

ORIGINAL ARTICLE

Prefrontal Regulation of Neuronal Activity in the Ventral Tegmental Area

Yong Sang Jo¹ and Sheri J.Y. Mizumori^{1,2}¹Department of Psychology, University of Washington, Seattle, WA, USA and ²Neuroscience Program, University of Washington, Seattle, WA 98195, USA

Address correspondence to Sheri J.Y. Mizumori, Email: mizumori@uw.edu

Abstract

Growing evidence indicates that midbrain dopamine (DA) cells integrate reward expectancy-related information from the prefrontal cortex to properly compute errors in reward prediction. Here we investigated how 2 major prefrontal subregions, the orbitofrontal cortex (OFC) and the medial prefrontal cortex (mPFC), contributed to DAergic prediction errors while rats performed a delay discounting task on a T-maze. Most putative DA cells in the task showed phasic responses to salient cues that predicted delayed rewards, but not to the actual rewards. After temporary inactivation of the OFC, putative DA cells exhibited strikingly reduced phasic responses to reward-predicting cues but increased responses to rewards. In contrast, mPFC inactivation significantly elevated DA responses to both predictive cues and rewards. In addition, OFC, but not mPFC, inactivation disrupted the activity of putative non-DA cells that encoded expected reward values during waiting periods. These results suggest that the 2 prefrontal subregions differentially regulate DAergic prediction errors and the OFC conveys value signals to midbrain dopamine systems.

Key words: decision-making, delay discounting, dopamine, orbitofrontal cortex, prefrontal cortex

Introduction

Midbrain dopamine (DA) cells encode differences between expected and actually received rewards and between previous and current estimates of future rewards (Montague et al. 1996; Schultz et al. 1997; Bayer and Glimcher 2005; Pan et al. 2005). These DAergic prediction errors play an essential role in assigning values to reward-predicting stimuli in reinforcement learning (Fiorillo et al. 2003; Tobler et al. 2005; Flagel et al. 2011). Such acquired or cached values are thought to be stored in a common neural currency in the striatum which, in turn, provides the value information to DA neurons in preparation for subsequent computation of prediction errors (Apicella et al. 1992; Suri and Schultz 2001; O'Doherty et al. 2004; Daw et al. 2005).

As a different valuation system, the prefrontal cortex (PFC) is also known to encode information about expected outcome values (Watanabe 1996; Padoa-Schioppa and Assad 2006; Roesch et al. 2006; Kim et al. 2008; Kennerley and Wallis 2009; Sul et al. 2010). The literature suggests that the PFC can flexibly adjust

outcome values in response to motivational or environmental changes even before experiencing the outcome (Corbit and Balleine 2003; Izquierdo et al. 2004; Ostlund and Balleine 2005; Jones et al. 2012; Takahashi et al. 2013). Due to this characteristic of updating values independent of reward prediction errors, little attention has been paid to the possibility that the PFC may directly influence computations of DAergic error signals. Recent findings indicate the orbitofrontal cortex (OFC) and medial PFC (mPFC) are the major PFC subregions that convey reward expectancies to DA cells (Takahashi et al. 2011; Jo et al. 2013). To further determine the differential contributions of the 2 structures to DAergic prediction errors, we used a within-cell design to analyze reward- and action-based neuronal activity in the ventral tegmental area (VTA) in a delay discounting task that required freely navigating rats to choose between 2 goal locations that were associated with different waiting periods before gaining access to chocolate milk reward. The decision-making task was intended to examine the animals' choice behavior and DA responses in a

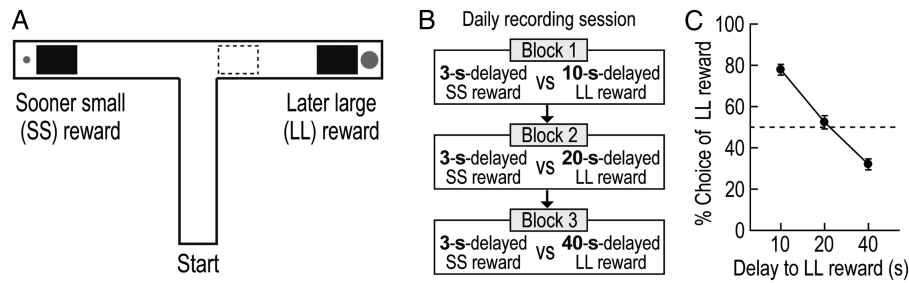


Figure 1. Choice performance in the delay-based decision-making task. (A) Illustration of the T-maze. Rectangular wooden barriers (black squares) were placed before the food cups on 2 goal arms to control animals' access to SS and LL rewards. When rats chose the goal arm associated with LL reward, an additional barrier (dashed rectangle) was placed at its entrance to prevent the animals from exiting the goal arm during the waiting period. (B) Daily behavioral recording procedures. Three different delays to LL reward were randomly ordered and tested in separate blocks of trials. The delay to SS reward remained unchanged. Each block consisted of forced-choice trials, followed by free-choice trials. (C) Choice preference for LL reward as a function of delay to LL reward. Error bars indicate SEM.

more naturalistic foraging environment than in Pavlovian and operant paradigms. Then we examined significant alternations in firing of putative DA and non-DA cells before and after temporary inactivation of each subregion within the same subjects. The results of this study provide valuable information about the distinct roles of the OFC and the mPFC in regulating VTA neuronal activity.

Materials and Methods

Subjects

Thirteen male Long-Evans rats (320–400 g; Simonson Labs, Gilroy, CA) were individually housed and initially allowed free access to food and water. Then food was restricted to maintain their body weights at 85% of free-feeding weights. All experiments were conducted during the light phase of a 12 h light/dark cycle (lights on at 7:00 am), in accordance with the University of Washington's Institutional Animal Care and Use Committee guidelines.

Behavioral Apparatus

An elevated T-maze (Fig. 1A), consisting of one start arm (the middle stem) and 2 goal arms (58 × 5.5 cm each), was located at the center of a circular curtained area. Each goal arm contained a metal food cup (0.7 cm in diameter × 0.6 cm deep) at the end, in front of which was a wooden barrier (10 × 4 × 15 cm) to control access to reward. The arm was hinged such that its proximal end closest to the maze center could be lowered by remote control if needed.

Presurgical Training

Each rat was placed on the maze and allowed to forage for chocolate milk drops scattered throughout the maze. Afterward, the rat was shaped to retrieve a reward delayed by 3 s (0.15 mL) only from 2 goal arms. Specifically, the animal was trained to run down the start arm and freely choose one goal arm. Upon arrival in front of the barrier, the rat had to wait for 3 s. As the barrier was removed by an experimenter who measured the elapsed time using a digital stopwatch, the rat was allowed to approach and consume the reward. After rebaiting the food cup and putting the barrier back in place, the experimenter gently guided the rat to the start arm for the next trial. Once the rat could finish 16 trials within 20 min, it underwent stereotaxic surgery.

Surgery

Tetrodes were made from 20 μm lacquer-coated tungsten wires (California Fine Wire, Grover Beach, CA) and final impedance of

each wire was adjusted to 0.2–0.4 MΩ (tested at 1 kHz). Six individually drivable tetrodes were chronically implanted in the right hemisphere dorsal to VTA (5.3 mm posterior to bregma, 0.8 mm lateral to midline, and 7.0 mm ventral to the brain surface). Some rats also received bilateral implantation of guide cannulae (25 gauge) aimed at the OFC (3.2 mm anterior, 3.1 mm lateral, and 4.6 mm ventral to bregma) and the mPFC (3.2 mm anterior, 0.7 mm lateral, and 3.3 mm ventral to bregma). A 33-gauge dummy cannula was inserted into each guide cannula to prevent blockade.

Delay Discounting Task

After a week of recovery, 2 groups of rats performed slightly different decision-making tasks on the maze. Seven rats implanted with recording electrodes alone performed a delay-based decision-making task in which they were required to choose between a sooner small (SS) reward and a later large (LL) reward (Fig. 1B). The delay to SS reward (0.05 mL) was held constant at 3 s throughout the experiments, whereas 3 different delays (10, 20 and 40 s) before LL reward (0.3 mL) were used to test possible changes in choice preference as a function of delay to LL reward. Thus, a daily testing session consisted of 3 blocks of trials to which the 3 delays were randomly assigned and only one delay was presented in a given block. To inform rats as to which delay was imposed before LL reward, each block began with 10 forced-choice trials followed by 6 or 8 free-choice trials. During the forced-choice trials, 5 SS and 5 LL reward trials were pseudorandomly ordered and only one goal arm was made available in each trial by lowering the other goal arm. Both goal arms were presented during the free-choice trials in which animals' choice performance was measured. As the rats selected and entered the goal arm associated with LL reward, an additional barrier was located at its entrance to confine them in the chosen arm during the longer delays. The 3 blocks were separated by an interblock interval of 5–10 min during which the animals were placed on a holding area adjacent to the maze. The location of SS and LL rewards in the goal arms remained constant within each rat but was counterbalanced across rats.

The second group of 6 rats implanted with bilateral cannulae as well as recording electrodes was trained in a modified version of the delay discounting task in which the delay to LL reward remained unvaried at 10 s across blocks (Fig. 5C). In a recording session, the location of LL and SS rewards was randomly selected and the rats were required to choose between a 3-s-delayed SS reward and a 10-s-delayed LL reward in the first block of trials. After either muscimol (MUS; a GABA receptor agonist) or saline (SAL) was injected into one of the PFC subregions, the second block of trials was tested. A total of 8 drug testing sessions were given per subregion of each rat.

Single-Unit Recording

Using a Cheetah data acquisition system (Neuralynx, Bozeman, MT), neural activity was monitored daily as described previously (Norton et al. 2011; Jo et al. 2013). Briefly, unit signals were amplified, filtered at 600–6000 Hz, and digitized at 16 kHz. The animal's head position was also recorded at 30 Hz by tracking 2 light-emitting diodes mounted on the headstage. To increase the chance of recording from putative DA cells, single-units displaying low discharge frequencies and wide spike waveforms were prescreened and recorded subsequently. If such units were not detected, tetrodes were advanced in 40 μm increments, up to 160 μm per day. High-firing units were recorded only if a DA-like unit was present on the same tetrode. All tetrodes were advanced at the end of each recording session to find new cells. For recording sessions involving drug infusion, however, tetrodes were lowered in an attempt to hold the same units for 2 consecutive days. In this way, the same units could be tested with either SAL or MUS injections into one PFC subregion. During the recording session, 3 salient events such as delay onset (DO), delay termination (DT), and reward were inserted into the data stream online. Specifically, timestamps for DO and DT were marked when an experimenter pressed the stopwatch buttons. The time of reward encounters was automatically detected by "lick-detectors" (custom designed by Neuralynx) when the animals first licked the chocolate milk in the food cups.

Intracranial Microinjection

A 33-gauge injection cannula extending 1 mm below the tip of the guide cannula was connected by polyethylene tubing (PE 20) to a 10- μL Hamilton syringe (Jo et al. 2007). Using a microinfusion pump (KD Scientific, Holliston, MA), MUS (1 $\mu\text{g}/\mu\text{L}$ dissolved in SAL) or its vehicle was bilaterally infused at a volume of 0.3 μL /side while rats were under light gas anesthesia by isoflurane. The injection cannula was left in place for an additional 1 min to ensure proper diffusion from its tip. Then they were returned to their home cage and the behavioral recording resumed in 30 min.

Histology

After completion of all experiments, small marker lesions were made by passing a 10 μA current for 10 s through 2 wires of each tetrode. As previously described (Jo and Lee 2010), all rats were transcardially perfused and their brains were stored in a 10% formalin/30% sucrose solution and cut in coronal sections (40 μm) on a freezing microtome. The sections were stained with cresyl violet and examined under light microscope to reconstruct tetrode tracks through the VTA and cannula tip locations in the PFC. Data recorded only from the VTA were analyzed.

Unit Classification

Single units (>2:1 signal-to-noise ratio) were isolated by clustering various spike waveform parameters using Offline sorter (Plexon, Dallas, TX). For some units that were recorded more than one day, the session in which the units were most clearly isolated from other units and background noise was used for analysis. Putative DA cells in the VTA were identified by a cluster analysis that was performed on 2 distinct waveform features of DA cells, spike duration, and amplitude ratio, as previously described (Roesch et al. 2007; Takahashi et al. 2009; Takahashi et al. 2011). To verify whether the classified DA cells were sensitive to D2 agonists such as quinpirole, a subset of these cells was additionally recorded

and their spontaneous activity was compared before and after quinpirole injection (0.4 mg/kg, s.c.).

Data Analysis

Spontaneous firing properties of putative DA cells were calculated from data collected while rats were placed in a holding area prior to the first block of trials and between blocks; these included mean firing rate and the percentage of spikes that occurred in bursts. A burst was defined as successive spiking with an interspike interval of <80 ms followed by an interspike interval of >160 ms (Grace and Bunney 1984). To examine DAergic phasic firing, peri-event time histograms (PETHs; bin width, 50 ms) were constructed separately for the 5 s period around DT and reward events in each block of all trials. Both forced-choice and free-choice trials were included for the generation of PETHs to increase the sample size of each reward condition (see Supplementary Fig. 2, 6, and 7 for the results shown separately for the 2 trial types). A putative DA cell was considered to show phasic responses to one of the 2 events or both if it passed the following 2 criteria: 1) its peak firing was observed within the 200 ms epoch around DT or within the 100 ms epoch after reward and 2) its average activity within the corresponding epochs was $\geq 200\%$ of its mean firing rate over each block. To determine a linear relationship between DA activity and delay to LL reward, Spearman's rank correlation coefficient was calculated in each time bin of the PETHs. The significance of correlation was estimated using a permutation test in which firing rates of each bin were randomly shuffled across blocks for 1000 times. A 99% confidence interval was calculated from correlation coefficients of the shuffled data.

Using PETHs, a putative non-DA cell was also categorized as reward-related if it met the following 3 criteria: 1) its highest firing occurred in a 400-ms window after reward (–50 to 350 ms), 2) the average activity during the window was $\geq 200\%$ of its mean firing for the block of all trials, and 3) its late delay activity (–1.5 s from DT) was significant higher than the early delay activity (1.5 s from DO) during at least one of 4 delays as tested by a Wilcoxon Signed Rank test.

Statistical Analysis

Analysis of variances (ANOVAs) with repeated measures were mainly used to test statistical significance of neural activity. Significant differences in firing between 2 reward conditions in a given block were analyzed with *t*-tests. Pearson's correlation tests were performed to establish a relationship between 2 variables. Two-tailed *P* values <0.05 were considered statistically significant. Data are expressed as mean \pm SEM.

Results

Choice Behavior

Seven rats were trained to choose between SS and LL rewards in a delay discounting task on an elevated T-maze (Fig. 1A). To investigate the animals' choice performance as a function of delay to LL reward, 3 different lengths of delay (10, 20, and 40 s) were randomly imposed prior to LL reward in separate blocks of trials (Fig. 1B). However, the delay to SS reward was kept constant at 3 s. The elapsed time since the animal's arrival at the barrier (i.e., DO) was measured using a digital stopwatch. The barrier was manually removed by an experimenter at the time of DT, which caused slight variations in delay length. In a total of 42 behavioral recording sessions (3–12 sessions per rat), SS reward was

delayed by 3.27 ± 0.26 s (mean \pm SD) and LL reward was delayed by 10.19 ± 0.36 , 20.25 ± 0.47 , or 40.34 ± 0.58 s.

Choice performance indicated that the value of LL reward was discounted by the delay preceding it (Fig. 1C). Specifically, rats exhibited a strong preference to a 10 s-delayed LL reward. However, the animals were indifferent when given an option between a SS reward and a 20 s-delayed LL reward. Further extension of the delay to LL reward to 40 s reversed choice preference such that they chose SS reward more frequently. A Pearson's correlation test demonstrated a significant negative relationship between choice performance and delay length prior to LL reward ($r = -0.94$, $P < 0.001$).

Responses of Putative DA Cells During the Delay Discounting Task

Single-unit activity was recorded from 364 neurons in the VTA of rats performing the task (Fig. 2A,B). Of these neurons, 116 cells were classified as putative DAergic by clustering a group of cells showing a wide spike duration and low amplitude ratio (Fig. 2C). Previous studies reported that putative DA cells identified by this cluster analysis encode reward prediction errors (Roesch et al. 2007; Takahashi et al. 2009; Takahashi et al. 2011). After daily recording sessions, a subset of 67 neurons was further tested to verify whether putative DA cells were inhibited by quinpirole, a D2 receptor agonist (Fig. 2D). The cells identified as putative DAergic ($n = 18$) were suppressed (ranged from -97 to -15%) after injecting quinpirole, whereas putative non-DA cells showed heterogeneous responses. *T*-tests revealed that quinpirole injections significantly inhibited spontaneous activity of putative DA cells

($t_{(17)} = 10.43$, $P < 0.001$), but significantly elevated firing rates of putative non-DA cells ($t_{(48)} = 3.61$, $P = 0.001$).

Accumulated evidence indicates that DA cells that are initially activated by the receipt of primary reward become transiently responsive to both a reward-predicting cue and the actual reward in the middle of Pavlovian conditioning, and then exhibit phasic responses only to the predictive cue after learning (Pan et al. 2005; Clark et al. 2010; Cohen et al. 2012). Consistent with the findings, putative DA cells in the current task showed a response shift in time from reward encounters to DT when the barriers on the goal arms were removed. For instance, a representative putative DA cell in the left column of Figure 3A responded to both DT and reward. In the following day, the same cell showed phasic responses only to DT but no longer to reward (Fig. 3A, right column), which indicated that the removal of barriers became a reliable predictor of upcoming rewards. The DT phasic activity persisted until the end of all experiments without an additional shift from DT to other preceding salient events such as the initial encounter with barriers at the time of DO (see Supplementary Fig. 1 for DA responses to DO). Overall, 25 putative DA cells (21.6%) responded to both reward and DT (Fig. 3B) and the fraction of these cells decreased as recording sessions progressed ($r = -0.79$, $P = 0.001$; Fig. 3C). In contrast, most putative DA cells (81/116, 69.8%) were phasically excited only at the time of DT and their fraction increased across recording sessions ($r = 0.58$, $P < 0.05$). The remaining putative DA cells (10/116, 8.6%) did not display task-related activity. Putative DA cells that responded only to reward were not found in the study, presumably because of the extensive pretraining before the recording sessions.

The population histograms of all task-relevant DA cells indicated that DT, but not reward, responses were influenced by delay length (Fig. 3D). Specifically, when Spearman's rank correlation coefficients for individual neuronal responses across 3 blocks of trials were calculated per time bin of 50 ms, significant negative correlations were found around the time of DT prior to LL reward (-50 to 100 ms). No significant relationships were found around the time of DT before SS reward or around the time of reward. These results were also supported by the magnitude of population responses (Fig. 3E). A repeated-measures ANOVA performed on average reward activity found no significant effects of block ($F_{(2,210)} = 1.88$, $P = 0.16$) or reward size ($F_{(1,105)} = 0.29$, $P = 0.59$), but an interaction between the variables was significant ($F_{(2,210)} = 3.84$, $P = 0.02$). In contrast, an ANOVA comparing average DT activity demonstrated significant effects of block ($F_{(2,210)} = 15.52$, $P < 0.001$; Fig. 3E) and reward size ($F_{(1,105)} = 113.67$, $P < 0.001$) and a significant interaction ($F_{(2,210)} = 24.98$, $P < 0.001$). Interestingly, *post hoc* comparisons showed that DT responses prior to LL reward were consistently greater than those before SS reward across blocks (P values < 0.001), even during the block where rats preferred a 3 s-delayed SS reward over a 40 s-delayed LL reward. These results indicated that although putative DA cells encoded the discounted value of LL reward by decreasing DT activity with increasing delays to LL reward, they did not reflect the animals' relative preference of 2 available outcomes within each block of trials (Hollon et al. 2014).

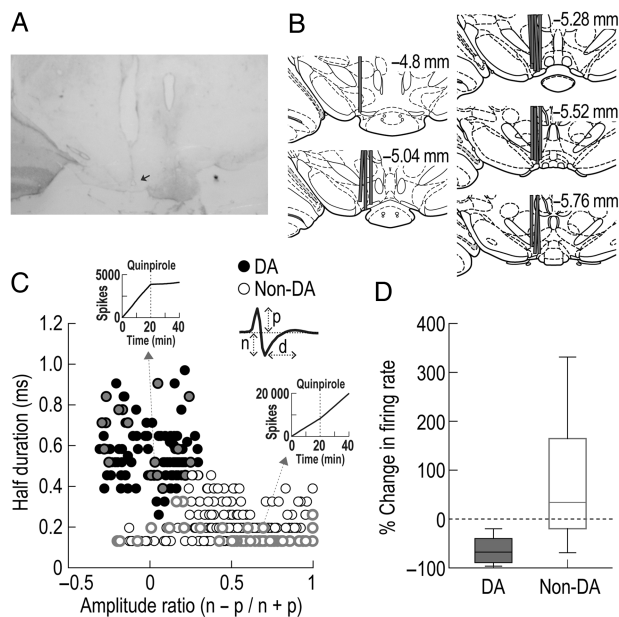


Figure 2. Histological verification of recording sites and classification of putative DA cells in the VTA. (A) A Nissl-stained section showing the final location of a tetrode tip in the VTA. The arrow shows an electrolytic lesion made after the final recording session. (B) Reconstruction of all tetrode tracks. (C) Cluster analysis for all VTA neurons. Putative DA cells in black were identified using 2 waveform features: half spike duration (d) and the amplitude ratio of the first positive peak (p) and negative valley (n). Inset cumulative sum plots depict the effects of quinpirole on spontaneous firing of a putative DA and a putative non-DA cell. (D) Box plots showing changes in spontaneous firing after quinpirole injections. VTA neurons in gray (C) were tested with quinpirole, a D2 receptor agonist, after daily behavioral recording.

Responses of Putative non-DA Cells During the Delay Discounting Task

GABAergic neurons in the VTA have been suggested to encode the value of expected outcomes (Cohen et al. 2012). According to standard economic theories, the subjective value of a delayed reward increases as the delay to the reward shortens while waiting and reaches its peak at the time of reward delivery

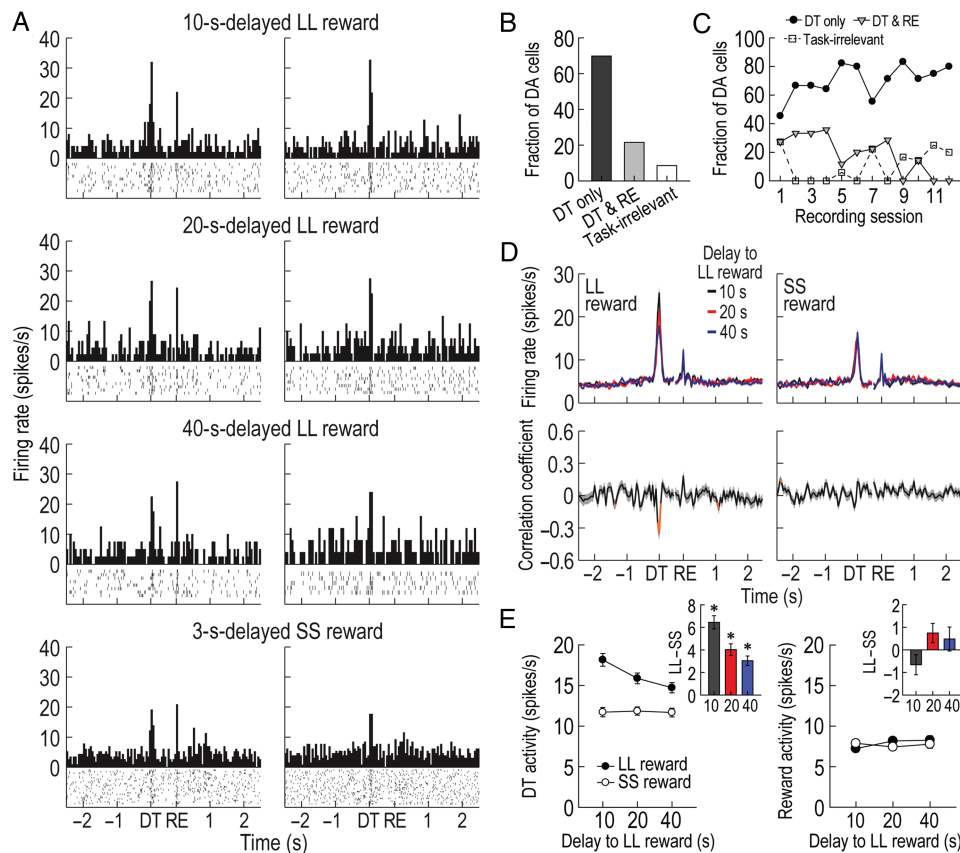


Figure 3. Phasic DA activity in the delay discounting task. (A) A representative putative DA cell recorded for 2 consecutive days. The DA cell initially responded to both DT and reward (RE) in the left column (bin width, 50 ms), but the same cell no longer exhibited phasic responses to reward on the next day (right column). Histograms are aligned to DT and RE. Neuronal responses to both forced-choice and free-choice trials were combined since the results of the 2 trial types were similar (see [Supplementary Fig. 2](#)). (B,C) Fractions of putative DA neurons showing phasic activity at different times. Most putative DA cells were excited only at the time of DT. Task-irrelevant DA cells were not included in the subsequent analyses. (D) Population activity of all task-relevant DA cells. Correlation coefficients for individual neuronal responses across 3 blocks of trials were calculated per each time bin. Orange data points that fell outside the 99% confidence interval obtained from a permutation test for at least 2 consecutive bins were considered significantly correlated (see Materials and Methods). (E) Average DT and reward responses. Inset bar graphs show differential firing between LL and SS reward conditions within blocks of trials (*t*-test, **P* < 0.001). Shaded areas and error bars indicate SEM.

(Samuelson 1937; Frederick et al. 2002; Rangel et al. 2008). In the current study, 16 reward-related non-DA cells showed such firing patterns by continuously ramping up their activity from DO to reward even though they were briefly inhibited at the time of DT when putative DA cells were phasically excited (Fig. 4A; see [Supplementary Results](#) for other groups of putative non-DA cells). In an early phase of 4 different delays (the 1.5 s epoch after DO), reward-related non-DA cells displayed no differential responses across different delays or between SS and LL rewards (Fig. 4B) as indicated by no significant effects of block (ANOVA with repeated measures, $F_{2,30} = 0.01$, $P = 0.99$) or reward size ($F_{1,15} = 0.89$, $P = 0.36$) and no significant interaction between the variables ($F_{2,30} = 2.93$, $P = 0.07$). In the late phase of the delays (the 1.5 s epoch before DT), however, their ramping responses prior to LL reward became higher than those before SS reward (Fig. 4C). The final responses were also more highly elevated with shorter delays to LL reward as significant negative correlations (-1 to -0.1 s before DT) were observed even for the individual neurons (Fig. 4A). A repeated-measures ANOVA confirmed a significant effect of reward size ($F_{1,15} = 40.5$, $P < 0.001$). Although no effect of block ($F_{2,30} = 1.33$, $P = 0.28$) was found, there was a significant interaction between the factors ($F_{2,30} = 5.62$, $P = 0.008$). *Post hoc* comparisons demonstrated a significant difference between 10 and 40 s-delayed LL rewards ($P = 0.04$), but no differences in the other pairwise

comparisons (P values > 0.1). When the rats subsequently obtained reward, the putative non-DA cells showed robust activation that lasted longer for LL reward than for SS reward. A repeated-measures ANOVA revealed a significant effect of reward size ($F_{1,15} = 36.04$, $P < 0.001$; Fig. 4D), whereas neither a significant effect of block ($F_{2,30} = 1.11$, $P = 0.34$) nor an interaction between the variables ($F_{2,30} = 0.51$, $P = 0.6$) was found. These firing patterns suggest that reward-related non-DA cells represent expected reward values in the task.

PFC Contributions to Choice Performance

To test whether the PFC modulated choice behavior and VTA neural activity in the task, a new group of 6 rats were implanted with 2 sets of bilateral cannulae targeted at the OFC and the mPFC (Fig. 5A,B) as well as recording electrodes in the VTA (see [Supplementary Fig. 3A,B](#)). The animals were trained in a modified version of the delay discounting task in which they had to choose between a 3-s-delayed SS reward and a 10-s-delayed LL reward (Fig. 5C). In each testing session, baseline behavioral performance and neural activity were measured in the first blocks of trials. After either MUS or SAL was injected into one subregion, the second block was given to compare significant changes from the baseline. The location of the 2 rewards was pseudorandomly switched between sessions.

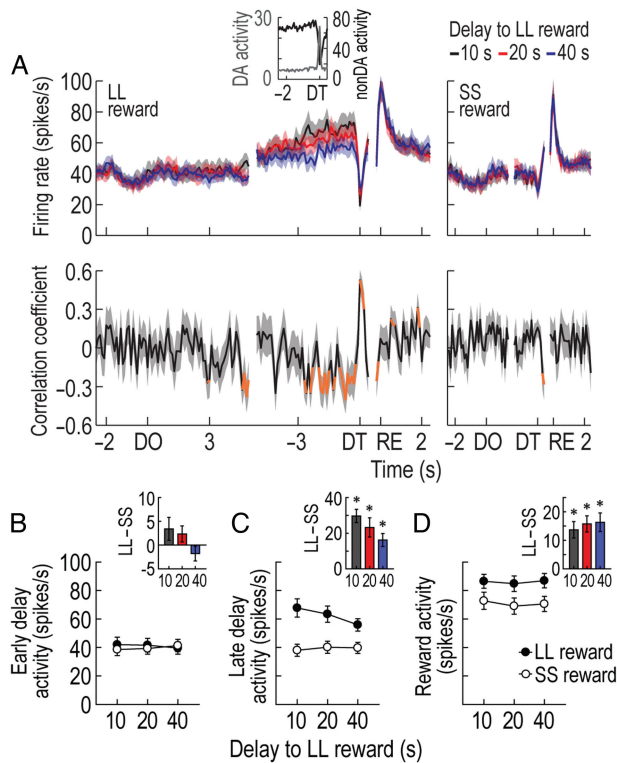


Figure 4. Activity of putative non-DA cells in the delay discounting task. (A) Population histograms (bin width, 100 ms) of reward-responsive non-DA cells and correlation coefficients over the course of delays. Data are aligned to DO, DT, and reward (RE). Significant correlation coefficients for >2 consecutive bins were depicted in orange. The inset graph (bin width, 50 ms) shows that when reward-responsive non-DA cells were abruptly inhibited to a low firing rate at the end of the 10-s delay to LL reward, putative DA cells reached the peak of their phasic responses. (B,C) Average responses during an early phase (1.5 s after DO) and a late phase (1.5 s before DT) of all delays. Inset bar graphs indicate differential firing between LL and SS reward conditions within blocks of trials (t-test, $P < 0.001$). (D) Average reward activity. Shaded areas and error bars show SEM.

These rats normally showed a strong preference for LL reward over SS reward (Fig. 5D). When drugs were injected into the OFC during 42 behavioral recording sessions (20 SAL and 22 MUS sessions; 2–5 SAL and 2–5 MUS sessions per rat), a repeated-measures ANOVA found significant effects of drug ($F_{1,5} = 10.9$, $P = 0.02$) and block ($F_{1,5} = 12.29$, $P = 0.01$) and a significant interaction between drug and block ($F_{1,5} = 51.86$, $P < 0.001$). *Post hoc* comparisons demonstrated that MUS injections significantly decreased the preference for LL reward relative to both baseline performance in the first block and the choice performance after SAL injections (P values < 0.01), whereas behavioral performance was comparable between blocks in SAL sessions ($P = 0.79$). This result was consistent with previous reports indicating OFC dysfunction induces impulsive choice on a T-maze (Rudebeck et al. 2006).

In contrast, drug injections into the mPFC during 38 sessions (18 SAL and 20 MUS sessions; 1–4 SAL and 2–4 MUS sessions per rat) did not cause significant changes in choice performance (Fig. 5E). This observation was confirmed by the results of no effects of drug ($F_{1,5} = 2.47$, $P = 0.18$) or block ($F_{1,5} = 0.44$, $P = 0.54$) and no significant interaction between the variables ($F_{1,5} = 0.75$, $P = 0.43$). The relative larger error bar for the MUS-injected blocks prompted a closer investigation which found a response bias to one spatial location (Fig. 5F). For example, mPFC-inactivated rats rapidly ran down the start arm (see Supplementary Fig. 4

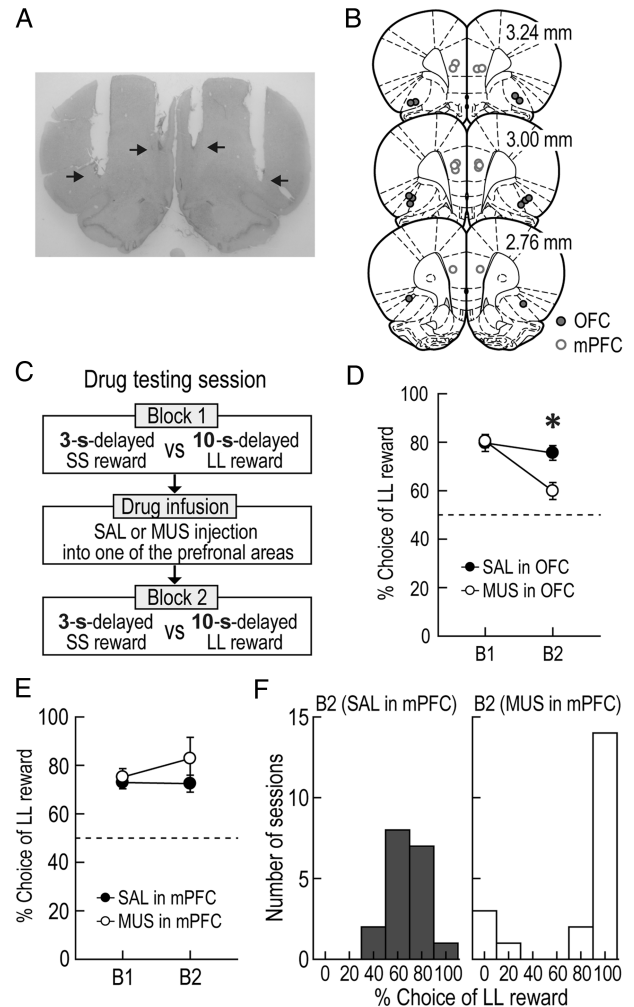


Figure 5. Effects of prefrontal inactivation on choice performance. (A) A Nissl-stained photomicrograph for cannula placement in the OFC and the mPFC. The arrows indicate the tips of injection cannula. (B) Illustration of all microinjection sites. (C) Drug testing procedures. The goal arms associated 3-s-delayed SS and 10-s-delayed LL rewards were randomly selected on each testing day. After the first block of trials (B1) was tested without any drugs, either SAL or MUS was infused into one of the PFC subregions prior to the second blocks (B2). (D,E) Changes in choice preference for LL reward after OFC (D) and mPFC (E) inactivation ($P < 0.01$). (F) Distributions of behavioral performance tested after injecting drugs into the mPFC. Error bars represent SEM.

for the animals' movement) and persistently chose one of the goal arms (only LL reward in 14 behavioral recording sessions and only SS reward in 3 sessions) even though they sampled both reward sizes in the preceding forced-choice trials. The abnormal choice behavior may result from a deficit in responding according to the previously acquired knowledge or rules of the task after mPFC inactivation (Stefani and Moghaddam 2005; Jung et al. 2008).

Prefrontal Regulation of DA Activity

A total of 176, 221, 123, and 155 cells were recorded with SAL and MUS injections into the OFC and the mPFC, respectively. Of these VTA cells, 45, 54, 39, and 52 neurons were identified as putative DAergic in SAL/OFC, MUS/OFC, SAL/mPFC, and MUS/mPFC sessions, respectively, based on the cluster analysis (see Supplementary Fig. 3). To first examine the effects of prefrontal

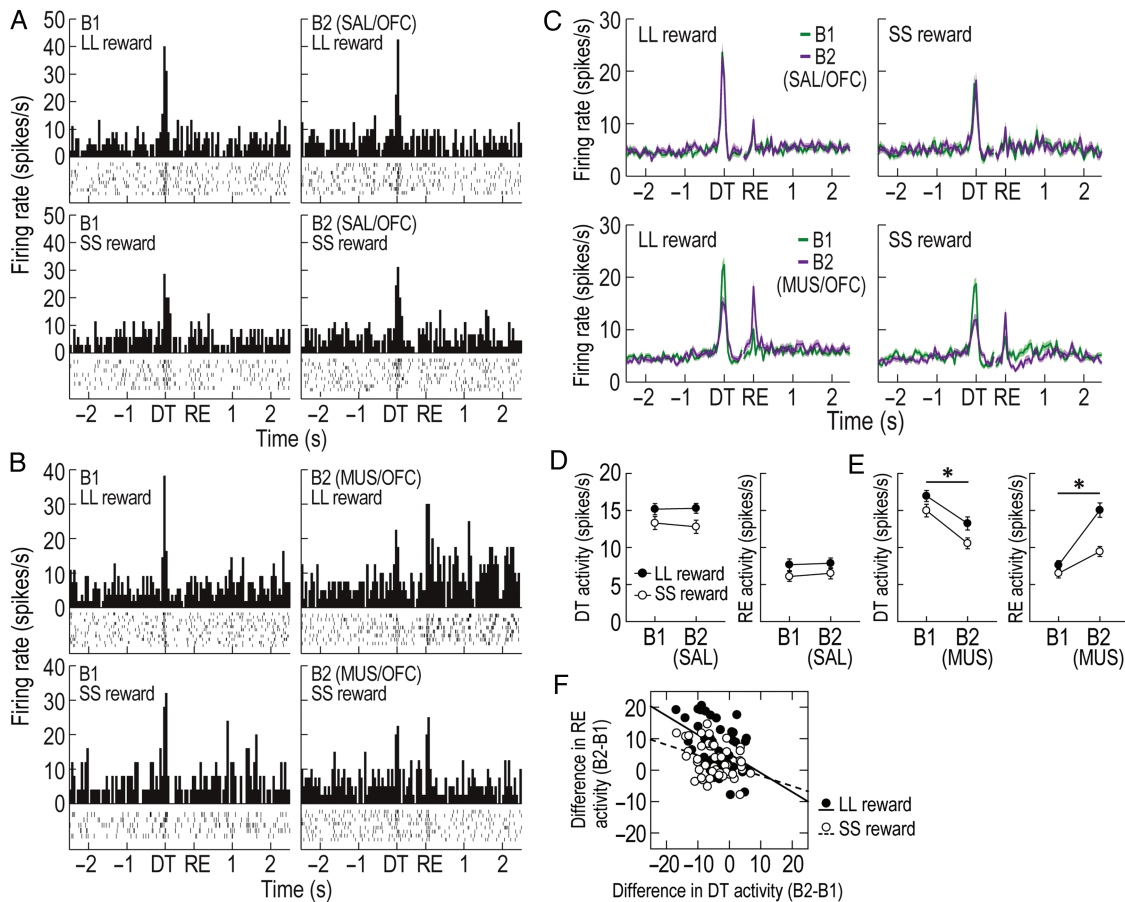


Figure 6. Effects of OFC inactivation on putative DA cells. (A,B) Examples of representative putative DA cells recorded with SAL (A) and MUS (B) injections. Histograms (bin width, 50 ms) are aligned to DT and reward (RE). (C) Population responses of all task-related DA cells. (D,E) Average DT and reward responses in the first (B1) and second blocks (B2) of SAL/OFC (D) and MUS/OFC (E) sessions ($P < 0.001$). (F) Correlations between altered DT and reward responses across blocks in MUS/OFC sessions. Shaded areas and error bars indicate SEM.

inactivation on spontaneous activity of putative DA cells, mean firing rates and percentages of spikes in bursts were measured while rats were not engaged in the task (see [Supplementary Table 1](#)). Paired *t*-tests found that the 2 measures were significantly reduced after MUS injections into the OFC ($t_{(53)}$ values >6.84 , P values <0.001) and the mPFC ($t_{(51)}$ values >3.89 , P values <0.001), but no alternations in spontaneous activity were found in SAL sessions irrespective of the PFC subregions (t values <1.66 , P values >0.1). It is noteworthy that compared with mPFC inactivation, OFC inactivation more severely decreased the percentage of spikes in burst ($t_{(104)} = 2.15$, $P = 0.04$), which is in line with anatomical evidence showing that the OFC sends stronger projections to VTA DA cells than the mPFC ([Watabe-Uchida et al. 2012](#)).

Next, it was analyzed whether OFC manipulations significantly altered DAergic prediction errors during choice performance in the task. In SAL/OFC and MUS/OFC sessions, 38 and 51 putative DA cells, respectively, exhibited task-relevant responses to DT or reward. As seen from representative examples and population responses (Fig. 6A–C), MUS, but not SAL, injections into the OFC caused a marked reduction in DA activity at the time of DT but, conversely, an elevation in response to reward as if the removal of barriers at the time of DT failed to properly predict an upcoming reward and therefore the reward was unexpectedly encountered. Repeated-measures ANOVAs separately performed on DT and reward activity during MUS/OFC sessions demonstrated significant effects of block ($F_{1,50}$ values >44.77 , P values

<0.001 ; Fig. 6E) and reward size ($F_{1,50}$ values >12.72 , P values <0.001). A significant interaction between the factors was found for reward activity ($F_{1,50} = 24.76$, $P < 0.001$), but not for DT activity ($F_{1,50} = 0.7$, $P = 0.41$). During SAL/OFC sessions, both DT and reward responses significantly differed between 2 reward sizes ($F_{1,37}$ values >8.06 , P values <0.01 ; Fig. 6D), but no effect of block ($F_{1,37}$ values <0.78 , P values >0.38) and no interactions between the variables ($F_{1,37}$ values <0.78 , P values >0.38) were found. To further determine a relationship between changes in DT and reward responses in MUS/OFC sessions, differential firing of the 2 responses across blocks was compared within individual neurons (Fig. 6F). Significant inverse correlations for both LL and SS reward conditions (Pearson's correlation, r values less than -0.36 , P values <0.01) indicated that the less putative DA cells responded at the time of DT, the stronger they were excited by reward. These alternations after OFC dysfunction suggest that the OFC may convey the information about expected rewards to DA cells ([Takahashi et al. 2011](#)). However, it seems unlikely that the OFC is the only source for outcome expectancies, because although reduced after OFC inactivation, the DT activity was still stronger in anticipation of LL over SS rewards (Fig. 6E). Other value signals may be fed to DA cells by different brain structures such as the ventral striatum ([Roesch et al. 2009](#); [Day et al. 2011](#); [Clark et al. 2012](#)).

When mPFC function was manipulated with SAL and MUS, 32 and 46 putative DA cells, respectively, showed task-relevant

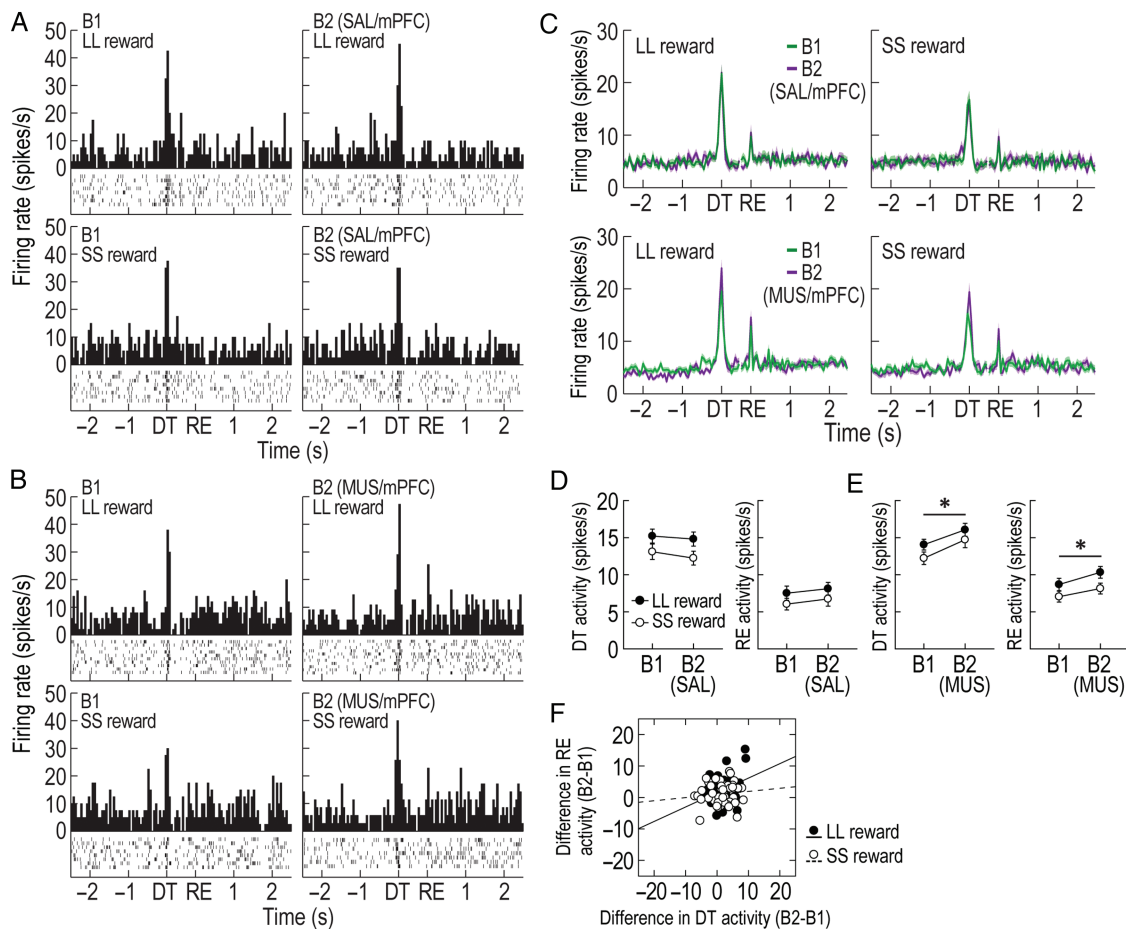


Figure 7. Effects of mPFC inactivation on putative DA cells. (A,B) Examples of representative putative DA cells recorded with SAL (A) and MUS (B) injections. Histograms (bin width, 50 ms) are aligned to DT and reward (RE). (C) Population responses of all task-related DA cells. (D,E) Average DT and reward responses in the first (B1) and second blocks (B2) of SAL/mPFC (D) and MUS/mPFC (E) sessions ($P < 0.01$). (F) Correlations between altered DT and reward responses across blocks in MUS/mPFC sessions. Shaded areas and error bars show SEM.

responses. MUS infusions into the mPFC increased phasic DA responses to both DT and reward, whereas SAL infusions did not alter the 2 responses (Fig. 7A–C). Repeated-measures ANOVAs revealed that both DT and RE responses in MUS/mPFC sessions significantly differed before and after mPFC inactivation ($F_{1,44}$ values > 11.77 , P values < 0.01 ; Fig. 7E) and between SS and LL rewards ($F_{1,44}$ values > 11.4 , P values < 0.01), but interactions between the variables were not significant ($F_{1,44}$ values < 0.58 , P values > 0.45). During SAL/mPFC sessions, the 2 responses differed between 2 reward sizes ($F_{1,31}$ values > 7.2 , P values < 0.05 ; Fig. 7D), but neither effects of block ($F_{1,31}$ values < 2.67 , P values < 0.11) nor interactions ($F_{1,31}$ values < 0.3 , P values < 0.58) were found. As a representative putative DA cell in the Figure 7B displayed elevated activity at the time of both DT and reward after mPFC inactivation, such overall increases in the 2 responses was evident in the population of putative DA neurons in LL reward trials (Pearson's correlation, $r = 0.34$, $P = 0.02$, Fig. 7F), but not in SS reward trials ($r = 0.11$, $P = 0.48$). Nevertheless, these results suggest that mPFC inactivation disinhibits phasic DA responses.

Prefrontal Regulation of non-DA Activity

Alternations in the activity of putative non-DA cells after PFC manipulations were also investigated. Similar to putative DA cells, putative non-DA cells also showed significant decreases in

spontaneous activity when MUS was infused into both PFC subregions (see [Supplementary Table 2](#)). Paired t-tests demonstrated significant differences in firing before and after MUS injections into the OFC and the mPFC (t values > 3.21 , P values < 0.01), whereas no significant changes in spontaneous activity were found in SAL sessions regardless of the subregions (t values < 0.33 , P values > 0.74).

As previously defined, 10, 11, 9, and 13 reward-related non-DA cells that signaled expected reward values were found during SAL/OFC, MUS/OFC, SAL/mPFC, and MUS/mPFC sessions, respectively. MUS injections into the OFC strikingly disrupted the development of ramping activity during the delays prior to both LL and SS rewards (Fig. 8A). This observation was supported by significant effects of block (repeated-measures ANOVA, $F_{1,10} = 10.35$, $P < 0.01$; Fig. 8B) and reward size ($F_{1,10} = 12.02$, $P < 0.01$), but no significant interaction between the factors ($F_{1,10} = 0.25$, $P = 0.63$). Moreover, a planned comparison revealed that the decreased delay responses after OFC inactivation was not different between SS and LL reward conditions ($t_{(10)} = 2.04$, $P = 0.07$). The reward activity was also slightly decreased, but a repeated-measures ANOVA failed to find an effect of block ($F_{1,10} = 1.97$, $P = 0.19$). Instead, an effect of reward size was only significant ($F_{1,10} = 50.72$, $P < 0.001$) without an interaction between the variables ($F_{1,10} = 0.01$, $P = 0.91$). After MUS injections into the mPFC (Fig. 8C) or SAL injections into either subregion, no distinct alternations were observed relative to the baseline firing.

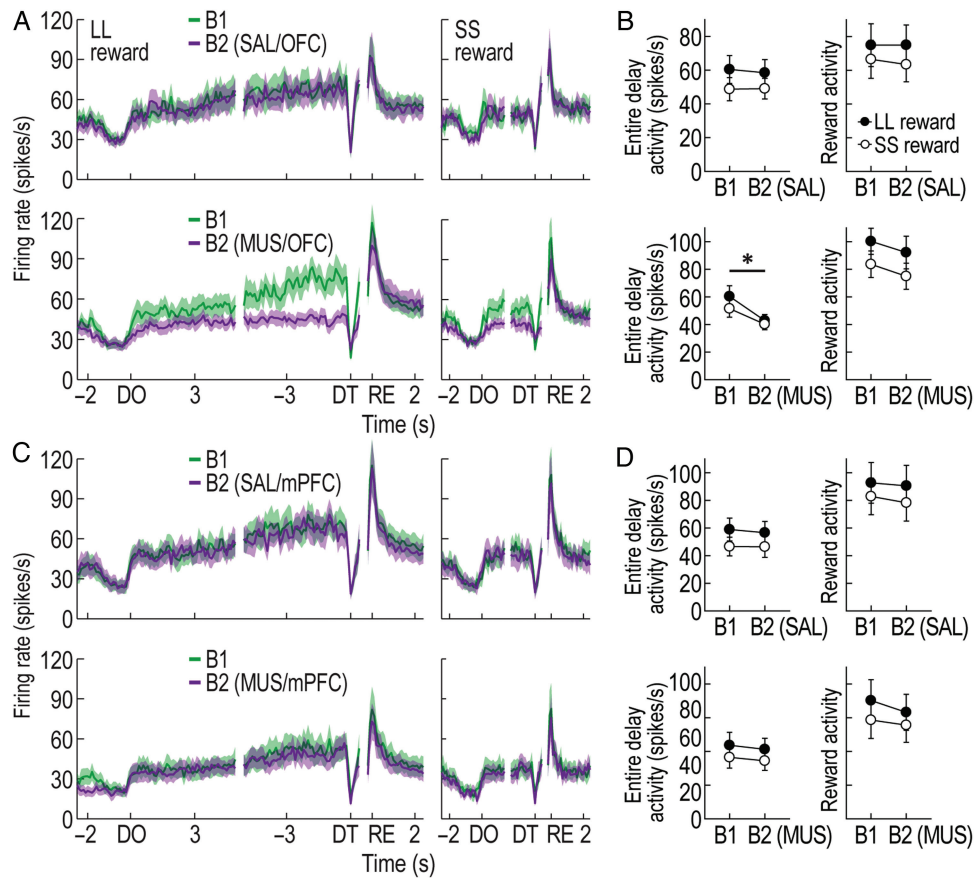


Figure 8. Effects of prefrontal inactivation on putative non-DA cells. (A) Population histograms (bin width, 100 ms) of all reward-responsive non-DA cells recorded with OFC manipulations. Data are aligned to DO, DT, and reward (RE). (B) Average activity during the entire delays and reward activity in the first (B1) and second blocks (B2) of SAL/OFC and MUS/OFC sessions ($P < 0.01$). (C) Population activity of all reward-responsive non-DA cells recorded with mPFC manipulations. (D) Average activity during the entire delays and reward activity with SAL or MUS injections into the mPFC. All graphs represent mean \pm SEM.

Repeated-measures ANOVAs found that delay and reward activity in these sessions significantly differed between 2 reward sizes (F values >7.29 , P values <0.05 ; Fig. 8D). There were no effects of block (F values <4.31 , P values >0.06) and no interactions between the variables (F values <1.73 , P values >0.21). These results provide compelling evidence that the OFC, but not the mPFC, is necessary for signaling expected reward values by putative non-DA cells. Additionally, the effects of PFC inactivation on other groups of putative non-DA cells can be found in [Supplementary Figures 7 and 8](#).

Movement Patterns After PFC Manipulations

The alterations in firing patterns of putative DA and non-DA cells after OFC and mPFC dysfunction might be due to variations in behavioral activity or attention to the wooden barrier at the time of DT, rather than the loss of PFC inputs to the VTA. To examine this possibility, animals' movement velocities were averaged over trials and compared before and after MUS infusion (see [Supplementary Fig. 4](#)). Marked differences in velocity were observed during the decision phase of the task. For instance, while traveling down the start arm, OFC-inactivated rats showed less-motivated behavior as reflected in a slower moving velocity than their baseline velocity in the first block, which resulted in a significant increase in choice latency from the start of each trial to the arrival in front of the barrier on a chosen goal arm (repeated-measures ANOVA, block, $F_{1,21} = 23.98$, $P < 0.001$).

In contrast, mPFC-inactivated rats showed a significant increase in velocity. As a result, their choice latency was significantly shorter than that in the first block ($F_{1,19} = 9.07$, $P = 0.007$). However, the average velocities around the times of DT and reward became comparable between the 2 blocks regardless of PFC manipulations. Repeated-measures ANOVAs also found no significant effect of block in approach latency from DT to reward acquisition in either OFC- or mPFC-inactivated sessions (F values <0.77 , P values >0.39). These findings indicated that the inactivation of the 2 prefrontal subregions did not affect movement patterns when putative DA and non-DA cells displayed task-relevant phasic responses and peak value signals, respectively.

In addition, to assess whether PFC manipulations disrupted the ability to attend to the removal of barriers, the distance between the center of the rat's head and one end wall of the wooden barrier facing towards the maze center was measured at the time of DT in each trial (see [Supplementary Fig. 5](#)). Wilcoxon signed-rank tests showed that the distance distributions for either LL or SS reward conditions not significantly different across blocks in OFC-inactivated sessions (P values >0.86) or mPFC-inactivated sessions (P values >0.82). In $>87\%$ of trials per block, the animals' head was located within 5 cm of the barrier wall, which suggest that irrespective of PFC inactivation, the rats were mostly in a position to easily detect the removal of the barrier at the time of DT. Overall, the behavioral and attentional variables were not responsible for the changes in DA activity induced by PFC inactivation.

Discussion

The current study characterized how putative DA and non-DA cells in the VTA responded in a delay discounting task. Consistent with prediction error signals observed in Pavlovian conditioning paradigms (Schultz et al. 1997; Pan et al. 2005; Clark et al. 2010), putative DA cells initially showed phasic responses to reward, but such responses were no longer present as recording sessions progressed. Instead, these neurons phasically responded when wooden barriers were removed at the time of DT, which indicates that the removal of the wooden barriers served as a reward-predicting cue in the task. The phasic responses did not further shift to other salient events preceding DT, such as the first arrival in front of the barriers at the time of DO. Considering that the manual removal of the barriers caused noticeable variations in length of the same delay periods between trials, it was likely that encountering the barriers did not acquire reward-predictive value, because this event failed to provide precise information about timing of DT and reward (Montague et al. 1996; Pan et al. 2005). Among putative non-DA cells, reward-related neurons gradually ramped up their firing over the course of waiting periods in anticipation of delayed rewards and then reached their peaks at the time of obtaining the rewards. The peak responses were significantly influenced by the size of the encountered rewards no matter how long the preceding delay was. These firing patterns support the view that putative non-DA, presumably GABAergic, cells in the VTA encode the value of expected rewards (Cohen et al. 2012).

Previous primate studies suggest that DA cells signal the subjective value of delayed rewards when cues predicting the future outcomes were presented prior to waiting periods (Fiorillo et al. 2008; Kobayashi and Schultz 2008). In line with the literature, DA activity at the time of DT decreased as the length of delay before LL reward increased. In the absence of reward, a prediction error is computed as a difference between the expected reward values of the current situation or state (e.g., around the time of DT) and the value of the previous state (e.g., before or around the time of DO). Since the current state value exerts a positive influence on the error signal, it is likely that the graded DT activity may result from differential current state values estimated around the time of DT as reflected by different levels of elevated firing of reward-related non-DA cells (Fig. 4A,C). Thus, the present study suggests that phasic DA responses triggered by predictive cues encode the discounted value of future rewards regardless of whether the predictive cues are presented before or after waiting periods.

OFC Contribution to Generating Expected Reward Value

The OFC has long been known to encode the relative value of expected reward (Tremblay and Schultz 1999; O'Doherty et al. 2001; Gottfried et al. 2003; Izquierdo et al. 2004; Padoa-Schioppa and Assad 2006; Roesch et al. 2006). However, it was recently proposed that the OFC is critical for discriminating various states of the behavioral task rather than representing values per se, based on multiple neuronal responses correlated with not only value but also other specific information about outcomes such as odor, magnitude, spatial location, and temporal delay to delivery (Schoenbaum et al. 2011; Takahashi et al. 2011; Wilson et al. 2014). Although the exact nature of OFC role remains elusive, the present study found that OFC inactivation disrupted expected reward values signaled by reward-related non-DA cells during delays. Consequently, DAergic prediction errors were severely altered as if delayed rewards were less expected than before OFC

inactivation. In the absence of OFC inputs, specifically, putative DA cells were less excited by the reward-predicting cue at the time of DT, but they exhibited greater phasic responses to actual rewards. The observed changes in firing of putative DA and non-DA cells were not attributable to other variations in behavior after OFC dysfunction, since the rats' location at the end of waiting periods and their movement velocities around the times of DT and reward were not significantly different before and after MUS injections into the OFC. These results provide compelling evidence that the OFC is necessary for proper prediction error signaling by contributing to accurately estimating the value of delayed rewards in the current state.

This view can also account for the significantly decreased preference for LL reward shown by OFC-inactivated rats. Previous rodent studies indicate that OFC neurons start representing the relative value of potential outcomes from the time a decision is made (Sul et al. 2010; Takahashi et al. 2013). Indeed, we now have preliminary evidence that neuronal activity in the OFC of rats performing the same task carried value signals at possible choice areas of the T-maze (e.g., the start arm and the junction of the 3 arms) and these signals persisted until the end of delay periods (Jo and Mizumori, unpublished data). Thus, it is highly likely that the disrupted value representation after OFC inactivation may cause the animals to choose between the 2 goal arms in a more unbiased way.

MPFC Regulation of DA Activity

Previous work indicates that the mPFC carries value signals, although weaker than the OFC (Sul et al. 2010) and mPFC inactivation can alter expected reward values represented by non-DA cells in a spatial working memory task (Jo et al. 2013). However, mPFC inactivation in the delay discounting task resulted in no changes in activity of reward-related non-DA cells during waiting periods for upcoming rewards. The fact that mPFC inactivation caused perseverative choice biases to one location after mPFC inactivation (Fig. 5D) implies that the mPFC is critical for processing other information such as knowledge of task rules or the passage of temporal delay, rather than for valuation of delayed outcomes at least in the current task (Stefani and Moghaddam 2005; Jung et al. 2008; Kim et al. 2013). Thus, it appears that the mPFC conveys different information to the VTA depending on task demands. Since OFC inactivation disrupted value signals by reward-related non-DA cells, the OFC is the primary prefrontal source of reward expectancies in the delay discounting task (Roesch et al. 2006; Takahashi et al. 2011).

Despite no effects on reward-related non-DA cells, mPFC dysfunction increased phasic DA responses to both reward-predicting cues at the time of DT and actual rewards. The elevated prediction errors did not result from the general disinhibition after reduced inputs from the mPFC, because mPFC inactivation lowered spontaneous firing of both putative DA and non-DA cells (presumably due to the lack of excitatory mPFC projections; see Supplementary Tables 1 and 2). Instead, among 3 major components that are required for DA cells to compute prediction errors (i.e., currently available reward and 2 future reward values estimated in the current and the previous states) (Montague et al. 1996; Schultz et al. 1997; Pan et al. 2005), it is possible that the mPFC may send to the VTA the previous state value that was held in its working memory. Since the previous state value negatively modulates DAergic prediction errors, the mPFC information may be fed to DA cells via GABAergic local interneurons in the VTA or other GABAergic projections from different brain areas such as the ventral striatum (Sesack and Grace 2010). A

reduction of these inhibitory afferents after mPFC dysfunction can lead to increases in phasic DA activity. In line with this hypothesis, there is evidence indicating that mPFC neurons temporarily store representations about past information in value-guided decision-making (Sul et al. 2010). Alternatively, it is also possible that the mPFC may convey temporal information to midbrain DA systems. It is well-known that DAergic prediction errors are influenced by the precise temporal relationships between reward-predicting cues and reward delivery (Montague et al. 1996; Fiorillo et al. 2008; Kobayashi and Schultz 2008). For example, DA cells exhibit phasic responses to a well-expected reward if it is delivered at a different time than scheduled. Therefore, if mPFC-inactivated rats perceived the lengths of delay to be shorter than the actual delay due to the inability to keep track of the elapsed time during waiting periods, DAergic prediction errors at the times of DT and reward should increase. Indeed, a recent study found that the mPFC is crucial for processing time intervals (Kim et al. 2013).

Conclusion

It has long been reported that both the OFC and mPFC form anticipatory neural responses to forthcoming reward during reward-guided behavior (Pratt and Mizumori 2001; Miyazaki et al. 2004; Roesch et al. 2006; Sul et al. 2010). Such expectancy signals are thought to be essential for calculating DAergic prediction errors. The current study revealed that the 2 prefrontal subregions differentially modulate DA as well as non-DA cells in the VTA. Along with previous observations (Takahashi et al. 2011), our findings strongly suggest that the OFC is the primary prefrontal area that provides midbrain DA systems with expected outcome values. On the other hand, although it is less clear, the mPFC may send different kinds of value or temporal information rather than current value of expected outcomes. Further experiments involving single-unit recording from the prefrontal areas are warranted to better understand the exact coding of each subregion in the present behavioral task.

Funding

The current study was supported by NIMH Grant MH 58755 to S.J.Y.M.

Notes

We thank David Thompson and Wambura Fobbs for their assistance in data collection. *Conflict of Interest:* None declared.

References

Apicella P, Scarnati E, Ljungberg T, Schultz W. 1992. Neuronal activity in monkey striatum related to the expectation of predictable environmental events. *J Neurophysiol.* 68:945–960.

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

Clark JJ, Hollon NG, Phillips PE. 2012. Pavlovian valuation systems in learning and decision making. *Curr Opin Neurobiol.* 22:1054–1061.

Clark JJ, Sandberg SG, Wanat MJ, Gan JO, Horne EA, Hart AS, Akers CA, Parker JG, Willuhn I, Martinez V, et al. 2010. Chronic microstimulators for longitudinal, subsecond dopamine detection in behaving animals. *Nat Methods.* 7:126–129.

Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N. 2012. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature.* 482:85–88.

Corbit LH, Balleine BW. 2003. The role of prelimbic cortex in instrumental conditioning. *Behav Brain Res.* 146:145–157.

Daw ND, Niv Y, Dayan P. 2005. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci.* 8:1704–1711.

Day JJ, Jones JL, Carelli RM. 2011. Nucleus accumbens neurons encode predicted and ongoing reward costs in rats. *Eur J Neurosci.* 33:308–321.

Fiorillo CD, Newsome WT, Schultz W. 2008. The temporal precision of reward prediction in dopamine neurons. *Nat Neurosci.* 11:966–973.

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

Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CA, Clinton SM, Phillips PE, Akil H. 2011. A selective role for dopamine in stimulus-reward learning. *Nature.* 469:53–57.

Frederick S, Loewenstein G, O'Donoghue T. 2002. Time discounting and time preference: a critical review. *J Econ Lit.* 40:351–401.

Gottfried JA, O'Doherty J, Dolan RJ. 2003. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science.* 301:1104–1107.

Grace AA, Bunney BS. 1984. The control of firing pattern in nigral dopamine neurons: burst firing. *J Neurosci.* 4:2877–2890.

Hollon NG, Arnold MM, Gan JO, Walton ME, Phillips PE. 2014. Dopamine-associated cached values are not sufficient as the basis for action selection. *Proc Natl Acad Sci U S A.* 111:18357–18362.

Izquierdo A, Suda RK, Murray EA. 2004. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J Neurosci.* 24:7540–7548.

Jo YS, Lee I. 2010. Disconnection of the hippocampal-perirhinal cortical circuits severely disrupts object-place paired associative memory. *J Neurosci.* 30:9850–9858.

Jo YS, Lee J, Mizumori SJ. 2013. Effects of prefrontal cortical inactivation on neural activity in the ventral tegmental area. *J Neurosci.* 33:8159–8171.

Jo YS, Park EH, Kim IH, Park SK, Kim H, Kim HT, Choi JS. 2007. The medial prefrontal cortex is involved in spatial memory retrieval under partial-cue conditions. *J Neurosci.* 27:13567–13578.

Jones JL, Esber GR, McDannald MA, Gruber AJ, Hernandez A, Mireni A, Schoenbaum G. 2012. Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science.* 338:953–956.

Jung MW, Baeg EH, Kim MJ, Kim YB, Kim JJ. 2008. Plasticity and memory in the prefrontal cortex. *Rev Neurosci.* 19:29–46.

Kennerley SW, Wallis JD. 2009. Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. *Eur J Neurosci.* 29:2061–2073.

Kim J, Ghim JW, Lee JH, Jung MW. 2013. Neural correlates of interval timing in rodent prefrontal cortex. *J Neurosci.* 33:13834–13847.

Kim S, Hwang J, Lee D. 2008. Prefrontal coding of temporally discounted values during intertemporal choice. *Neuron.* 59:161–172.

Kobayashi S, Schultz W. 2008. Influence of reward delays on responses of dopamine neurons. *J Neurosci.* 28:7837–7846.

Miyazaki K, Miyazaki KW, Matsumoto G. 2004. Different representation of forthcoming reward in nucleus accumbens and medial prefrontal cortex. *Neuroreport.* 15:721–726.

- Montague PR, Dayan P, Sejnowski TJ. 1996. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci.* 16:1936–1947.
- Norton AB, Jo YS, Clark EW, Taylor CA, Mizumori SJ. 2011. Independent neural coding of reward and movement by pedunculo-pontine tegmental nucleus neurons in freely navigating rats. *Eur J Neurosci.* 33:1885–1896.
- O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science.* 304:452–454.
- O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C. 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci.* 4:95–102.
- Ostlund SB, Balleine BW. 2005. Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. *J Neurosci.* 25:7763–7770.
- Padoa-Schioppa C, Assad JA. 2006. Neurons in the orbitofrontal cortex encode economic value. *Nature.* 441:223–226.
- Pan WX, Schmidt R, Wickens JR, Hyland BI. 2005. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J Neurosci.* 25:6235–6242.
- Pratt WE, Mizumori SJ. 2001. Neurons in rat medial prefrontal cortex show anticipatory rate changes to predictable differential rewards in a spatial memory task. *Behav Brain Res.* 123:165–183.
- Rangel A, Camerer C, Montague PR. 2008. A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci.* 9:545–556.
- Roesch MR, Calu DJ, Schoenbaum G. 2007. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci.* 10:1615–1624.
- Roesch MR, Singh T, Brown PL, Mullins SE, Schoenbaum G. 2009. Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. *J Neurosci.* 29:13365–13376.
- Roesch MR, Taylor AR, Schoenbaum G. 2006. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron.* 51:509–520.
- Rudebeck PH, Walton ME, Smyth AN, Bannerman DM, Rushworth MF. 2006. Separate neural pathways process different decision costs. *Nat Neurosci.* 9:1161–1168.
- Samuelson PA. 1937. A note on measurement of utility. *Rev Econ Stud.* 4:155–161.
- Schoenbaum G, Takahashi Y, Liu TL, McDannald MA. 2011. Does the orbitofrontal cortex signal value? *Ann N Y Acad Sci.* 1239:87–99.
- Schultz W, Dayan P, Montague PR. 1997. A neural substrate of prediction and reward. *Science.* 275:1593–1599.
- Sesack SR, Grace AA. 2010. Cortico-basal ganglia reward network: microcircuitry. *Neuropsychopharmacology.* 35:27–47.
- Stefani MR, Moghaddam B. 2005. Systemic and prefrontal cortical NMDA receptor blockade differentially affect discrimination learning and set-shift ability in rats. *Behav Neurosci.* 119:420–428.
- Sul JH, Kim H, Huh N, Lee D, Jung MW. 2010. Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron.* 66:449–460.
- Suri RE, Schultz W. 2001. Temporal difference model reproduces anticipatory neural activity. *Neural Comput.* 13:841–862.
- Takahashi YK, Chang CY, Lucantonio F, Haney RZ, Berg BA, Yau HJ, Bonci A, Schoenbaum G. 2013. Neural estimates of imagined outcomes in the orbitofrontal cortex drive behavior and learning. *Neuron.* 80:507–518.
- Takahashi YK, Roesch MR, Stalnaker TA, Haney RZ, Calu DJ, Taylor AR, Burke KA, Schoenbaum G. 2009. The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron.* 62:269–280.
- Takahashi YK, Roesch MR, Wilson RC, Toreson K, O'Donnell P, Niv Y, Schoenbaum G. 2011. Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nat Neurosci.* 14:1590–1597.
- Tobler PN, Fiorillo CD, Schultz W. 2005. Adaptive coding of reward value by dopamine neurons. *Science.* 307:1642–1645.
- Tremblay L, Schultz W. 1999. Relative reward preference in primate orbitofrontal cortex. *Nature.* 398:704–708.
- Watabe-Uchida M, Zhu L, Ogawa SK, Vamanrao A, Uchida N. 2012. Whole-brain mapping of direct inputs to midbrain dopamine neurons. *Neuron.* 74:858–873.
- Watanabe M. 1996. Reward expectancy in primate prefrontal neurons. *Nature.* 382:629–632.
- Wilson RC, Takahashi YK, Schoenbaum G, Niv Y. 2014. Orbitofrontal cortex as a cognitive map of task space. *Neuron.* 81:267–279.