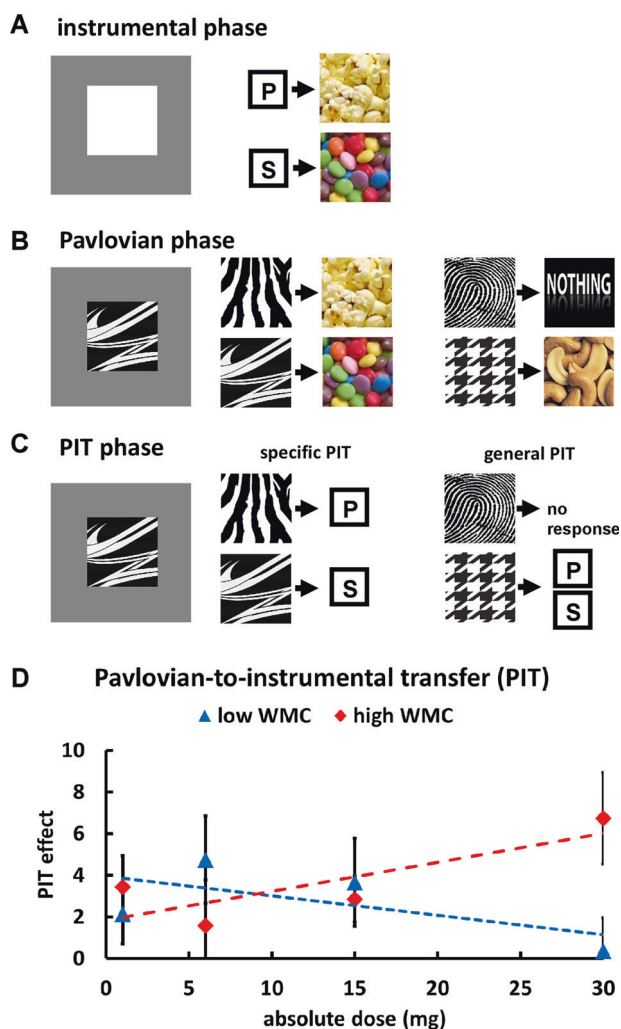


Fig. 1 Task design and results for Pavlovian-to-instrumental transfer (PIT) task. **a** In the initial instrumental conditioning phase, participants learned to associate distinct behavioral responses (key presses) with distinct primary reinforcers (popcorn and smarties). **b** During Pavlovian conditioning, participants associated four different predictive cues with one of four outcome types (popcorn, smarties, cashew nuts, and nothing). **c** In the final transfer-test PIT phase, we measured key presses to the Pavlovian cues. **d** In individuals with high baseline WMC, stronger D1R stimulation increased PIT (independently of PIT type), indicating enhanced cue reactivity. No significant drug effects occurred in low WMC individuals.

(a happy smiley with a “+100” sign), the other with punishment (a sad smiley with a “-100” sign). At the start of each trial, one of the stimuli was selected by the computer (indicated by a white frame), and within 1500 ms participants had to predict whether the selected picture was associated with a reward or loss by pressing the left or right arrow key (key-response assignment was counterbalanced across participants). Following choice, the actual outcome was presented for 1500 ms (Fig. 2a). Participants were instructed that the displayed outcomes did not depend on their performance (i.e., whether they made a correct or wrong prediction) but only on the outcome that was deterministically associated with the picture (face or landscape) selected by the computer. The stimulus–outcome contingencies (i.e., assignment of faces and landscapes to rewards and punishments) reversed after 5–9 consecutive correct predictions. On such reversal trials, participants thus experienced unexpected punishment (after selection of a previously rewarded stimulus) or unexpected reward (after selection of a previously punished stimulus).

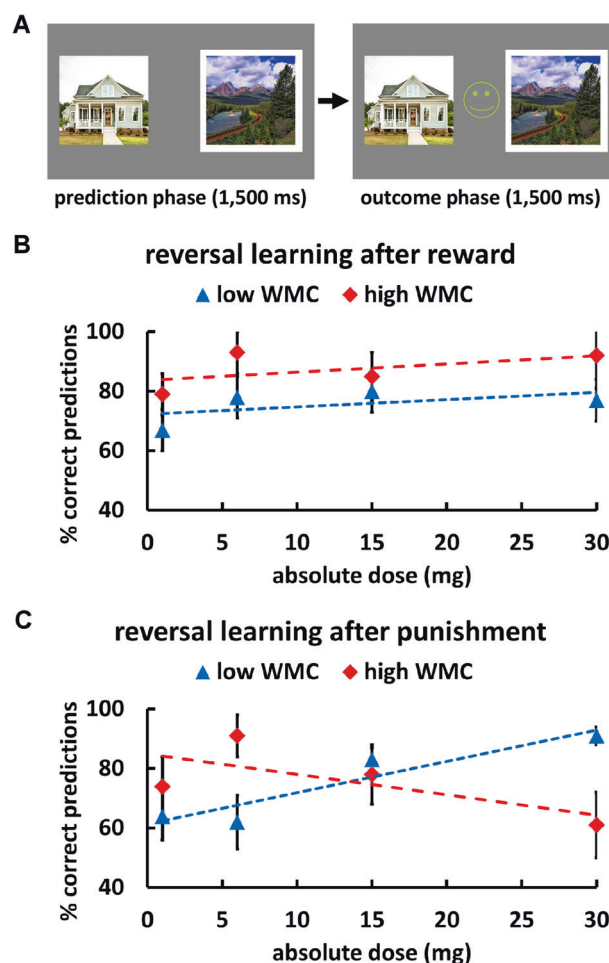


Fig. 2 Task design and results for the reversal learning task. **a** In each trial, one of two stimuli (a house and a landscape) was selected by the computer (indicated by a white frame) and participants had to predict whether the selected stimulus was associated with a reward (a happy smiley) or with punishment (a sad smiley). After participants’ prediction, the outcome was presented. Stimulus–outcome associations reversed after variable numbers of correct predictions. **b** D1R stimulation had no impact on predictions of the correct outcome following reversal trials with unexpected rewards. **c** Following unexpected punishment, increasing D1R activation improved behavior selectively in individuals with low baseline WMC.

Accuracy on the trials directly following these unexpected outcomes (reversal trials) reflects how well participants updated Pavlovian stimulus–outcome associations after either unexpected rewards or unexpected punishments. Following a practice block of 20 trials, participants performed a total of 120 trials of the reversal learning task.

Data analysis

We used Matlab R2016b (MathWorks, Natick, MA) and IBM SPSS Statistics 22 to analyze the data. The alpha threshold was set to 5% (two-tailed) for all analyses. Data in the PIT and reversal learning task were analyzed with mixed generalized linear models (MGLMs) with Absolute dose as measure of drug effects and including (median-split) predictors for WMC as proxy for baseline dopamine levels (SI Methods). Cohen’s *d* is reported as measure of effect size. We also aimed to assess drug effects in an individualized fashion. We therefore also conducted all MGLMs with Relative dose (absolute dose divided by individual body weight in kg) or Plasma concentration (mean of pharmacokinetic

samples for PF-06412562 4 h and 8 h after drug administration) instead of Absolute Dose. While the predictor Absolute dose makes the (implausible) assumption that the impact of a given dose is constant for all participants, individualized measures such as Relative dose and Plasma concentration might more realistically model the effective impact of a given dose on an individual's metabolism.

Finally, to assess the robustness of our findings, we recomputed the (non-Bayesian) MGLMs also as Bayesian mixed models using the brms package in R 3.6.1 (SI Methods).

RESULTS

D1R stimulation increases cue reactivity

In the PIT task, participants successfully learned key-outcome associations in the instrumental phase (mean = 99% correct key-snack associations in the explicit test after the first half and at the end of the instrumental phase) as well as cue-outcome associations in the Pavlovian phase (mean = 97% correct cue-snack associations in the explicit tests in the Pavlovian phase). MGLMs on these measures with predictors for Absolute dose, WMC, and the interaction showed no significant results, all $t < 1.56$, all $p > 0.12$, suggesting that the D1 agonist did not affect instrumental or Pavlovian reward learning. Also in the non-cued test phase after Pavlovian conditioning (where no cues were displayed), we observed no drug effects on the number of key presses, all $t < 1$, all $p > 0.70$; mean = 12.9 key presses during the 10 s presentation time for the white square in that phase.

To test the impact of the D1 agonist on Pavlovian-to-instrumental transfer, we regressed key presses during cue presentation in the transfer-test phase on predictors for Absolute dose (0, 6, 15, or 30 mg), WMC (low vs. high WMC), PIT type (specific vs. general), and CS (CS_{reward} vs. $CS_{\text{other reward}}/CS_{\text{no reward}}$). We observed a main effect of CS, $\beta = 0.078$, $t(464) = 2.89$, $p = 0.004$, $d = 0.28$, which was modulated by a PIT type \times CS interaction, $\beta = -0.056$, $t(464) = 3.02$, $p = 0.003$, $d = 0.27$, suggesting that specific PIT (more key presses for CS_{reward} than $CS_{\text{other reward}}$) was significantly stronger than general PIT (more key presses for CS_{reward} than $CS_{\text{no reward}}$). The WMC \times CS interaction was not significant, $\beta = -0.043$, $t(464) = 1.57$, $p = 0.114$, $d = 0.09$, providing no evidence that PIT varied as function of baseline WMC in the placebo group. Importantly, the D1 agonist modulated PIT as a function of baseline WMC, Absolute dose \times WMC \times CS, $\beta = 0.004$, $t(464) = 2.37$, $p = 0.018$, $d = 0.18$, with this effect not significantly differing between specific and general PIT, PIT type \times Absolute dose \times WMC \times CS, $\beta = -0.001$, $t = 1.28$, $p = 0.20$, $d = 0.07$ (Fig. 1d and Table S3). The three-way Absolute dose \times WMC \times CS interaction was robust to modeling drug effects by Relative dose, $\beta = 0.217$, $t(464) = 2.66$, $p = 0.008$, $d = 0.18$, and showed a marginally significant effect for Plasma concentration, $\beta = 0.001$, $t(464) = 1.95$, $p = 0.051$, $d = 0.16$. This result pattern was also robust to employing Bayesian mixed models (SI Results). This suggests baseline-dependent effects of the D1 agonist on PIT.

To resolve the three-way interaction, we computed separate MGLMs for the low and high WMC groups. In the low WMC group there were no significant effects including the predictor Absolute dose, all $t < 1.44$, all $p > 0.15$, all $d < 0.14$. In the high WMC group, in contrast, we observed a significant Absolute dose \times CS interaction, $\beta = 0.004$, $t(228) = 2.57$, $p = 0.011$, $d = 0.24$. Again, this effect was robust to modeling drug effects by Relative dose, $\beta = 0.301$, $t(228) = 3.02$, $p = 0.003$, $d = 0.25$, and Plasma concentration, $\beta = 0.001$, $t(228) = 2.12$, $p = 0.035$, $d = 0.21$. This suggests that in high WMC individuals increasing doses of PF-06412562 enhanced cue reactivity. There was no evidence for dissociable drug effects on specific vs. general PIT, all $t < 1.38$, all $p > 0.16$, all $d < 0.14$. Taken together, our data suggest strong D1R activation to enhance cue reactivity in individuals with high WMC.

D1R stimulation improves reversal learning after unexpected punishment

We analyzed data with an MGLM, in which binary correct predictions (0 = incorrect, 1 = correct prediction) on reversal trials were regressed on predictors for Absolute dose, WMC, Valence (reward vs. punishment), and the interactions between these factors. Participants with high WMC showed better reversal learning performance than low WMC individuals in the baseline placebo group, $\beta = 0.464$, $t(1105) = 3.09$, $p = 0.002$, $d = 0.28$. We observed a marginally significant Absolute dose \times WMC \times Valence interaction, $\beta = 0.015$, $t(1105) = 1.94$, $p = 0.053$, $d = 0.17$, which was corroborated when using Relative dose, $\beta = 1.134$, $t(1105) = 2.26$, $p = 0.02$, $d = 0.21$, or Plasma concentration, $\beta = 0.002$, $t(1105) = 2.13$, $p = 0.03$, $d = 0.20$, as individualized measures of drug effects. Again, this result pattern was also robust to employing Bayesian mixed models (SI Results). Thus, the impact of the D1 agonist on reversal learning depended on both baseline WMC and outcome valence. To resolve this interaction effect, we computed separate MGLMs for low and high WMC groups, as for the PIT task. In the low WMC group, we found a marginally significant Absolute dose \times Valence interaction, $\beta = -0.019$, $t(557) = 1.93$, $p = 0.054$, $d = 0.17$, which was significant when modeling drug effects with Relative dose, $\beta = -1.486$, $t(557) = 2.23$, $p = 0.026$, $d = 0.20$, or Plasma concentration, $\beta = -0.003$, $t(557) = 1.97$, $p = 0.049$, $d = 0.17$. While the D1 agonist did not affect performance if a reversal was associated with a rewarding outcome, $\beta = -0.001$, $t < 1$, $p = 0.94$, $d = 0.01$ (Fig. 2b), following punishment participants with low baseline WMC showed more correct predictions with increasing doses of the D1 agonist, $\beta = 0.035$, $t(261) = 2.35$, $p = 0.02$, $d = 0.21$ (Fig. 2c). In contrast, in the high WMC group, we observed no significant main effect of Absolute dose or Absolute dose \times Valence interaction, both $\beta < 0.012$, both $t < 1$, both $p > 0.33$, $d = 0.08$. These findings suggest that D1R stimulation selectively improves reversal learning after punishment and in individuals with low dopamine baseline levels.

DISCUSSION

The current study investigated the causal impact of pharmacologically stimulating D1Rs on reactivity and flexibility to cues predicting motivationally relevant outcomes. Our results suggest a crucial function of D1Rs for both mediating the impact of Pavlovian cues on instrumental behavior and the flexible re-learning of cue-outcome associations.

In the PIT task, we observed that in individuals with high WMC (as proxy for baseline dopamine synthesis capacity) enhanced D1R activation increased the reactivity to reward-associated cues. Noteworthy, there was no evidence for dissociable D1R stimulation effects on specific versus general PIT. Specific and general PIT have been related to motivating different aspects of goal-directed behavior [44] and may be implemented by dissociable neural structures, NAcc shell and core, respectively [45, 46]. However, both NAcc shell and core contain D1Rs, which mediate the impact of learned stimulus-reward associations on behavior. Consistent with the anatomy, our findings suggest that D1Rs are indeed causally involved in triggering outcome-specific and unspecific PIT. Our result that stronger D1R stimulation increases PIT is consistent with animal findings that D1R activity enhances cue-induced reward seeking [15, 16] and support a theoretical account ascribing D1Rs a role for mediating instrumental responses triggered by reward-predicting stimuli [9]. However, given that this effect occurred only in individuals with high, rather than low, baseline dopamine synthesis capacity, strong levels of D1R activation appear necessary to enhance reactivity to reward-predicting cues. In any case, the current data provide first direct evidence for the hypothesized D1R involvement in human PIT.

Besides translating reward cues into goal-directed behavior, D1 neurotransmission is also involved in updating stimulus–outcome associations [18, 47]. Our reversal learning data concur in that increasing D1R activation in individuals with low baseline WMC improved reversal learning, particularly after unexpected punishment (we note, though, that drug effects on reversal learning were significant only when employing individualized measures of effective dose). Theoretical accounts assume that reward versus punishment learning depends on the balance between the D1-mediated direct “go”-pathway and the D2-mediated indirect “nogo”-pathway, whereby a dominance of the direct over the indirect pathway facilitates learning from rewards relative to punishment [9, 48]. Empirical findings support this model by showing that compared with punishment-based reversal learning, reward-based reversal learning is inhibited by the D2 agonist bromocriptine (which reduces the impact of the direct path) [14] as well as improved by the D2 antagonist sulpiride [13]. However, these results appear to be limited to individuals with high baseline dopamine levels, whereas in low baseline individuals strengthening the indirect “nogo” over the direct “go” pathway with bromocriptine impaired punishment compared with reward reversal learning [14]. Interestingly, our results mirror these previous findings [14], because we observed that strengthening the direct over the indirect pathway with a D1 agonist improved punishment compared to reward reversal learning. While our results thus appear consistent with the existing literature on dopaminergic modulation of outcome-specific reversal learning, it remains open how they can be reconciled with theoretical models on D1/D2 balance in reversal learning [9, 48]. One might speculate that under conditions of high baseline dopamine synthesis capacity the direct and indirect pathways are in balance, whereas for low baseline dopamine levels the direct pathway dominates over the indirect one, facilitating reward learning. In fact, in our sample low WMC individuals were numerically (albeit non-significantly) better in reward than punishment reversal learning under placebo. In this case, further upregulation of D1R activation in the direct path might *reduce* the dominance of the direct path due to overactivation of D1Rs (assuming an inverted u-shaped function), thereby relatively facilitating punishment over reward reversal learning. Furthermore, while the OpaL model focusses on the role of striatal D1Rs for reversal learning, reversal learning depends also on prefrontal circuits [49–52], leaving the possibility that the observed effects are caused by prefrontal, rather than striatal, dopaminergic activity. Enhancing the excitability of the prefrontal cortex with noninvasive brain stimulation was linked to improved reversal learning [50], and also animal studies found reversal learning to be affected by D1R stimulation in the prefrontal cortex [17, 18]. The D1 agonist may thus have facilitated learning from punishment by stimulating prefrontal, rather than striatal, D1Rs.

Some limitations of the current study need to be mentioned. First, because we used a systemic, rather than a locally specific, manipulation of dopaminergic activity, the current data do not allow deciding which brain regions were affected by the D1 agonist. Besides the striatal direct pathway, D1Rs are also expressed in the frontal cortex, such that also changes in frontal D1R activation might be responsible for the observed behavioral effect, given that frontal regions too are involved in PIT [53, 54] and reversal learning [49–52]. Related to this, our findings are also agonistic as to whether the observed effects might be caused by drug effects on any nonmeasured aspects of cognition, such as working memory or attention. We note, though, that at least in the PIT task we observed no impact of baseline WMC on PIT in the placebo group, suggesting that working memory processes might not directly contribute to the strength of PIT. A further limitation is that we used WMC as proxy for dopamine synthesis capacity instead of measuring baseline dopamine levels more directly using PET. This approach represents a commonly used procedure

in the field of pharmacological manipulations in humans [11, 13, 34, 55] and its validity is supported by [35]. We note that the finding that baseline WMC differentially affected drug effects on PIT and reversal learning is consistent with the assumption that the relationship between D1R activation and various aspects of cognition follows a variety of functions (rather than the canonical inverted-u) [27]. However, without a direct measure of dopamine levels it is not possible to determine the precise shape of the relationship between D1R activation and PIT or reversal learning. Therefore, such conclusions have to be taken with caution.

To conclude, the current findings provide first evidence in humans for a causal involvement of D1R activation in mediating reactivity and flexibility to reward cues. D1R stimulation affected both cue-induced reactivity and the flexible updating of cue–outcome associations in contexts where perseveration is maladaptive. Because dysfunctions in PIT and reversal learning belong to the core symptoms of various psychiatric disorders, including addiction, schizophrenia, or depression [1–3, 56], these findings also advance our understanding of the neurobiological foundations of these disorders.

FUNDING AND DISCLOSURE

PNT received financial support from Pfizer for conducting this study (WI203648) as well as from the Swiss National Science Foundation (Grants PP00P1 150739 and 100014_165884) and from the Velux Foundation (Grant 981). AS received an Emmy Noether fellowship (SO 1636/2-1) from the German Research Foundation. EF received an ERC advanced grant (295642). AS, CJB, EF, AJ, and PNT declare to have no conflicts of interest. RK, NdM, and WH are former full-time employees of Pfizer and are currently fulltime employees of Takeda (RK), Praxis Precision Medicines (NdM), and Virginia Tech (WH), respectively.

ACKNOWLEDGEMENTS

We are grateful to Poppy Watson, Sanne de Wit and Roshan Cools for providing us with versions of the PIT and reversal learning tasks that we reprogrammed in Cogent. We are also thank Barbara Fischer and Ali Sigaroudi for support with the medical examinations, Susan Mahoney for clinical operations support, Sofie Carellis, Cecilia Etterlin, Alina Frick, Geraldine Gvozdanovic, Amalia Stüssi, Karl Treiber for expert help with data collection and Nadine Enseleit for study monitoring.

AUTHOR CONTRIBUTIONS

AS, RK, NdM, WH, CJB, EF, AJ, and PNT designed research; AS and AJ performed research; AS analyzed data; AS and PNT wrote manuscript; all authors approved final version of manuscript.

ADDITIONAL INFORMATION

Supplementary Information accompanies this paper at (<https://doi.org/10.1038/s41386-020-0617-z>).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Garbusow M, et al. Pavlovian-to-instrumental transfer effects in the nucleus accumbens relate to relapse in alcohol dependence. *Addict Biol.* 2016;21:719–31.
2. Reddy LF, Waltz JA, Green MF, Wynn JK, Horan WP. Probabilistic reversal learning in schizophrenia: stability of deficits and potential causal mechanisms. *Schizophr Bull.* 2016;42:942–51.
3. Timmer MHM, Sescousse G, van der Schaaf ME, Esselink RAJ, Cools R. Reward learning deficits in Parkinson's disease depend on depression. *Psychol Med.* 2017;47:2302–11.
4. Schultz W. Neuronal reward and decision signals: from theories to data. *Physiol Rev.* 2015;95:853–951.

5. Salamone JD, Correa M. The mysterious motivational functions of mesolimbic dopamine. *Neuron*. 2012;76:470–85.
6. Ott T, Nieder A. Dopamine and cognitive control in prefrontal cortex. *Trends Cogn Sci*. 2019;23:213–34.
7. de Boer L, et al. Attenuation of dopamine-modulated prefrontal value signals underlies probabilistic reward learning deficits in old age. *Elife*. 2017;6:e26424.
8. Durstewitz D, Seamans JK. The computational role of dopamine D1 receptors in working memory. *Neural Netw*. 2002;15:561–72.
9. Collins AG, Frank MJ. Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychol Rev*. 2014;121:337–66.
10. Weber SC, et al. Dopamine D2/3- and mu-opioid receptor antagonists reduce cue-induced responding and reward impulsivity in humans. *Transl Psychiatry*. 2016;6:e850.
11. Swart JC, et al. Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. *Elife*. 2017;6:e22169.
12. Hebart MN, Glascher J. Serotonin and dopamine differentially affect appetitive and aversive general Pavlovian-to-instrumental transfer. *Psychopharmacology*. 2015;232:437–51.
13. van der Schaaf ME, et al. Establishing the dopamine dependency of human striatal signals during reward and punishment reversal learning. *Cereb Cortex*. 2014;24:633–42.
14. Cools R, et al. Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. *J Neurosci*. 2009;29:1538–43.
15. Glueck E, Ginder D, Hyde J, North K, Grimm JW. Effects of dopamine D1 and D2 receptor agonists on environmental enrichment attenuated sucrose cue reactivity in rats. *Psychopharmacology*. 2017;234:815–25.
16. Lex A, Hauber W. Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. *Learn Mem*. 2008;15:483–91.
17. Thompson JL, et al. Age-dependent D1-D2 receptor coactivation in the lateral orbitofrontal cortex potentiates NMDA receptors and facilitates cognitive flexibility. *Cereb Cortex*. 2016;26:4524–39.
18. Calaminus C, Hauber W. Guidance of instrumental behavior under reversal conditions requires dopamine D1 and D2 receptor activation in the orbitofrontal cortex. *Neuroscience*. 2008;154:1195–204.
19. Haluk DM, Floresco SB. Ventral striatal dopamine modulation of different forms of behavioral flexibility. *Neuropsychopharmacology*. 2009;34:2041–52.
20. Calaminus C, Hauber W. 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*. 2007;191:551–66.
21. de Boer L, et al. Dorsal striatal dopamine D1 receptor availability predicts an instrumental bias in action learning. *Proc Natl Acad Sci USA*. 2019;116:261–70.
22. Frank MJ, Fossella JA. Neurogenetics and pharmacology of learning, motivation, and cognition. *Neuropsychopharmacology*. 2011;36:133–52.
23. Cox SM, et al. Striatal D1 and D2 signaling differentially predict learning from positive and negative outcomes. *Neuroimage*. 2015;109:95–101.
24. Davoren JE, et al. Discovery and lead optimization of atropisomer D1 agonists with reduced desensitization. *J Med Chem*. 2018;61:11384–97.
25. Gray DL, et al. Impaired beta-arrestin recruitment and reduced desensitization by non-catechol agonists of the D1 dopamine receptor. *Nat Commun*. 2018;9:674.
26. Vijayraghavan S, Wang M, Birnbaum SG, Williams GV, Arnsten AF. Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nat Neurosci*. 2007;10:376–84.
27. Floresco SB. Prefrontal dopamine and behavioral flexibility: shifting from an "inverted-U" toward a family of functions. *Front Neurosci*. 2013;7:62.
28. Arnsten AF. Catecholamine influences on dorsolateral prefrontal cortical networks. *Biol Psychiatry*. 2011;69:e89–99.
29. Takahashi H, Yamada M, Suhara T. Functional significance of central D1 receptors in cognition: beyond working memory. *J Cereb Blood Flow Metab*. 2012;32:1248–58.
30. Muller U, von Cramon DY, Pollmann S. D1- versus D2-receptor modulation of visuospatial working memory in humans. *J Neurosci*. 1998;18:2720–8.
31. Cools R, D'Esposito M. Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biol Psychiatry*. 2011;69:e113–25.
32. Landau SM, Lal R, O'Neil JP, Baker S, Jagust WJ. Striatal dopamine and working memory. *Cereb Cortex*. 2008;19:445–54.
33. Bäckman L, et al. Effects of working-memory training on striatal dopamine release. *Science*. 2011;333:718.
34. Kimberg DY, D'Esposito M, Farah MJ. Effects of bromocriptine on human subjects depend on working memory capacity. *Neuroreport*. 1997;8:3581–5.
35. Cools R, Gibbs SE, Miyakawa A, Jagust W, D'Esposito M. Working memory capacity predicts dopamine synthesis capacity in the human striatum. *J Neurosci*. 2008;28:1208–12.
36. Soutschek A, et al. Dopaminergic D1 receptor stimulation affects effort and risk preferences. *Biol Psychiatry*. 2019. Epub 2019/11/02.
37. Carver CS, White TL. Behavioral-inhibition, behavioral activation, and affective responses to impending reward and punishment—the Bis Bas Scales. *J Pers Soc Psychol*. 1994;67:319–33.
38. Lehrl S. Mehrfachwahl-Wortschatz-Intelligenztest: MWT-B. Spitta, Balingen, 1999.
39. Patton JH, Stanford MS, Barratt ES. Factor structure of the Barratt Impulsiveness Scale. *J Clin Psychol*. 1995;51:768–74.
40. Lovibond PF, Colagiuri B. Facilitation of voluntary goal-directed action by reward cues. *Psychol Sci*. 2013;24:2030–7.
41. Watson P, Wiers RW, Hommel B, de Wit S. Working for food you don't desire: cues interfere with goal-directed food-seeking. *Appetite*. 2014;79:139–48.
42. Corbit LH, Janak PH, Balleine BW. General and outcome-specific forms of Pavlovian-instrumental transfer: the effect of shifts in motivational state and inactivation of the ventral tegmental area. *Eur J Neurosci*. 2007;26:3141–9.
43. Cools R, Altamirano L, D'Esposito M. Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia*. 2006;44:1663–73.
44. Cartoni E, Puglisi-Allegra S, Baldassarre G. The three principles of action: a Pavlovian-instrumental transfer hypothesis. *Front Behav Neurosci*. 2013;7:153.
45. Corbit LH, Balleine BW. Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of Pavlovian-instrumental transfer. *J Neurosci*. 2005;25:962–70.
46. Corbit LH, Balleine BW. The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. *J Neurosci*. 2011;31:11786–94.
47. Verharen JPH, Adan RAH, Vanderschuren L. Differential contributions of striatal dopamine D1 and D2 receptors to component processes of value-based decision making. *Neuropsychopharmacology*. <https://doi.org/10.1038/s41386-019-0454-0> (2019).
48. Frank MJ. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *J Cogn Neurosci*. 2005;17:51–72.
49. Xue G, et al. Common neural mechanisms underlying reversal learning by reward and punishment. *PLoS ONE*. 2013;8:e82169.
50. Wischniewski M, Zerr P, Schutter D. Effects of theta transcranial alternating current stimulation over the frontal cortex on reversal learning. *Brain Stimul*. 2016;9:705–11.
51. Dalton GL, Wang NY, Phillips AG, Floresco SB. Multifaceted contributions by different regions of the orbitofrontal and medial prefrontal cortex to probabilistic reversal learning. *J Neurosci*. 2016;36:1996–2006.
52. Ghods-Sharifi S, Haluk DM, Floresco SB. Differential effects of inactivation of the orbitofrontal cortex on strategy set-shifting and reversal learning. *Neurobiol Learn Mem*. 2008;89:567–73.
53. Homayoun H, Moghaddam B. Differential representation of Pavlovian-instrumental transfer by prefrontal cortex subregions and striatum. *Eur J Neurosci*. 2009;29:1461–76.
54. Keistler C, Barker JM, Taylor JR. Infralimbic prefrontal cortex interacts with nucleus accumbens shell to unmask expression of outcome-selective Pavlovian-to-instrumental transfer. *Learn Mem*. 2015;22:509–13.
55. van der Schaaf ME, Fallon SJ, Ter Huurne N, Buitelaar J, Cools R. Working memory capacity predicts effects of methylphenidate on reversal learning. *Neuropsychopharmacology*. 2013;38:2011–8.
56. Schlagenhauf F, et al. Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *Neuroimage*. 2014;89:171–80.