Dissociable dopaminergic control of saccadic target selection and its implications for reward modulation

Alireza Soltani[a,b,1], Behrad Noudoost[b], and Tirin Moore[a,b]

[a]Howard Hughes Medical Institute, and [b]Department of Neurobiology, Stanford University School of Medicine, Stanford, CA 94305

Edited by Peter H. Schiller, Massachusetts Institute of Technology, Cambridge, MA, and approved January 11, 2013 (received for review December 6, 2012)

To investigate mechanisms by which reward modulates target selection, we studied the behavioral effects of perturbing dopaminergic activity within the frontal eye field (FEF) of monkeys performing a saccadic choice task and simulated the effects using a plausible cortical network. We found that manipulation of FEF activity either by blocking D1 receptors (D1Rs) or by stimulating D2 receptors (D2Rs) increased the tendency to choose targets in the response field of the affected site. However, the D1R manipulation decreased the tendency to repeat choices on subsequent trials, whereas the D2R manipulation increased that tendency. Moreover, the amount of shift in target selection resulting from the two manipulations correlated in opposite ways with the baseline stochasticity of choice behavior. Our network simulation results suggest that D1Rs influence target selection mainly through their effects on the strength of inputs to the FEF and on recurrent connectivity, whereas D2Rs influence the excitability of FEF output neurons. Altogether, these results reveal dissociable dopaminergic mechanisms influencing target selection and suggest how reward can influence adaptive choice behavior via prefrontal dopamine.

computational modeling | decision making | oculomotor

As the primary means of exploring the visual environment, we shift our gaze several times each second via saccadic eye movements. Where we look depends not only on the physical salience of visual stimuli, but also on their reward value (1–3). For example, while shopping, your attention may be captured differently by colored sales tags upon realizing that they indicate different discounts (e.g., yellow tags for 20% vs. red tags for 40% off). Studies of value-based behavior have established that midbrain dopamine (DA) neurons signal different aspects of reward (4) and that neurons in many cortical structures receiving dopaminergic projections represent the reward value of visual stimuli (5). The frontal eye field (FEF), the area of prefrontal cortex (PFC) most directly involved in triggering saccadic eye movements, receives inputs from structures encoding reward value (6), and these inputs might be sufficient to control value-based, saccadic choice (target selection). It is also possible that this control operates via FEF projections to caudate in the basal ganglia, where DA-mediated plasticity modulates reward-dependent target selection (7). In addition, the direct dopaminergic inputs to the FEF (8) could provide a mechanism that independently modulates target selection based on reward signals. Determining how dopamine within the FEF influences saccades is thus important for understanding the mechanism by which reward modulates the selection of visual targets for eye movements.

To address these questions, we manipulated dopaminergic activity within the FEF of monkeys performing a saccadic choice task. In the task, monkeys freely selected between two identical visual targets appearing at varying temporal asynchronies. By systematically altering the target onset asynchrony (TOA), we could measure the bias in selecting either target, the change in selection probability due to previous choice, and the sensitivity of choice to the TOA. The selection of either target was rewarded after the saccade, so the bias in target selection corresponded to the delay in reward that the animal endured for selecting its preferred target. Dopaminergic FEF activity was manipulated either by blocking D1 receptors (D1Rs) with a selective antagonist or by stimulating D2 receptors (D2Rs) with a selective agonist. We recently reported that these two manipulations produce equivalent effects on saccadic target selection, both increasing the propensity of monkeys to make saccades to stimuli within the response fields (RFs) of affected FEF neurons (9). Both receptor subtypes are known to exert complex, modulatory effects on neural activity and behavior (10–13). For example, D1Rs exhibit dose-dependent, inverted U-shape effects on the persistent activity of PFC neurons (14, 15) and on working memory (16). To understand the complex modulatory effects of FEF dopamine on saccadic choice behavior we looked for differential influences of D1Rs and D2Rs on that behavior and simulated those influences using a plausible cortical network. We observed dissociable influences of D1R- and D2R-mediated FEF activity on saccadic choice that could be explained by dopaminergic modulation of synaptic plasticity and neural activity within different cortical layers.

Results

Experimental Findings. To quantify target selection, we used a task in which monkeys were trained to choose one of two stimuli as the targets of saccadic eye movements (Fig. 1*A*). During each experiment, we positioned one of the two targets (T_{in}) within the RFs of FEF neurons at the site of a drug infusion and the other target (T_{out}) at a diametrically opposite location. Targets were presented at varying TOAs over a wide range of values ($\sim \pm 200$ ms) and monkeys were rewarded regardless of which target they chose. Positive (negative) values of the TOA denote that T_{in} (T_{out}) appeared first, followed by the appearance of the second target TOA milliseconds later. To select the second target, the monkey thus had to wait (an amount of time equal to the TOA) until it appeared. All TOAs were presented with equal probability and were pseudorandomly interleaved such that on any given trial the monkey could not predict the TOA.

When targets appeared simultaneously, monkeys tended to choose one target more often than the other, exhibiting a bias. By varying the TOA, we could measure that bias in terms of the delay in onset at which the monkey began to choose the alternative target (17). We refer to that bias as the point of equal selection (PES), namely the TOA at which the monkey chooses the two targets with equal probability. The PES was determined from a logistic fit of the probability of choosing T_{in} as a function of the TOA (*Materials and Methods*). Positive values of the PES denote biases in favor of T_{out}, whereas negative values denote biases in favor of T_{in}.

Saccadic choice was measured during control trials and after the manipulation of either D1R- or D2R-mediated FEF activity via the local infusion of the selective D1R antagonist SCH23390 or the selective D2R agonist quinpirole. In total, we performed 34 experiments on two monkeys: 21 D1R antagonist infusions (14 in monkey A and 7 in monkey B) and 13 D2R agonist infusions (8 in monkey A and 5 in monkey B). The infusions in monkeys A and B were done in the left and right hemispheres, respectively. We analyzed data from blocks of trials before (control) and after the

Author contributions: A.S., B.N., and T.M. designed research; A.S. and B.N. performed research; A.S. analyzed data; and A.S., B.N., and T.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: asoltani@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1221236110/-/DCSupplemental.

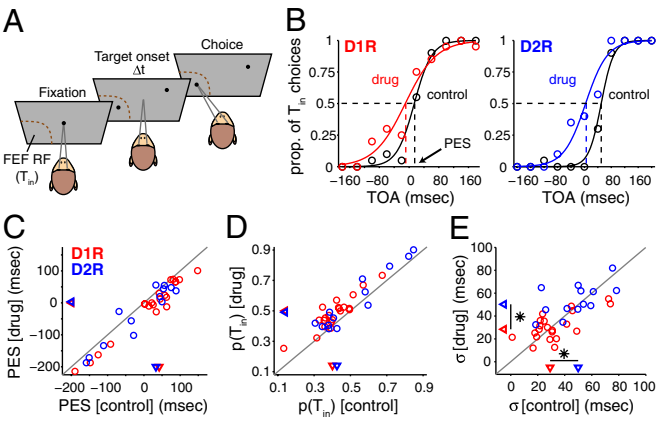

Fig. 1. Similar effects of manipulating D1R- and D2R-mediated FEF activity on saccadic choice behavior. (*A*) The saccadic choice task. In the task, two targets appeared on the display asynchronously (Δt), and the monkey was rewarded for making a saccadic eye movement to either one of them. One of the targets (T_{in}) appeared within the FEF RF. (*B*) Psychometric functions from two example experiments measuring the proportion of T_{in} choices across TOAs before (black) and after manipulation of D1R- (red) or D2R-mediated (blue) FEF activity. Solid curves show the logistic regression fit. The vertical dashed lines denote the TOAs yielding a 0.5 proportion of T_{in} choices (i.e., PES). (*C*–*E*) Distribution of PES values (*C*), overall choice probability [$p(T_{in})$] (*D*), and stochasticity of choice (σ) (*E*) before and after the dopaminergic manipulations. Triangles show the median of each distribution and asterisks denote significant differences ($P < 0.05$) between distributions.

drug infusion (drug). These data consisted of an average of 166 (control, SD = 48) and 173 (drug, SD = 42) trials in each of the D1R experiments and 201 (control, SD = 43) and 205 (drug, SD = 44) trials in each of the D2R experiments. As the results were consistent between the two monkeys and there were no effects of learning (Fig. S1 and *SI Text, Consistency of Experimental Results Between Two Monkeys* and *Learning Effects*), we performed our analyses using the combined data.

Fig. 1*B* shows the psychometric functions obtained in two example experiments in monkey A before and after a D1R antagonist or a D2R agonist infusion into the FEF. In both cases, the PES during control was slightly positive, indicating a small bias toward T_{out}. Following both the D1R and the D2R manipulations the PES was shifted leftward, indicating an increase in T_{in} selection. This pattern of results was consistent across all experiments in the two monkeys. Although there were no significant differences between the PESs measured during the separate D1R and D2R experiments, either before ($control_{D1R}$ vs. $control_{D2R}$, $z = -0.99$, $P = 0.3$) or after the drug infusion (D1R vs. D2R, $z = -0.64$, $P = 0.5$), within each experiment the PES was significantly reduced by both drug manipulations ($\Delta PES = -29.3 \pm 4.0$ ms for D1R, $z = -4.0$, $P < 0.0001$; $\Delta PES = -20.5 \pm 8.1$ ms for D2R, $z = -2.1$, $P < 0.04$) (Fig. 1*C*). Furthermore, the magnitude of the shift toward T_{in} choices was independent of the control PES for both drug manipulations ($r = 0.03$, $P = 0.9$ for D1R; $r = -0.09$, $P = 0.8$ for D2R), indicating a fixed increment in T_{in} preference. We observed no effects of either drug manipulation on saccadic latency or amplitude (*SI Text, Drug Effects on Saccade Metrics*). In addition to measuring the effect of the drug manipulations on the choice bias (PES), we also examined its effect on the overall probability of selecting T_{in}, or $p(T_{in})$. Consistent with the PES effects, we found that both blocking D1Rs [$\Delta p(T_{in}) = 0.071 \pm 0.010$, $z = -4.0$, $P < 0.0001$] and stimulating D2Rs [$\Delta p(T_{in}) = 0.046 \pm 0.016$, $z = -2.3$, $P < 0.02$] increased $p(T_{in})$ above that of control trials (Fig. 1*D*). Thus, saccadic target choice was shifted in favor of T_{in} following both dopaminergic manipulations.

In addition to measuring bias in the selection of targets, we also measured the stochasticity of the monkeys' choices, determined from σ of the fitted psychometric function (*Materials and Methods*). Larger values of σ correspond to greater stochasticity, i.e., a larger range of TOAs at which the choice is not determined solely by the TOA, whereas lower values of σ correspond to more deterministic choice behavior. We found that there were significant differences between the σ-values measured during the separate D1R and D2R manipulations both before ($control_{D1R}$ vs. $control_{D2R}$, $z = 2.4$, $P < 0.01$) and after the drug infusion (D1R vs. D2R, $z = 3.8$, $P < 0.0001$) (Fig. 1*E*). The differences might be due to the different experiments being carried out on different days. However, the key comparison is between control and drug values within each manipulation, carried out on the same day. For this comparison, we found that neither the D1R ($z = -0.43$, $P = 0.7$) nor the D2R ($z = -1.4$, $P = 0.2$) altered σ. Thus, the stochasticity of choice behavior was unaltered by either dopaminergic manipulation.

In contrast to the nearly identical effects of blocking D1Rs and stimulating D2Rs on target preference [$p(T_{in})$ and PES] and on the stochasticity of choice (σ), we found that the two manipulations altered choice behavior in very different ways. First, we found that the D1R and D2R manipulations exerted different effects on the tendency of monkeys to repeat choices on subsequent trials. We quantified the tendency to repeat choices with a repetition index (RI), positive values indicating a probability of repetition that exceeds the tendency due solely to the choice bias (*SI Text, Repetition Index as a Measure of Repetition in Choice*). Fig. 2*A* shows the distribution of RIs before and after the D1R and D2R manipulations. Although neither of the control RIs significantly differed from zero ($control_{D1R}$, $z = -1.4$, $P = 0.2$; $control_{D2R}$, $z = -0.73$, $P = 0.5$), both drug manipulations yielded nonzero RIs, yet in opposite ways. The D1R antagonist reduced RIs below zero ($RI = -0.032 \pm 0.009$, $z = -2.8$, $P < 0.005$), to values significantly less than control ($\Delta RI = -0.028 \pm 0.010$, $z = -2.4$, $P < 0.02$), indicating that monkeys became less likely to repeat target choices on subsequent trials. In contrast, the D2R agonist increased RIs above zero ($RI = 0.050 \pm 0.012$, $z = -2.8$, $P < 0.005$), to values significantly greater than control ($\Delta RI = 0.040 \pm 0.012$, $z = -2.6$, $P < 0.01$), indicating that monkeys became more likely to repeat target choices on subsequent trials. Consequently, although the control RIs did not differ before the drug infusion ($control_{D1R}$ vs. $control_{D2R}$, $z = 0.78$, $P = 0.4$), they did so after the drug infusion (D1R vs. D2R, $z = 4.0$, $P < 0.00005$). Thus, the two dopaminergic manipulations exerted opposite effects on the tendency of monkeys to repeat saccadic choices.

Second, we found that although the two dopaminergic manipulations increased T_{in} choices equally, the magnitude of their effects correlated with the choice stochasticity in opposite ways. Fig. 2*B* shows how the increases in T_{in} choices varied as a function of choice stochasticity measured during control trials. Following the D1R antagonist infusion, the increase in $p(T_{in})$ was positively correlated with the σ measured during control trials ($r = 0.61$, $P = 0.003$). This result indicates that during experiments in which

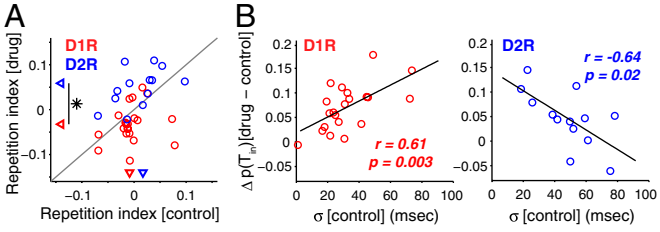

Fig. 2. Dissociable effects of blocking D1Rs and stimulating D2Rs on saccadic choice behavior. (*A*) Distribution of the repetition index values before and after the dopaminergic manipulations. Other conventions are as in Fig. 1. (*B*) Correlations between changes in $p(T_{in})$ due to drug manipulations and the stochasticity of choice (σ) measured during control trials. Correlation coefficients, significance, and linear fits (solid line) are shown for the two datasets.

monkeys exhibited greater stochasticity in their choices (larger σ-values, i.e., choices were less determined by the TOA), blocking D1Rs led to greater increases in T_{in} choices. In contrast, the increase in $p(T_{in})$ following stimulation of D2Rs was negatively correlated with the σ measured during control trials ($r = -0.64$, $P = 0.02$). Thus, during experiments in which monkeys exhibited greater stochasticity in their choices, D2R stimulation led to smaller increases in T_{in} choices. In addition, an analysis of covariance (ANCOVA) confirmed the contrasting relationships between stochasticity and increases in T_{in} selection produced by the D1R and D2R manipulations [$F(1, 30) = 15.7$, $P = 0.0004$]. Finally, we tested whether changes in choice bias and repetition were generated through the same mechanisms by computing the correlation between the two. The PES and RI were not correlated either before ($control_{D1R}$, $r = 0.069$, $P = 0.8$; $control_{D2R}$, $r = 0.13$, $P = 0.7$) or after the drug infusion (D1R, $r = 0.22$, $P = 0.3$; D2R, $r = 0.44$, $P = 0.1$). In addition, we found no significant correlation between changes in the PES and RI due to blocking D1Rs ($r = 0.28$, $P = 0.2$) or stimulating D2Rs ($r = -0.36$, $P = 0.2$).

Modeling Results. In an attempt to account for the observed dopaminergic effects on saccadic choice behavior, we constructed a biophysically plausible cortical network model of target selection and examined how changing model parameters altered choice behavior in the task. The model was composed of two FEF columns, consisting of pools of excitatory neurons within superficial (layers II and III) and deep (layers V and VI) cortical laminae (Fig. 3*A*). Each column contained two pools of excitatory pyramidal neurons, one in superficial layers and one in deep layers. In addition, a pool of inhibitory interneurons mediated mutual inhibition between the excitatory pools of the superficial layers. The excitatory pools within superficial and deep layers had RFs corresponding to the location of the two saccadic targets. Both superficial layer pools were driven by three types of input: background, visual, and value based (*SI Text, Computational Model*). These pools competed in a winner-take-all fashion to drive deep layer (output) pools, which in turn rendered a choice. The winner-take-all property was due to connectivity of excitatory and inhibitory pools in the superficial layers. In the deep layers, however, there was only weak recurrent excitation between neurons with similar selectivity. The excitatory pools in the deep layers sent outputs to the brainstem or the superior colliculus, driving target selection. Therefore, the activity of neural pools in the deep layers determined the network's choice on each trial. Specifically, we assumed that the network's choice on a given trial was the target that was represented by the deep layer pool whose activity reached 15 Hz first. Consequently, changes in the excitability or input efficacy of deep layer pools can also affect decision. The model also implemented dopaminergic modulation of synaptic plasticity and neural activity, providing a means by which to simulate the D1R and D2R manipulations (*Materials and Methods*).

The model selected between the two saccade targets with a probability that depended on the TOA. The probability of choosing T_{in} tended to be 1 or 0 when the absolute value of TOA was large (i.e., when the T_{in} appears long before or after T_{out}), due to a lack of visual input to one of the columns during the interval between target onsets and activity within that column being suppressed by visually driven activity within the other. On the other hand, when the TOA was close to zero, choice behavior was less determined by the TOA. This occurs because the time interval in which visual input differs is near zero and thus the outcome of the competition between T_{in} and T_{out} pools depends on other inputs. Fig. 3*A* (*Inset*) depicts the simulated response trajectories of pools within the superficial layers of T_{in} and T_{out} columns during trials in which either the first- or the second-appearing target was chosen (Fig. S2). In the latter case, the responses of the two pools tended to remain equal in the interval between target onsets (indicated by the gray circle), but could diverge later solely due to random fluctuations in the inputs. Notably, we found that the model's choice behavior could be fit as a sigmoid function of the TOA, similar to the experimental data

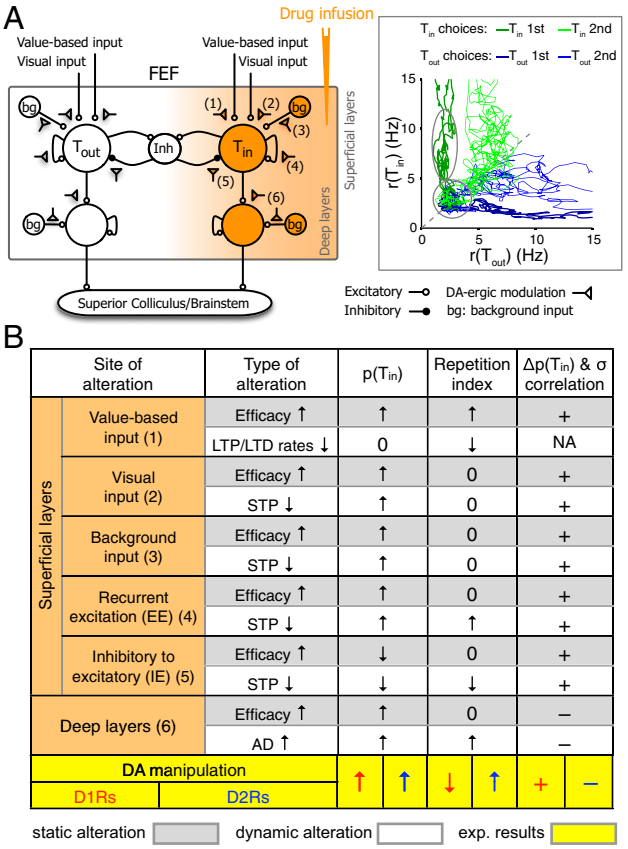

Fig. 3. Network architecture of the model used to simulate saccadic choice and the effects of manipulating different network elements on choice behavior. (*A*) The model comprised two FEF columns (T_{out} and T_{in}), consisting of excitatory neurons within superficial and deep layers. Inhibitory (Inh) interneurons mediated mutual inhibition between the excitatory pools within the superficial layers. Superficial layer pools were driven by value-based, visual, and background inputs and projected to deep layer pools. Deep layer pools projected to brainstem oculomotor structures to render a saccadic choice. Connections and neural activity in the two columns were modulated by DA. To reproduce the effects of the drug infusion, only elements within the T_{in} column were altered in the simulations (orange). *Inset* shows example activity trajectories of the T_{out} and T_{in} pools [$r(T_{in})$ vs. $r(T_{out})$] during trials on which the target appearing first (dark green or dark blue) or second (pale green or pale blue) was selected. The gray ellipse highlights the activity trajectories measured after the appearance of the first target and before the appearance of the second target (TOA epoch) on trials in which the first target (T_{in}) was chosen. Note that activity within the two pools diverges quickly during this epoch. The gray circle highlights the activity trajectory during the TOA epoch on trials in which the second target was chosen. Note that activity within the two pools remains equal during this interval. (*B*) Summary of choice behavior changes resulting from alterations to different network sites. The direction of alteration at each site was chosen such that it increased T_{in} selection (except for IE). Zero represents no change. For comparison, a summary of the experimental findings is shown at the bottom.

(Fig. S3*A*). Moreover, by modulating the background inputs and overall visual inputs we could account for the experimentally observed variability in PES and σ-values (Fig. S3*B* and *SI Text, Computational Model*).

After establishing that the model can qualitatively replicate saccadic choice behavior during the control experiments, we next studied how independently altering different elements of the network changes that behavior (*SI Text, Effects of Drug-Induced Alterations on the Model's Choice Behavior*). The model considered two classes of alterations, one static and one dynamic; the former corresponded to alterations to history-independent synaptic efficacy, whereas the latter corresponded to alterations to history-dependent

processes (e.g., short-term plasticity, STP). First, we found that static alterations to all types of input, recurrent connections, and the excitability of the output pool (via alterations of the efficacy of the deep layer background input) could alter the probability of T_{in} choices, $p(T_{in})$ or equivalently the PES (Fig. 3*B*). Increasing the efficacy increased $p(T_{in})$ for all sites except the inhibitory–excitatory (IE) connections, where increased efficacy decreased $p(T_{in})$. These results are expected as increases in the efficacy of the aforementioned sites (except IE connections) increase the activity in the T_{in} column and therefore increase the selection of the T_{in} target. Note that alterations in the efficacy of excitatory–inhibitory connections and of superficial to deep layer pools resulted in qualitatively similar effects to those of alterations in IE and the excitability of output pools, respectively. Second, static alteration to all sites produced changes in $p(T_{in})$ that depended on the value of σ (stochasticity in choice) (Fig. S4), resembling our experimental observations. Namely, changes in $p(T_{in})$ were larger when σ was larger for static alterations to all sites except the deep layers where these changes were smaller for larger σ-values (see *SI Text, Comparison of the Effects of Alterations to the Superficial Layers vs. the Deep Layers* for an intuitive explanation). Third, we found that altering value-based inputs was the only static alteration that could produce changes in the RI (Fig. S4). This is because only the efficacy of value-based input carried information about the choice on the previous trial. Fourth, dynamic alteration to all sites except value-based input [via changing the rates of long-term depression (LTD) and potentiation (LTP)] produced changes in $p(T_{in})$. Moreover, as with the static alterations, changes in $p(T_{in})$ depended on the value of σ. Fifth, dynamic alteration to all sites except visual and background inputs altered the RI (Fig. S5). Specifically, reduction in STP at recurrent connections resulted in changes in the RI that were positive at excitatory–excitatory (EE) and negative at IE connections. Note that STP reduces the strength of EE connections within the recently active column and results in alternation; therefore, reduction in STP increases repetition. In addition, increases in the rates of LTD and LTP and increases in afterdepolarization (AD, an increase in membrane potential that is dependent upon a preceding action potential for its initiation) both increased the RI (Fig. S5).

After exploring the effects of manipulating individual elements of the network, we used a combination of those alterations to reproduce our experimental findings (Fig. 4). Specifically, we sought to recapitulate the increase in $p(T_{in})$ observed with both the D1R and the D2R manipulations, opposite effects of the two manipulations on the RI, and the contrasting correlations between the $p(T_{in})$ increases and baseline σ. Based on modeling results summarized in Fig. 3*B*, one can notice that negative correlations between the $p(T_{in})$ increases and σ can be obtained only via alterations to activity in the deep layer pool, whereas any alterations to superficial layers resulted in positive correlation between the $p(T_{in})$ increases and σ. We reproduced the D1R effects with the network simulations by (*i*) increasing the efficacy of all inputs and both types of recurrent connections, (*ii*) decreasing the rates of LTD and LTP, and (*iii*) decreasing STP at all inputs and recurrent connections (EE and IE), but less strongly at EE connections (Fig. S6). The D2R effects were reproduced by (*i*) increasing the excitability of the output pool and (*ii*) increasing the AD within the output pool (Fig. S6). Both sets of network alterations resulted in increases in $p(T_{in})$ that were comparable to those observed experimentally (~0.08) (Fig. 4*A*). In contrast, the two sets of alterations produced opposite changes to the RI. The first set of alterations produced D1R-like decreases in the RI (ΔRI = −0.04), whereas the second set produced D2R-like increases in the RI (ΔRI = 0.04) (Fig. 4*B*). The D1R-like decrease in the RI was achieved via imposing weaker decreases in STP at EE connections compared with other sites. The D2R-like increase in the RI was achieved via increases in the AD within the output pool. Finally, similar to the experimental results, the first set of alterations yielded an increase in $p(T_{in})$ that was positively correlated with σ during control trials,

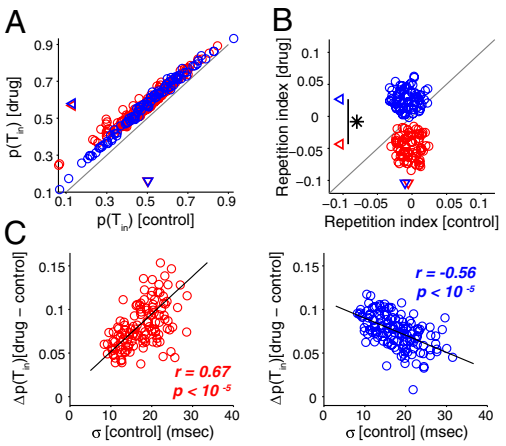

Fig. 4. Replication of experimental results with two sets of alterations to network elements. (*A*–*C*) Distributions of $p(T_{in})$ (*A*) and RIs (*B*) before and after the two sets of network alterations and correlations (*C*) between alteration-induced changes in $p(T_{in})$ and stochasticity of choice (σ). Red and blue symbols indicate D1R-like and D2R-like effects, respectively. Other conventions are as in Fig. 2.

whereas the second set of alterations produced an increase in $p(T_{in})$ that was negatively correlated with σ (Fig. 4*C*).

Discussion

We observed dissociable influences of DA-mediated FEF activity on saccadic target selection. Increases in target selection were achieved either by blocking D1Rs or by stimulating D2Rs. However, the former manipulation decreased the tendency to repeat choices on subsequent trials, whereas the latter increased that tendency. Moreover, the amount of shift in choice resulting from blocking D1Rs was positively correlated with baseline stochasticity of choice, whereas the amount of shift due to stimulating D2Rs was negatively correlated with stochasticity. A simple prediction of these results is that decreases in target selection would result from stimulating D1Rs or blocking D2Rs, yet the dissociable effects on repetition and choice stochasticity would remain. Our simulations using a plausible cortical model of target selection reproduced the experimental effects and pointed to the biophysical mechanisms underlying DA's influence on target selection. Specifically, our model predicts that D1Rs influence target selection mainly through their effects on the strength of inputs to the FEF and on recurrent connectivity within superficial layers, whereas D2Rs influence the excitability of FEF output neurons.

It is significant that our cortical network model was able to reproduce the experimental results via a unique (in terms of loci and relative strength) set of alterations to multiple network elements (Fig. S6). Nonetheless, it is also important to consider whether the alterations needed to achieve the modeling results are consistent with the known prefrontal distributions of D1Rs and D2Rs and their effects on neural activity (11). Notably, it is known that D1Rs modulate the efficacy of inputs (18, 19) and both excitatory and inhibitory recurrent connections (20, 21), modulate STP at inputs (18) and recurrent connections (22–24), and influence both LTD and LTP within the PFC (25–28). These properties of D1Rs are consistent with the network alterations necessary to reproduce the experimental results, namely increases in the efficacy of inputs and recurrent connections, decreases in STP, and decreases in the rates of LTD and LTP. The requirement of stronger modulation of STP in inhibitory interneurons is also compatible with the observation that GABAergic activity within the FEF contributes to target selection (17). Moreover, the decrease in rates of LTD and LTP is consistent with a recent finding that D1R-mediated activity within the lateral prefrontal cortex contributes to learning of novel visuomotor associations (29).

In addition, D2Rs are expressed largely within deep layers of cortex (30, 31), where in the FEF, layer V pyramidal neurons provide the primary output to the superior colliculus and brainstem oculomotor nuclei (32, 33). In contrast, D1Rs are expressed throughout cortical layers (30). Furthermore, prefrontal D2Rs are known to enhance the excitability of layer V pyramidal neurons (34) and it was recently reported that AD is also enhanced by D2R agonists in a subtype of layer V pyramidals (35), thereby prolonging activity of these neurons for hundreds of milliseconds. Consistent with these properties, an increase in the excitability and AD of the deep layer pool was precisely the alterations to the model network required to reproduce our experimental D2R effects. Thus, not only was the model able to reproduce the dissociable effects of the D1R and D2R manipulations on target selection, but also the alterations to the model required to reproduce those effects were consistent with the known properties of D1Rs and D2Rs and their distributions across cortical layers.

Role of Dopamine in Saccadic Target Selection. Saccadic target selection is determined in part by the reward value of potential targets (1–3). Given the clear role of the FEF in target selection, it is surprising that few studies have explored the contribution of FEF dopamine to this behavior. It has been shown that reward-dependent modulation of saccadic target selection relies on dissociable effects of D1Rs and D2Rs on neural activity in the striatum, presumably through modulation of long-term synaptic plasticity in the caudate (7, 36). Both caudate and FEF neurons have been shown to exhibit modulation of activity by the reward value of saccadic targets (37). Thus, it is important to determine the relative contributions of these two areas to value-based target selection. Similar to the caudate studies, our study also demonstrates dissociable roles of D1R- and D2R-mediated FEF activity on target selection. However, it is possible that dopaminergic modulations within these two areas could contribute to reward control of target selection under different conditions: modulation within the FEF for when the reward values are changing or are unpredictable (e.g., timing of the second target in our experiment) and modulation within the caudate for when the rewards for different targets are fixed and predictable (e.g., small and big rewards) and a bias is desirable (7, 36).

The dissociable dopaminergic effects we observed suggest potential mechanisms that may underlie different aspects of adaptive choice behavior. First, we found opposing effects of the D1R and D2R manipulations on repetition (RI), namely decreases and increases in repetition for the two manipulations, respectively. The observed decrease in repetition following the infusion of a D1R antagonist is consistent with previous results showing increases in repetitive behavior (perseveration) following the infusion of a D1R agonist into prefrontal cortex (16). One might suggest that variation in the degree of repetition may be related to the exploration–exploitation trade-off observed in adaptive choice behavior. This trade-off reflects the balance between maximizing reward based on current knowledge and testing alternative actions to acquire new knowledge (38). A recent study found that DA levels in mice influence the degree of exploitation (39). Another study in humans found evidence of a correlation between exploration and prefrontal DA (40). In the context of this trade-off, the D1R effects we observed might be viewed as an increase in exploration, whereas the D2R effects might be viewed as an increase in exploitation.

Second, although both manipulations equally increased the probability of T_{in} selection, those increases depended on the baseline stochasticity in choice in opposite ways. For the D1R antagonist, larger increases in T_{in} selection were associated with greater stochasticity in choice with respect to the TOA. In contrast, for the D2R agonist, larger increases in T_{in} selection were associated with smaller stochasticity. This dissociation suggests a possible mechanism for adjusting behavior according to the sensitivity to relevant information. Given that greater stochasticity reflects lower sensitivity to visual information, specifically the TOA, DA could adjust choice bias according to that sensitivity. Differential adjustments of choice bias according to sensitivity could be beneficial when multiple cues exist, but only one carries reward information. When behavior relies on the relevant cue and cue sensitivity is high, the degree of bias adjustment should be small (i.e., inversely proportional to sensitivity). On the other hand, when behavior relies on an irrelevant cue, and sensitivity to that cue is high, then the degree of bias adjustment should be large (i.e., proportional to sensitivity). The latter case may be particularly important for escaping local maxima in reward maximization (38).

Previous modeling work has helped elucidate how DA contributes to working memory via D1R-mediated differential changes to NMDA and AMPA currents (41) and differential dopaminergic modulation of NMDA currents in excitatory and inhibitory neurons (42). Here, we were able to pinpoint possible neural mechanisms through which D1Rs and D2Rs differentially alter saccadic target selection by virtue of their effects in different cortical layers. Our model suggests how dopaminergic modulation of the afferents to the FEF could alter reward-dependent choice. The predictions of this model could be tested in experiments in which reward delivery is probabilistic and therefore the animal's choice is determined by the integration of reward history (43, 44). One might predict, for example, that after blocking D1Rs within the FEF, the form and time constant of reward integration would be altered such that the impact of previous rewards on current choices could be increased or decreased. Such an outcome could further clarify the contribution of FEF dopamine to reward-dependent choice behavior.

Materials and Methods

Experimental Procedures. Two monkeys (*Macaca mulatta*) were trained on a saccadic choice task. All experimental procedures were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals, the Society for Neuroscience Guidelines and Policies, and the Stanford University Animal Care and Use Committee. For detailed general and surgical procedures see *SI Text*, Experimental Procedures.

Dopaminergic Manipulation of FEF Activity. We used a microinjectrode system for simultaneous microstimulation and microinfusion of drugs (45). The center of the RF of the FEF site under study was defined by the endpoint of the saccades evoked by its electrical microstimulation. We positioned one of the targets (T_{in}) in the RF of the FEF site and the other one in the opposite hemifield (T_{out}). Small volumes (0.5–1 μL) of the selective D1R antagonist SCH23390 or the D2R agonist quinpirole were delivered into the FEF with infusion rates less than 100 nL/min. Volumes of this size diffuse ~1–3 mm within the cortex (46) and thus affect neurons within only a few columns of the FEF. Both drugs were obtained from Sigma-Aldrich. The acidic PH of the D1R antagonist solution was adjusted to 5.5–6.0 before the infusion, whereas the D2R agonist required no PH adjustment. Given the limitations to the number of possible repeated infusions in a single cortical region, due to the risk of damage (45), we chose the two (of four possible) dopaminergic manipulations most likely to yield interpretable results.

Data Analysis. To quantify target selection, for each experiment we measured the psychometric function by computing the probability of selecting T_{in} as a function of the TOA, for trials before and after drug infusion. The psychometric function was then fitted by a sigmoid (logistic function), which yielded two parameters: a bias parameter (PES) that determined the TOA for which the two targets were selected with equal frequency and a measure of stochasticity in choice (σ, often referred to as the temperature) with respect to the TOA. For one of the D2R manipulation experiments in monkey B, the choice probability plateaued at about 0.2 and 0.8 for the minimum and maximum values of the TOA, respectively. To get a better fit for this experiment, we bounded our logistic function between 0.2 and 0.8; however, our results were unaffected by this choice of fitting.

Unless otherwise mentioned, we used the Wilcoxon signed-rank test for comparison between control and drug experiments and the Wilcoxon rank sum test for comparisons between the two drug conditions (for which the z and P values are reported). For correlation measures we reported the Pearson correlation coefficient and its significance value. Unless otherwise mentioned, data are expressed as a mean plus or minus the SEM.

Model Implementation. To simulate the superficial layers we implemented a mean-field reduction of a detailed spiking network model (47). The details of this implementation have been described previously (47, 48). The deep layer pools were simulated using a firing-rate model with a realistic response function. In addition, we incorporated short-term (STP) and long-term synaptic plasticity in the inputs, STP in connections between pools in the superficial layers, and the effect of afterdepolarization on neural activity in the deep layer pools (*SI Text, Computational Model*).

Simulation of Dopaminergic Manipulations. We assumed that reward harvest is signaled globally by the phasic activity of midbrain DA and this signal results in an elevation of prefrontal DA for a few hundred milliseconds, during which short-term and long-term plasticity are modulated by DA (see *SI Text, Computational Model* for more details). In addition, the drug manipulation could alter connections, synaptic plasticity, or neural excitability of the column that was infused with drug, in two different ways. First, the drug manipulation could change the synaptic efficacy of a given pathway or the neural excitability of a pool (static alterations). Second, the drug manipulation could alter STP of a given pathway, DA-dependent rates of LTD and LTP of value-based inputs, or the activity-dependent change in neural excitability between trials (dynamic alterations). These alterations could affect the network through the following pathways: background, visual, and value-based inputs to the superficial layers; the strength of connections between excitatory pools, between inhibitory and excitatory pools, and from superficial to deep layer pools; and the excitability of the deep layer pools (see *SI Text, Computational Model and Effects of Drug-Induced Alterations on the Model's Choice Behavior* for more details).

ACKNOWLEDGMENTS. We thank D. S. Aldrich for technical assistance and K. L. Clark and M. Zirnsak for helpful comments on the manuscript. This work was supported by National Institutes of Health Grant EY014924.

1. Navalpakkam V, Koch C, Rangel A, Perona P (2010) Optimal reward harvesting in complex perceptual environments. *Proc Natl Acad Sci USA* 107(11):5232–5237.
2. Markowitz DA, Shewcraft RA, Wong YT, Pesaran B (2011) Competition for visual selection in the oculomotor system. *J Neurosci* 31(25):9298–9306.
3. Schütz AC, Trommershäuser J, Gegenfurtner KR (2012) Dynamic integration of information about salience and value for saccadic eye movements. *Proc Natl Acad Sci USA* 109(19):7547–7552.
4. Schultz W (2002) Getting formal with dopamine and reward. *Neuron* 36(2):241–263.
5. Soltani A, Wang X-J (2008) From biophysics to cognition: Reward-dependent adaptive choice behavior. *Curr Opin Neurobiol* 18(2):209–216.
6. Huerta MF, Krubitzer LA, Kaas JH (1987) Frontal eye field as defined by intracortical microstimulation in squirrel monkeys, owl monkeys, and macaque monkeys. II. Cortical connections. *J Comp Neurol* 265(3):332–361.
7. Hikosaka O (2007) Basal ganglia mechanisms of reward-oriented eye movement. *Ann N Y Acad Sci* 1104:229–249.
8. Berger B, Trottier S, Verney C, Gaspar P, Alvarez C (1988) Regional and laminar distribution of the dopamine and serotonin innervation in the macaque cerebral cortex: A radioautographic study. *J Comp Neurol* 273(1):99–119.
9. Noudoost B, Moore T (2011) Control of visual cortical signals by prefrontal dopamine. *Nature* 474(7351):372–375.
10. Missale C, Nash SR, Robinson SW, Jaber M, Caron MG (1998) Dopamine receptors: From structure to function. *Physiol Rev* 78(1):189–225.
11. Seamans JK, Yang CR (2004) The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog Neurobiol* 74(1):1–58.
12. Robbins TW, Arnsten AFT (2009) The neuropsychopharmacology of fronto-executive function: Monoaminergic modulation. *Annu Rev Neurosci* 32:267–287.
13. Monte-Silva K, et al. (2009) Dose-dependent inverted U-shaped effect of dopamine (D2-like) receptor activation on focal and nonfocal plasticity in humans. *J Neurosci* 29(19):6124–6131.
14. Williams GV, Goldman-Rakic PS (1995) Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. *Nature* 376(6541):572–575.
15. Vijayraghavan S, Wang M, Birnbaum SG, Williams GV, Arnsten AFT (2007) Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nat Neurosci* 10(3):376–384.
16. Zahrt J, Taylor JR, Mathew RG, Arnsten AF (1997) Supranormal stimulation of D1 dopamine receptors in the rodent prefrontal cortex impairs spatial working memory performance. *J Neurosci* 17(21):8528–8535.
17. Schiller PH, Tehovnik EJ (2003) Cortical inhibitory circuits in eye-movement generation. *Eur J Neurosci* 18(11):3127–3133.
18. Law-Tho D, Hirsch JC, Crepel F (1994) Dopamine modulation of synaptic transmission in rat prefrontal cortex: An in vitro electrophysiological study. *Neurosci Res* 21(2):151–160.
19. Urban NN, González-Burgos G, Henze DA, Lewis DA, Barrionuevo G (2002) Selective reduction by dopamine of excitatory synaptic inputs to pyramidal neurons in primate prefrontal cortex. *J Physiol* 539(Pt 3):707–712.
20. Gao WJ, Krimer LS, Goldman-Rakic PS (2001) Presynaptic regulation of recurrent excitation by D1 receptors in prefrontal circuits. *Proc Natl Acad Sci USA* 98(1):295–300.
21. Gao W-J, Wang Y, Goldman-Rakic PS (2003) Dopamine modulation of perisomatic and peridendritic inhibition in prefrontal cortex. *J Neurosci* 23(5):1622–1630.
22. Seamans JK, Durstewitz D, Christie BR, Stevens CF, Sejnowski TJ (2001) Dopamine D1/D5 receptor modulation of excitatory synaptic inputs to layer V prefrontal cortex neurons. *Proc Natl Acad Sci USA* 98(1):301–306.
23. González-Burgos G, Krimer LS, Urban NN, Barrionuevo G, Lewis DA (2004) Synaptic efficacy during repetitive activation of excitatory inputs in primate dorsolateral prefrontal cortex. *Cereb Cortex* 14(5):530–542.
24. Young CE, Yang CR (2005) Dopamine D1-like receptor modulates layer- and frequency-specific short-term synaptic plasticity in rat prefrontal cortical neurons. *Eur J Neurosci* 21(12):3310–3320.
25. Otani S, Blond O, Desce JM, Crépel F (1998) Dopamine facilitates long-term depression of glutamatergic transmission in rat prefrontal cortex. *Neuroscience* 85(3):669–676.
26. Gurden H, Takita M, Jay TM (2000) Essential role of D1 but not D2 receptors in the NMDA receptor-dependent long-term potentiation at hippocampal-prefrontal cortex synapses in vivo. *J Neurosci* 20(22):RC106.
27. Huang Y-Y, Simpson E, Kellendonk C, Kandel ER (2004) Genetic evidence for the bidirectional modulation of synaptic plasticity in the prefrontal cortex by D1 receptors. *Proc Natl Acad Sci USA* 101(9):3236–3241.
28. Matsuda Y, Marzo A, Otani S (2006) The presence of background dopamine signal converts long-term synaptic depression to potentiation in rat prefrontal cortex. *J Neurosci* 26(18):4803–4810.
29. Puig MV, Miller EK (2012) The role of prefrontal dopamine D1 receptors in the neural mechanisms of associative learning. *Neuron* 74(5):874–886.
30. Lidow MS, Goldman-Rakic PS, Gallager DW, Rakic P (1991) Distribution of dopaminergic receptors in the primate cerebral cortex: Quantitative autoradiographic analysis using [3H]raclopride, [3H]spiperone and [3H]SCH23390. *Neuroscience* 40(3):657–671.
31. Santana N, Mengod G, Artigas F (2009) Quantitative analysis of the expression of dopamine D1 and D2 receptors in pyramidal and GABAergic neurons of the rat prefrontal cortex. *Cereb Cortex* 19(4):849–860.
32. Stanton GB, Goldberg ME, Bruce CJ (1988) Frontal eye field efferents in the macaque monkey: I. Subcortical pathways and topography of striatal and thalamic terminal fields. *J Comp Neurol* 271(4):473–492.
33. Segraves MA, Goldberg ME (1987) Functional properties of corticotectal neurons in the monkey's frontal eye field. *J Neurophysiol* 58(6):1387–1419.
34. Wang Y, Goldman-Rakic PS (2004) D2 receptor regulation of synaptic burst firing in prefrontal cortical pyramidal neurons. *Proc Natl Acad Sci USA* 101(14):5093–5098.
35. Gee S, et al. (2012) Synaptic activity unmasks dopamine D2 receptor modulation of a specific class of layer V pyramidal neurons in prefrontal cortex. *J Neurosci* 32(14):4959–4971.
36. Nakamura K, Hikosaka O (2006) Role of dopamine in the primate caudate nucleus in reward modulation of saccades. *J Neurosci* 26(20):5360–5369.
37. Ding L, Hikosaka O (2006) Comparison of reward modulation in the frontal eye field and caudate of the macaque. *J Neurosci* 26(25):6695–6703.
38. Sutton RS, Barto AG (1998) *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA).
39. Beeler JA, Daw N, Frazier CRM, Zhuang X (2010) Tonic dopamine modulates exploitation of reward learning. *Front Behav Neurosci* 4:1–14.
40. Frank MJ, Doll BB, Oas-Terpstra J, Moreno F (2009) Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat Neurosci* 12(8):1062–1068.
41. Durstewitz D, Seamans JK, Sejnowski TJ (2000) Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *J Neurophysiol* 83(3):1733–1750.
42. Brunel N, Wang XJ (2001) Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J Comput Neurosci* 11(1):63–85.
43. Sugrue LP, Corrado GS, Newsome WT (2004) Matching behavior and the representation of value in the parietal cortex. *Science* 304(5678):1782–1787.
44. Soltani A, Wang X-J (2006) A biophysically based neural model of matching law behavior: Melioration by stochastic synapses. *J Neurosci* 26(14):3731–3744.
45. Noudoost B, Moore T (2011) A reliable microinjectrode system for use in behaving monkeys. *J Neurosci Methods* 194(2):218–223.
46. Hupé JM, Chouvet G, Bullier J (1999) Spatial and temporal parameters of cortical inactivation by GABA. *J Neurosci Methods* 86(2):129–143.
47. Wong K-F, Wang X-J (2006) A recurrent network mechanism of time integration in perceptual decisions. *J Neurosci* 26(4):1314–1328.
48. Soltani A, Wang X-J (2010) Synaptic computation underlying probabilistic inference. *Nat Neurosci* 13(1):112–119.

Supporting Information

Soltani et al. 10.1073/pnas.1221236110

SI Text

Experimental Procedures

Two male rhesus monkeys (*Macaca mulatta*, 9 and 10 kg) were used in these experiments. Each animal was surgically implanted with a titanium head post, two recording chambers, and a scleral eye coil. Surgery was conducted using aseptic techniques under general anesthesia (isoflurane) and analgesics were provided during post-surgical recovery. Structural magnetic resonance imaging was performed to locate the arcuate sulcus in one of the monkeys for the placement of a recording chamber in a subsequent surgery. A craniotomy was performed in the chamber on each animal, allowing access to the frontal eye field (FEF). All experimental procedures were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals, the Society for Neuroscience Guidelines and Policies, and the Stanford University Animal Care and Use Committee.

At the beginning of each experimental session, we mapped the saccade vector elicited via microstimulation at the FEF site under study with the use of a separate behavioral paradigm (1). In this paradigm, the monkey was required to fixate on a visual stimulus (0.5° diameter circle) for 500 ms, after which a 100-ms stimulation train was delivered on half the trials. The endpoint of the evoked saccade vector was used to define the response field (RF) of the FEF site. Electrical microstimulation consisted of a 100-ms train of biphasic current pulses (0.25 ms, 250 Hz) delivered with a Grass stimulator (S88) and two Grass stimulation isolation units (PSIU-6) (Grass Instruments). Current amplitude was measured via the voltage drop across a 1-kΩ resistor in series with the return lead of the current source.

During each experimental session, monkeys performed saccadic choice tasks. In a given experiment, at least two blocks of saccadic choice trials were collected, one before FEF inactivation and one following it. Each block consisted of at least 10 trials per target onset asynchrony (TOA). The saccade targets were 0.5° diameter white circles on a gray background presented on a 29° × 39° colorimetrically calibrated CRT monitor (Mitsubishi Diamond Pro-2070SB-BK) with medium short persistence phosphors. The eccentricity (and luminance) of targets was always equal. The T_{in} target was always in the same hemifield for each monkey, namely the hemifield contralateral to the drug manipulation site in the FEF. The fixation point was a 0.25° circle presented at the center of the screen. Data acquisition and behavioral monitoring were controlled by the CORTEX system. Eye movement was monitored with a scleral search coil and digitized at 500 Hz (CNC Engineering). The spatial resolution of eye position measurements was less than 0.1°.

Computational Model

Model Inputs. All neural pools received a background input that mimicked the large barrage of inputs received by cortical neurons due to spontaneous firing of neighboring cortical neurons with different RFs. In addition, the excitatory pools in the superficial layers received visual and value-based inputs when a visual target was presented in their RFs. The visual inputs were similar for the two pools but their timing was determined by the target onset times and thus was different for the two pools by the TOA. The value-based inputs, computed in another cortical area (as described in ref. 2), carried information about the reward values associated with each target.

Static Alterations. To incorporate static alterations of connections within the network (say the visual input) due to the drug manipulation of the T_{in} column, we assumed that the efficacy of connections within the T_{in} column was changed as follows (without any changes to connections within the T_{out} column),

$$J_{vis,\ in} = (1+\lambda_{vis}) \times J_{vis,\ in}, \qquad \text{[S1]}$$

where λ_{vis} determined the strength of alteration of the visual input. Similar equations were used for static alterations of other connections.

Dynamic Alterations. The dynamic alteration of the value-based input was implemented by changing the potentiation and depression rates for the inputs to the T_{in} column, while keeping the rates for the other column intact (see below). The dynamic alteration of other connections was implemented by changing the strength of short-term plasticity (STP) of those connections within the T_{in} column, while keeping STP for the other column intact.

The STP in a given connection was implemented by setting the efficacy of that connection in the two columns ($J_{vis,\ in}$ and $J_{vis,\ out}$) equal to

$$\begin{aligned} J_{vis,\ in} &= \bar{J}_{vis}\left(1+d_{vis,\ in}(t)\right) \\ J_{vis,\ out} &= \bar{J}_{vis}\left(1+d_{vis,\ out}(t)\right), \end{aligned} \qquad \text{[S2]}$$

where $\bar{J}_{vis}$ determined the overall efficacy of the visual input, and $d_{vis,in}(t)$ and $d_{vis,out}(t)$ determined the amount of short-term depression in the input to T_{in} and T_{out} pools, respectively. Due to activation of the visual input, the depression factors $d_{vis,in}(t)$ and $d_{vis,out}(t)$ were decreased at the end of each trial by the amounts of $\delta d_{vis,in}$ ($d_{vis,in}(t) \rightarrow d_{vis,in}(t) - \delta d_{vis,in}$) and $\delta d_{vis,out}$ ($d_{vis,out}(t) \rightarrow d_{vis,out}(t) - \delta d_{vis,out}$), respectively, and exponentially decayed back to zero during the intertrial interval,

$$\begin{aligned} d_{vis,in}(t+1) &= d_{vis,in}(t) \times e^{-\left(\frac{1}{\tau_{vis}}\right)} \\ d_{vis,out}(t+1) &= d_{vis,out}(t) \times e^{-\left(\frac{1}{\tau_{vis}}\right)}, \end{aligned} \qquad \text{[S3]}$$

where τ_{vis} determined the time constant of STP in the visual input. Therefore, $\delta d_{vis,in}$ and $\delta d_{vis,out}$ controlled the strength of STP in the visual input to T_{in} and T_{out} pools, respectively. The dynamic alterations of a given connection were implemented by changing the strength of STP of that connection within the T_{in} column, $\delta d_{vis,in}$, while keeping STP for the other column intact. In the case of both static and dynamic alterations, the efficacies of connections for T_{in} and T_{out} columns (say for the visual input), $J_{vis,in}$ and $J_{vis,out}$, were set to

$$\begin{aligned} J_{vis,in} &= (1+\lambda_{vis}) \times \bar{J}_{vis}\left(1+d_{vis,in}(t)\right) \\ J_{vis,out} &= \bar{J}_{vis}\left(1+d_{vis,out}(t)\right). \end{aligned} \qquad \text{[S4]}$$

We used a similar approach to implement the effect of afterdepolarization (AD) on neural activity in the deep layers. The AD increases the membrane potential but it depends upon a preceding action potential for its initiation. Because our model is a firing-rate model, we simulated the effect of afterdepolarization on neural activity by trial-to-trial changes in the

neural excitability. We assumed that due to AD, the excitability of a given pool in the deep layers increased after each trial if that pool was highly active (i.e., if the corresponding target was the choice on that trial). We incorporated AD by setting the efficacy of the background input in the deep layers (which determined the excitability of neural pools in the deep layers), $J_{bg(d),\,in}$ and $J_{bg(d),\,out}$, equal to

$$\begin{aligned} J_{bg(d),\,in} &= \bar{J}_{bg(d)}\left(1+g_{bg,\,in}(t)\right) \\ J_{bg(d),\,out} &= \bar{J}_{bg(d)}\left(1+g_{bg,\,out}(t)\right), \end{aligned} \quad \textbf{[S5]}$$

where $\bar{J}_{bg(d)}$ determined the overall efficacy of the background input in the deep layers, and $g_{bg,in}(t)$ and $g_{bg,out}(t)$ determined the amount AD in T_{in} and T_{out} pools, respectively. On each trial, only the deep layer pool that determined the choice on that trial was highly active, increasing the AD for that pool. Therefore, we assumed that if T_{in} was selected on a given trial, only the factor $g_{bg,in}(t)$ was increased at the end of that trial by the amount of $\delta g_{bg,in}$ ($g_{bg,\,in}(t) \rightarrow g_{bg,\,in}(t)+\delta g_{bg,\,in}$). During the intertrial interval both $g_{bg,in}(t)$ and $g_{bg,out}(t)$ exponentially decayed back to zero with time constant $\tau_{bg(d)}$ (similar to Eq. **S3**). For simplicity, we set the τ for STP and AD equal to 0.5 (trials), but we obtained qualitatively similar results for a wide range of τ-values. The dynamic alteration of the output layers was implemented by changing the strength of AD for the T_{in} column, $\delta g_{bg,in}$, while keeping AD for the other column intact.

We also included dopamine (DA)-dependent plasticity in the value-based inputs to the FEF. This plasticity was implemented by multiplying the value-based input to T_{in} and T_{out} excitatory pools by the synaptic strengths, c_{in} and c_{out}, respectively,

$$\begin{aligned} J_{val,\,in} &= \bar{J}_{val} \times c_{in}(t) \\ J_{val,\,out} &= \bar{J}_{val} \times c_{out}(t), \end{aligned} \quad \textbf{[S6]}$$

where $\bar{J}_{val}$ determined the overall strength of the value-based input. At the end of each trial, the synaptic strengths were updated according to a plasticity rule that depended on the activity of pre- and postsynaptic neurons and the presence or absence of DA (2, 3). We assumed that the harvest of reward on a given trial is signaled by the phasic activity of midbrain DA. Because in our experiment presynaptic neurons that provide value-based input are active on every trial and every trial is rewarded (so DA is released on every trial), the only factor that determines the direction of plasticity is the choice of the network. Specifically, at the end of a trial on which T_{in} was selected, the synaptic strengths were updated as

$$\begin{aligned} c_{in}(t+1) &= c_{in}(t)+q_{p,\,in}\left(1-c_{in}(t)\right) \\ c_{out}(t+1) &= c_{out}(t)-q_{d,\,\mathrm{out}}c_{out}(t), \end{aligned} \quad \textbf{[S7]}$$

where q_p and q_d determined the rates of potentiation (long-term potentiation, LTP) and depression (long-term depression, LTD), respectively. On the other hand if T_{out} was selected, the synaptic strengths were updated as

$$\begin{aligned} c_{in}(t+1) &= c_{in}(t)-q_{d,\,in}c_{in}(t) \\ c_{out}(t+1) &= c_{out}(t)+q_{p,\,out}\left(1-c_{out}(t)\right). \end{aligned} \quad \textbf{[S8]}$$

The dynamic alteration of the value-based input was implemented by changing the potentiation and depression rates for the inputs to the T_{in} column, $q_{p,in}$ and $q_{d,in}$, while keeping the rates for the other column intact.

Simulation of Day-to-Day Variability in Behavior. To account for experimentally observed day-to-day changes in the point of equal selection (PES) and σ-values, we assumed that the observed changes in the PES (bias in target selection) and σ (stochasticity of choice)-values were due to changes in the overall visual inputs or the background inputs rather than differences in connections within the two columns. We multiplied background and visual inputs to the superficial layers in the two columns by two factors,

$$\begin{aligned} J_{bg(s),\,in} &= \bar{J}_{bg(s)}\left(1+b_{in}\right), & J_{vis,\,in} &= \gamma \times \bar{J}_{vis} \\ J_{bg(s),\,out} &= \bar{J}_{bg(s)}\left(1+b_{out}\right), & J_{vis,\,out} &= \gamma \times \bar{J}_{vis}, \end{aligned} \quad \textbf{[S9]}$$

where $J_{bg(s),in}$ and $J_{bg(s),out}$ are the background inputs to the superficial T_{in} and T_{out} pools, and for each day of the experiment b_{in} and b_{out} were selected uniformly from an interval [−0.002, 0.006] and γ was selected from a normal distribution, $N(1,0.15)$ ($0.75 < \gamma < 1.25$). We found that changes in the overall visual and background inputs, which respectively mimic variability in excitability of the two columns and changes in the subject's engagement or attention, can produce a wide range of independent PES and σ-values (Fig. S3*B*). On the one hand, increases in γ decreased the value of σ because the network's choice behavior became less stochastic with stronger inputs. On the other hand, differential background input to the two superficial pools shifted target selection in favor of the target that received a stronger background input. Therefore, these two changes could account for day-to-day variability in bias and stochasticity of choice.

Dopaminergic Signaling and Manipulations. In this work, we assumed that the harvest of reward is signaled globally by the phasic activity of midbrain DA and this signal results in an elevation of prefrontal DA for a few hundred milliseconds, during which short-term and long-term plasticities are modulated by DA. More specifically, on each trial the phasic DA signal modulates the value-based input to the FEF, STP in visual and background inputs and in recurrent connections, and the AD in the deep layers. On the other hand, the tonic DA signal affects the efficacy of different inputs and recurrent connections in the network. However, because every trial is rewarded in our experiment, the effects of phasic DA signal are similar to those of the tonic DA signal. Future experiments will be needed to determine which one of these two signals contributes to target selection in the prefrontal cortex (PFC) more strongly.

As suggested previously (4, 5), even though DA neurons respond transiently to reward or behaviorally relevant stimuli, DA levels in target structures rise slowly, perhaps due to the slow filtering dynamics of DA release (6). Similarly, it has been argued that a DA concentration large enough to induce plasticity in the PFC of an intact animal can be sustained even through background activity of DA neurons (7). Moreover, despite many speculations on different roles of phasic and tonic DA activity in the modulation of activity in the PFC and striatum, there are great similarities between them. For example, whereas paired stimulation of ventral tegmental area and auditory tones results in plasticity in the auditory cortex (8), stimulation of substantia nigra paired with postsynaptic activity results in DA-dependent plasticity [through D1 receptor (D1R) activation] in corticostriatal synapses (9).

Consistency of Experimental Results Between Two Monkeys

Here we provide our main experimental results computed for each monkey separately and show that results from the two monkeys are statistically similar and thus can be combined in the analysis (Fig. S1).

First, the average changes in the PES due to the D1R manipulation for monkey A and monkey B were ΔPES = −27.9 and −32.3 ms, respectively, and were not statistically distinguishable ($P = 0.5$). The average changes in the PES due to the D2 receptor (D2R) manipulation for monkeys A and B were ΔPES = −30.6 and −4.7 ms,

respectively, and were not statistically distinguishable ($P = 0.07$). Second, changes in σ-values due to the D1R ($P = 0.7$) and D2R ($P = 0.6$) manipulations were similar between the two monkeys. Third, the average changes in the repetition index (RI) due to the D1R manipulation for monkeys A and B were $\Delta RI = 0.032$ and 0.020, respectively, and were not statistically distinguishable ($P = 0.6$). The average changes in the RI due to the D2R manipulation for monkeys A and B were $\Delta RI = -0.039$ and -0.043, respectively, and were not statistically distinguishable ($P = 0.8$). Finally, an analysis of covariance (ANCOVA) for changes in $p(T_{in})$ due the D1R manipulation revealed significant effects of σ ($P < 0.006$), but no difference between the two monkeys ($P = 0.3$), and the relationship between changes in $p(T_{in})$ and σ was similar between the two monkeys ($P = 0.9$). A similar ANCOVA for changes in $p(T_{in})$ due the D2R manipulation revealed marginally significant effects of σ ($P = 0.1$), but no difference between the two monkeys ($P = 0.7$), and the relationship between changes in $p(T_{in})$ and σ ($P = 0.7$) was similar between the two monkeys.

Learning Effects

The saccadic choice task required very little training. Monkeys fully learned to perform the current task within 1 wk of training. Considering the short amount of training, it is unlikely that the monkeys' choice behavior was affected by their experience in the task. However, to examine any effects of learning on the choice behavior, we performed three analyses. First, we examined whether the PES or σ during the control experiments changed over the course of the experiment. We found no effects of training (i.e., the experiment order) on the PES: D1R experiments in monkey 1, $P = 0.5$; D2R experiments in monkey 1, $P = 0.1$; D1R experiments in monkey 2, $P = 1.0$; and D2R experiments in monkey 2, $P = 0.9$. Similarly, we found no effects of training on the stochasticity in choice, σ: D1R experiments in monkey 1, $P = 0.8$; D2R experiments in monkey 1, $P = 0.7$; D1R experiments in monkey 2, $P = 0.6$; and D2R experiments in monkey 2, $P = 0.5$. Therefore, neither choice bias nor the stochasticity in choice changed over the course of the experiment. Second, we computed the PES and σ for trials at the beginning (first half) and at the end (second half) of each experiment and we found no difference between these values: $P = 0.2$ and $P = 0.1$ (before D1R manipulation), $P = 0.1$ and $P = 0.6$ (after D1R manipulation), $P = 1.0$ and $P = 0.5$ (before D2R manipulation), and $P = 0.2$ and $P = 0.5$ (after D2R manipulation) for the PES and σ, respectively. Third, we computed the overall choice probability, $p(T_{in})$, for the first and the second half of trials in each experiment. We found no statistical difference between $p(T_{in})$ in these two sets of trials: $P = 0.2$ (before D1R manipulation), $P = 0.2$ (after D1R manipulation), $P = 0.9$ (before D2R manipulation), and $P = 0.2$ (after D2R manipulation). Based on these three analyses, there were no clear effects of learning on target selection.

Drug Effects on Saccade Metrics

To examine whether the drug manipulations affected metrics of saccades, we computed the latency and amplitude of saccades toward the affected part of space (T_{in} target). First, we calculated the mean latencies of saccades to the T_{in} target for different TOA values (note that the saccade latency depends on the TOA) before and after the drug infusions. A two-way analysis of variance (ANOVA) on saccade latency for the D1R manipulations showed a significant TOA effect ($F = 3.99$, $P = 0.0001$), but no drug effect ($F = 3.11$, $P = 0.2$) and no TOA–drug interaction ($F = 1.35$, $P = 0.2$). Similarly, a two-way ANOVA on saccade latency for the D2R manipulations showed a significant TOA effect ($F = 11.29$, $P = 0.001$), but no drug effect ($F = 2.80$, $P = 0.2$) and no TOA–drug interaction ($F = 1.11$, $P = 0.4$). Second, we calculated mean amplitude of saccades to the T_{in} target for different TOA values, before and after the drug infusions. A two-way ANOVA on saccade amplitude for the D1R manipulations showed no significant TOA effect ($F = 1.32$, $P = 0.2$), no drug effect ($F = 1.69$, $P = 0.2$), and no TOA–drug interaction ($F = 0.53$, $P = 0.9$). Similarly, a two-way ANOVA on saccade amplitude for the D2R manipulations showed no TOA effect ($F = 0.81$, $P = 0.7$), no drug effect ($F = 1.98$, $P = 0.3$), and no TOA–drug interaction ($F = 1.01$, $P = 0.4$). Thus, there are no clear effects of the drug manipulations on saccadic latency or amplitude.

Repetition Index as a Measure of Repetition in Choice

To examine the relationship between target selection on two consecutive trials and how the drug manipulations altered this relationship, we computed the probability that a given target was selected on two consecutive trials, $p(Stay)$. Before the drug infusion, $p(Stay)$ was different from chance (0.5) for both manipulations ($control_{D1R}$, $P = 0.02$; $control_{D2R}$, $P = 0.008$). However, the D1R manipulation reduced $p(Stay)$ ($P < 0.005$), whereas the D2R manipulation increased $p(Stay)$ ($P < 0.003$). Moreover, $p(Stay)$ was different for both manipulations after the drug infusion (D1R vs. D2R, $P < 0.0001$), whereas it was similar before the infusion ($control_{D1R}$ vs. $control_{D2R}$, $P = 0.1$).

Nevertheless, to remove the tendency to repeat due solely to the choice bias, we also computed the RI, which is the probability that a given target was selected on two consecutive trials ($p(Stay)$) minus the probability of choosing the same target on consecutive trials by chance, given the observed overall $p(T_{in})$:

$$RI = p(Stay) - p(T_{in}) \times p(T_{in}) - (1 - p(T_{in})) \times (1 - p(T_{in})). \quad \textbf{[S10]}$$

Therefore, RI measures the tendency to repeat choices on consecutive trials after removing that tendency due to the overall preference.

Effects of Drug-Induced Alterations on the Model's Choice Behavior

In the following sections, we describe the effects of drug-induced alterations of individual elements of the model on its choice behavior. We quantified the effects of both static and dynamic alterations of different sites of the network on the bias in target selection (measured via PES), stochasticity of choice (measured via σ), and the influence of previous outcome on choice (via the RI). Note that measuring changes in target selection via $p(T_{in})$ gave results similar to those measured via the PES. To examine how the effects of drug-induced alterations depend on σ, we repeated each simulation for five different values of γ, producing different values of σ (increasing γ reduced σ-values).

Effects of Alterations of the Value-Based Input. In our model, the value-based input to the FEF carried information about reward values in terms of the return of a given target (i.e., reward harvested per selection of a target), as observed in other areas of the prefrontal cortex (2). Because the monkeys' choice on every trial was rewarded, the return for both target options was equal to one. However, the efficacy of value-based afferents to the FEF underwent DA-dependent plasticity, which enabled these synapses to estimate the income (i.e., reward harvested per trial) for each target (Eqs. **S6**–**S8**) (in the absence of such plasticity, the value-based input would carry signals that were identical for the two targets). The income for a given target approached one when that target was selected consecutively and approached zero when it was repeatedly not selected. The static alteration of the value-based input was implemented by multiplying the strength of the value-based input to the T_{in} column by $(1 + \lambda_{val})$, $-0.4 \leq \lambda_{val} \leq 0.4$. This implementation is consistent with the effects of DA on the efficacy of inputs to pyramidal neurons in the PFC (10, 11). The dynamic alteration of the value-based input was implemented by changing the DA-dependent depression and potentiation rates of the value-based afferents to the FEF. This implementation is

consistent with known effects of DA on LTP and LTD in the PFC (7, 12–17) and in the striatum (18, 19). For simplicity, we set the rates of LTP and LTD equal to each other: $0 \leq q_{p,in}$ $(= q_{d,in}) \leq 0.5$ and $q_{p,out}$ $(= q_{d,out}) = 0.25$.

First, we found that the static manipulation of the value-based input shifted target selection and changed stochasticity and repetition in choice (Fig. S4*A*). More specifically, we found that increasing λ_{val} resulted in a decrease of the PES (i.e., more selection of T_{in}) but this decrease was faster for smaller values of γ. Therefore, increasing λ_{val} reduced the PES faster when σ was larger. In addition, increasing λ_{val} increased the RI, and the RI was larger for larger values of σ (lower values of γ). Simultaneous increases in σ and positive values for the RI can be explained by noting that due to the DA-dependent plasticity of value-based input to the FEF, the reward value of the target that was selected (not selected, respectively) on the previous trial was increased (respectively, decreased), which biased the choice toward the previously selected target. This bias toward the previously selected target corresponds to a reduced dependence of choice on the TOA, and thus a greater λ_{val} produces increases in both the RI and σ.

Second, we found that the dynamic manipulation of the value-based input did not shift choice, whereas it affected stochasticity and repetition in choice (Fig. S5*A*). More specifically, increasing the rates of LTD and LTP ($q_{p,in}$ and $q_{d,in}$) in the value-based inputs did not change the PES, whereas it increased σ and the RI. These results are expected as an increase in the rates of LTD and LTP increased the influence of choice on the previous trial on the choice during the current trial, which therefore increased repetition and simultaneously made the choice behavior more stochastic. Moreover, larger LTP and LTD rates of the value-based input to the T_{in} column did not shift the choice behavior, because every trial was rewarded in the experiment and therefore, a larger increase in the input to T_{in} due to a larger potentiation after T_{in} selection was offset by a larger decrease in this input due to a larger depression after T_{out} selection.

Effects of Alterations of the Visual Input. In the task the visual targets were identical, but their efficacy in driving neural pools in the two columns could be different due to heterogeneity of the visual pathways to the FEF. Therefore, we assumed that the superficial pools in the two columns could receive different inputs that could arrive at different times due to the TOA. The static alteration of the visual input was implemented by multiplying the strength of the visual input to the T_{in} column by $(1 + \lambda_{vis})$, $-0.25 \leq \lambda_{vis} \leq 0.25$, whereas the overall strength of visual inputs was scaled by γ. Therefore, T_{in} and T_{out} columns received visual inputs proportional to $\gamma(1 + \lambda_{vis})$ and γ, respectively. The dynamic alteration was implemented by changing $\delta d_{vis,in}$ (strength of STP in the visual input) for visual input to the T_{in} column ($0 \leq \delta d_{vis,in} \leq 0.25$), while keeping $\delta d_{vis,out} = 0.125$. This implementation is consistent with the known effects of DA on STP in the PFC (10, 20–22).

First, we found that the static manipulation of the value-based input induced shifts in choice and changed stochasticity of choice, but did not affect repetition (Fig. S4*B*). More specifically, we found that increasing λ_{vis} resulted in a decrease of the PES (i.e., more selection of T_{in}), and this decrease was faster for smaller values of γ. Moreover, the value of σ decreased as λ_{vis} or γ increased because the network's choice behavior became less stochastic with stronger inputs [the overall visual input to the network was proportional to $\gamma(1 + \lambda_{vis}/2)$; Eqs. **S1** and **S9**]. Therefore, increasing λ_{vis} resulted in a decrease of the PES that was faster for larger values of σ. The result is interesting because it shows that selective alteration of the efficacy of visual input (e.g., via drug manipulation) can cause a shift in target selection that is larger for larger values of σ. In other words, the more sensitive the network is to the visual input, the less the model's choice can be shifted by manipulation of the efficacy of this input (this is also true for other types of alteration to the superficial layers; see *SI Text, Comparison of the Effects of Alterations to the Superficial Layers vs. the Deep Layers* for an intuitive explanation).

In addition, we found a larger change in the PES for negative values of λ_{vis} than for its positive values. This can be explained by noting that the PES is approximately proportional to the ratio of differential visual inputs to the overall visual inputs (to the two columns), $\lambda_{vis}/(1 + \lambda_{vis}/2)$. Finally, we found that the static alteration of the visual input did not affect the RI strongly. This is expected, because this alteration increased only the efficacy of the visual input and did not change the dependence of choice on a given trial on the choice on the preceding trials.

Second, we found that the dynamic alteration of the visual input affected the model's choice behavior similar to the static alteration (Fig. S5*B*). More specifically, decreasing $\delta d_{vis,in}$ resulted in a decrease in the PES that was stronger for larger values of σ. Moreover, a decrease in STP reduced stochasticity of choice (smaller value of σ) but did not affect repetition. The latter can be explained by noting that STP in the visual input was independent of the choice on a given trial (because the visual inputs to both pools were active on every trial), therefore, it did not change the dependence between the choices on two consecutive trials.

Effects of Alterations of the Superficial Layers Background Input (or Neural Excitability). The background input set the excitability of neural pools within superficial and deep layers of the network. Therefore, we examined how alterations of this input changed the model's choice behavior. We modeled these alterations in a manner qualitatively similar to that of alterations of the visual input; however, because of the larger efficacy of the background input to the superficial layers, we used much smaller values for $\lambda_{bg(d)}$ and $\delta d_{bg,in}$ [$-0.005 \leq \lambda_{bg(d)} \leq 0.005$ and $0 \leq \delta d_{bg,in} \leq 0.0075$ whereas $\delta d_{bg,out} = 0.0037$].

For both static and dynamic alteration of the background input, we found results qualitatively similar to those for alterations of the visual input (Figs. S4*C* and S5*C*). This is expected as changes in the background input within superficial layers, which sets the excitability of the superficial pools, have qualitatively similar effects on the choice behavior to those of changes in the efficacy of the visual input.

Effects of Alterations of the Recurrent Connections. To study the effects of drug manipulation of recurrent connections on the model's choice behavior, we examined the alterations of two sets of connections: connections between pyramidal neurons in the superficial layers and connections from inhibitory to excitatory neurons in the superficial layers. The effects of alterations of connections from excitatory to inhibitory neurons were similar to those effects for connections from inhibitory to excitatory neurons. Also, as we argue in the next section, the effects of alteration of connections between excitatory pools in superficial and deep layers were qualitatively similar to the effects of alteration of neural excitability of pools in the deep layers.

The static alteration of connections between excitatory pools in superficial layers [excitatory–excitatory (EE) connections] was implemented by multiplying the strength of this connection in the T_{in} column by $(1 + \lambda_{EE})$, $-0.045 \leq \lambda_{EE} \leq 0.045$. Similarly, the static alteration of connections from inhibitory to excitatory pools [inhibitory–excitatory (IE) connections] was implemented by multiplying the strength of this connection in the T_{in} column by $(1 + \lambda_{IE})$, $-0.15 \leq \lambda_{IE} \leq 0.15$. Note that because the strength of recurrent connections had a nonlinear influence on dynamics of the network, we adopted a smaller range for alterations of these connections. The dynamic alteration of recurrent connections was implemented by changing $\delta d_{EE,in}$ and $\delta d_{IE,in}$: $0 \leq \delta d_{EE,in} \leq 0.075$ and $0 \leq \delta d_{IE,in} \leq 0.15$, while keeping $\delta d_{EE,out} = 0.0375$ and $\delta d_{IE,out} = 0.075$. However, we assumed that STP occurred only for connections that were active on a given trial, that is, for connections in the column that won the competition on a given trial and

determined the choice. Therefore, δd_{EE}s and δd_{IE}s determined how strongly the outcome on previous trials changed the choice behavior of the network on future trials. The implementation of the static and dynamic alterations is consistent with the known effects of DA on the synaptic efficacy (23, 24) and STP in recurrent connection in the PFC (25–28).

First, we found that the static alteration of both EE and IE connections affected the choice behavior very similar to the static alteration of the visual and background inputs (Fig. S4 *D* and *E*). Importantly, alterations of the efficacy recurrent connections caused a shift in target selection that was larger for larger values of σ. However, the alteration of IE connections biased the choice behavior in the opposite direction to that of other alterations (stronger IE connections resulted in less selection of T_{in}). Interestingly, the static alteration of the background input and EE and IE connections resulted in a U-shape change in σ, which showed that minimum stochasticity of choice could be achieved when the background input and recurrent connections in the two columns were equal.

Second, we found that the dynamic alteration of EE connections gave qualitatively similar results to those of the visual input, except that it affected repetition in choice more strongly (Fig. S5*D*). More specifically, we found that decreasing $\delta d_{EE,in}$ reduced the RI. This was expected because STP in EE connections reduced the strength of excitatory connections between neurons selective to the recently selected target and therefore increased the probability of selecting the other target. Similarly, dynamic alteration of the IE connections introduced repetition in target selection that became stronger as $\delta d_{IE,in}$ decreased (Fig. S5*E*). The last result showed that DA-dependent reductions in STP of connections between inhibitory and pyramidal neurons could provide a mechanism for the experimentally observed decreases in repetition.

Effects of Alterations of the Neural Excitability in the Deep Layers. Neural pools in the deep layers (i.e., output layers) contributed to decision-making processes in a different manner than neural pools in superficial layers, which initiated competition between the two columns selective to the two targets. More specifically, the deep layer pools relayed the signal in the superficial layers to the superior colliculus and brainstem to move the eyes and therefore affected the decision processes by biasing the readout of the signal generated by the superficial layers. This biasing mechanism is especially effective when decision time is short, as in our target selection experiment. In addition, because main inputs to the deep layer pools originated from those in the superficial layers, changes in the excitability of deep layer pools had qualitatively similar effects to changes in the strength of connections from the superficial to the deep layers. Therefore, here we show only how drug-induced changes in the excitability of the neural pool in the deep layers affected the model's choice behavior.

The static alteration of the excitability of deep layer pools was implemented by changing the background inputs to these pools, specifically, by multiplying the background inputs to the pool in the T_{in} column by $(1 + \lambda_{bg(d)})$, $-0.02 \leq \lambda_{bg(d)} \leq 0.04$. This implementation is consistent with the effect of DA (D2R-mediated activity) on excitability of the deep layers cortical neurons (29). The dynamic alteration of the excitability of the deep layer pools was implemented by changing $\delta g_{bg,in}$ ($0 \leq \delta g_{bg,in} \leq 0.04$), while keeping $\delta g_{bg,out} = 0.02$. This implementation is consistent with the effects of DA (D2R-mediated activity) on afterdepolarization in the deep layers of cortex (30).

First, we found that static alteration of the deep layer pool induced shifts in choice and changed stochasticity of choice, but did not affect repetition (Fig. S4*F*). More specifically, we found that increasing the excitability of the T_{in} pool in the deep layers [via increasing $\lambda_{bg(d)}$] decreased the PES (i.e., more selection of T_{in}) and increased σ. In stark contrast to other static alterations, the decrease in the PES due to increase in excitability was faster for larger values of γ (corresponding to smaller values of σ). This indicated that drug manipulation of the output layers could shift the choice more strongly when the network's choice was less stochastic (see *SI Text*, *Comparison of the Effects of Alterations to the Superficial Layers vs. the Deep Layers* for an intuitive explanation).

Second, we found that the dynamic alteration of neural excitability of the T_{in} pool in the deep layers changed the model's choice behavior similar to the static alteration of neural excitability (Fig. S5*F*). In addition, increasing $\delta g_{bg,in}$ increased the RI. Overall, the increase in neural excitability of the deep (output) layer pool provided a mechanism for increasing repetition and for shifting the choice behavior more strongly when the network's choice was less stochastic.

Comparison of the Effects of Alterations to the Superficial Layers vs. the Deep Layers

We found that for all alterations to the superficial layers the amount of shift in choice was larger for larger values of σ (except for the dynamic alteration of the value-based input, which did not shift choice), whereas alterations of the deep layer pools shifted choice more strongly for smaller values of σ. These contrasting results can be explained as follows.

Alteration of the superficial layers (input) in the network affects target selection via changing the competition between the two excitatory pools in these layers. In the case where the network is very sensitive to the TOA (i.e., small value of σ), a small difference in the onset time of the visual input results in rapid divergence of the activity in the superficial pools. To shift target selection via alteration of the superficial layer pools, a strong biased input (e.g., background, value-based, or visual) or similarly large biased recurrent connections are required to avoid the first appearing target to win the competition and determine the choice. On the other hand, when the network is not sensitive to the TOA (i.e., large value of σ) a small biased input can easily shift target selection. Therefore, for a given amount of biased input the amount of shift is larger when the network is less sensitive to the TOA, which explains why the amount of shift due to alterations of the superficial layer pools is larger for larger values of σ.

Alteration of the deep layers (output) in the network affects target selection via biasing the readout of the signal generated by the superficial layers. In the case where the network is very sensitive to the TOA, the activity of the two deep layer pools diverges very rapidly (due to their inputs from the superficial pools). Therefore, increased input efficacy or the excitability of the deep layer pools can easily trigger choice and thus shift target selection. On the other hand, when the network is not sensitive to the TOA, the activity of the deep layer pools diverges very slowly while it stays close to the baseline activity. As a result, changes in the excitability of the deep layer pools are not effective in shifting the choice and mostly increase the stochasticity. These two cases demonstrate why the amount of shift due to alterations of the deep layer pools is larger for smaller values of σ.

1. Moore T, Fallah M (2001) Control of eye movements and spatial attention. *Proc Natl Acad Sci USA* 98(3):1273–1276.
2. Soltani A, Wang X-J (2006) A biophysically based neural model of matching law behavior: Melioration by stochastic synapses. *J Neurosci* 26(14):3731–3744.
3. Soltani A, Lee D, Wang X-J (2006) Neural mechanism for stochastic behaviour during a competitive game. *Neural Netw* 19(8):1075–1090.
4. Cohen JD, Braver TS, Brown JW (2002) Computational perspectives on dopamine function in prefrontal cortex. *Curr Opin Neurobiol* 12(2):223–229.
5. Durstewitz D, Seamans JK (2002) The computational role of dopamine D1 receptors in working memory. *Neural Netw* 15(4–6):561–572.
6. Montague PR, et al. (2004) Dynamic gain control of dopamine delivery in freely moving animals. *J Neurosci* 24(7):1754–1759.

7. Goto Y, Yang CR, Otani S (2010) Functional and dysfunctional synaptic plasticity in prefrontal cortex: Roles in psychiatric disorders. *Biol Psychiatry* 67(3):199–207.
8. Bao S, Chan VT, Merzenich MM (2001) Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature* 412(6842):79–83.
9. Reynolds JN, Hyland BI, Wickens JR (2001) A cellular mechanism of reward-related learning. *Nature* 413(6851):67–70.
10. Law-Tho D, Hirsch JC, Crepel F (1994) Dopamine modulation of synaptic transmission in rat prefrontal cortex: An in vitro electrophysiological study. *Neurosci Res* 21(2):151–160.
11. Urban NN, González-Burgos G, Henze DA, Lewis DA, Barrionuevo G (2002) Selective reduction by dopamine of excitatory synaptic inputs to pyramidal neurons in primate prefrontal cortex. *J Physiol* 539(Pt 3):707–712.
12. Law-Tho D, Desce JM, Crepel F (1995) Dopamine favours the emergence of long-term depression versus long-term potentiation in slices of rat prefrontal cortex. *Neurosci Lett* 188(2):125–128.
13. Otani S, Blond O, Desce JM, Crépel F (1998) Dopamine facilitates long-term depression of glutamatergic transmission in rat prefrontal cortex. *Neuroscience* 85(3):669–676.
14. Gurden H, Takita M, Jay TM (2000) Essential role of D1 but not D2 receptors in the NMDA receptor-dependent long-term potentiation at hippocampal-prefrontal cortex synapses in vivo. *J Neurosci* 20(22):RC106.
15. Otani S, Daniel H, Roisin M-P, Crepel F (2003) Dopaminergic modulation of long-term synaptic plasticity in rat prefrontal neurons. *Cereb Cortex* 13(11):1251–1256.
16. Huang Y-Y, Simpson E, Kellendonk C, Kandel ER (2004) Genetic evidence for the bidirectional modulation of synaptic plasticity in the prefrontal cortex by D1 receptors. *Proc Natl Acad Sci USA* 101(9):3236–3241.
17. Matsuda Y, Marzo A, Otani S (2006) The presence of background dopamine signal converts long-term synaptic depression to potentiation in rat prefrontal cortex. *J Neurosci* 26(18):4803–4810.
18. Reynolds JNJ, Wickens JR (2002) Dopamine-dependent plasticity of corticostriatal synapses. *Neural Netw* 15(4–6):507–521.
19. Calabresi P, Picconi B, Tozzi A, Di Filippo M (2007) Dopamine-mediated regulation of corticostriatal synaptic plasticity. *Trends Neurosci* 30(5):211–219.
20. Young CE, Yang CR (2005) Dopamine D1-like receptor modulates layer- and frequency-specific short-term synaptic plasticity in rat prefrontal cortical neurons. *Eur J Neurosci* 21(12):3310–3320.
21. Seamans JK, Durstewitz D, Christie BR, Stevens CF, Sejnowski TJ (2001) Dopamine D1/D5 receptor modulation of excitatory synaptic inputs to layer V prefrontal cortex neurons. *Proc Natl Acad Sci USA* 98(1):301–306.
22. Gonzalez-Burgos G, Kroener S, Seamans JK, Lewis DA, Barrionuevo G (2005) Dopaminergic modulation of short-term synaptic plasticity in fast-spiking interneurons of primate dorsolateral prefrontal cortex. *J Neurophysiol* 94(6):4168–4177.
23. Gao WJ, Krimer LS, Goldman-Rakic PS (2001) Presynaptic regulation of recurrent excitation by D1 receptors in prefrontal circuits. *Proc Natl Acad Sci USA* 98(1): 295–300.
24. Gao W-J, Wang Y, Goldman-Rakic PS (2003) Dopamine modulation of perisomatic and peridendritic inhibition in prefrontal cortex. *J Neurosci* 23(5):1622–1630.
25. Seamans JK, Gorelova N, Durstewitz D, Yang CR (2001) Bidirectional dopamine modulation of GABAergic inhibition in prefrontal cortical pyramidal neurons. *J Neurosci* 21(10):3628–3638.
26. Seamans JK, Yang CR (2004) The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog Neurobiol* 74(1):1–58.
27. Gorelova N, Seamans JK, Yang CR (2002) Mechanisms of dopamine activation of fast-spiking interneurons that exert inhibition in rat prefrontal cortex. *J Neurophysiol* 88(6):3150–3166.
28. González-Burgos G, Krimer LS, Urban NN, Barrionuevo G, Lewis DA (2004) Synaptic efficacy during repetitive activation of excitatory inputs in primate dorsolateral prefrontal cortex. *Cereb Cortex* 14(5):530–542.
29. Wang Y, Goldman-Rakic PS (2004) D2 receptor regulation of synaptic burst firing in prefrontal cortical pyramidal neurons. *Proc Natl Acad Sci USA* 101(14):5093–5098.
30. Gee S, et al. (2012) Synaptic activity unmasks dopamine D2 receptor modulation of a specific class of layer V pyramidal neurons in prefrontal cortex. *J Neurosci* 32(14): 4959–4971.

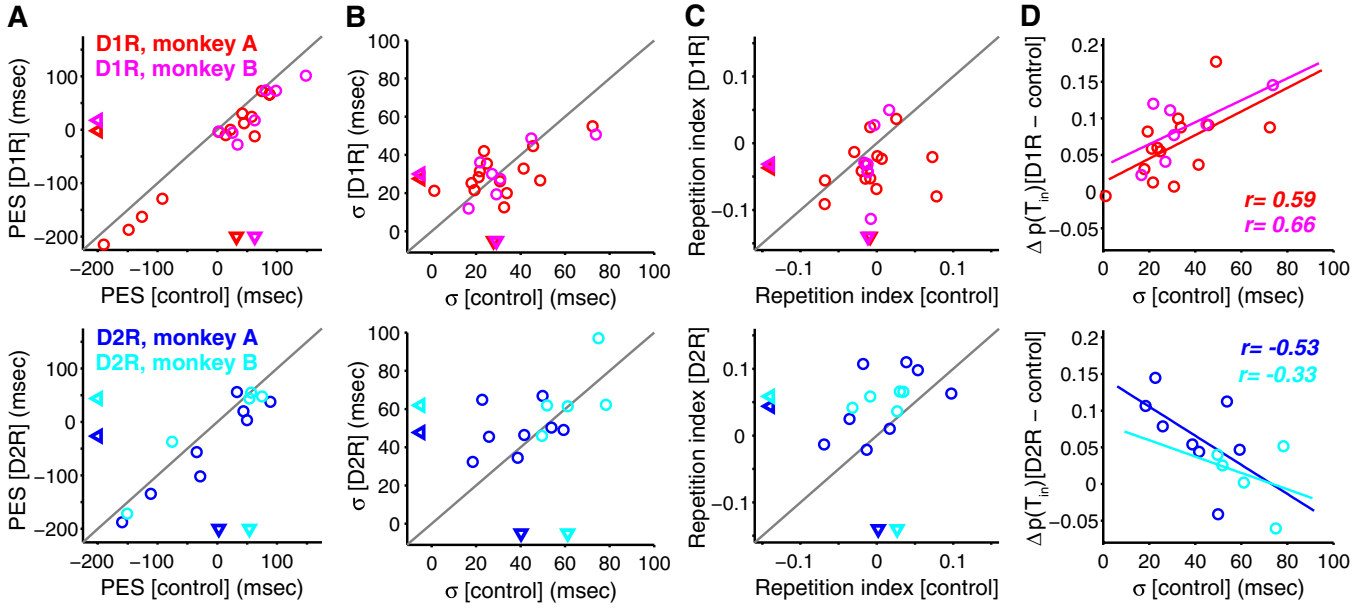

Fig. S1. Main experimental results shown separately for individual monkeys. (*A*–*D*) Plotted are the PES (*A*), stochasticity in choice (*B*), and repetition index (*C*) before and after the drug infusion and the correlation between changes in $p(T_{in})$ due to drug manipulations and the stochasticity of choice (σ) measured during control trials (*D*). Conventions are similar to those in Figs. 1 and 2. All of the main effects were consistent between the two monkeys.

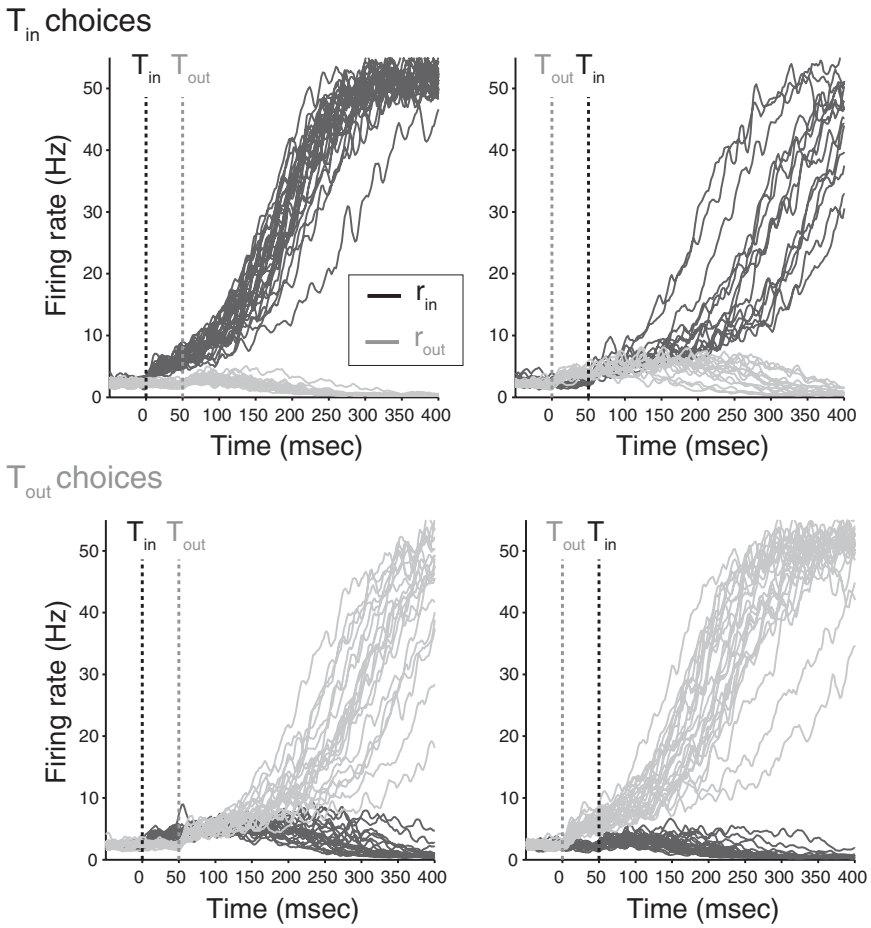

Fig. S2. Example of neural activity in the superficial layers of the model during trials in which T_{in} appeared 50 ms before the T_{out} target (*Left*) and trials in which T_{in} appeared 50 ms after the T_{out} target (*Right*), separately for trials in which T_{in} (*Upper*) or T_{out} (*Lower*) was selected. On most trials where the first target was selected, the activity within the two pools diverged quickly after the appearance of the first target and before the appearance of the second target (i.e., the TOA epoch indicated by the two vertical dashed lines). However, on some trials where the second target was selected the activity within the two pools remained equal during the TOA epoch and diverged only later.

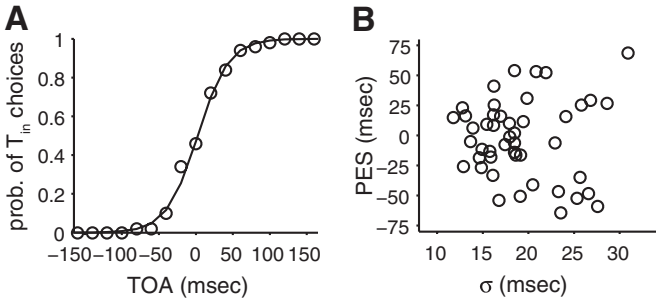

Fig. S3. Model's choice behavior during control trials. (*A*) An example of the model's psychometric function during control trials of the saccadic choice task. (*B*) Uncorrelated PES and σ-values generated through changes in the background input and the overall visual input.

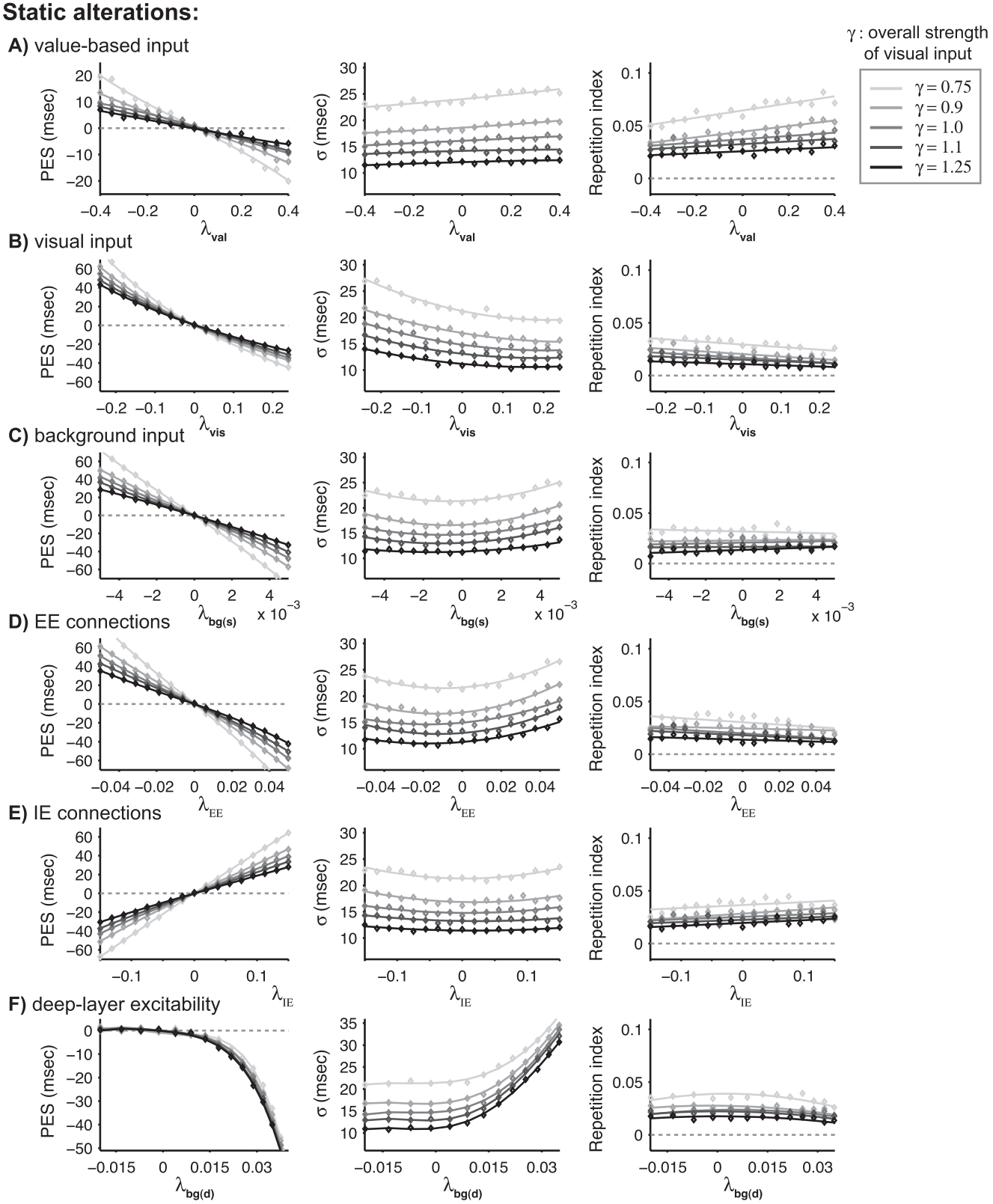

Fig. S4. Effects of the static alterations of individual model elements on its choice behavior. The model's choice behavior quantified by the PES, σ, and the RI is plotted as a function of the efficacy of individual model elements in the T_{in} column: (*A*) efficacy of the value-based input, λ_{val}; (*B*) efficacy of the visual input, λ_{vis}; (*C*) efficacy of the background input to the superficial layer pool, $\lambda_{bg(s)}$; (*D*) efficacy of the EE connections, λ_{EE}; (*E*) efficacy of the IE connections, λ_{IE}; and (*F*) efficacy of the background input to the deep layer pool, $\lambda_{bg(d)}$. Different shades of gray correspond to simulations with different values of the overall visual inputs (quantified by γ, Eq. **S9**), indicated in *Inset*. The curves on each plot show the results of polynomial regression fits for each γ-value (linear fit for the RI). Note that the static alterations at all sites except the deep layers resulted in changes that were larger when γ was smaller (corresponding to larger values of σ). The results presented here are summarized in Fig. 3*B* (static alterations).

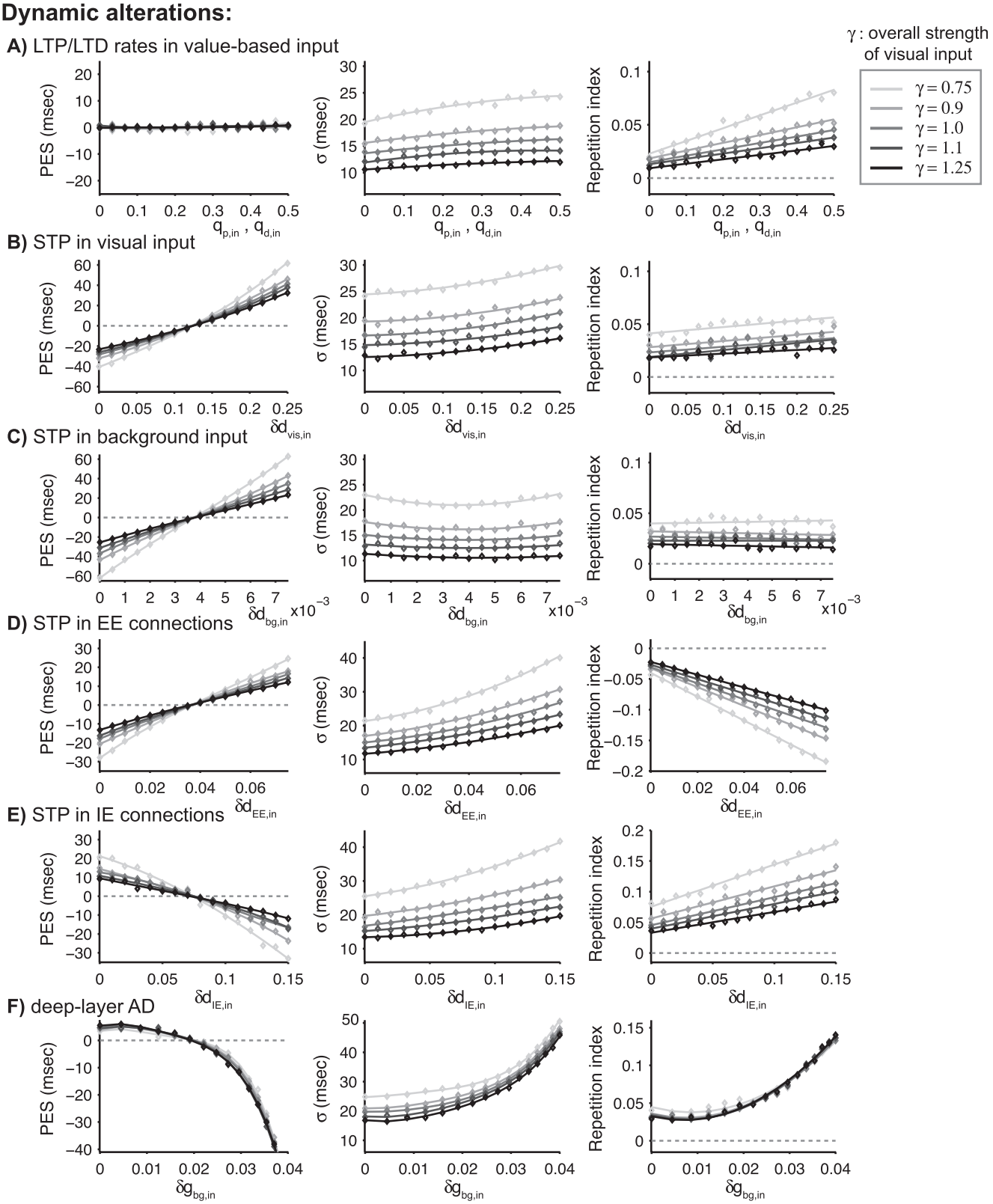

Fig. S5. Effects of the dynamic alterations of individual model elements on its choice behavior. The model's choice behavior quantified by the PES, σ, and the RI is plotted as a function of the model's parameters in the T_{in} column: (*A*) the rates of LTP and LTD of the value-based input, $q_{p,in}$, $q_{d,in}$; (*B*) the strength of STP in the visual input, $\delta d_{vis,in}$; (*C*) the strength of STP in the background input to the superficial layers, $\delta d_{bg,in}$; (*D*) the strength of STP in the EE connections, $\delta d_{EE,in}$; (*E*) the strength of STP in the IE connections, $\delta d_{IE,in}$; and (*F*) the efficacy of afterdepolarization in the deep layer pool, $\delta g_{bg,in}$. Conventions are the same as in Fig. S4. Note that the dynamic alterations at all sites except the deep layers resulted in changes that were larger when γ was smaller (corresponding to larger σ). The results presented here are summarized in Fig. 3*B* (dynamic alterations).

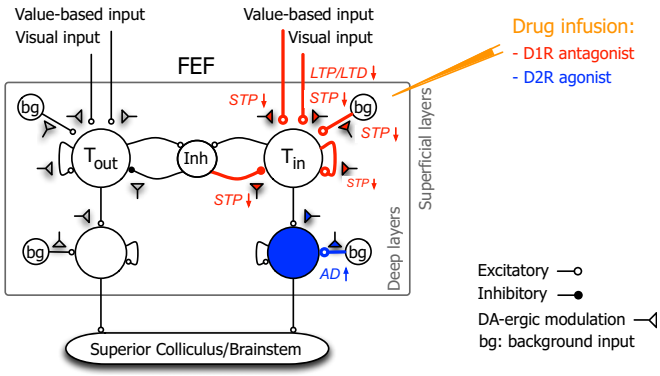

Fig. S6. Illustration of alterations to the network model that mimic the effects of the D1R and D2R manipulations in our experiment.