Behavioural Brain Research 280 (2015) 92–100

Contents lists available at ScienceDirect

Behavioural Brain Research

journal homepage: www.elsevier.com/locate/bbr

Research report

Dopamine bioavailability in the mPFC modulates operant learning performance in rats: An experimental study with a computational interpretation

Maximiliano Rapanelli[a], Luciana Romina Frick[a], Anabel M.M. Miguelez Fernández[b], B. Silvano Zanutto[b,c,*]

[a] Department of Psychiatry, Yale University School of Medicine, New Haven, CT 06519, USA
[b] Instituto de Biología y Medicina Experimental, Consejo Nacional de Investigaciones Científicas y Técnicas (IByME – CONICET), Vuelta de Obligado 2490, C1428ADN Ciudad de Buenos Aires, Argentina
[c] Instituto de Ingeniería Biomédica, Facultad de Ingeniería, Universidad de Buenos Aires, Paseo Colon 850, Ciudad de Buenos Aires, Argentina

HIGHLIGHTS

- We studied how DA bioavailability in the mPFC affects operant conditioning learning.
- Systemic and local administration of COMT inhibitor improves learning performance.
- Administration of recombinant COMT in the mPFC produces the opposite effect.
- We interpreted these results by means of a computational theory.
- Simulated and experimental results show that learning is influenced by DA clearance.

ARTICLE INFO

Article history:
Received 1 July 2014
Received in revised form
12 November 2014
Accepted 20 November 2014
Available online 27 November 2014

Keywords:
Prefrontal cortex
Dopamine
COMT
Neural network
Computational model
Entacapone

ABSTRACT

Dopamine encodes reward and its prediction in reinforcement learning. Catechol-O-methyltransferase (COMT) activity in the medial prefrontal cortex (mPFC) has been shown to influence cognitive abilities by modifying dopamine clearance. Nevertheless, it is unknown how COMT in the mPFC influences operant learning. Systemic entacapone (50 mg/kg), as well as local entacapone (3 pg) and recombinant COMT (17 μg) in the mPFC were administered to male Long Evans rats prior to training in an operant conditioning task. We found that systemic and local administration of the COMT inhibitor entacapone significantly improves learning performance. Conversely, recombinant COMT administration totally impaired learning. These data have been interpreted through a computational model where the phasic firing of dopaminergic neurons was computed by means of a temporal difference algorithm and dopamine bioavailability in the mPFC was simulated with a gating window. The duration of this window was selected to simulate the effects of inhibited or enhanced COMT activity (by entacapone or recombinant COMT respectively). The model accounts for an improved performance reproducing the entacapone effects, and a detrimental impact on learning when the clearance is increased reproducing the recombinant COMT effects. The experimental and computational results show that learning performance can be deeply influenced by COMT manipulations in the mPFC.

Abbreviations: COMT, catechol-O-methyltransferase; mPFC, medial prefrontal cortex; lPFC, dorsolateral prefrontal cortex; DA, dopamine; LTP, long term potentiation; LTD, long term depression; VTA, ventral tegmental area; TD, temporal difference algorithm; BG, basal ganglia; PMC, pre-motor cortex; STM, short-term memory activity; CS, conditioned stimulus; US, unconditioned stimulus; SNc, substantia nigra pars compacta; DOPAC, 3,4-dihydroxyphenylacetic acid.

* Corresponding author at: Instituto de Biología y Medicina Experimental, Consejo Nacional de Investigaciones Científicas y Técnicas (IByME – CONICET), Vuelta de Obligado 2490, C1428ADN Ciudad de Buenos Aires, Argentina. Tel.: +54 11 47832869x235; fax: +54 11 4786 2564.
E-mail address: silvano@fi.uba.ar (B.S. Zanutto).

http://dx.doi.org/10.1016/j.bbr.2014.11.031
0166-4328/

1. Introduction

The medial prefrontal cortex (mPFC) in rats and the dorsolateral prefrontal cortex (lPFC) in primates (humans and non-human) are key areas for cognition, working memory and goal directed behaviors [1–3]. Lesions in rat mPFC have been shown to induce severe executive dysfunctions [4]. Changes in dopamine (DA) levels in the mPFC modulate long-term potentiation (LTP) and depression (LTD) by a complex network of signal transduction [5]. In mammals, it was found that phasic activation of dopaminergic neurons is sufficient to drive behavioral conditioning [6] and, in that sense, phasic activation of DA neurons could signal rewarding events and prediction errors at the PFC [7,8].

The main responsible for DA clearance in the mPFC is catechol-O-methyltransferase (COMT) [9–11]. The role of COMT in humans has been extensively studied in cognition, particularly due to polymorphisms that can positively or negatively influence cognitive abilities depending on the catabolic activity of the enzyme [12,13]. Previous studies in animals using attentional set shifting and delayed spatial win-shift tasks showed that COMT inhibition improves learning [14,15]. Genetically modified mice that overexpress COMT exhibit disrupted attentional set-shifting abilities and impaired working and recognition memory, whereas COMT knock-out mice show improved working memory performances [16]. Additionally, Tunbridge et al. [14] observed that COMT inhibition increases evoked DA release without affecting norepinephrine levels in the mPFC. Furthermore, some studies point toward the PFC as the key site for COMT's actions in tasks with fixed reward contingencies [17]. Therefore, targeting COMT in animal models can provide useful evidence on how manipulation of this enzyme in the mPFC impacts on learning.

In order to understand the role of dopamine in reward prediction and its complex interaction with neural networks in the PFC, many computational models of the ventral tegmental area (VTA) and the PFC have been proposed [18–20]. The temporal difference (TD) algorithm [21] accurately describes firing rates profiles of dopamine VTA neurons [21,22], however the connection between these profiles and DA bioavailability in the PFC is not well described.

Our group proposed a neurobehavioral-based computational theory to simulate operant learning that predicts relevant features of operant conditioning and category learning [24–26], well suited to interpret the effect that DA availability in PFC has on operant learning. Neurophysiological studies recording simultaneous neural activity during simple operant learning showed different time courses of learning-related activity in the PFC and striatum [27]. This evidence induced some authors [27] to hypothesize that rewarded associations are first identified by the basal ganglia (BG) and then its output "trains" slower learning mechanisms in the PFC. We showed in a previous version of the present neural network model [24] that time courses of learning-related activity in PFC and striatum are task-dependent. When the model was trained in a cognitively demanding task (like delayed matching to sample), changes in BG activity were concurrent with changes in PFC activity. Instead, when presented with a simple operant task (visual discrimination) changes in BG activity "lead" changes in PFC [24]. This raises the question of whether simple operant learning is controlled only by BG or if PFC is also involved.

Thus, in this work we aimed to study how manipulation of DA clearance in the rat mPFC affects operant learning. To this end, we administered a selective COMT inhibitor (entacapone) or exogenous COMT protein (recombinant COMT) during learning an operant conditioning task. The experimental data was interpreted with a computational model where the role of the DA bioavailability in the mPFC is simulated depending not only on the firing of dopamine VTA neurons but also on DA clearance.

2. Materials and methods

2.1. Experimental procedures

All experimental procedures were approved by the ethics committee of the IByME-CONICET (A2008) and were conducted according to the NIH Guide for Care and Use of Laboratory Animals.

2.2. Animals

Two months old male *Long Evans* rats (250–300 g) were provided by the IBYME-CONICET, maintained on a 12/12 h light/dark cycle (lights on at 8 am) with food and tap water available *ad libitum*.

2.3. Surgery

The anterior cingulate cortex plays a critical role in stimulus-reinforcement learning and reward-guided selection of actions [28] and it receives major dopaminergic inputs from the VTA [29]. Therefore this area within mPFC was targeted.

Rats were anesthetized with ketamine (75 mg/kg, i.p.)/xylazine (20 mg/kg, i.p.), placed in a stereotaxic instrument (David Kopf, USA), a longitudinal incision was made on the scalp and the bone was then exposed. Next, stainless steel cannulae (Small Parts, USA) were bilaterally implanted in the anterior cingulate cortex in mPFC (Cg1) *via* stereotaxic guided craniotomy (anteroposterior, +2.52 mm from bregma; laterolateral, ±0.2 mm from midline; dorsoventral and −1.4 mm from skull surface) [30].

Three stainless screws were fixed on the skull using a screwdriver. The screws and the cannulae were fixed to the skull with dental cement. Rats were treated with analgesics and antibiotics in the tap water for 5 days after the surgery and they were allowed to recover for a week before training procedures began.

2.4. Operant conditioning task

Rats were trained in an operant conditioning task previously described by our group [31–33]. All behavioral procedures were performed during the light phase of the light/dark cycle, using a standard operant chamber (MED associates Inc., USA) equipped with an input (DIG 710/711) and output (DIG 720/721/722) card for data acquisition and processing; one automated retractable lever; white house light; contextual red light; white noise (random signal with a flat power spectral density) and an automated feeder. Rats were food restricted to maintain ∼80% of their *ad libitum* body weight for 3 days before training and throughout the experiments. During the period of food restriction and experimental procedures, animals were fed with rat chow plus the 45 mg dustless precision pellets (BioServ, USA) used for training procedures to avoid palatability issues during the experiments.

Animals were placed in the training room for 10 min prior to every session. In habituation sessions rats were placed in the operant chamber, exposed to contextual red light and white noise and fed with 25 pellets given randomly by the automated feeder. Only one habituation session of 20 min was performed. In training sessions, before the training procedure started the operant chamber had the lever retracted, the house light on, white noise, and a red context light that remained on at all times. An operant conditioning task training session with a fixed ratio of 1 consisted of 25 trials. Each trial began when the lever was extended into the chamber and the house light was turned off. If the animal pressed the lever within 60 s, the lever retracted and a pellet of 45 mg was delivered as a reward. Pellet delivery was within 1 s of lever pressing and it was coupled with the activation of a white light inside the feeder for 2 s. When the trial finished, the white house light was turned on and the lever remained retracted for 20 s. The action of pressing the

lever was considered a correct response. Animals were able to press the lever only once per trial. When the animals did not respond during the trial, no reward was given. The percentage of responses in a session was calculated by counting the total number of lever presses in a training session and then divided by 25 (maximum number of lever presses that can be performed in a session) and then multiplied by 100. Response latency was measured from the insertion of the lever into the chamber to a lever pressing response. If no response was performed during the trial, latency of response was considered 60 s. Training criteria was to reach 100% of correct responses with a latency under 5 s for 3 consecutive sessions. Only animals that reached these criteria were included for data analysis. Experimental groups were as follows: Control ($n = 18$), Systemic entacapone ($n = 18$) $\text{Vehicle}_{\text{entacapone}}$ ($n = 8$), Entacapone in the mPFC ($n = 8$), $\text{Vehicle}_{\text{recombinant COMT}}$ ($n = 8$) and recombinant COMT administered in the mPFC ($n = 8$). To avoid differences in the behavioral performance we tested all groups in the same session each day.

2.5. Treatments

Entacapone (Novartis, USA) was administered in a high dose (50 mg/kg), to assure a good penetration of blood–brain barrier [34,35]. For oral administration tablets were crushed and suspended in water. Control group was orally administered with water only. For intra-mPFC microinjection, entacapone (Trust & We, China) was dissolved in PBS (pH 4)/Triton X-100 (0.1%), whereas recombinant COMT (Sigma–Aldrich, USA) was dissolved in PBS. Both entacapone (3 pg) and recombinant COMT (17 μg) were injected in a 1 μl volume 10 min prior to each training session. Then, cannulated animals were microinjected in the mPFC with PBS (pH 4)/Triton X-100 (0.1%) (control group of local administration of entacapone: $\text{Vehicle}_{\text{entacapone}}$ group) or with PBS (pH 7) (control group of local administration of recombinant COMT: $\text{Vehicle}_{\text{recombinant COMT}}$ group). Injections were manually performed using an injector coupled to a polyimide tube attached to a micropipette (0.1–2.5 μl) at 0.2 μl/min to allow proper diffusion of the volume. The injector did not protrude below the tip of the cannulae.

2.6. Histology

Animals were anesthetized with 100 mg/kg of ketamine and 20 mg/kg of xilazine and perfused transcardially with 300 ml of saline solution followed by 300 ml of 4% formaline/PBS solution with a peristaltic pump. Then, brains were coronally sectioned with a vibratome at 40 μM and slices were stained with a solution of 0.40% of hematoxilin and 0.48% of eosin. Images were taken with a Leica DM500 microscope at 100× magnification.

2.7. Sample preparation and measurements of cathecolamines

For this procedure we used rats that received the same treatments as the ones used in the behavioral task, but they were not trained [Control ($n = 4$), Systemic entacapone ($n = 4$), Entacapone in the mPFC ($n = 4$), recombinant COMT ($n = 4$)]. Rats were sacrificed after being treated and the brains were immediately removed. Once obtained, the brains were cut in half through the midline and mPFC was dissected based on anatomical characteristics described in a rat brain atlas [30]. Samples were quickly homogenized in 0.2 N of percloric acid. Then, samples were centrifuged and then seeded in a Nova-Pak C18 column of a Varian 5000 HPLC coupled to an electrochemichal detector (LC-4C, BAS) with a +0.7 V electrode potential (with respect to the Ag/AgCl reference electrode). Peaks were quantified using DATA Jet Integrator (Spectra-Physics).

2.8. Statistical Analysis

Statistical analysis was performed using GraphPad Prism 6.00 (GraphPad Software, San Diego, CA, USA). Values are expressed as means ± SEM and compared using one way ANOVA or two-way repeated measures ANOVA followed by Tukey's *post hoc* test. Differences among experimental conditions were considered statistically significant if $p < 0.05$.

2.9. Computational model

The computational theory used in the present study is based on a previously described neural network that is capable of learning several tasks including operant conditioning [24–26]. Briefly, the network is structured in three layers (Fig. 1). The first layer of the network receives sensory input from the environment (conditioned stimulus (*CS*) and unconditioned stimulus (*US*)) and generates short-term memory traces (STM) for those stimuli. Biologically, this is not the result of a single structure information processing, but instead it arises from the interaction between sensory cortices, associative cortices (PFC, inferotemporal cortex) and subcortical structures (amygdala, hippocampus). The second layer also involves the interaction of many structures. It represents mainly the mPFC (lPFC in monkeys), which conveys information from other PFC areas, inferotemporal cortex, cingulate and parietal cortex. It allows further filtering of task relevant stimuli, which are also actively maintained in working memory. These clusters will then be associated with the proper response (in the third layer) according to the contingencies of the task. Thus, although PFC is engaged in working memory processes, different areas within PFC represent different kinds of information [36].

Three possible responses are included (even though the model still works well with more responses), but only one of them is associated with reward (the other ones stand for any response unrelated to the task). Responses are initially executed randomly, simulating a motivational state generated by deprivation. Reinforcement information reaches the VTA and the locus coeruleus mainly from frontal structures related to its processing, as the mPFC. However, the VTA and the locus coeruleus process such information in a different way in order to produce different patterns of neural responses and/or learning mechanisms. Once the information about reinforcement reaches the VTA, DA is released and those conditioned stimuli traces at that time active can be associated with reinforcement. As learning proceeds, the probability of being reinforced increases and the synaptic weights in VTA will represent the association between conditioned stimuli and reinforcement. Consequently, every time that a *CS* is presented, the VTA will fire strongly releasing DA over the PFC and the BG, pre-motor cortex (PMC). In order to generate this process the VTA/substantia nigra pars compacta (SNc) is modeled by a TD model with eligibility traces like the one in Pan et al. [37].

Although the TD model reproduces the phasic firing pattern of dopaminergic neurons, it cannot simulate the prolonged excitatory effect on target neurons in mPFC and in motor-related structures [38,39]. In our model, a gating window ($W_t^{\delta} = 1$) simulates DA postsynaptics effects. When the DA firing is close to baseline $W_t^{\delta} = 0.5$ and when bursts of DA neurons occur, if the prediction error at time t ($\delta(t)$) exceeds a threshold (θ_{hebb}) then $W_t^{\delta} = 1$. The gating window duration (T) represents the time period in which DA is released in the synaptic cleft of mPFC neurons [21], and depends on $\delta(t)$. To explore the impact of DA availability on the learning process, we modeled T as depending on the presence of entacapone or recombinant COMT.

The length of T was chosen based in neurophysiological data where VTA's neuronal activity was registered [22,23], showing

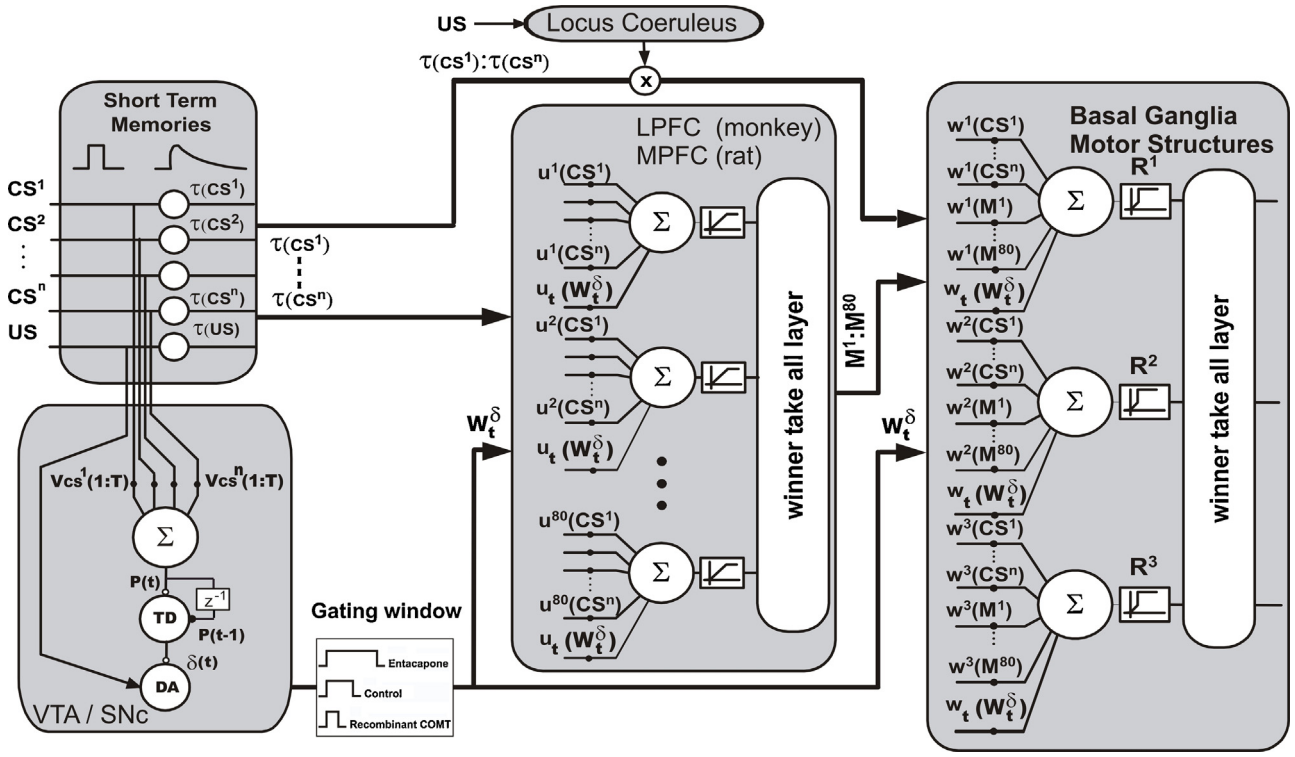

Fig. 1. Scheme of the neural network model. The first layer generates short-term memories of the stimuli as a result of the interaction between different structures such as ventromedial PFC, inferotemporal cortex, posterior parietal cortex, hippocampus and amygdala. We used 80 neurons in the mPFC, 3 in the BG-PMC and a TD (λ_{TD}) model in the VTA/SNc. See Appendix for full details.

that the peristimulus time histogram for these neurons presented increased activity, reaching its maximum approximately between 100 ms and 200 ms. The exact duration and dynamics of evoked DA levels in the extracellular space of mPFC *in vivo* is unknown. Yavich et al. [10] studied evoked DA overflow after electrical stimulation of the medial forebrain bundle and described DA concentration dynamics in PFC in wild type and COMT-KO mice. However, since the stimulation protocol they selected is out of the range of physiological firing rates recorded by Schultz [22] we did not considered absolute values but we did accounted for the DA dynamics profile. The authors found that DA concentration in PFC rises three times faster than it falls, therefore we have chosen the basal $T = 400$ ms. In the same work it is shown that once DA levels are maximal, its elimination is approximately two times longer in COMT-KO than in wild type mice. Assuming that the clearance with entacapone treatment is higher than in wild type, but smaller than COMT-KO mice, we have chosen T values of 600 ms for entacapone and 300 ms for recombinant COMT.

Further information and the equations of the model can be found in Appendix.

3. Results

3.1. Systemic entacapone improves animal performance in operant conditioning learning

In the first set of the experiments, animals were divided in the following groups: Control (vehicle) or Systemic entacapone (50 mg/kg). Animals were trained in an operant conditioning task, in which a food reward is obtained after pressing a lever (this being the correct response to get food). Seven sessions of 25 trials each were performed, and each trial lasted 60 s.

Initially, we found a statistically significant main effect on the treatment factor in the learning performance [$F_{6,36} = 120.7$; $p < 0.001$] and in the latency [$F_{6,36} = 61.33$; $p < 0.001$]. Next, we detected that there was a main effect of session factor in mean percentage of responses [$F_{6,36} = 338.3$; $p < 0.001$] and in the latency [$F_{6,36} = 152.2$; $p < 0.001$]. Additionally, there was a significant interaction between the factors [$F_{6,36} = 15.53$; $p < 0.001$] and latency time [$F_{6,36} = 9.312$; $p < 0.001$]. Then, we proceed to perform the *post hoc* comparisons.

Entacapone administration significantly improved learning performance (Fig. 2A). A significant increment in the percentage of correct responses (lever press) was observed in the first [$t(4.2)$, $p < 0.001$] (Fig. 2A), second [$t(14.1)$, $p < 0.001$] (Fig. 2A) and third [$t(9.4)$, $p < 0.001$] (Fig. 2A) sessions in animals treated with Systemic entacapone with respect to the controls, and these increments reached values of 15.25%, 51% and 31% respectively (Fig. 2A). No differences in the number of responses were found in the fourth, fifth, sixth and seventh sessions between both groups (Fig. 2A). Additionally, latency was substantially reduced in animals systemically treated with entacapone in the first, second and third sessions (Fig. 2B). In the first session, latency of entacapone-treated animals was decreased by 12.4 s with respect to control group [$t(3.52)$, $p < 0.01$] (Fig. 2B), whereas, in the second session, they yielded a decrement of 35 s in latency compared with the control group [$t(9.94)$, $p < 0.001$] (Fig. 2B). Latency of entacapone-treated animals was 34 s lower than the one observed in control animals in the third training session [$t(9.65)$, $p < 0.001$] (Fig. 2B). From the fourth to seventh training session no significant differences were found (Fig. 2B).

3.2. Local administration of entacapone in the mPFC also improves animal performance in operant conditioning learning

Next, to evaluate if the mPFC is the brain region involved in this event, local administration of entacapone was performed. Entacapone microinjected in the mPFC exerted a similar effect on animal

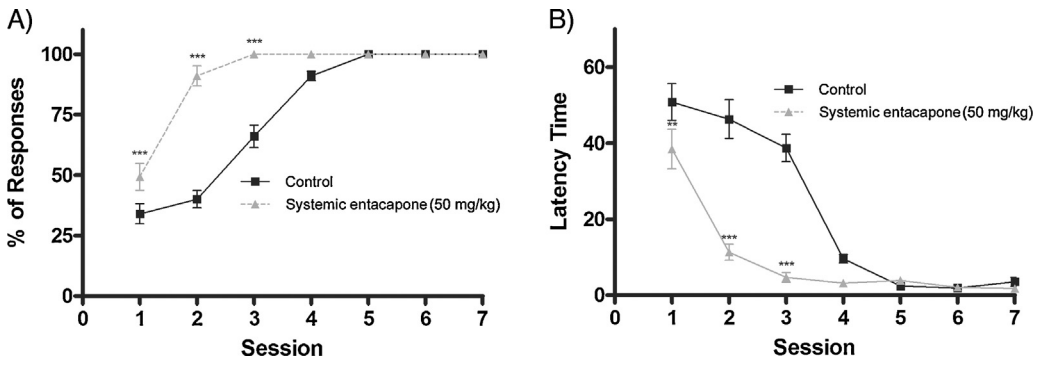

Fig. 2. Systemic entacapone administration improves learning in an operant conditioning task. Percentage of responses is expressed as the mean ± SEM of responses in a 25 trials training session (panel A). Latency is expressed as the mean ± SEM of the time elapsed between lever presentation and lever pressing of each 60 s trial (panel B). Control ($n = 18$) and Systemic entacapone ($n = 18$). Statistical differences were analyzed using two-way repeated measures ANOVA. $^{**}p < 0.01$; $^{***}p < 0.001$.

performance in the operant conditioning task compared with systemic administration of the drug, and this effect was even more pronounced.

Statistical analysis showed that, in the first session, animals that received a local administration of entacapone in the mPFC had an increment of 57.3% in the percentage of responses with respect to sham group [$t(10.88)$, $p < 0.001$] (Fig. 3A). Entacapone-mPFC group exhibited an increment of 63% of responses in the second session compared to sham animals [$t(11.96)$, $p < 0.001$] (Fig. 3A). In the third session animals administered with entacapone also exhibited an increment of about 40% of responses in comparison with the respective control [$t(7.52)$, $p < 0.001$] (Fig. 3A). No differences were found between both experimental groups in the learning performance from the fourth to the last training session (Fig. 3A). Animals that were microinjected with entacapone in the mPFC presented decreased latency in the first session compared with sham controls, being lower by 31.2 s [$t(7.09)$, $p < 0.001$] (Fig. 3B). Sessions two and three of entacapone-treated animals showed decreases of 37.9 [$t(8.612)$, $p < 0.001$] (Fig. 3B) and 36.09 s [$t(8.19)$, $p < 0.001$] (Fig. 3B), respectively.

Interestingly, comparison between systemic and local administration of entacapone revealed a significant difference in the first session, showing a 32% of increase in the number of responses when the drug is directly administered in the mPFC [$t(32.05)$, $p < 0.001$] (Figs. 2A and 3A). In this session, latency was also statistically significant with a difference of 18.8 s between both groups [$t(4.27)$, $p < 0.001$] (Figs. 2B and B). This is probably due to the limited ability of entacapone to cross the blood–brain barrier, further considering the high dose required for systemic treatment with this drug.

3.3. Recombinant COMT administration in the mPFC impairs animal performance in operant conditioning learning

Conversely to entacapone, the administration of recombinant COMT into the mPFC impaired learning in the operant conditioning task. No significant differences between COMT and sham groups were found during the first two sessions of the training (Fig. 4A). However, in the third session, animals that received the recombinant COMT protein in the mPFC presented a 30.4% decrease of the number of responses compared with the sham group [$t(5.77)$, $p < 0.001$] (Fig. 4A). Besides, this effect was also observed in the fourth and fifth session where decrements in the percentage of responses in the COMT group with respect to sham group reached 57% [$t(10.82)$, $p < 0.001$] (Fig. 4A) and 32% [$t(6.11)$, $p < 0.001$] (Fig. 4A) respectively. Lower values in the percentage of responses in COMT-treated animals persisted in the sixth session, where mean value reflected a 16.6% decrease with respect to controls [$t(3.15)$, $p < 0.01$] (Fig. 4A). There was no significant difference in the seventh session between both groups.

COMT treatment also exerted a general increase of the latency to press the lever. As observed for learning performance, in the first and second sessions mean values of latency remained unaltered among sham and COMT groups (Fig. 4B). Nevertheless, we observed that animals that were injected with recombinant COMT in the mPFC had increments in their latencies of response of 20 s

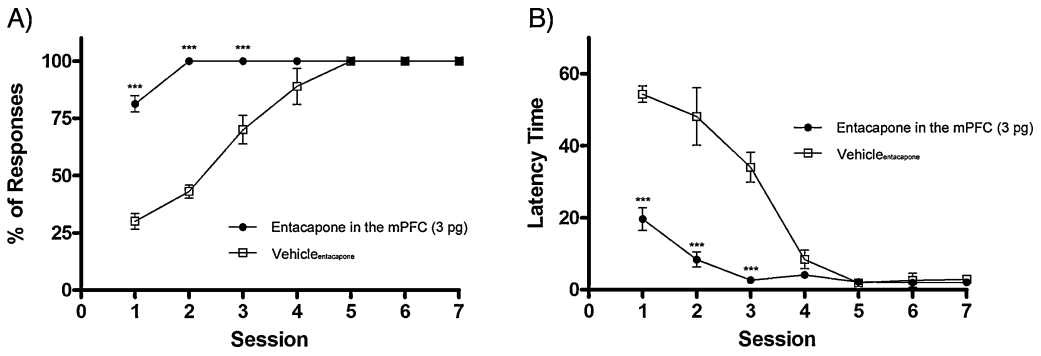

Fig. 3. Local administration of entacapone in the mPFC improves performance in operant conditioning learning. Percentage of responses is expressed as the mean ± SEM of responses in a 25 trials training session (panel A). Latency is expressed as the mean ± SEM of the time elapsed between lever presentation and lever pressing of each 60 s trial (panel B). Vehicle ($n = 8$) and Entacapone in the mPFC ($n = 8$). Statistical differences were analyzed using two-way repeated measures ANOVA. $^{***}p < 0.001$.

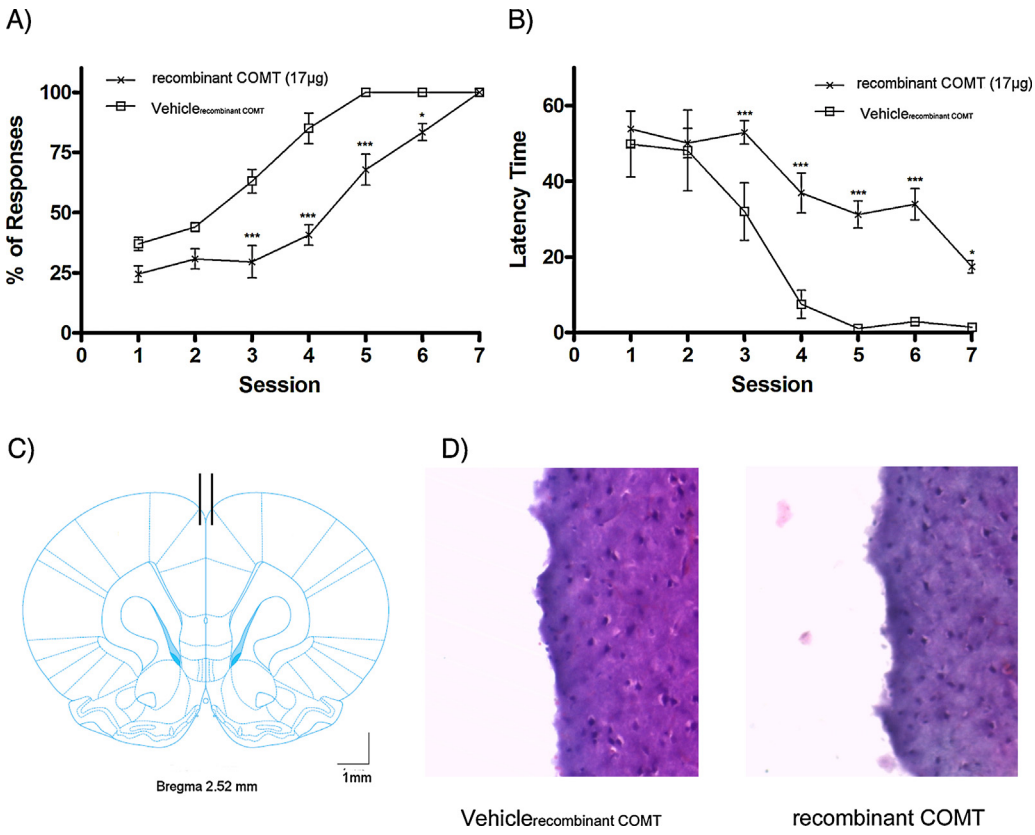

Fig. 4. Local administration of recombinant COMT in the mPFC impairs learning in an operant conditioning task. Percentage of responses is expressed as the mean ± SEM of responses in a 25 trials training session (panel A). Latency is expressed as the mean ± SEM of the time elapsed between lever presentation and lever pressing of each 60 s trial (panel B). Recombinant COMT administered in the mPFC ($n = 8$) and Vehicle$_{\text{recombinant COMT}}$ ($n = 8$). Cannula placement map (panel C; modified from Paxinos and Watson [30]). Hematoxylin/eosin staining of the site of injection of recombinant COMT (right, panel D) and Vehicle$_{\text{recombinant COMT}}$ (left, panel D) shows there was no infiltration of inflammatory cells. Statistical differences were analyzed using two-way repeated measures ANOVA. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

for the third session [$t(3.89)$, $p < 0.001$] (Fig. 4B), 28.3 s for the fourth session [$t(5.52)$, $p < 0.001$] (Fig. 4B), 28.9 s for the fifth session [$t(5.63)$, $p < 0.001$] (Fig. 4B) and 32.6 s for the sixth session [$t(6.35)$, $p < 0.001$] (Fig. 4B). Although animals reached 100% of responses in the seventh session, the latency of COMT group was still higher by 2.85 s with respect to sham group [$t(6.35)$, $p < 0.001$] (Fig. 4B). To rule out the possibility that the deleterious effect on cognition observed when we microinjected recombinant COMT is due to an inflammatory response, we looked around the site of injection (Fig. 4C) for infiltration of inflammatory cells. As a result, we found that the injection of recombinant COMT did not induce an inflammatory response in the mPFC (Fig. 4D).

3.4. Effect of COMT manipulations on DA metabolites levels in the mPFC

As a control of the effectiveness with which COMT activity in the mPFC was affected by the different treatments, we quantified DA and 3,4-dihydroxyphenylacetic acid (DOPAC) by HPLC in the mPFC of animals that were not trained but received the same treatments as those used for behavioral procedures. Statistical analysis of DA levels showed no difference among different groups [$F_{3,10} = 2.248$; $p > 0.05$] (Fig. 5). Statistical analysis also showed there was a difference in the levels of DOPAC within the mPFC among groups [$F_{3,10} = 29.38$; $p < 0.001$]. DOPAC levels were significantly increased in the mPFC when entacapone was locally administered in this area or systemically (Fig. 5). Also, we found that recombinant COMT administration decreased the levels of DOPAC in this area (Fig. 5). These results showed that our treatments successfully manipulated COMT activity within the mPFC.

3.5. Computational modeling of in vivo COMT manipulations in the mPFC during operant conditioning learning

We interpreted COMT manipulations in the mPFC with a computational model. In the model (Fig. 1), each trial started with the presentation of a sample stimulus (S1) lasting 400 ms. After presentation of the *CS*, the animal is able to execute three different responses (R1, R2 or R3), but only response R1 (pressing the lever) gives reward. A block using the TD algorithm computes the phasic firing of dopaminergic neurons, simulating the changes in the firing pattern in the VTA and the SNc. Information on the prediction error of being rewarded after repeated presentation of *CS* and *US* is then used to modulate learning.

For each experimental condition, that is Entacapone, Control and recombinant COMT we performed 100 simulations with gating window durations (T) of 600 ms, 400 ms and 300 ms respectively. Here, the variable T represents the time period in which DA is released in the synaptic space of mPFC (see Section 2.9 for full details). The computational model showed that lengthening the gating window improved learning performance in the second and third session (Fig. 6). In this case, fewer sessions were required to learn the operant conditioning task, reproducing the effect of entacapone administration. On the contrary, shortening T (mimicking the effect of enhanced COMT activity) results in an impaired learning performance. This deleterious effect can be observed between the third and seventh session (Fig. 6).

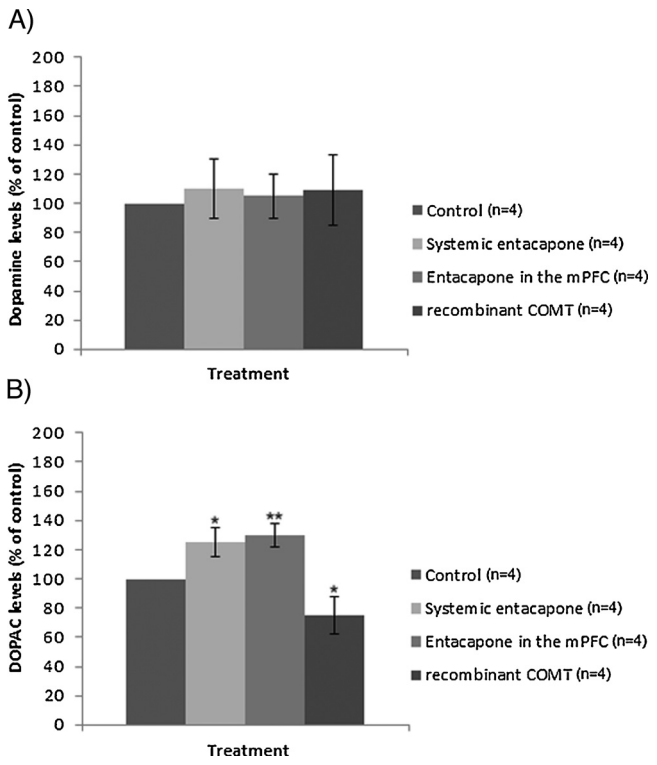

Fig. 5. DOPAC and DA levels in the mPFC after systemic administration of entacapone and local administration of entacapone or recombinant COMT. There was no change in DA levels between different groups (panel A). DOPAC levels were increased when entacapone was orally administered or locally injected in the mPFC (panel B), whereas microinjection of recombinant COMT administration in the mPFC decreased DOPAC levels in this area (panel B). Values are expressed as percentage of the control group and represent the mean ± SEM. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$. One way ANOVA followed Tukey's *post hoc* test.

Dopamine level was used to trigger gating windows during which synaptic changes take place. When the gating window $W_t^{\delta} = 1$, hebbian learning was applied in order to update both mPFC and BG-PMC neurons synaptic weights. Consequently, synaptic weight changes in the mPFC were increased when entacapone treatment was simulated using a longer gating window (600 ms) than the control group (400 ms) (Fig. 7). Instead, shortening the gating window to 300 ms (representing recombinant COMT group) diminished the percentage of synaptic weight modifications during operant conditioning learning (Fig. 7).

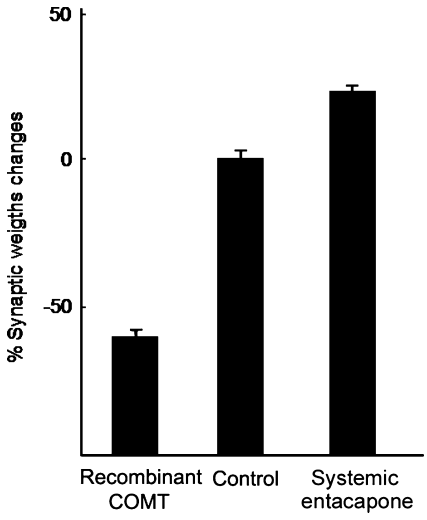

Fig. 7. Synaptic weights modifications during learning. Entacapone and recombinant COMT simulations were used for synaptic weight calculations in the mPFC during operant conditioning learning. The changes are expressed with respect to the control group.

4. Discussion

In this work, we studied how COMT activity in the mPFC influences operant learning performance in rats, and the results were interpreted with a computational theory.

Systemic administration of COMT inhibitor entacapone modified learning performance: animals treated with this drug showed improved learning performance as well as decreased latency of response during the first three training sessions. Our results are complementary to previous evidence where COMT inhibition improved performance of delayed spatial win-shift task and set shifting [14,15]. Based on the dual state theory proposed by Durstewitz and Seamans [40] one plausible explanation for the positive effects on cognition is that COMT inhibitors lead the network to D2 state (dopamine D2 receptor activation) enhancing cognitive flexibility and updating the information in the PFC. This effect is mediated by an increment in evoked release of DA without modifying basal levels [14,15,41]. In this sense, knock-out mice for COMT showed better performances in working memory and spatial tasks [42,43], confirming the key role of this enzyme in learning. However, the relationship between increased levels of DA and cognitive and learning performances is not linear. This can be clearly

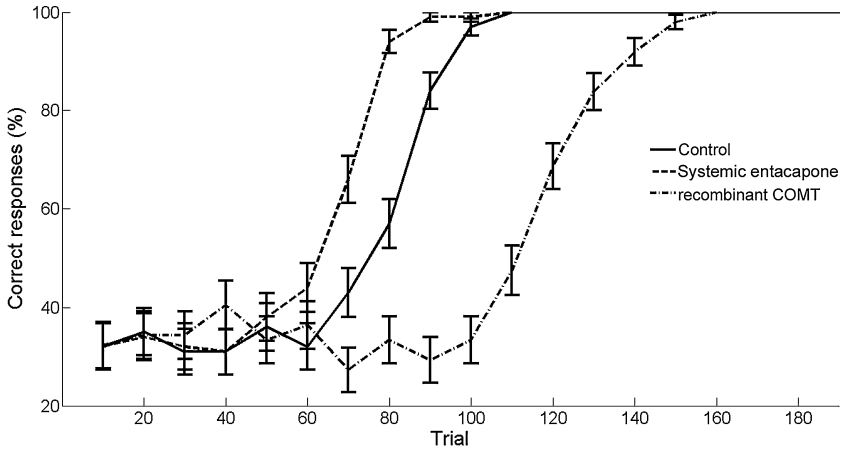

Fig. 6. Simulated performance on an operant conditioning task using the computational model. Systemic entacapone, Control and recombinant COMT conditions were simulated by varying the dopamine gating window duration to 600 ms, 400 ms and 300 ms respectively.

observed in dopamine transporter (DAT) knock-out mice which have increased extracellular DA levels. These animals displayed hyperactivity, stereotypies, perseverative patterns and distorted exploratory behaviors [44–46].

Considering the evidence discussed above, it was interesting to evaluate the differences between the systemic effects of this drug and the local effects in brain areas relevant to operant conditioning learning. To study this issue, we locally injected the rats with entacapone in the mPFC during learning an operant conditioning task. Treated animals presented an improved learning performance: in the first session they reached 82.5% of correct responses compared with 24.0% of correct responses in the sham group, and in the second session they reached 100% of correct responses, whereas control animals needed four sessions to completely learn the task. In addition, latency of response for these animals was significantly shorter than those in the first three sessions. On the other hand, our results showed that learning was impaired and latency was increased when animals were locally administered with recombinant COMT in the mPFC. These findings are complementary to previous works showing that overexpression of human COMT impairs working memory and learning in mice [16,47].

We discard that the results here reported were due to an increase in locomotor activity or vigor, since latencies of the first session do not significantly differ among intra-session blocks [$F_{4,310} = 1.064$, $p = 0.3745$] (Supplemental Figure 1). Therefore differences in latencies of response found here respond to the different treatments. Additionally entacapone *per se* or in combination with levodopa/carbidopa has been shown not to increase locomotor activity [48,49] and we did not observe differences in the weight and food intake between the control groups and the groups with COMT manipulations (data not shown). The results presented here indicate that learning performance in an operant conditioning is strongly influenced by COMT in the mPFC.

The results from the computational model showed a pronounced influence of the DA bioavailability in the mPFC – simulated by the gating window duration (T) – on the neural ensemble: when the length of the gating window was increased the learning performance was enhanced. On the contrary, shortening the gating window, *i.e.* mimicking the effect of recombinant COMT treatment produces an impaired learning performance. Analysis of synaptic weight changes in the mPFC showed that simulations with longer T values presented bigger changes compared to the control group, while reducing the gating window duration had the opposite effect (*i.e.* a decrease in synaptic modifications). In the simulations, an increased (decreased) DA bioavailability enables the increase (decrease) of synaptic plasticity during the learning process. Based on these computational results and their direct relationship with learning performance, we theorize that enhanced DA clearance reduces cortical synaptic plasticity, whereas a prolonged DA availability increases synaptic plasticity in the mPFC, improving learning performance in operant conditioning learning.

Goal directed behaviors sometimes become habitual responses [50,51]. The striatum is typically associated to habituation and motor control, whereas PFC to goal directed behavior. Different computational models studies investigated how these two independent controllers – the striatum and the PFC – are integrated [24,52]. Daw et al. [52] developed a computational model that integrates these structures focusing in how goal directed behaviors proceed as habitual responses. Using a set of computational methods of reinforcement learning, the authors [52] identify a trade-off between PFC and the striatum and suggest a Bayesian principle of arbitration between them according to uncertainty, so each controller is deployed when it should be most accurate. From a different perspective, in a previous work [24] we found that time courses of learning-related activity in PFC and striatum depend on the task's complexity using a simpler version of our neural network model. When training this computational model in a simple operant task, activity in PFC increases later and less abruptly than in BG [24]. However, the experiments and simulations here presented demonstrate that PFC indeed plays a role in controlling the speed of learning in a simple operant conditioning task.

Taken together, the experimental and simulated results showed that learning performance and plasticity could be deeply influenced by COMT manipulations in the mPFC. The connection between the computational theory and the experimental results provides a quantitative framework for future experiments to the further understanding of operant learning.

Conflict of interest

The authors have declared that no competing interests exist.

Acknowledgments

We are very grateful to Sergio Lew for the computational theory simulation, and helpful discussions and comments. Funding was from the Agencia Nacional de Promocion Cientifica y Tecnologica (PICT 2012 1519), Consejo Nacional de Investigaciones Cientificas y Tecnicas (PIP 112 201101 01054) and Universidad de Buenos Aires (UBACYT 20020130100485BA).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bbr.2014.11.031.

References

[1] Watanabe M, Sakagami M. Integration of cognitive and motivational context information in the primate prefrontal cortex. Cereb Cortex 2007;17:i101–9, http://dx.doi.org/10.1093/cercor/bhm067.
[2] Kennerley SW, Wallis JD. Reward-dependent modulation of working memory in lateral prefrontal cortex. J Neurosci 2009;29:3259–70, http://dx.doi.org/10.1523/JNEUROSCI.5353-08.2009.
[3] Funahashi S, Bruce CJ, Goldman-Rakic PS. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J Neurophysiol 1989;61:331–49.
[4] Miner LAH, Ostrander M, Sarter M. Effects of ibotenic acid-induced loss of neurons in the medial prefrontal cortex of rats on behavioral vigilance: evidence for executive dysfunction. J Psychopharmacol 1997;11:169–78, http://dx.doi.org/10.1177/026988119701100210.
[5] Seamans JK, Yang CR. The principal features and mechanisms of dopamine modulation in the prefrontal cortex. Prog Neurobiol 2004;74:1–58, http://dx.doi.org/10.1016/j.pneurobio.2004.05.006.
[6] Tsai HC, Zhang F, Adamantidis A, Stuber GD, Bonci A, de Lecea L, et al. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. Science 2009;324:1080–4, http://dx.doi.org/10.1126/science.1168878.
[7] Goto Y, Otani S, Grace AA. The Yin and Yang of dopamine release: a new perspective. Neuropharmacology 2007;53:583–7, http://dx.doi.org/10.1016/j.neuropharm.2007.07.007.
[8] Schultz W. Getting formal with dopamine and reward. Neuron 2002;36:241–63, http://dx.doi.org/10.1016/S0896-6273(02)00967-4.
[9] Käenmäki M, Tammimäki A, Myöhänen T, Pakarinen K, Amberg C, Karayiorgou M, et al. Quantitative role of COMT in dopamine clearance in the prefrontal cortex of freely moving mice. J Neurochem 2010;114:1745–55, http://dx.doi.org/10.1111/j.1471-4159.2010.06889.x.
[10] Yavich L, Forsberg MM, Karayiorgou M, Gogos JA, Männistö PT. Site-specific role of catechol-O-methyltransferase in dopamine overflow within prefrontal cortex and dorsal striatum. J Neurosci 2007;27:10196–209, http://dx.doi.org/10.1523/JNEUROSCI.0665-07.2007.
[11] Mundorf ML, Joseph JD, Austin CM, Caron MG, Wightman RM. Catecholamine release and uptake in the mouse prefrontal cortex. J Neurochem 2001;79:130–42, http://dx.doi.org/10.1046/j.1471-4159.2001.00554.x.
[12] Tan HY, Chen Q, Goldberg TE, Mattay VS, Meyer-Lindenberg A, R Weinberger D, et al. Catechol-O-methyltransferase Val158Met modulation of prefrontal–parietal–striatal brain systems during arithmetic and temporal transformations in working memory. J Neurosci 2007;27:13393–401, http://dx.doi.org/10.1523/JNEUROSCI.4041-07.2007.
[13] Goldman D, Weinberger DR, Malhotra AK, Goldberg TE. The role of COMT Val158Met in cognition. Biol Psychiatry 2009;65:e1–4, http://dx.doi.org/10.1016/j.biopsych.2008.07.032.

[14] Tunbridge EM, Bannerman DM, Sharp T, Harrison PJ. Catechol-o-methyltransferase inhibition improves set-shifting performance and elevates stimulated dopamine release in the rat prefrontal cortex. J Neurosci 2004;24:5331–5, http://dx.doi.org/10.1523/JNEUROSCI.1124-04.2004.

[15] Lapish CC, Ahn S, Evangelista LM, So K, Seamans JK, Phillips AG. Tolcapone enhances food-evoked dopamine efflux and executive memory processes mediated by the rat prefrontal cortex. Psychopharmacology (Berlin) 2009;202:521–30, http://dx.doi.org/10.1007/s00213-008-1342-1.

[16] Papaleo F, Crawley JN, Song J, Lipska BK, Pickel J, Weinberger DR, et al. Genetic dissection of the role of catechol-O-methyltransferase in cognition and stress reactivity in mice. J Neurosci 2008;28:8709–23, http://dx.doi.org/10.1523/JNEUROSCI.2077-08.2008.

[17] Tunbridge EM, Huber A, Farrell SM, Stumpenhorst K, Harrison PJ, Walton ME. The role of catechol-o-methyltransferase in reward processing and addiction. CNS Neurol Disord Drug Targets 2012;11:306–23, http://dx.doi.org/10.2174/187152712800672409.

[18] Cohen JD, Braver TS, Brown JW. Computational perspectives on dopamine function in prefrontal cortex. Curr Opin Neurobiol 2002;12:223–9, http://dx.doi.org/10.1016/S0959-4388(02)00314-8.

[19] O'Reilly RC, Herd SA, Pauli WM. Computational models of cognitive control. Curr Opin Neurobiol 2010;20:257–61, http://dx.doi.org/10.1016/j.conb.2010.01.008.

[20] Durstewitz D, Seamans JK, Sejnowski TJ. Neurocomputational models of working memory. Nat Neurosci 2000;3:1184–91, doi:10.1038/81460.

[21] Sutton RS. Learning to predict by the method of temporal difference. Mach Learn 1988;3:9–44.

[22] Schultz W. Predictive reward signal of dopamine neurons. J Neurophysiol 1998;80:1–27.

[23] Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science 1997;275:1593–9, http://dx.doi.org/10.1126/science.275.5306.1593.

[24] Lew SE, Rey HG, Gutnisky DA, Zanutto BS. Differences in prefrontal and motor structures learning dynamics depend on task complexity: a neural network model. Neurocomputing 2008;71:2782–93, http://dx.doi.org/10.1016/j.neucom.2007.09.010.

[25] Rapanelli M, Lew SE, Frick LR, Zanutto BS. Plasticity in the rat prefrontal cortex: linking gene expression and an operant learning with a computational theory. PLoS ONE 2010;5:e8656, http://dx.doi.org/10.1371/journal.pone.0008656.

[26] Lew SE, Zanutto BS. A computational theory for the learning of equivalence relations. Front Hum Neurosci 2011;5:113, http://dx.doi.org/10.3389/fnhum.2011.00113.

[27] Pasupathy A, Miller EK. Different time courses of learning-related activity in the prefrontal cortex and striatum. Nature 2005;433:873–6, http://dx.doi.org/10.1038/nature03287.

[28] Schweimer J, Hauber W. Involvement of the rat anterior cingulate cortex in control of instrumental responses guided by reward expectancy. Learn Mem 2005;12(3):334–42, http://dx.doi.org/10.1101/lm.90605.

[29] Fallon JH, Loughlin SE, *Substantia nigra* In: Paxinos G, editor. The rat nervous system. London: Academic Press; 1995. p. 215–37.

[30] Paxinos G, Watson C. The rat brain in stereotaxic coordinates. 6th ed. San Diego: Elsevier Academic Press; 2007.

[31] Rapanelli M, Frick LR, Bernardez-Vidal M, Zanutto BS. Different MK-801 administration schedules induce mild to severe learning impairments in an operant conditioning task: role of buspirone and risperidone in ameliorating these cognitive deficits. Behav Brain Res 2013;257:156–65, http://dx.doi.org/10.1016/j.bbr.2013.09.043.

[32] Rapanelli M, Frick LR, Zanutto BS. Modulation of endothelial and neuronal nitric oxide synthases during learning of an operant conditioning task. J Neurochem 2010;113:725–34, http://dx.doi.org/10.1111/j.1471-4159.2010.06640.x.

[33] Rapanelli M, Frick LR, Zanutto BS. Differential gene expression in the rat hippocampus during learning of an operant conditioning task. Neuroscience 2009;163:1031–8, http://dx.doi.org/10.1016/j.neuroscience.2009.07.037.

[34] Brannan T, Prikhojan A, Yahr MD. Peripheral and central inhibitors of catechol-O-methyl transferase: Effects on liver and brain COMT activity and L-DOPA metabolism. J Neural Transm Gen Sect 1997;104:77–87, http://dx.doi.org/10.1007/BF01271296.

[35] Forsberg M, Lehtonen M, Heikkinen M, Savolainen J, Järvinen T, Männistö PT. Pharmacokinetics and pharmacodynamics of entacapone and tolcapone after acute and repeated administration: a comparative study in the rat. J Pharmacol Exp Ther 2003;304:498–506, http://dx.doi.org/10.1124/jpet.102.042846.

[36] O'Reilly RC, Noelle DC, Braver TS, Cohen JD. Prefrontal cortex and dynamic categorization tasks: representational organization and neuromodulatory control. Cereb Cortex 2002;12:246–57, http://dx.doi.org/10.1093/cercor/12.3.246.

[37] Pan WX, Schmidt R, Wickens JR, Hyland BI. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. J Neurosci 2005;25:6235–42, http://dx.doi.org/10.1523/JNEUROSCI.1478-05.2005.

[38] Garris PA, Wightman RM. Different kinetics govern dopaminergic transmission in the amygdala, prefrontal cortex, and striatum: an in vivo voltammetric study. J Neurosci 1994;14:442–50.

[39] Gonon F. Prolonged and extrasynaptic excitatory action of dopamine mediated by D1 receptors in the rat striatum in vivo. J Neurosci 1997;17:5972–8.

[40] Durstewitz D, Seamans JK. The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. Biol Psychiatry 2008;64:739–49, http://dx.doi.org/10.1016/j.biopsych.2008.05.015.

[41] Li YH, Wirth T, Huotari M, Laitinen K, MacDonald E, Männistö PT. No change of brain extracellular catecholamine levels after acute catechol-O-methyltransferase inhibition: a microdialysis study in anaesthetized rats. Eur J Pharmacol 1998;356:127–37, http://dx.doi.org/10.1016/S0014-2999(98)00524-X.

[42] Tammimäki A, Männistö PT. Effect of genetic modifications in the synaptic dopamine clearance systems on addiction-like behaviour in mice. Basic Clin Pharmacol Toxicol 2011;108:2–8, http://dx.doi.org/10.1111/j.1742-7843.2010.00647.x.

[43] Babovic D, O'Tuathaigh CM, O'Connor AM, O'Sullivan GJ, Tighe O, Croke DT, et al. Phenotypic characterization of cognition and social behavior in mice with heterozygous versus homozygous deletion of catechol-O-methyltransferase. Neuroscience 2008;155:1021–9, http://dx.doi.org/10.1016/j.neuroscience.2008.07.006.

[44] Gainetdinov RR, Wetsel WC, Jones SR, Levin ED, Jaber M, Caron MG. Role of serotonin in the paradoxical calming effect of psychostimulants on hyperactivity. Science 1999;283:397–401, http://dx.doi.org/10.1126/science.283.5400.397.

[45] Spielewoy C, Gonon F, Roubert C, Fauchey V, Jaber M, Caron MG, et al. Increased rewarding properties of morphine in dopamine-transporter knockout mice. Eur J Neurosci 2000;12:1827–37, http://dx.doi.org/10.1046/j.1460-9568.2000.00063.x.

[46] Ralph RJ, Paulus MP, Fumagalli F, Caron MG, Geyer MA. Prepulse inhibition deficits and perseverative motor patterns in dopamine transporter knockout mice: differential effects of D1 and D2 receptor antagonists. J Neurosci 2001;21:305–13.

[47] Suzuki G, Harper KM, Hiramoto T, Funke B, Lee M, Kang G, et al. Over-expression of a human chromosome 22q11.2 segment including TXNRD2, COMT and ARVCF developmentally affects incentive learning and working memory in mice. Hum Mol Genet 2009;18:3914–25, http://dx.doi.org/10.1093/hmg/ddp334.

[48] Männistö PT, Lang A, Rauhala P, Vasar E. Beneficial effects of co-administration of catechol-O-methyltransferase inhibitors and L-dihydroxyphenylalanine in rat models of depression. Eur J Pharmacol 1995;274:229–33, http://dx.doi.org/10.1016/0014-2999(94)00776-4.

[49] Törnwall M, Tuomainen P, Männistö PT. Neurochemical and psychomotor interactions of new selective COMT inhibitors with clorgyline and nomifensine in levodopa-treated rats and mice. Arch Int Pharmacodyn Ther 1992;320:5–20.

[50] Aarts H, Dijksterhuis A. Habits as knowledge structures: automaticity in goal-directed behavior. J Pers Soc Psychol 2000;78:53–63, http://dx.doi.org/10.1037/0022-3514.78.1.53.

[51] Yin HH, Knowlton BJ. The role of the basal ganglia in habit formation. Nat Rev Neurosci 2006;7:464–76, http://dx.doi.org/10.1038/nrn1919.

[52] Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci 2005;8:1704–11, http://dx.doi.org/10.1038/nn1560.