nature neuroscience

Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value

Lung-Hao Tai[1,7], A Moses Lee[1,2,7], Nora Benavidez[1,3], Antonello Bonci[4–6] & Linda Wilbrecht[1,4]

In changing environments, animals must adaptively select actions to achieve their goals. In tasks involving goal-directed action selection, striatal neural activity has been shown to represent the value of competing actions. Striatal representations of action value could potentially bias responses toward actions of higher value. However, no study to date has demonstrated the direct effect of distinct striatal pathways in goal-directed action selection. We found that transient optogenetic stimulation of dorsal striatal dopamine D1 and D2 receptor–expressing neurons during decision-making in mice introduced opposing biases in the distribution of choices. The effect of stimulation on choice was dependent on recent reward history and mimicked an additive change in the action value. Although stimulation before and during movement initiation produced a robust bias in choice behavior, this bias was substantially diminished when stimulation was delayed after response initiation. Together, our data suggest that striatal activity is involved in goal-directed action selection.

© 2012 Nature America, Inc. All rights reserved.

npg

Animals are faced with the challenge of optimally selecting actions in changing environments. Theoretically, these decisions can be implemented by estimating the value of different actions and then choosing the action of greatest value. Recently, the striatum has been implicated as an important neural structure that may mediate this process of action selection. Electrophysiological recordings in primate and rodent striatum have identified signals correlating with the expected outcomes of actions and measures of motivation for particular responses[1,2]. Other studies have shown that the striatum encodes representations of the value of actions in free choice tasks[3,4]. Striatal activity also parallels the learning of rewarded responses on the basis of previous experience[5,6] and is essential for the acquisition and execution of goal-directed behaviors[7].

Two dorsal striatal pathways have been proposed to mediate opposing influences on the selection of actions[8–10]. Activity in the direct pathway has been shown to facilitate actions, whereas activity in the indirect pathway has been demonstrated to inhibit behaviors[11,12]. The direct pathway is comprised of medium spiny neurons (MSNs) that express the dopamine D1 receptor (D1R), whereas the indirect pathway is comprised of MSNs that express the dopamine D2 receptor (D2R)[13–15]. It has been proposed that the D1R- and D2R-expressing MSNs encode a representation of the value of actions and generate a response bias toward actions of higher value[16,17]. Although this model is consistent with numerous studies, other studies have suggested that these populations of neurons may merely have a permissive role in learning associations[18] or only modulate the vigor of actions without affecting the animals' choice behavior[19,20]. No study to date has definitely demonstrated the role of these two populations of neurons in the context of reward-based decision-making[13].

To directly investigate the role of striatal activity in action selection, we created a probabilistic switching task in which mice chose to enter a reward port located to either their left or right side. To inform their choices in the task, mice relied on recent reward history to assess whether water would be delivered from one of two reward ports. Critically, we selected a left versus right choice design on the basis of previous studies that showed that unilateral striatal manipulations can affect lateralized body movements[11,21], neurons in the striatum encode actions for movements contralateral or ipsilateral to the recording site[22–24], and brainstem motor programs mediating orienting and approach behaviors[25] are regulated by the basal ganglia[16,25–27]. From these previous findings, we reasoned that unilateral manipulations to each striatal hemisphere could affect the selection of spatially lateralized responses in the context of this task. We developed a computational model that assigns a value for each action given a particular history of rewards to predict the distribution of left versus right choices. We then used optogenetic techniques to examine the effect of transient unilateral stimulation of the D1R and D2R-expressing striatal neurons during an epoch in the task when animals were choosing to approach a left or a right port.

We found that transient activation of D1R-expressing striatal neurons biased choices toward the port contralateral to the side of stimulation and transient activation of D2R-expressing striatal neurons biased choices toward the ipsilateral port. However, rather than giving rise to a stereotyped and consistent motor response, optical stimulation induced a bias in the likelihood of choice responses that was dependent on the previous history of rewards for each choice. In the context of our model, we found that optical stimulation mimics a fixed additive shift in action value[28,29]. The effect of

[1]Ernest Gallo Clinic and Research Center, Emeryville, California, USA. [2]Medical Scientist Training Program, Neuroscience Graduate Program, University of California, San Francisco, California, USA. [3]Cognitive Science Program, University of California, Berkeley, California, USA. [4]Department of Neurology, University of California, San Francisco, California, USA. [5]Intramural Program, National Institute of Drug Abuse, Baltimore, Maryland, USA. [6]Solomon H. Snyder Department of Neuroscience, Johns Hopkins School of Medicine, Baltimore, Maryland, USA. [7]These authors contributed equally to this work. Correspondence should be addressed to L.W. (lwilbrecht@gallo.ucsf.edu).

Received 25 April; accepted 17 July; published online 19 August 2012; doi:10.1038/nn.3188

stimulation was only effective in a narrow temporal window before and during the early initiation of movements. Striatal activation also sped and slowed the initiation of responses in a manner similar to changes in action value. Together, these data are consistent with the hypothesis that the striatum encodes the value of actions and indicate that activity in these pathways can be used to bias the selection and vigor of competing goal-directed actions.

RESULTS

To study action selection during decision making in mice, we trained adult BAC transgenic mice ($n = 28$) expressing Cre recombinase under either dopamine D1R or D2R regulatory elements on a spatial two-alternative forced-choice probabilistic switching task. Briefly, the task required animals to initiate a trial with a nose poke into a central port followed by movement to a left or a right port to obtain reward. The rewarded port delivered a water reward for 75% of correct responses, and this port was periodically switched across blocks. The length of each block was randomly distributed between 7–23 trials and the switch only took place after a rewarded trial (**Fig. 1a**). A Go cue signaled when mice could approach either choice port. Thus, the only information provided to guide the animals' choice behavior was the expectation of reward based on the outcome of previous trials. After initial training, mice took, on average, 2.20 ± 0.04 (right to left) and 2.22 ± 0.05 (left to right) trials (±s.e.m., $n = 28$) to switch their behavioral responses following reversals in action-outcome contingencies between blocks (**Fig. 1b**). When we analyzed the effects of reward history in the previous two trials on upcoming choice, we found that mice implemented a win-stay, lose-shift strategy in which rewards served as evidence to continue responding at a port and the lack of reward served as evidence to switch (**Fig. 1c**). The mice's choice probability generally tracked the reward probability at each port for various reward histories (**Supplementary Fig. 1a**).

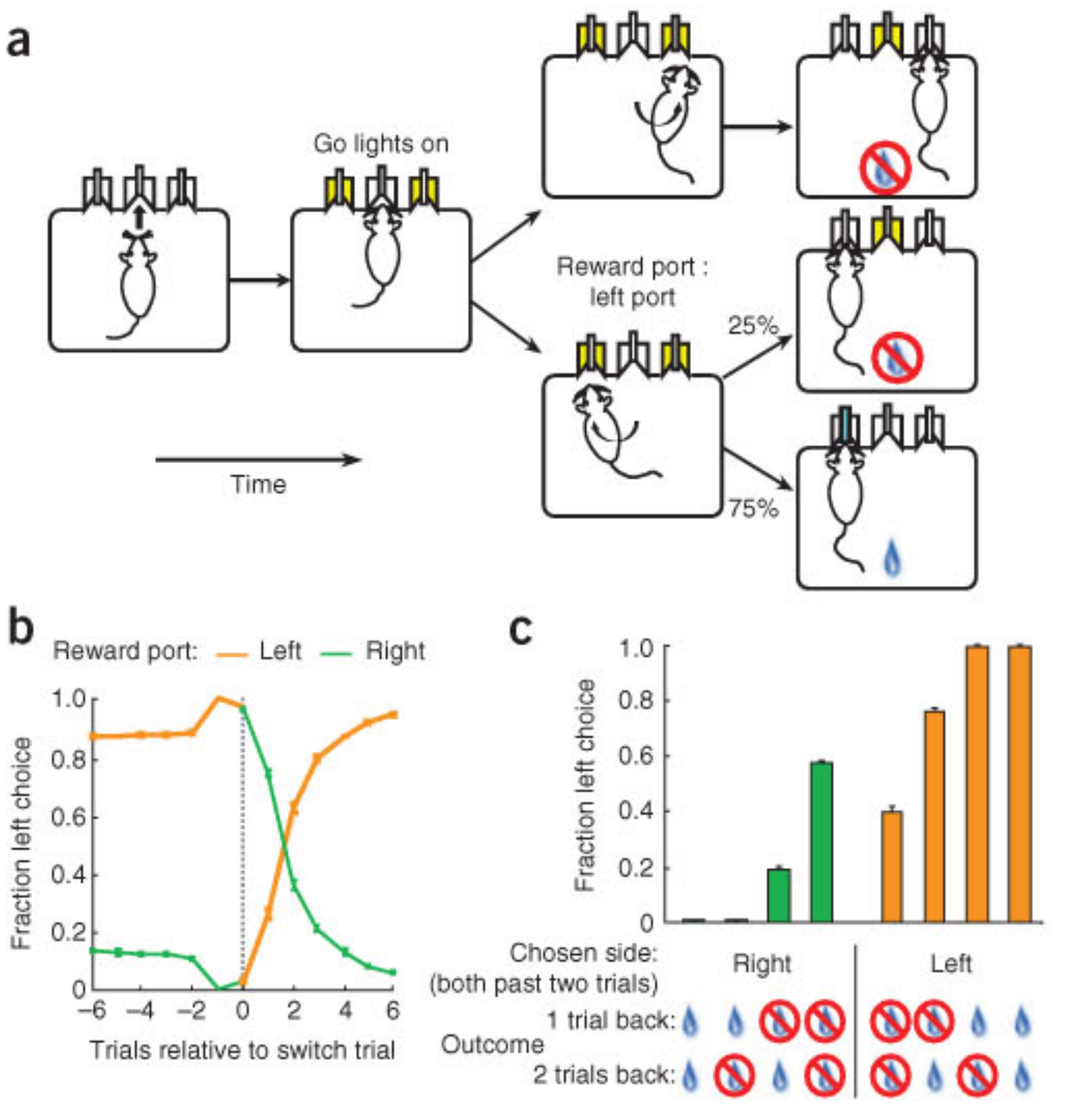


Figure 1 Design of the probabilistic switching task and mouse performance. (**a**) Sequence of events in a probabilistic switching task. Mice learned to initiate a trial at the center port and to choose a left or right peripheral port for water reinforcement. Only one peripheral port was rewarded at a time. In 25% of trials, neither port was rewarded. The rewarded port was switched only after a rewarded trial. (**b**) Fraction of choices for left port ($n = 28$ subjects) for trials before and after a switch of the rewarded port (at trial 0). (**c**) Fraction of choices for the left port from one subject for reward histories in which two consecutive choices to either the left or the right port were made during the previous two trials. Data from mixed choice histories are not shown for brevity. All error bars represent s.e.m.

Action value estimates based on previous reward history

We generated a quantitative model to describe the mice's behavior in the task. Prior electrophysiological studies have found that the activity of striatal neurons correlates with estimates of trial-to-trial values of actions[3,4,30]. These estimates of value assume the softmax decision rule[31], which describes the tendency of decision makers to have more variable responses when the alternatives are more similar in value. In the case of two alternatives under the softmax rule, the contribution of previous reward history to the value of each action can be estimated by multivariate logistic regression[30]. We fit a regression model in which the probability of choices at each port in the upcoming trial was determined by the animals' previous reward history (**Fig. 2** and Online Methods).

The regression analysis revealed that the contribution of prior rewards declined with the passage of trials (**Fig. 2a**). Rewards in the previous three trials had a significant effect on choices in the upcoming trial, serving as evidence for the animal to stay at the rewarded port, as indicated by the positive regression coefficients ($P < 4 \times 10^{-6}$; **Fig. 2a**), whereas the lack of reward at the chosen port in the previous trial significantly promoted switching in the following trial, as represented by the negative regression coefficient ($P = 1.1 \times 10^{-5}$; **Fig. 2a**). The large regression coefficients for the previous two trials validated our initial attempts to analyze choice behavior by segregating trials on the basis of the animals' reward history in the previous two trials (**Fig. 1c**). From the model, we generated dynamic estimates of action values and probabilities for responding at each port on the basis of the animals' recent reward history (**Fig. 2a**,**b**,**e**). These action value estimates were the sum of regression coefficients corresponding to the previous reward history for each side (**Fig. 2a**) determining the distribution of choices of the subject (**Fig. 2b**). The choice probabilities predicted by an identically generated regression model using 70% of the data recapitulated the actual distribution of choices in the remaining 30% of the data, confirming the predictive validity of our model (**Supplementary Fig. 1e**,**f**).

Stimulation of distinct striatal neurons biases choice

To independently study the activity of D1R-expressing and D2R-expressing striatal neurons in our task, we bilaterally injected an adeno-associated virus (AAV) into the dorsal striatum that enabled Cre-dependent viral expression of channelrhodopsin (ChR2) and yellow fluorescent protein (eYFP) in transgenic mice expressing Cre under regulatory elements for either the dopamine D1R (D1-Cre mice) or D2R (D2-Cre mice). We then chronically implanted an optic fiber in or just above the dorsal medial striatum of each hemisphere. Our Cre-dependent strategy targeted ChR2 expression to the direct or indirect pathways of the basal ganglia (**Supplementary Fig. 2**). In D2-Cre mice, 37.7% of putative MSNs expressed eYFP (consistent with indirect pathway targeting) and 38.7% of the choline acetyltransferase (ChAT)-expressing neurons co-expressed ChR2-eYFP (**Supplementary Fig. 3**). Estimates suggest that 80–97.7% of neurons in the striatum are MSNs and 0.3–2% are cholinergic[32,33]. We therefore estimate in D2-Cre mice we transduced a ratio of MSNs to cholinergic neurons of roughly 50:1. By simultaneously recording and optically stimulating striatal neurons, we were able to confirm that optically driven striatal neuronal firing in our system was time-locked to stimulation

© 2012 Nature America, Inc. All rights reserved.
npg

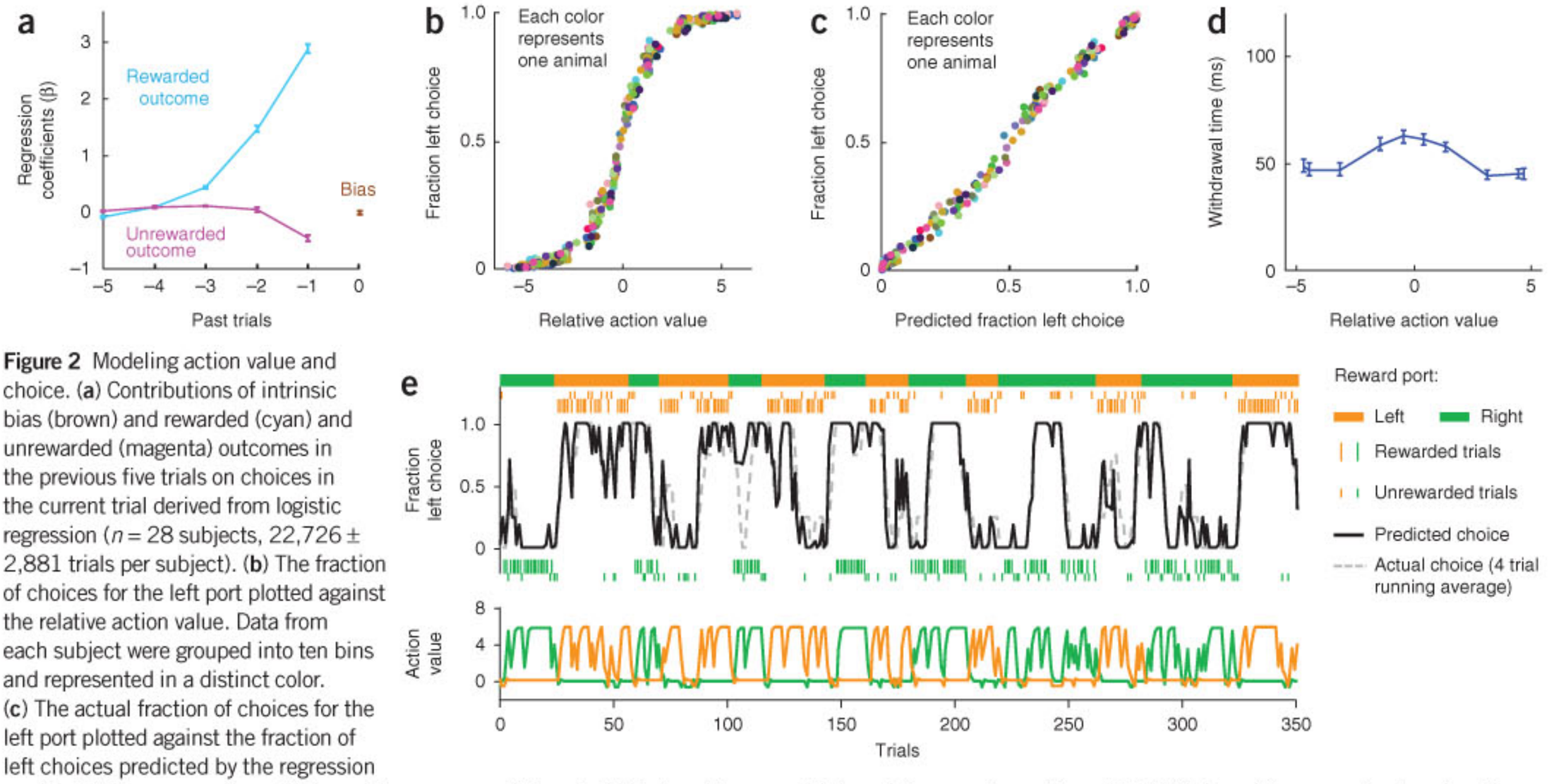


Figure 2 Modeling action value and choice. (**a**) Contributions of intrinsic bias (brown) and rewarded (cyan) and unrewarded (magenta) outcomes in the previous five trials on choices in the current trial derived from logistic regression ($n = 28$ subjects, 22,726 ± 2,881 trials per subject). (**b**) The fraction of choices for the left port plotted against the relative action value. Data from each subject were grouped into ten bins and represented in a distinct color. (**c**) The actual fraction of choices for the left port plotted against the fraction of left choices predicted by the regression model. (**d**) Average median withdrawal time measured from Go light signal to nose withdrawal from center port ($n = 28$). Withdrawal time was shorter when the relative action value for either port was higher. (**e**) Example data from 14 trial blocks. Top, right reward blocks are represented in green and left reward blocks in orange. The dashed line indicates the subject's probability of choosing the left port averaged across four trials and the black line indicates the predicted probability of choice on the basis of action value estimates (bottom). Long ticks correspond with rewarded trials and short tics represent unrewarded trials. All error bars represent s.e.m.

© 2012 Nature America, Inc. All rights reserved.

(**Supplementary Fig. 4**). Optically evoked activity could be a result of direct ChR2 activation or potentially indirect activation of neurons downstream of ChR2-expressing cells.

We then sought to determine whether activation of specific neural populations in the striatum could affect the choice behavior of animals[34]. Optical stimulation was delivered at a decision point in

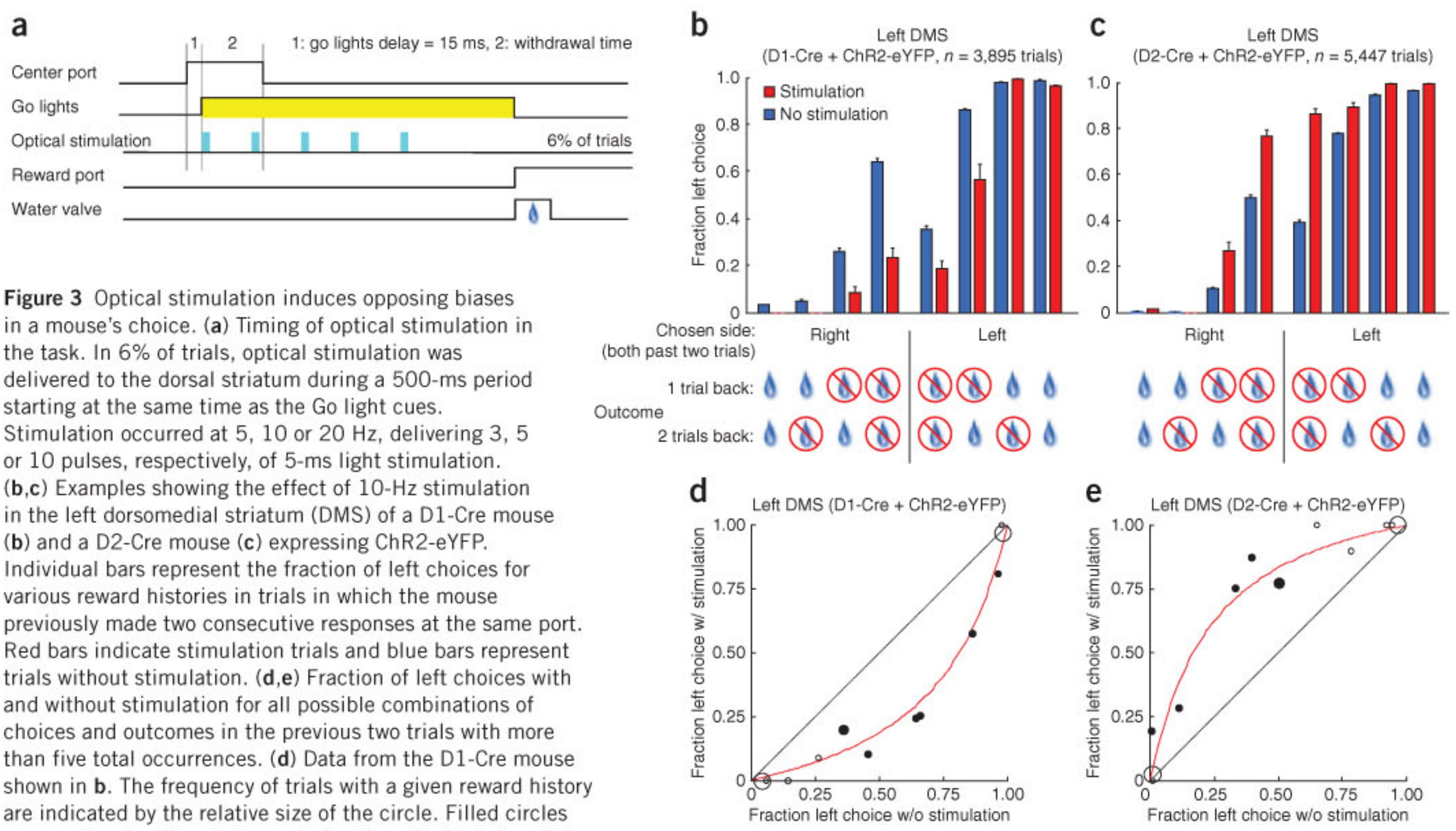


Figure 3 Optical stimulation induces opposing biases in a mouse's choice. (**a**) Timing of optical stimulation in the task. In 6% of trials, optical stimulation was delivered to the dorsal striatum during a 500-ms period starting at the same time as the Go light cues. Stimulation occurred at 5, 10 or 20 Hz, delivering 3, 5 or 10 pulses, respectively, of 5-ms light stimulation. (**b**,**c**) Examples showing the effect of 10-Hz stimulation in the left dorsomedial striatum (DMS) of a D1-Cre mouse (**b**) and a D2-Cre mouse (**c**) expressing ChR2-eYFP. Individual bars represent the fraction of left choices for various reward histories in trials in which the mouse previously made two consecutive responses at the same port. Red bars indicate stimulation trials and blue bars represent trials without stimulation. (**d**,**e**) Fraction of left choices with and without stimulation for all possible combinations of choices and outcomes in the previous two trials with more than five total occurrences. (**d**) Data from the D1-Cre mouse shown in **b**. The frequency of trials with a given reward history are indicated by the relative size of the circle. Filled circles represent a significant change in fraction of left choice with stimulation ($P < 0.05$, Fisher's exact test). The red curve relates the probabilities of choice with and without stimulation for a fixed odds ratio (odds ratio = $e^{-1.33 \pm 0.20}$). (**e**) Data from the D2-Cre mouse shown in **c** (odds ratio = $e^{1.45 \pm 0.18}$). All error bars represent s.e.m.

the task, when the Go cue signals the animal to make their choice (**Fig. 3a**). At this point in the task, the choice ports are equidistant in egocentric space to the left and right of the mouse, and this period marks the time at which animals must select and initiate an action to acquire a potential reward (**Fig. 1a**). Stimulation was delivered at 5, 10 or 20 Hz for 500 ms in 6% of total trials interspersed at random. The stimulation parameters that we used reflect physiologically relevant activity found in awake, freely moving striatal recordings in mice[23,35]. The presence of stimulation and nonstimulation trials in the same session allowed us to make highly controlled within subject comparisons to determine the effect of stimulation. Stimulation sessions were interspersed with training sessions without stimulation, and the hemisphere that was stimulated alternated with every stimulation session. There were no differences in animals' choice behavior across sessions with identical stimulation parameters; thus, data from these sessions were pooled for analysis. To rule out any possible effect of the optical stimulation or viral expression, we also injected mice with a control virus expressing Cre-dependent eYFP. These control mice were subject to identical training regimens and stimulation parameters as experimental mice.

Given that the mice's responses were dependent on their reward history, we analyzed choice in stimulation and nonstimulation trials on the basis of the history of reward and choice over the previous two trials. Activation of D1R-expressing neurons with optical stimulation induced a bias in choice for the port contralateral to the hemisphere in which the light was delivered. The stimulation-induced bias was greater after unrewarded trials when the animals' responses tended to be more variable (**Fig. 3b**). A small bias was seen following rewarded trials (**Fig. 3b**).

Striatal stimulation in D2-Cre mice induced a bias to the port ipsilateral to the site of stimulation (**Fig. 3c**), opposite to the direction observed in D1-Cre mice. Again, stimulation induced bias was greater after unrewarded trials when the animals' responses were more variable (**Fig. 3c**).

To further investigate the relationship between stimulation and reward history, we plotted the probability of a left choice for a given reward history in trials with and without stimulation (**Fig. 3d,e**). The plot revealed 'bowing' of data points off the unity line in opposite directions for D1- or D2-Cre mice. This bowing can be quantified as the odds of choice without stimulation scaled by a fixed factor called the odds ratio.

Optical stimulation mimics a change in action value

In mice performing the switching task without stimulation, recent reward history exerted a strong effect on a mouse's upcoming choice (**Fig. 1**). We quantified these effects by generating estimates of the value of actions (**Fig. 2b**), which could be used to guide the animals' selection of future actions[3,4,30]. Similar to rewards, striatal stimulation also had the ability to bias mice's choices. Given that neurons in striatum may encode the value of actions[3,4], we hypothesized that striatal stimulation may mimic the effects that a change in valuation of actions (on the basis of reward history) would have on the mice's upcoming choice.

Using our previous estimates of the relative action value for various reward and choice histories (**Fig. 2b**), we plotted the probability of choices made with and without stimulation for individual subjects in the same sessions (**Figs. 4** and **5** and **Supplementary Figs. 5** and **6**). We found that striatal stimulation shifted the sigmoid choice probability curve along the relative action value axis. In this way, stimulation can be interpreted as mimicking a change in the relative valuation for selecting the left versus right port by a fixed factor. In individual mice, optical stimulation of D1R-expressing striatal neurons appeared to increase the value of the contralateral port (**Fig. 4a**), whereas stimulation of the D2R-expressing neurons resembled a decrease in the value for the contralateral port (**Fig. 5a**). In both individual and group data, the magnitude of the shift in relative action value consistently increased with the frequency of stimulation and the direction of the shift induced by stimulation showed

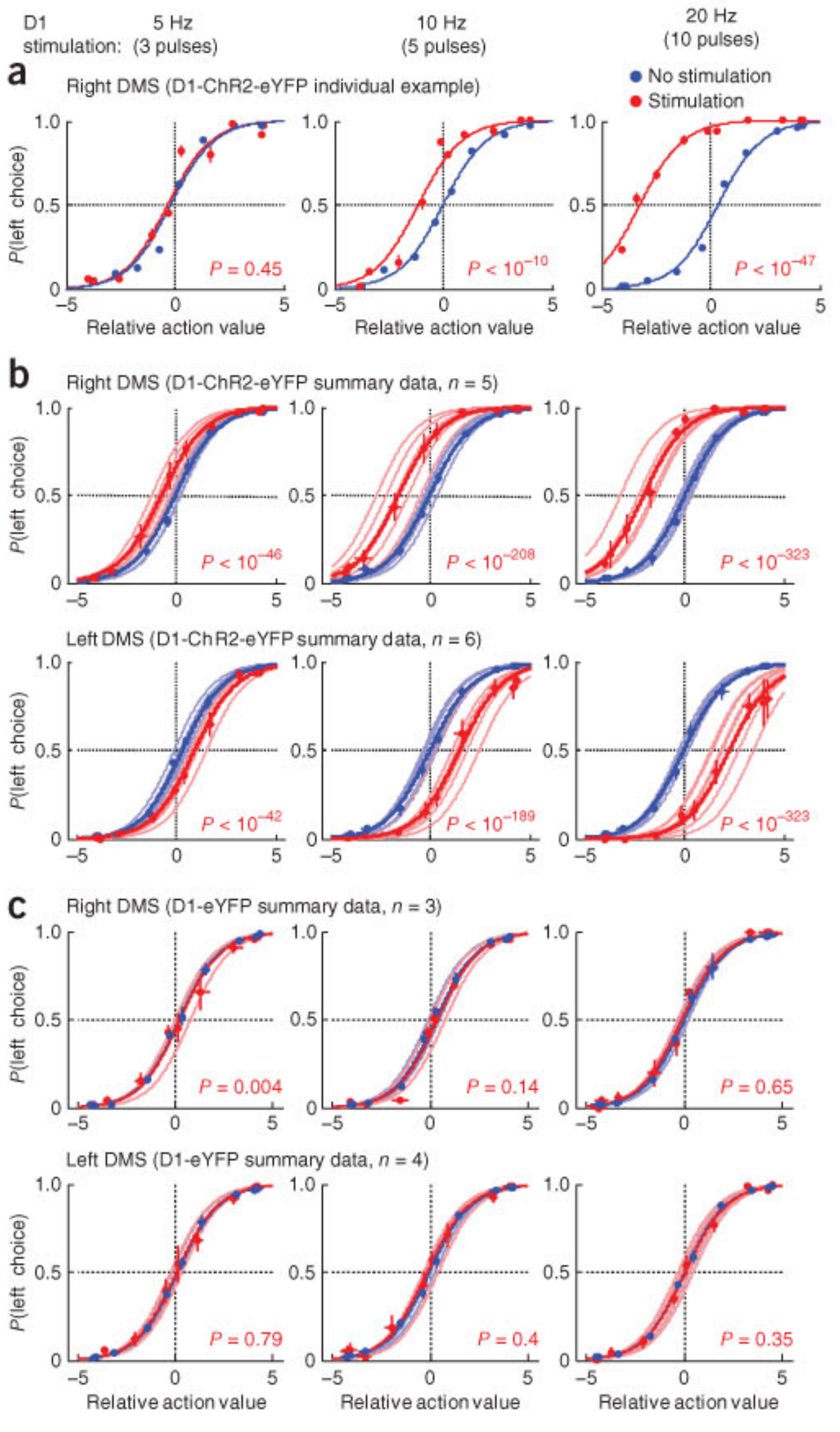


Figure 4 Dorsal striatal D1R-expressing neuron activation mimics an increase in relative action value for contralateral choice. Fraction of choices for the left port on trials with different relative action value estimates in D1-Cre mice in the presence (red) or absence (blue) of optical stimulation. (**a**) Representative data from one mouse transduced with AAV-EF1α-DIO-ChR2-eYFP with optical stimulation in the right hemisphere. Logistic regression was used to fit the data from trials with (red line) and without stimulation (blue line). A leftward shift in the logistic curve represents a bias for the left reward port. (**b**) Summary data for probability of choice for the left port and relative action value pooled from all D1-Cre mice expressing ChR2-eYFP and stimulated on either the right hemisphere (top) or left hemisphere (bottom). A logistic regression fit for the no stimulation (thick blue line) and stimulation (thick red line) trial data was produced for data pooled from all mice. Curves representing the estimated probability of choice for given relative action values in individual animals are plotted in light blue (no stimulation) or light red (with stimulation). Reported *n* refers to number of stimulation sites (one per hemisphere). (**c**) Summary data from all D1-Cre mice expressing eYFP alone and stimulated on either the right hemisphere (top) or left hemisphere (bottom). *P* values reported for *t* tests: H_0:$\beta_{stim} = 0$ (distance between thick red and blue lines). All error bars represent s.e.m.

© 2012 Nature America, Inc. All rights reserved.

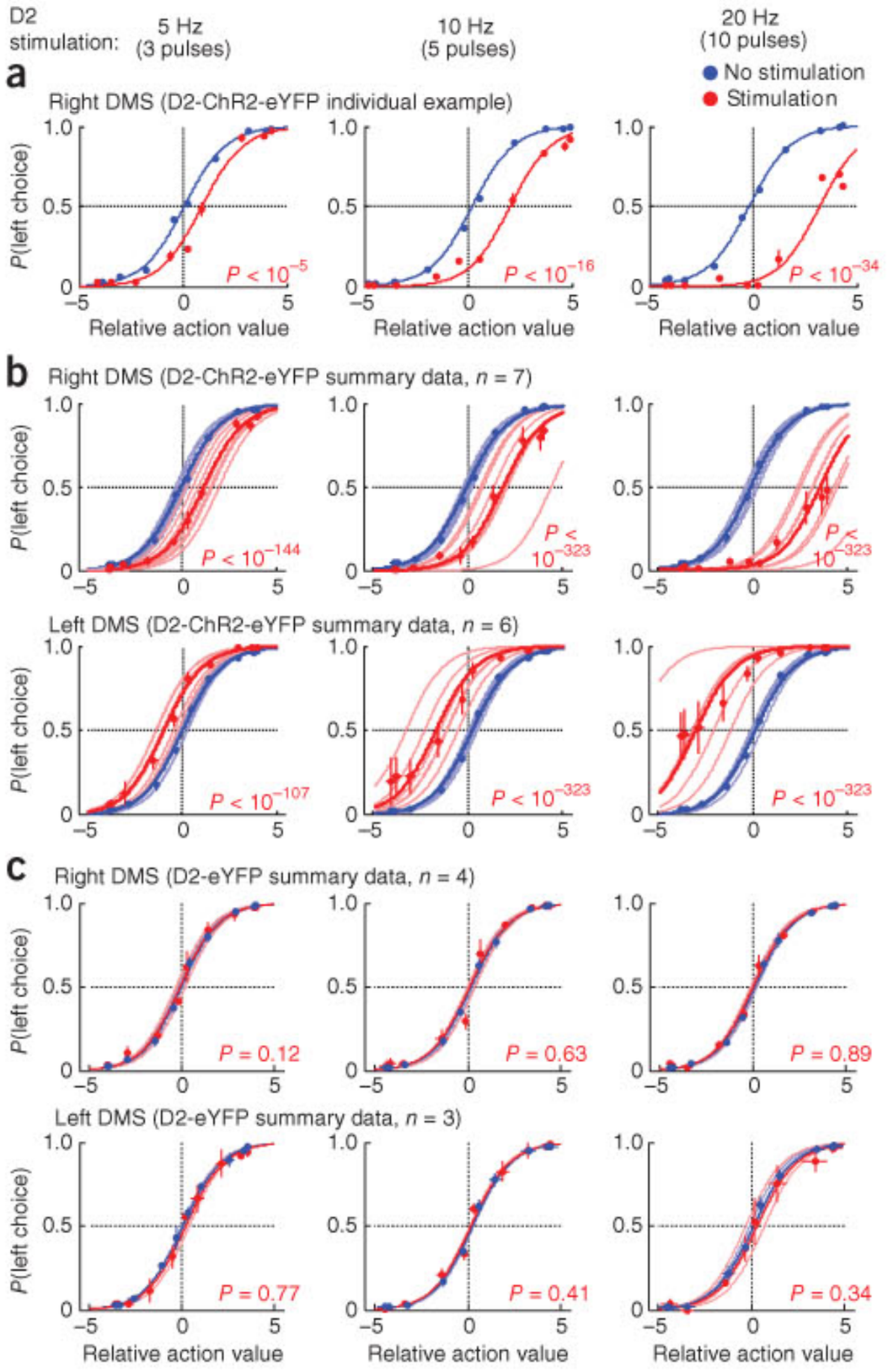


Figure 5 Dorsal striatal D2R-expressing neuron activation mimics a decrease in relative action value for contralateral choice. Fraction of choices for the left port on trials with different relative action value estimates in D2-Cre mice in the presence (red) or absence (blue) of optical stimulation. (**a**) Representative data from one mouse transduced with AAV-EF1α-DIO-ChR2-eYFP with optical stimulation in the right hemisphere. Logistic regression was used to fit the data from trials with (red line) and without stimulation (blue line). A rightward shift in the logistic curve represents a bias for the right reward port. (**b**) Summary data for the probability of choice for the left port and relative action value pooled from all D2-Cre mice expressing ChR2-eYFP and stimulated on either the right hemisphere (top) or left hemisphere (bottom). A logistic regression fit for the no stimulation (thick blue line) and stimulation (thick red line) trial data was produced for data pooled from all mice. Curves representing the estimated probability of choice for given relative action values in individual mouse are plotted in light blue (no stimulation) or light red (with stimulation). Reported *n* refers to number of stimulation sites (one per hemisphere). (**c**) Summary data from all D2-Cre mice expressing eYFP alone and stimulated on either the right hemisphere (top) or left hemisphere (bottom). *P* values reported for *t* tests: H_0:$\beta_{stim} = 0$ (distance between thick red and blue lines). All error bars represent s.e.m.

a consistent relationship with the stimulated hemisphere (**Figs. 4b**, **5b** and **6a** and **Supplementary Figs. 5** and **6**). Notably, control mice did not show a bias in their distribution of choices between trials with or without optical stimulation. In either D1-Cre or D2-Cre control mice expressing only eYFP without ChR2, no shift in action value was observed (**Figs. 4c** and **5c**).

In addition to its role in action selection, activity in the striatum correlates with reaction time[1,36]. In our task, the time a mouse takes to withdraw from the center port after trial initiation can be used to measure the latency to initiate movement. Without optical stimulation, the withdrawal time for a particular response was substantially shorter when the relative action value for that response was high (**Fig. 2d**). Stimulation of D1R-expresing neurons decreased center port withdrawal time when the value of the contralateral port was greater (**Fig. 6b**). However, when the ipsilateral port was valued more highly, stimulation of D1R-expressing neurons slowed withdrawal time (**Fig. 6b**). In contrast with D1-Cre mice, striatal optogenetic stimulation in the D2-Cre mice increased center port withdrawal time when the value of the contralateral port was greater (**Fig. 6b**). However, when the ipsilateral port was valued more highly, stimulation of D2-expressing neurons sped withdrawal times (**Fig. 6b**). Changes in withdrawal time with optical stimulation were still apparent when we controlled for the eventual choice response of the mice (**Supplementary Fig. 7a**,**b**). Notably, we found that, after rewarded trials when optical stimulation was not likely to affect port choice, stimulation was still observed to affect withdrawal time.

Effective time window for optical stimulation

Previous reports have found that phasic striatal activity before response initiation correlates with trial-by-trial estimates of action value[3,4]. In the experiments described above, light pulses were delivered during a 500-ms epoch coinciding with onset of a Go cue, after which mice initiated their motor responses (**Fig. 3a**).

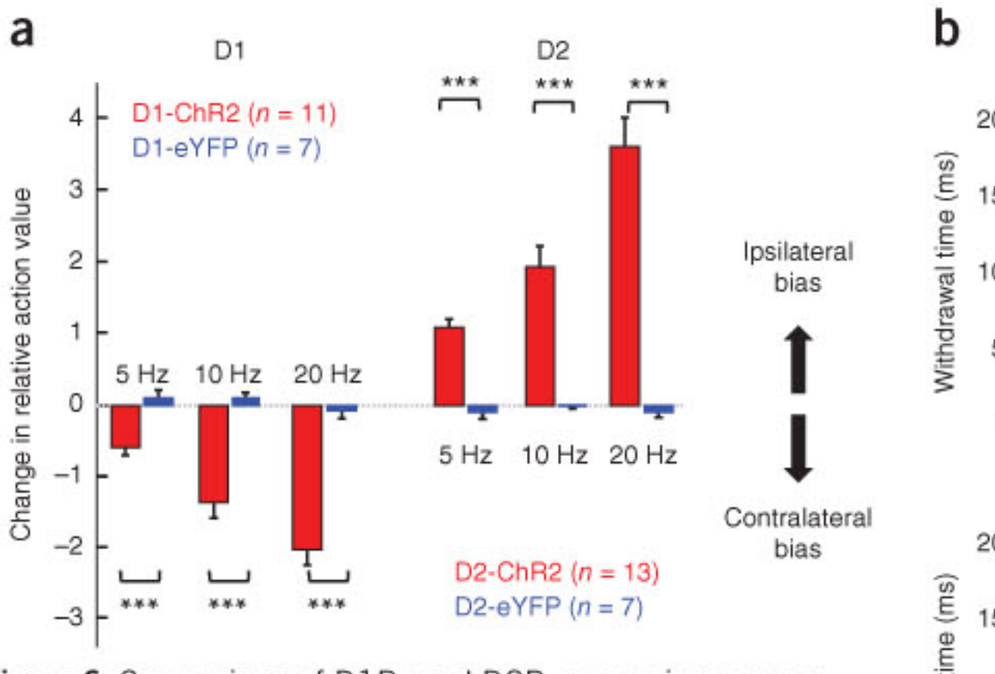
We next performed two experiments to investigate the effect of striatal activity before and after response initiation. Response initiation in our task can be measured as the latency to withdraw from the center port. The median withdrawal times for mice in this study were approximately 50 ms, and the majority of response withdrawal times could be found in a window spanning approximately 30–140 ms (50 ± 25% confidence interval; **Supplementary Fig. 1c**).

To determine whether stimulation before movement initiation was sufficient to bias the mice's choice behavior, we trained a new cohort of mice in a protocol in which two optical pulses separated by 50 ms (20 Hz) were delivered before an auditory Go cue (**Fig. 7a**). Using this second protocol, we confirmed that it is possible to bias the mice's choice behavior with striatal activation of both D1R-expressing and D2R-expressing striatal neurons before movement initiation (**Fig. 7b–d** and **Supplementary Fig. 8a**,**b**).

In a third protocol, we examined the effect of delaying stimulation until after movement initiation. In a subset of mice trained in the original protocol, 10-Hz optical stimulation was delivered at either 0 ms or 150 ms following the presentation of the Go cue light (**Fig. 8a**). The bias induced by optical stimulation delayed by 150 ms was significantly weaker than the bias induced without delay ($P < 0.005$; **Fig. 8b–d**). These data suggest that the effect of striatal activity on choice behavior decayed with time after the onset of the Go cue.

Following the switching task, we carried out two more experiments to grossly measure the effect of transient and prolonged stimulation on the mice's locomotor behavior outside the context of the task. In a 10-cm-diameter cylindrical environment, transient bursts of 20-Hz stimulation (identical to the maximum burst used in our switching task) did not produce a significant change in head or body orientation ($P > 0.4$; **Supplementary Fig. 9c**,**d**). However, prolonged stimulation of the D1R-expressing striatal neurons for 60 s at 5, 10 and 20 Hz induced a graded increase in contralateral rotations and prolonged stimulation

© 2012 Nature America, Inc. All rights reserved.
npg



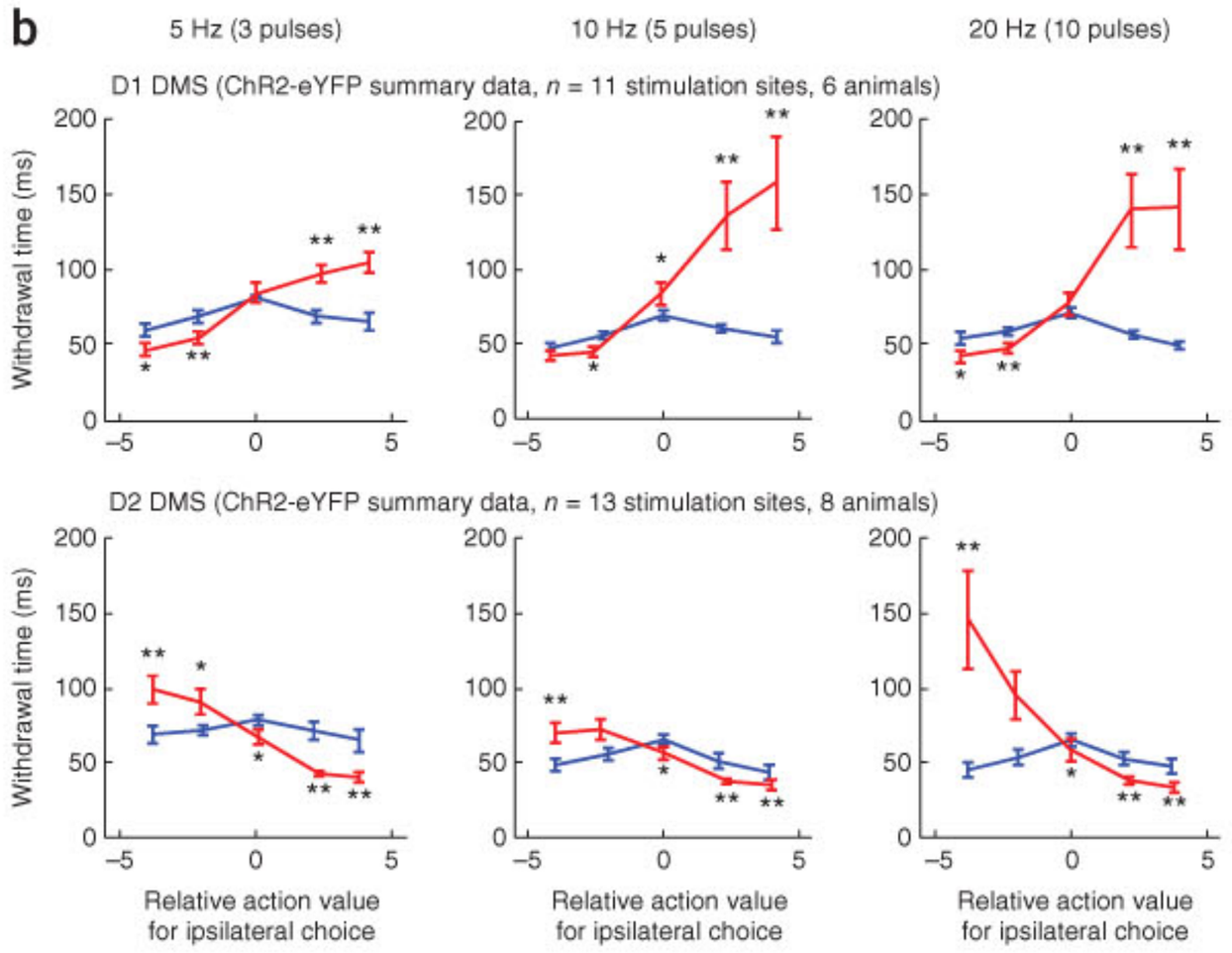
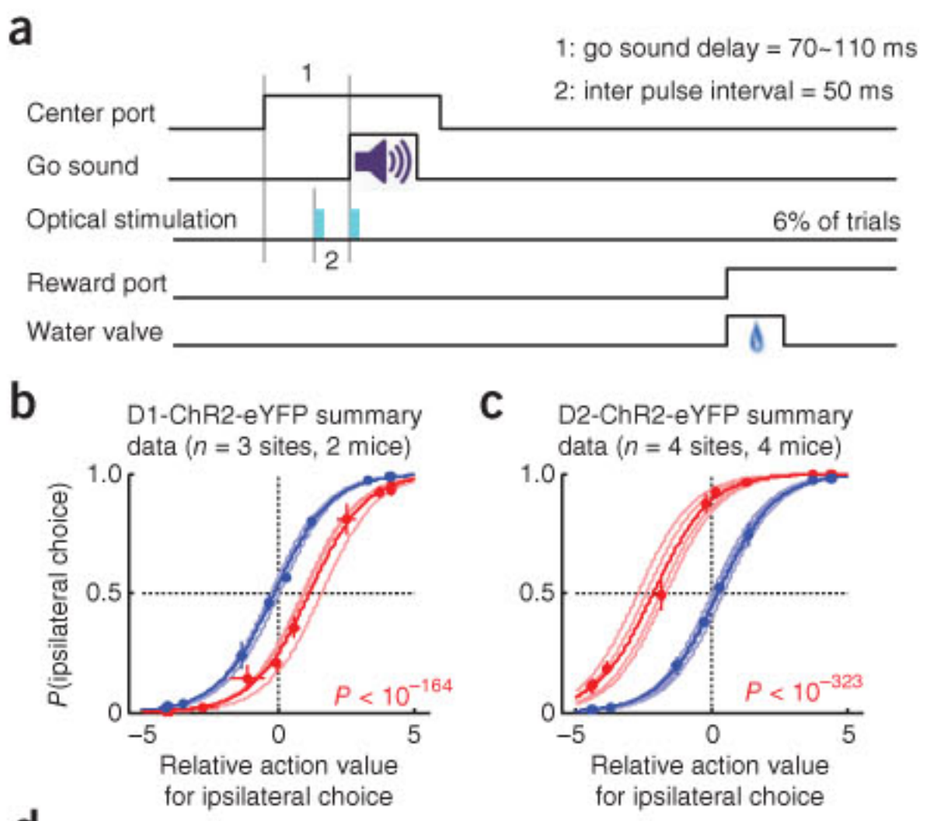


Figure 6 Comparison of D1R- and D2R-expressing neuron stimulation. (**a**) Estimated change in the relative action value for choosing the port ipsilateral versus contralateral to the site of stimulation averaged across individuals in a group. Positive changes in relative action values correspond to an ipsilateral bias and negative changes correspond to a contralateral bias. Estimates of relative action change were derived from logistic regression analysis (Online Methods) for 5-, 10- and 20-Hz stimulation sessions. Reported *n* refers to number of stimulation sites. The D1-Cre data set consisted of six stimulated mice expressing ChR2-eYFP and four control mice expressing only eYFP. The D2-Cre data set consisted of eight stimulated mice expressing ChR2-eYFP and five control mice expressing only eYFP. (**b**) The median time taken to withdraw from the center port averaged across individual D1-Cre mice (top) and D2-Cre mice (bottom) expressing ChR2-eYFP in trials without stimulation (blue) or with stimulation (red) and across different relative action values for choosing the port ipsilateral versus contralateral to the site of stimulation. Positive relative action values correspond to trials in which the value of the port ipsilateral to the site of stimulation is greater than the contralateral port. Median times to withdraw are plotted for 5- (left), 10- (middle) and 20-Hz (right) stimulation sessions. All error bars represent s.e.m. $^{*}P < 0.05$, $^{**}P < 0.01$, Wilcoxon signed-rank test. $^{***}P < 0.001$, Wilcoxon rank-sum test.

© 2012 Nature America, Inc. All rights reserved.

npg

of D2R-expressing neurons induced a decrease in contralateral rotations in an open field ($P > 0.4$; **Supplementary Fig. 9e,f**). Together, these data are consistent with previous reports[11] and the known organization of distinct striatal pathways in the basal ganglia[8,10,13].



DISCUSSION

Our data suggest that activity in distinct populations of striatal neurons exert opposing biases on the selection of goal-directed responses. Activation of D1R-expressing striatal neurons increased the occurrence of choices for the port contralateral to the side of stimulation (**Figs. 3**, **4** and **6**), whereas stimulation of D2R-expressing striatal neurons increased the occurrence of ipsilateral choices (**Figs. 3**, **5** and **6**). The effect of stimulating each population of neurons was not deterministic, but was dependent on the animals' recent reward history (**Fig. 3**). On closer inspection, the magnitude of this bias mimicked an additive change in the relative value of actions estimated using a simple model based on the mice's history of rewards and choices (**Figs. 2b** and **4–6**). Stimulation also

Figure 7 Significant bias was induced by stimulation limited to a 50-ms period before a Go cue. (**a**) The timing of events in sessions in which a Go cue sound was delayed 70–110 ms after initiation of trial with center port entry. In these trials, two 5-ms optical pulses separated by 50 ms were delivered just prior and at the onset of the Go cue and before withdrawal from the center port. (**b**,**c**) Summary data for probability of choice for the ipsilateral port on trials with different relative action value estimates pooled from D1-Cre mice (**b**) and D2-Cre mice (**c**) expressing ChR2-eYFP with confirmed eYFP expression sites. Positive relative action values correspond to trials in which the value of the port ipsilateral to the site of stimulation was greater than the contralateral port. A logistic regression fit for the no stimulation (thick blue line) and stimulation (thick red line) trial data was produced for data pooled from all subjects. Curves representing the estimated probability of ipsilateral choice for given relative action values in individual animals are plotted in light blue (no stimulation) or light red (with stimulation). *P* values reported for *t* tests: H_0:$\beta_{stim} = 0$ (distance between thick red and blue lines). (**d**) Change in estimated relative action value for choosing the port ipsilateral versus contralateral to the site of stimulation averaged across individuals stimulated with two optical pulses before movement initiation. All error bars represent s.e.m.

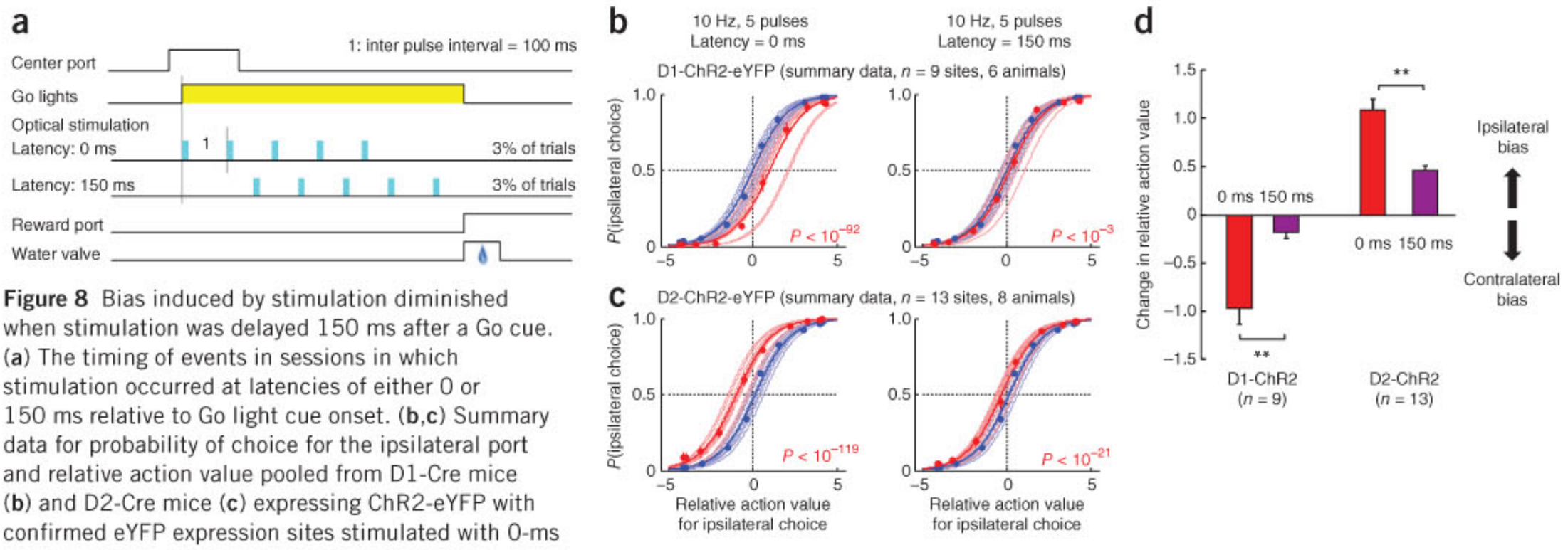


Figure 8 Bias induced by stimulation diminished when stimulation was delayed 150 ms after a Go cue. (**a**) The timing of events in sessions in which stimulation occurred at latencies of either 0 or 150 ms relative to Go light cue onset. (**b**,**c**) Summary data for probability of choice for the ipsilateral port and relative action value pooled from D1-Cre mice (**b**) and D2-Cre mice (**c**) expressing ChR2-eYFP with confirmed eYFP expression sites stimulated with 0-ms latency (left) or 150-ms latency (right) relative to Go light onset. A logistic regression fit for the no stimulation (thick blue line) and stimulation (thick red line) trial data was produced for data pooled from D1-Cre or D2-Cre mice. Curves representing the estimated probability of choice for given relative action values in individual animals are plotted in light blue (no stimulation) or light red (with stimulation). *P* values reported for *t* tests: H_0:$\beta_{stim} = 0$ (distance between thick red and blue lines). (**d**) Estimated change in relative action value for choosing the port ipsilateral versus contralateral to the side of stimulation averaged across individuals stimulated at 10 Hz with 0- and 150-ms latency in the same session. Estimates of relative action value change were derived from logistic regression analysis (Online Methods). Reported *n* refers to number of stimulation sites. $^{**}P < 0.005$, Wilcoxon signed-rank test. All error bars represent s.e.m.

© 2012 Nature America, Inc. All rights reserved.
npg

altered the latency to movement initiation, as measured by the withdrawal time in a manner that was dependent on the relative action value (**Fig. 6b**).

Qualitatively, the effects of stimulating D1R- and D2R-expressing neurons match existing accounts of the opposing functions of the direct and indirect pathway in the basal ganglia. This lateralized effect may be a result of the presence of dense ipsilateral descending connectivity from basal ganglia nuclei and the role that downstream efferent structures have in controlling contralateral movement[25–27]. The ability of D2R-expressing neurons to promote choices to the ipsilateral port is consistent with previous findings suggesting that action selection involves inhibition of competing alternatives through the indirect pathway in a manner that allows or even facilitates focal promotion of desired actions by the direct pathway[37]. The proposed occurrence of targeted inhibition from basal ganglia pallidal outputs has been suggested to coordinate agonist and antagonist musculature involved in limb and visuomotor movements in primates[38,39]. Competition between the opposing actions of orienting to the left and right may also be regulated in the downstream targets of the basal ganglia. The direction of the bias that we observed is largely consistent with previous reports in which D1R- and D2R-expressing neurons are selectively activated using optogenetic techniques[11], as well as the effect of pharmacological manipulations to the dorsal striatum in rodents[21].

Optogenetic stimulation did not induce a uniform effect across all trials, but was dependent on the mice's previous reward history. Stimulation induced a larger bias on choice when mice had greater variability in their responses following unrewarded trials (**Figs. 3–5**). However, striatal stimulation had a weak effect on choice behavior after recently rewarded trials, when mice were likely to return to a port at which water had just been delivered. Thus, the effect of stimulation could be overruled if the alternative response had a high incentive value after being recently rewarded.

The effect of optical stimulation was also time dependent. Stimulation was effective when limited to two 5-ms pulses in an epoch of the task before the mice's movement initiation (**Fig. 7**), and it became significantly weaker if delayed by 150 ms after trial initiation ($P < 0.005$; **Fig. 8**). Together, these data suggest that striatal activation may need to take place in a 'decision window' to alter the action selection of the animal. This is consistent with a number of recording studies in the striatum of rodents[22,24] and in primates[3,4,36]. Finally, outside of the task, transient stimulation at the maximal experimental level did not induce head or body orientation (**Supplementary Fig. 9**). The interaction of stimulation with reward history, the presence of an effective decision window and the lack of similar motor output outside of the task indicate that striatal stimulation did not dictate a motor action that deterministically affected choice. We conclude that striatal activity alone is not sufficient to drive the motor responses of mice in this task, but may also require the temporally coincident activity of other neural structures to orchestrate a complex process of action selection.

Given that the magnitude of the stimulation bias was dependent on whether previous actions had been rewarded, we hypothesized that the striatal activation may act similarly to the influence of rewards over choice. In the context of this task, animals are required to adaptively and flexibly switch their actions across blocks as a result of changing reward contingencies. At a theoretical level, this process of goal-directed action selection can be modeled as a dynamic comparison between the value of actions and bias for selection of the option of highest value[40]. More specifically, we conjectured that the additional activity induced by optical stimulation mimicked the effects resulting from a change in the value of competing actions.

To examine this directly, we estimated the value of actions on the basis of various reward and choice histories, assuming the softmax decision rule[4,31,41]. We then re-examined quantitative features of the bias introduced by striatal activation and found that stimulation mimicked a fixed additive shift in the value of the contralateral choice (**Figs. 4–6** and **Supplementary Figs. 5**, **6** and **8**) without altering sensitivity to reward (**Supplementary Fig. 10**). In this way, activation of D1R-expressing neurons mimics an increase in the value of the contralateral choice, and activation of D2R-expressing neurons mimics a decrease of value of the contralateral choice. The nonlinear features of the softmax rule were sufficient to explain the gross tendency for stimulation to have a larger effect in a range where responses are most variable (**Figs. 3–5**).

Although many descriptions of striatal function have focused on its role in action selection, others have suggested that the basal ganglia merely regulates the vigor of responses without altering which response is selected[19,20]. We found that, in addition to biasing the mice's action selection, striatal activity can also alter the vigor of responses by speeding up or slowing down the initiation of movements. We found that stimulation of D1R-expressing neurons reduced movement latencies in trials when the value of the contralateral port was greater and congruent with the direction of bias caused by stimulation (**Fig. 6b**). This is consistent with the canonical view that the direct pathway promotes movements[8,10]. However, latencies following stimulation of D1R-expressing neurons were markedly slower in trials in which the ipsilateral port was valued more highly. The slowed response is perhaps a result of the incongruence between bias caused by stimulation and intrinsic valuation in the action selection systems. Similarly, stimulation of D2R-expresing neurons slowed movements when the port contralateral to the site of stimulation was of greater value and it sped movements when the port ipsilateral to the stimulation site was of greater value. This data is consistent with the idea that action selection can be facilitated by the suppression of alternate reponses[38]. Our data suggest that the balance of activity in striatal populations reflecting the relative value of approaching each port affects the speed of movement initiation. These findings are consistent with evidence that striatal activity correlates with reaction times[1,36].

Our data is consistent with electrophysiological studies that have suggested that neural activity in the striatum more often represents the value of actions than pure motor variables[3,4]. Striatal activity has been associated with response bias for rewarded actions[1] and successful switching following action contingency reversals[36,42]. In these studies, neurons primarily encode a bias for contralateral responses when decisions are reported as saccades or locomotor approach[1,22,36]. Lesions of the dorsomedial striatum can also impair the learning of the contingencies between actions and their outcomes in various reversal tasks[7,43,44]. Taken together, these findings are consistent with our interpretation that striatal stimulation biases both the selection and vigor of actions on the basis of the value of choices.

We attempted to reproduce physiological conditions using optogenetics. The stimulation patterns that we used parallel data from awake *in vivo* striatal recordings in mice[23,35]. Other studies have also identified striatal activity in a similar decision point in a reward-based spatial task in rodents[24]. The bias induced by stimulation scaled over a range of frequency parameters, supporting the robustness of our results. However, questions still remain regarding whether optogenetic stimulation mimics physiological patterns *in vivo* given the large number of neurons that are synchronously recruited. In the D2-Cre mice, we also infected a small proportion of cholinergic neurons, which may contribute to our behavioral effects. An additional feature and caveat of our study is that our optical stimulation was delivered unilaterally in the context of a task in which competing responses are lateralized to the left and right. This is both a feature and caveat because we predict that our method will likely not cause behavioral bias in alternative task designs in which responses take other forms, such as up versus down. Lastly, our results do not exclude the possibility that D1R-expressing and D2R-expressing neurons may serve other roles outside of action selection. Numerous studies have found evidence that striatal activity also correlates with a process that evaluates the outcomes of actions in rodents[24,45] and in primates[4], and our data do not exclude these aspects of striatal function.

Striatal neurons receive massively convergent input from cortical and thalamic sources. This integration of diverse information may be used to generate a representation of action value that can be used to mediate goal-oriented action selection. In this framework, the updating of action values may correspond with dopamine-dependent plasticity of inputs into the striatum in concert with shifts in goal and reward-related activity from the cortex[41]. Striatal neurons may alter action selection by regulating tonic inhibition from the globus pallidus and the substantia nigra pars reticulata onto the brainstem and motor thalamus[46,47]. This provides a mechanism by which rewards modulate the responses of premotor structures for particular actions[46,48]. In this way, the cortico-basal ganglia system may instantiate the computations necessary for goal-oriented, highly flexible behavior[49,50].

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank G. Stuber, F. Ambroggi, D. Niell, K. Berger, C. Wu and H. Peckler for assistance with experiments, V. Kharazia for immunostaining and histological quantification, K. Deisseroth (Stanford University) for advice and reagents, and C. Niell for comments on the manuscript. This work was supported by the State of California, the US National Institutes of Health (1RC2NS069350; R01MH087542 and R01DA029150), and the P. Royer and K. Clayton Family.

AUTHOR CONTRIBUTIONS

L.-H.T., A.M.L. and L.W. designed the study. L.-H.T., A.M.L. and N.B. collected behavioral data. L.-H.T. and A.M.L. analyzed and modeled data. A.M.L. and N.B. processed tissue. L.-H.T., A.M.L., L.W. and A.B. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/nn.3188.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Lauwereyns, J., Watanabe, K., Coe, B. & Hikosaka, O. A neural correlate of response bias in monkey caudate nucleus. *Nature* **418**, 413–417 (2002).
2. Cromwell, H.C. & Schultz, W. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J. Neurophysiol.* **89**, 2823–2838 (2003).
3. Samejima, K., Ueda, Y., Doya, K. & Kimura, M. Representation of action-specific reward values in the striatum. *Science* **310**, 1337–1340 (2005).
4. Lau, B. & Glimcher, P.W. Value representations in the primate striatum during matching behavior. *Neuron* **58**, 451–463 (2008).
5. Barnes, T.D., Kubota, Y., Hu, D., Jin, D.Z. & Graybiel, A.M. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature* **437**, 1158–1161 (2005).
6. Pasupathy, A. & Miller, E.K. Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* **433**, 873–876 (2005).
7. Balleine, B.W., Delgado, M.R. & Hikosaka, O. The role of the dorsal striatum in reward and decision-making. *J. Neurosci.* **27**, 8161–8165 (2007).
8. Albin, R.L., Young, A.B. & Penney, J.B. The functional anatomy of basal ganglia disorders. *Trends Neurosci.* **12**, 366–375 (1989).
9. Gerfen, C.R. The neostriatal mosaic: multiple levels of compartmental organization. *Trends Neurosci.* **15**, 133–139 (1992).
10. DeLong, M.R. Primate models of movement disorders of basal ganglia origin. *Trends Neurosci.* **13**, 281–285 (1990).
11. Kravitz, A.V. *et al.* Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature* **466**, 622–626 (2010).
12. Lobo, M.K. & Nestler, E.J. The striatal balancing act in drug addiction: distinct roles of direct and indirect pathway medium spiny neurons. *Front. Neuroanat.* **5**, 41 (2011).
13. Kreitzer, A.C. & Berke, J.D. Investigating striatal function through cell type–specific manipulations. *Neuroscience* **198**, 19–26 (2011).
14. Kreitzer, A.C. & Malenka, R.C. Striatal plasticity and basal ganglia circuit function. *Neuron* **60**, 543–554 (2008).
15. Shuen, J.A., Chen, M., Gloss, B. & Calakos, N. Drd1a-tdTomato BAC transgenic mice for simultaneous visualization of medium spiny neurons in the direct and indirect pathways of the basal ganglia. *J. Neurosci.* **28**, 2681–2685 (2008).
16. Hikosaka, O., Nakamura, K. & Nakahara, H. Basal ganglia orient eyes to reward. *J. Neurophysiol.* **95**, 567–584 (2006).
17. Isoda, M. & Hikosaka, O. Cortico-basal ganglia mechanisms for overcoming innate, habitual and motivational behaviors. *Eur. J. Neurosci.* **33**, 2058–2069 (2011).

© 2012 Nature America, Inc. All rights reserved.
npg

18. Brainard, M.S. & Doupe, A.J. Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature* **404**, 762–766 (2000).
19. Turner, R.S. & Desmurget, M. Basal ganglia contributions to motor control: a vigorous tutor. *Curr. Opin. Neurobiol.* **20**, 704–716 (2010).
20. Desmurget, M. & Turner, R.S. Motor sequences and the basal ganglia: kinematics, not habits. *J. Neurosci.* **30**, 7685–7690 (2010).
21. Schwarting, R.K. & Huston, J.P. The unilateral 6-hydroxydopamine lesion model in behavioral brain research. Analysis of functional deficits, recovery and treatments. *Prog. Neurobiol.* **50**, 275–331 (1996).
22. Thorn, C.A., Atallah, H., Howe, M. & Graybiel, A.M. Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron* **66**, 781–795 (2010).
23. Kubota, Y. *et al.* Stable encoding of task structure coexists with flexible coding of task events in sensorimotor striatum. *J. Neurophysiol.* **102**, 2142–2160 (2009).
24. Kim, H., Sul, J.H., Huh, N., Lee, D. & Jung, M.W. Role of striatum in updating values of chosen actions. *J. Neurosci.* **29**, 14701–14712 (2009).
25. Felsen, G. & Mainen, Z.F. Neural substrates of sensory-guided locomotor decisions in the rat superior colliculus. *Neuron* **60**, 137–148 (2008).
26. Grillner, S., Hellgren, J., Menard, A., Saitoh, K. & Wikstrom, M.A. Mechanisms for selection of basic motor programs–roles for the striatum and pallidum. *Trends Neurosci.* **28**, 364–370 (2005).
27. Grillner, S., Wallen, P., Saitoh, K., Kozlov, A. & Robertson, B. Neural bases of goal-directed locomotion in vertebrates–an overview. *Brain Res. Rev.* **57**, 2–12 (2008).
28. Sugrue, L.P., Corrado, G.S. & Newsome, W.T. Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* **6**, 363–375 (2005).
29. Gold, J.I. & Shadlen, M.N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
30. Lau, B. & Glimcher, P.W. Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* **84**, 555–579 (2005).
31. Sutton, R.S. & Barto, A.G. *Reinforcement Learning: an Introduction* (MIT Press, Cambridge, Massachusetts, 1998).
32. Kawaguchi, Y., Wilson, C.J., Augood, S.J. & Emson, P.C. Striatal interneurones: chemical, physiological and morphological characterization. *Trends Neurosci.* **18**, 527–535 (1995).
33. Rymar, V.V., Sasseville, R., Luk, K.C. & Sadikot, A.F. Neurogenesis and stereological morphometry of calretinin-immunoreactive GABAergic interneurons of the neostriatum. *J. Comp. Neurol.* **469**, 325–339 (2004).
34. Salzman, C.D., Britten, K.H. & Newsome, W.T. Cortical microstimulation influences perceptual judgements of motion direction. *Nature* **346**, 174–177 (1990).
35. Jin, X. & Costa, R.M. Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* **466**, 457–462 (2010).
36. Watanabe, K. & Hikosaka, O. Immediate changes in anticipatory activity of caudate neurons associated with reversal of position-reward contingency. *J. Neurophysiol.* **94**, 1879–1887 (2005).
37. Redgrave, P., Prescott, T.J. & Gurney, K. The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience* **89**, 1009–1023 (1999).
38. Mink, J.W. The basal ganglia: focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* **50**, 381–425 (1996).
39. Jiang, H., Stein, B.E. & McHaffie, J.G. Opposing basal ganglia processes shape midbrain visuomotor activity bilaterally. *Nature* **423**, 982–986 (2003).
40. Rangel, A., Camerer, C. & Montague, P.R. A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* **9**, 545–556 (2008).
41. Schultz, W., Dayan, P. & Montague, P.R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
42. Kimchi, E.Y. & Laubach, M. The dorsomedial striatum reflects response bias during learning. *J. Neurosci.* **29**, 14891–14902 (2009).
43. Castane, A., Theobald, D.E. & Robbins, T.W. Selective lesions of the dorsomedial striatum impair serial spatial reversal learning in rats. *Behav. Brain Res.* **210**, 74–83 (2010).
44. Ragozzino, M.E. The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. *Ann. NY Acad. Sci.* **1121**, 355–375 (2007).
45. Kravitz, A.V., Tye, L.D. & Kreitzer, A.C. Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nat. Neurosci.* **15**, 816–818 (2012).
46. Ikeda, T. & Hikosaka, O. Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron* **39**, 693–700 (2003).
47. Lo, C.C. & Wang, X.J. Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat. Neurosci.* **9**, 956–963 (2006).
48. Kable, J.W. & Glimcher, P.W. The neurobiology of decision: consensus and controversy. *Neuron* **63**, 733–745 (2009).
49. Redgrave, P., Prescott, T.J. & Gurney, K. Is the short-latency dopamine response too short to signal reward error? *Trends Neurosci.* **22**, 146–151 (1999).
50. Hikosaka, O. & Isoda, M. Switching from automatic to controlled behavior: cortico-basal ganglia mechanisms. *Trends Cogn. Sci.* **14**, 154–161 (2010).

© 2012 Nature America, Inc. All rights reserved.

ONLINE METHODS

Animals. C57BL/6J BAC transgenic mice expressing Cre recombinase under the regulatory elements for the D1 and D2 receptor (D1-Cre and D2-Cre ER43) were obtained from Mutant Mouse Regional Resource and bred in our colony. All of the animals that we used were adults (25–30 g) group housed under a reverse 12-h light/dark cycle (light onset at 10:00 a.m.) until surgery. Mice were given food and water *ad libitum* before water deprivation in preparation for training. All procedures were approved by the Ernest Gallo Clinic and Research Center Animal Care and Use Committee.

Construct and virus preparation. Plasmids encoding the DNA sequences for pAAV-EF1α -DIO-ChR2(H124R)-eYFP or pAAV-EF1α-DIO-eYFP were obtained from K. Deisseroth (Stanford University). Amplification and purification of plasmids was performed using a standard plasmid maxiprep kit (Qiagen) and confirmed by sequencing. EF1α-DIO-CHR2(H124R)-eYFP and EF1α-DIO-eYFP cassettes were packaged in AAV vectors and serotyped with AAV5 coat proteins by the viral core at University of North Carolina. The final concentration was $1–2\times10^{12}$ viral particles per ml.

Implantable chronic optical fibers and optic cables construction. Optical stimulators were constructed according to published protocols[51]. Briefly, optical fibers were constructed by attaching a 200-μm, 0.37 NA optical fiber (Thor Labs) with epoxy resin into a metal ferrule that had previously been cut and scored. Fiber-ferrule units were then cut and polished. Only implants with efficiency greater than 70% and comparable efficiencies (±10%) were used. Optical-patch cables were constructed from 62.5-μm core diameter optic fiber (Thor labs) that were connected to a ferrule on one end and a fiber-optic connector for physical contact (FCPC) connector on the other end. Cables were covered in furcation tubing to protect the fiber and to prevent light from escaping through the optic-patch cord. The ferrule at the end of the optic patch cord was fitted with a zirconium sleeve to interface with the chronic implant. The FCPC connector was coupled to a 473-nm diode-pumped solid-state laser (200 mW). The laser driver current was adjusted to yield 20-mW output from the patch cable.

Stereotaxic AAV injection and optical implant surgery. Animals were anesthetized with either ketamine (150 mg per kg of body weight) and xylazine (50 mg per kg) or 2% isofluorane (vol/vol) gas anesthesia. Animals were placed in a stereotaxic frame and 26-gauge microinjection needles were inserted through a burr hole bilaterally into the dorsomedial striatum (coordinates from bregma: 0.75 mm anterior, ±1.5 mm medial-lateral, −3.0 ventral) of D1-Cre and D2-Cre mice to deliver 1.0 μl of either AAV-EF1-DIO-ChR2(H124R)-eYFP or AAV-EF1-DIO-eYFP (~1,012 IU ml^{-1}). Injections were performed using a 1-μl Hamilton syringe through a hydraulic pump (Harvard Instruments) and took place over 10 min followed by 10 min of recovery. The length of the optic fiber protruding from the implant was cut to be 2 mm. The tip of the fiber optic from the implant was then inserted through the same burr hole as was used previously for the virus injection and was lowered 2 mm ventral to the dura. The implant was cemented to the skull using dental cement. Mice were monitored until recovery from surgery and then returned to their home cage where they were housed individually.

Anesthetized *in vivo* optrode recording. Recordings under conditions of anesthesia were made with a custom silicon probe (model A1x16-5mm50-413, NeuroNexus Technologies) that was attached with a 200-μm core diameter optic fiber connected to a FCPC connector with epoxy resin. Mice were placed under ketamine (150 mg per kg) and xylazine anesthesia (50 mg per kg, intraperitoneal) 1 month after injection of AAV-EF1-DIO-ChR2(H124R)-eYFP into the dorsomedial striatum. A craniotomy was performed above the injection site and the optrode was lowered. Data were acquired using commercial systems (Plexon). Optical stimulation was delivered during recording as continuous trains of stimulation to assess the fidelity of optically induced firing as well as using the same stimulation protocols that we used during our behavioral task. After each recording, the probe was moved to a new recording tract in the same animal. Following recordings, animals were killed for assessment of track location and viral expression. Single units were identified with principal component analysis (Offline Sorter, Plexon). The data were then imported into Matlab (MathWorks) for subsequent analysis.

Probabilistic switching (two-alternative spatial choice) task. Mice were trained on a two-alternative spatial choice task in which the location of a water reward was periodically switched at random intervals. The initiation port was located in the middle of one wall, and two choice ports were located 63.5 mm to the left and right of the initiation port (center to center; **Fig. 1a**). An infrared photodiode/phototransistor pair was placed on either side of the port to report the times of port entry and exit (Island Motion). The water valves (Neptune Research) were calibrated to deliver a volume of water (2 μl) for rewarded choices.

Mice initiated each trial by entering the center port, triggering Go lights instructing animals that water was potentially available. Mice then chose a left or right peripheral port for water reinforcement (**Fig. 1a**). Only one peripheral port was rewarded at a time and, on 25% of trials, neither port was rewarded. The length of trial blocks was dependent on the number of rewards obtained in each block, and this number of rewards was randomly set between 7–23 rewards. After the set amount of rewards was obtained, the rewarded side was switched to the opposite port. This structure prevented the subjects from predicting the timing of the block switch.

Logistic regression analysis of behavior choices. The contribution of past rewards or lack of rewards on the subject's current choice was analyzed on a trial-by-trial basis using the following logistic regression model[30]

$$\log\left(\frac{P_L(i)}{1-P_L(i)}\right)=\sum_{j=1}^{n}\beta_j^{\text{Reward}}(Y_L(i-j)-Y_R(i-j))+\sum_{j=1}^{n}\beta_j^{\text{No Reward}}(N_L(i-j)-N_R(i-j))+\beta_0$$

where $P_L(i)$ is the probability of selecting the left port in the i-th trial. The variables $Y_L(i)$ or $Y_R(i)$ represent whether a reward was delivered (1 or 0) at the left or right port in the i-th trial, respectively. $N_L(i)$ or $N_R(i)$ represent the lack of reward (1 or otherwise 0) at the chosen left or right port in the i-th trial, respectively. n indicates the number of past trials that were included in the model (n = 5). The regression coefficients β_j^{Reward} and $\beta_j^{\text{No Reward}}$ represent the contribution of past rewards and lack of rewards, respectively, and β_0 indicates the intrinsic bias of the animal.

We modeled the contribution of optical stimulation on the subject's current choice as a dummy variable characterized by the β_{stim} coefficient

$$\log\left(\frac{P_L(i)}{1-P_L(i)}\right)=\sum_{j=1}^{n}\beta_j^{\text{Reward}}(Y_L(i-j)-Y_R(i-j))+\sum_{j=1}^{n}\beta_j^{\text{No Reward}}(N_L(i-j)-N_R(i-j))+\beta_0+\beta_{\text{stim}}\cdot X_{\text{stim}}(i)$$

The variables $X_{\text{stim}}(i)$ represents whether stimulation was delivered (1) or not (0) in the i-th trial. These estimated coefficients are the shifts in action value characterized in **Figure 6**.

Optical stimulation in the behavior task. Optical stimulation was delivered at the start of the Go signal. Stimulation was delivered at 5, 10 and 20 Hz for 500 ms with the frequency pseudo-randomly chosen before each session. Stimulation trials occurred in 6% of total trials. Stimulation sessions were performed every other day, interspersed by training sessions. The hemisphere that was stimulated was alternated across stimulation sessions. The infrequent occurrence of stimulation trials was to prevent any plastic or compensating adaptations from occurring during the course of a session, and the relatively long interval of days between stimulation sessions was to prevent any systemic biases in responses from arising across stimulation sessions.

To determine whether stimulation had to occur in a specific time window relative to movement initiation, we performed additional experiments. First, a subset of mice that underwent stimulation at 5, 10 and 20 Hz underwent an additional set of stimulation sessions in which two types of stimulation trials were present, each for 3% of trials. In one set of stimulation trials, 10-Hz stimulation was delivered at the same time that the Go light appeared, 15 ms after initiation

© 2012 Nature America, Inc. All rights reserved. npg

of a trial by a center poke (as in earlier experiments). In a second set of stimulation trials, 10-Hz stimulation was delayed by 150 ms from the appearance of the Go light. This 150-ms delayed stimulation time corresponds to a point at which movement was initiated in >90% of trials.

To look at the effects of stimulation exclusively before the onset of the Go cue, we also trained new cohort of mice (n = 2 D1-Cre and 4 D2-Cre mice) on a variant of the stimulation experiment in which two optical pulses (interpulse interval = 50 ms) were delivered before and coincident with a Go sound which was delayed by 70–100 ms after initiation of a trial by a center poke (**Fig. 7a**). This was to confine stimulation to an epoch of the task before the initiation of the animal's movement as determined by the time of withdrawal from the center poke. Stimulation trials occurred in 6% of total trials. Stimulation sessions were performed every other day interspersed by training sessions.

Optical stimulation outside of the switching task context. To measure the effect of transient stimulation on locomotor behavior outside the context of the task, we placed mice in a 10-cm-diameter cylindrical environment (**Supplementary Fig. 9**). Optical stimulation was delivered at 20 Hz for 500 ms and repeated every 60 s. Each subject received 50 stimulations per stimulation site. The body orientation at stimulation onset, and 0.5 and 1 s before and after stimulation onset (five time points) was measured from video recording frames by a blind observer using custom software.

To measure the effect of prolonged stimulation on locomotor behavior outside the context of the task, we placed the mice in an open field arena. Baseline locomotion was measured over a period of 60 s. Optical stimulation was then delivered at 5, 10 or 20 Hz for 60 s, and was followed by a 60-s recovery period (**Supplementary Fig. 9**). Rotation was quantified using Anymaze software (Stoelting) and summed for each 60-s period (baseline, stimulation and recovery).

Statistical analysis. Fisher's exact test was used to determine whether the probability of choices with and without stimulation were significantly different. Logistic regression was used to fit data for trials with different reward histories with and without stimulation. t test was used to determine whether the change in relative action value caused by striatal activation (β_{stim}) was significantly different from zero for a given stimulation condition. Wilcoxon signed-rank and rank-sum tests were used to determine whether the changes in relative action value between different stimulation conditions or between groups of subject expressing ChR2-eYFP and eYFP were significantly different. Standard errors for probabilities of choice were calculated from binomial statistics. α was set at 0.05.

Histology and reconstruction of optical stimulation sites. Viral expression of ChR2-eYFP and eYFP was confirmed by histology after stimulation experiments (**Supplementary Fig. 2a**,**b**). Paraformaldehyde-fixed coronal sections were stained to identify cell bodies (Neurotrace) alongside AAV-driven expression of YFP. In the cases reported, fiber implant tracks could be identified and were found to be located in or directly above the medial half of the dorsal striatum (**Supplementary Fig. 2c**).

To determine whether ChR2-eYFP was expressed in MSNs as well as cholinergic neurons of the striatum, we permeablized sections (coronal, 50 µm) in 50% alcohol for 10 min, rinsed them in phosphate-buffered saline (PBS), blocked them in 10% normal donkey serum (vol/vol) for 30 min and incubated them for 48 h in a mixture of primary antibodies: rabbit polyclonal antibody to GFP (for YFP, 1:10,000, Abcam, ab290), goat polyclonal antibody to ChAT (1:500, Millipore, AB144p) and mouse monoclonal antibody to Kv2.1 (1:200; University of California Davis and National Institute of Mental Health NeuroMab, 75-014) at 4 °C[52]. Sections were then rinsed in PBS and incubated in 2% normal donkey serum for 10 min, then incubated for 4–6 h in mixture of secondary antibodies (all made in donkey): Alexa Fluor488–conjugated antibody to rabbit (1:300, Life Technologies, A-11008), CF647 antibody to goat (1:300, Biotium, 20048) and CF555 antibody to mouse (1:300, Biotium, 20037). Sections were rinsed in PBS, mounted and coverslipped with VECTASHIELD mounting media (Vector Laboratories). Multi-channel images of ChAT-positive cells were acquired with a Zeiss LSM 510 META laser confocal microscope (Zeiss), using a 63×/1.4 NA PlanApo objective, 488-, 543- and 633-nm excitation lines and factory recommended detector settings.

51. Sparta, D.R. *et al.* Construction of implantable optical fibers for long-term optogenetic manipulation of neural circuits. *Nat. Protoc.* **7**, 12–23 (2012).
52. Ariano, M.A. *et al.* Striatal potassium channel dysfunction in Huntington's disease transgenic mice. *J. Neurophysiol.* **93**, 2565–2574 (2005).

© 2012 Nature America, Inc. All rights reserved.
npg