LETTER

doi:10.1038/nature17639

Opponent and bidirectional control of movement velocity in the basal ganglia

Eric A. Yttri[1] & Joshua T. Dudman[1]

For goal-directed behaviour it is critical that we can both select the appropriate action and learn to modify the underlying movements (for example, the pitch of a note or velocity of a reach) to improve outcomes. The basal ganglia are a critical nexus where circuits necessary for the production of behaviour, such as the neocortex and thalamus, are integrated with reward signalling[1] to reinforce successful, purposive actions[2]. The dorsal striatum, a major input structure of basal ganglia, is composed of two opponent pathways, direct and indirect, thought to select actions that elicit positive outcomes and suppress actions that do not, respectively[3,4]. Activity-dependent plasticity modulated by reward is thought to be sufficient for selecting actions in the striatum[5,6]. Although perturbations of basal ganglia function produce profound changes in movement[7], it remains unknown whether activity-dependent plasticity is sufficient to produce learned changes in movement kinematics, such as velocity. Here we use cell-type-specific stimulation in mice delivered in closed loop during movement to demonstrate that activity in either the direct or indirect pathway is sufficient to produce specific and sustained increases or decreases in velocity, without affecting action selection or motivation. These behavioural changes were a form of learning that accumulated over trials, persisted after the cessation of stimulation, and were abolished in the presence of dopamine antagonists. Our results reveal that the direct and indirect pathways can each bidirectionally control movement velocity, demonstrating unprecedented specificity and flexibility in the control of volition by the basal ganglia.

Purposive action requires selection of a goal (for example, go left) and execution parameters (for example, how fast to go). For example, in bird song selection of both discrete, sequential actions (syllables) as well as the pitch can be controlled by reinforcement in cortico-basal ganglia pathways[8,9]. The striatum is a major input nucleus in basal ganglia[1] and the direct and indirect pathway are primarily composed of two molecularly distinct[10] populations of medium spiny projection neurons (MSNs): direct striatonigral (dMSN) and indirect striatopallidal (iMSN) neurons. Sustained activation of dMSNs increases movement, whereas sustained activation of iMSNs reduces movement[11]. As a result, the balance of activity-dependent plasticity at cortical synapses onto dMSNs and iMSNs is thought to underlie the selection of successful goal-directed actions[3,5,12]. While it is known that stimulation of direct pathway neurons can support self-stimulation[13] and bias concomitant choice behaviour[14], there is little direct evidence that MSN activity is sufficient to produce persistent, specific changes in subsequent actions.

We trained mice expressing channelrhodopsin-2 (ChR2) in either dMSNs or iMSNs to perform self-paced, bimanual forelimb movements while head-fixed to obtain a water reward (Fig. 1a; Supplementary Videos 1–4). These single, discrete movements provided a reliable, repeatable behaviour from which we could extract movement parameters (Fig. 1b–d). To determine whether activity in MSNs during a voluntary action is sufficient to control movement parameters, we administered closed-loop photostimulation to the dorsomedial striatum during the fastest third of movements. Stimulation intensity was adjusted to be subthreshold for direct effects on movement, but sufficient to modulate activity to a similar magnitude as endogenous modulation of striatal activity during limb movements (Fig. 1e, f and Extended Data Fig. 1). Stimulation onset occurred within 15 ms of the beginning of a movement and persisted for 450 ms (comparable to movement duration; 505 ms; Fig. 1c, d). To maintain motivation to perform the task independent of stimulation, all movements that crossed the criterion amplitude threshold elicited a delayed liquid reward.

We first asked whether photostimulation of dMSNs during the fastest third of movements could alter the velocity of subsequent movements. Indeed, brief dMSN stimulation was sufficient to produce a significant increase in the peak velocity ($1.4\,\mathrm{cm\,s^{-1}}$ increase from $29.7\,\mathrm{cm\,s^{-1}}$; $P < 7 \times 10^{-5}$; Fig. 2 and Extended Data Fig. 2) of all limb movements. Other movement parameters that were not targeted for closed-loop stimulation such as the amplitude, duration, and tortuosity remained unaltered ($P > 0.7$). This is despite the fact that mice were capable of rapidly adjusting movement parameters to changing reward contingencies (Extended Data Fig. 3). By contrast, iMSNs stimulation during the fastest third of limb movements produced a significant reduction in peak velocity ($-1.1\,\mathrm{cm\,s^{-1}}$; $P < 7 \times 10^{-4}$). The effect of iMSN stimulation had its maximal effect on velocity; movement duration and tortuosity were not significantly altered ($P > 0.3$). Prolonged tonic activation of dMSNs tends to be pro-kinetic in that it evokes generalized increases in voluntary movement ('response vigor'[15]), whereas tonic activation of iMSNs tends to decrease voluntary movement[11]. However, we found that neither brief dMSN nor iMSN stimulation during the fastest movements produced a change in the rate of trial initiation or the rate of licking during reward anticipation and consumption (Fig. 2b and Extended Data Table 1). These results thus demonstrate that closed-loop activation of MSNs is sufficient to produce sustained changes in movement parameters without generalized changes in movement or motivation.

We next examined the effect of successive stimulation on limb movement velocity. If stimulation merely altered the velocity of the current movement, then repeated stimulation should produce an immediate, but constant effect. However, stimulation drove a steady change in velocity that accumulated over the course of several trials (Fig. 2d), apparent in individual sessions (Fig. 2a and Extended Data Fig. 2). We also found that unstimulated movements (trials with subthreshold velocity) were changed to a similar extent. dMSN stimulation produced a $0.9\,\mathrm{cm\,s^{-1}}$ increase ($P = 0.014$) in velocity on unstimulated movements whereas iMSN stimulation produced a $-1.0\,\mathrm{cm\,s^{-1}}$ decrease ($P = 0.001$) in the velocity of unstimulated movements. Moreover, there was no change in variance of the distribution of velocities throughout the session (F-test, $P > 0.5$ for both groups, Extended Data Fig. 4). Together these observations argue that selective stimulation produced a gradual, accumulating shift in the entire distribution of velocities, rather than a change restricted to the stimulated subset (for example, making only fast, stimulated limb movements even faster). These cumulative changes in behaviour may be contrasted with previous reports of optogenetic stimulation that have observed transient effects

[1]Janelia Research Campus, Howard Hughes Medical Institute, 19700 Helix Drive, Ashburn, Virginia 20147, USA.

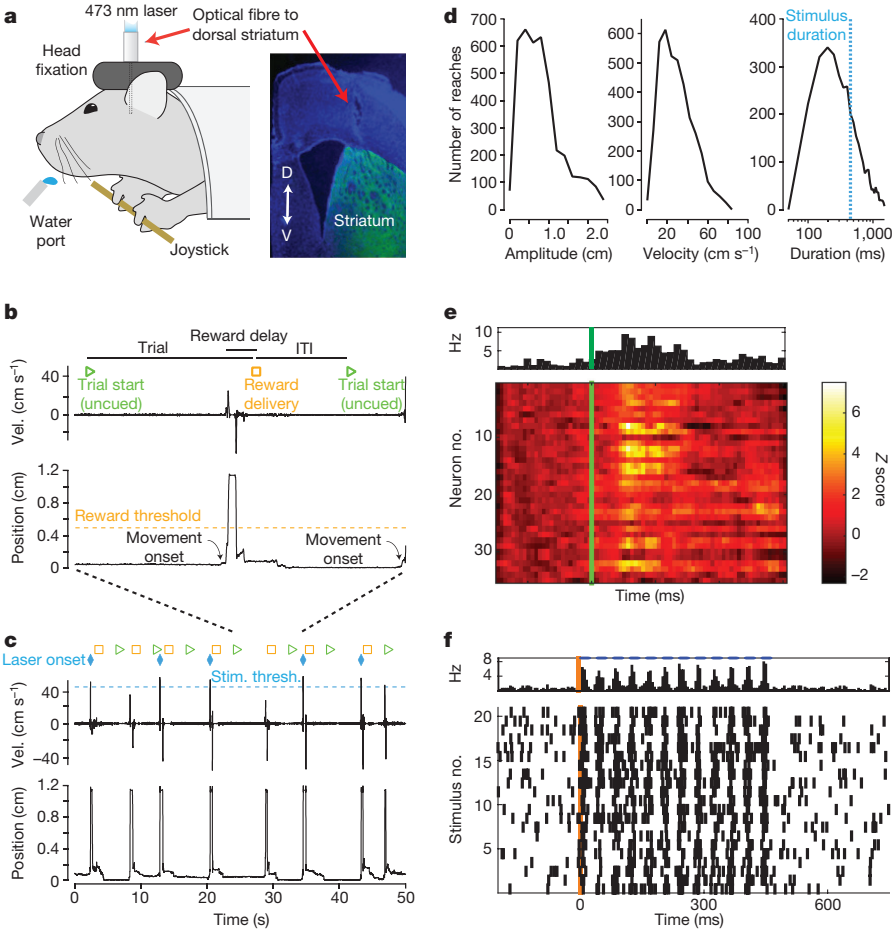


Figure 1 | Paradigm for closed-loop stimulation in dorsomedial striatum. a, Mice were head-fixed in front of a side-mounted joystick and a water port. Optical fibres were chronically implanted. Tips were positioned in the dorsomedial striatum and coupled to a 473 nm laser. Insert shows fibre position; D, dorsal; V, ventral. Fluorescent image is from iMSN neurons expressing ChR2–YFP. **b**, To receive liquid reward, mice made forelimb movements with the joystick (either a pull or push) past the criterion distance. Reward delivered 1 s after threshold crossing. Inter-trial intervals (ITIs) were 3 s (uncued). **c**, Instantaneous velocity and position of joystick for seven trials (green triangle indicates trial start). Velocity threshold for closed-loop optical stimulation and time of stimulation onset indicated by the blue dashed line and diamonds, respectively. Yellow squares indicate reward. **d**, Histograms of movement amplitude, peak velocity, and duration for all eight mice (45 sham sessions). **e**, Average response (z-scored change from baseline firing rate) of striatal units aligned to movement onset from a single session. Population average shown above. **f**, Raster plot of population activity during photostimulation from a single session.

confined to the stimulated trial[13,14] or concomitant with stimulus delivery[11].

If stimulation of the fastest movements produces a persistent change in the selection of movement parameters, the change should persist without stimulation. We plotted the velocity of movements made during the block of trials immediately following the stimulation block. In this recovery block, no stimulation was delivered. We found that stimulation-induced changes in the distribution of velocities persisted for tens of trials before gradually returning to the pre-stimulation baseline during the recovery block (Fig. 2a, d; paired t-test, $P = 0.64$ and 0.90, dMSN and iMSN, respectively). Importantly, this return to the pre-stimulation distribution had a similar time course whether it required a decrease or increase in the mean velocity following dMSN or iMSN stimulation, respectively.

We have shown that dMSN and iMSN have opponent roles in the reinforcement of movement parameters with unprecedented specificity. Stimulation induced changes are signed—dMSN stimulation increases a kinematic parameter of movement (velocity) whereas iMSN stimulation decreases the same property. However, there is a limitation to this simple opponency for learning: reinforcement should, in principle, alter behaviour so as to increase a reinforcing outcome regardless of the sign of the behavioural change[16]. It should be possible, for example, to learn to move more slowly to obtain more reward. Our data are also consistent with an alternative possibility: dMSN stimulation may be sufficient to drive changes towards movements that elicit stimulation independent of the sign (positive/negative) of the change. To distinguish between these alternatives, we stimulated MSNs during the slowest, rather than the fastest, third of limb movements. This stimulation protocol produced the opposite effects for both dMSN and iMSN stimulation (Fig. 2e, f). Under these conditions, stimulation of dMSN was sufficient to produce a cumulative decrease in velocity ($-1.1\ \mathrm{cm\ s^{-1}}$, $P = 0.008$). Conversely, iMSN stimulation produced an accumulating increase in velocity ($0.9\ \mathrm{cm\ s^{-1}}$, $P = 0.012$). Thus, the direct and indirect pathways of the basal ganglia are opponent pathways that are also sufficient for bidirectional changes in a continuous parameter that specifies purposive movement.

Models of the basal ganglia in which reinforcement learning acts to select amongst mutually exclusive actions can explain a broad array of empirical results in the learning literature[12]. However, such models cannot readily account for reinforcement acting on a continuous parameter of movement such as velocity[12] (see Supplementary Discussion). By contrast, a learning rule in which closed-loop stimulation provides a pathway-specific, signed learning signal that determines the mean of the velocity distribution could reproduce our data (Fig. 3a and Methods). Owing to the bidirectional behavioural changes observed, this learning rule makes a specific prediction: stimulation on every trial

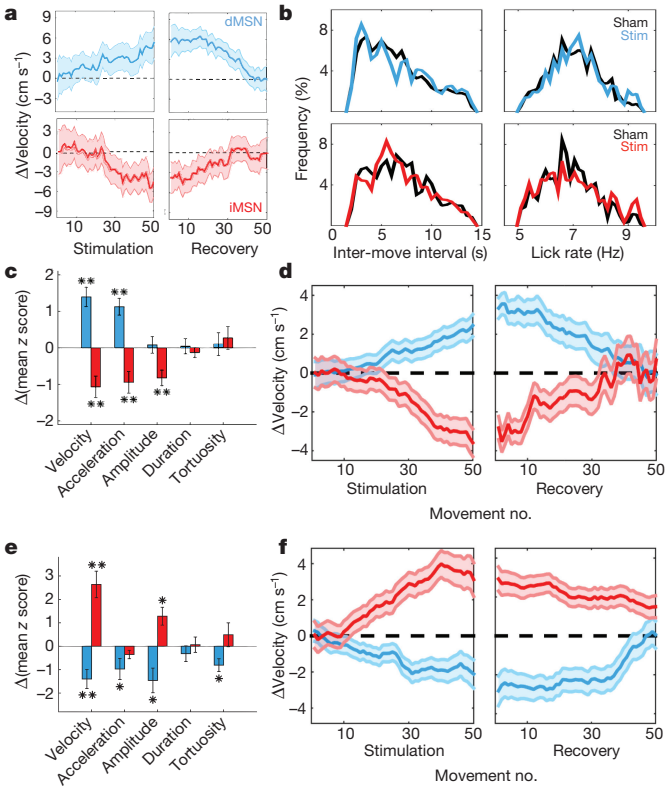


Figure 2 | Closed-loop stimulation produces opponent, bidirectional control of movement velocity. a, Difference in peak velocity between stimulation and sham session (Δvelocity) for sessions in which dMSN (upper, blue throughout) or iMSN (lower, red throughout) were stimulated on the fastest third of 50 trials during stimulation and no stimulus was delivered during recovery. Example session shown. **b**, Histograms of inter-movement-interval (left) and lick rate during reward consumption (right) for sham (black; 25 sessions in 4 dMSN mice, 20 sessions in 4 iMSN mice) and stimulation (coloured; 22 sessions in dMSN mice, 26 sessions in iMSN mice) sessions. **c**, Population average of change in movement parameters when fastest third of reaches were stimulated. **d**, Population average Δvelocity as a function of movement (trial) number when fastest third of reaches were stimulated. **e**, **f**, Same as **c**, **d**, but for sessions in which stimulation occurred on the slowest third of movements. $*P<0.05$; $**P<0.005$, two tailed t-test. Shaded area indicates standard error of the mean. Data are from 16 stimulation and 18 sham sessions in the same 4 dMSN mice, 20 stimulation, 16 sham sessions in the same 4 iMSN mice.

or at random throughout a session should produce no net change in velocity. Consistent with this prediction, each simulated stimulation protocol failed to produce a detectable change in movement velocity ($P>0.2$ for all conditions, Fig. 3 and Extended Data Fig. 5).

As formulated, this learning rule would induce a persistent change in velocity following stimulation. Extinction formulated as a fixed decay in synaptic weight[12] would not produce symmetric recovery as observed (Fig. 2 and Supplementary Discussion). To account for this feature of the data, we assumed a homeostatic component and refer to the rule as 'mean shift with homeostasis' (MeSH). Thus, the mean velocity of movement is determined by a set point that opposes learned changes and restores velocity towards baseline during recovery. When incorporated into the learning rule, we found that simulations closely reproduced the data during stimulation and recovery epochs. Selective stimulation that biased the reward-based feedback steadily drove velocity towards (dMSN) or away (iMSN) from the threshold that elicited stimulation (Fig. 3a). Upon cessation of stimulation, the homeostatic component drove recovery to the pre-stimulation baseline within 50 trials (Fig. 3).

MeSH assumes an explicit interaction between reward signalling, putatively carried by midbrain dopaminergic inputs to the dorsal striatum[6], and exogenous activation of MSNs. In contrast to this prediction, previous work has suggested that intracranial self-stimulation supported by striatal stimulation is independent of dopamine receptor activation[13]. However, the movement-related activity of striatal populations and our use of brief 450 ms stimulation both differ from the sustained, post-movement stimulation used previously. Thus, we next asked if dopamine was necessary for stimulation to elicit changes in movement velocity. We found that a low concentration of D1 and D2 receptor antagonists (SCH23390, 0.02 mg kg^{-1} and sulpiride, 25 mg kg^{-1}) injected before a behavioural session[13] eliminated persistent changes in velocity following closed-loop stimulation of either dMSN or iMSN (Fig. 3b, d) while largely sparing normal task performance (Fig. 3c, all dMSN parameters $P>0.2$; all iMSN parameters $P>0.3$). Dopamine antagonists significantly reduced the magnitude of the stimulation effect for both dMSN stimulation (92% decrease; 1.3 cm s^{-1}, $P<1\times10^{-8}$) and iMSN stimulation (109% increase; 1.2 cm s^{-1}, $P=1.5\times10^{-7}$).

Our results suggest that stimulation engages a dopamine-dependent, bidirectional plasticity in striatal MSNs. While a learning rule that acts directly on a parameter specifying the velocity distribution can account for our behavioural results, MeSH is abstract and it is unclear how it could be implemented in corticostriatal circuits critical for goal-directed, instrumental behaviour[17]. Thus, we implemented a simplified corticostriatal circuit model with the following key features (Extended Data Fig. 6).

While an action-value formulation implies that movements of different velocities are represented as a set of distinguishable neural states, MeSH implies a continuous representation of speed. There is little empirical evidence for representation of specific velocity ranges in cortical activity[18]. By contrast, there is substantial evidence that cortical[19,20] and striatal[21] representations of forelimb movements are monotonically tuned to speed. Consistent with the anatomy of the corticostriatal pathway[1], we assume that the speed of a movement is determined by both cortical and basal ganglia output (Extended Data Fig. 6). In combination with monotonic tuning this implies that the mean movement velocity is proportional to the average weight of corticostriatal synapses (Methods).

Dopamine and spike-timing dependent plasticity (STDP) has been described in the striatum[22,23]. STDP can result in bidirectional plasticity with the balance of potentiation and depression adapted to the range of population activity (that is, BCM-type plasticity[24]) in the presence of variable spike trains[25]. We assume a balance of potentiation and depression such that movements made at the average speed produce no net change. We posit that photostimulation enhances both potentiation and depression consistent with our observation that selective stimulation produces biased changes whereas non-selective stimulation produces no net change (Figs 2 and 3; Supplementary Discussion). Balanced synaptic plasticity that is enhanced by stimulation is sufficient to produce bidirectional changes in the average corticostriatal synapse weight during selective photostimulation. When incorporated into a corticostriatal circuit model, this plasticity rule produces opponent, bidirectional, and symmetric changes in movement speed (Fig. 4a, b and Extended Data Fig. 6).

Finally, we sought to validate our circuit model with electrophysiological recordings from dorsomedial striatum. Individual units (putative MSNs) recorded from mice performing the task were monotonically tuned to movement velocity (Fig. 4c), confirming previous observations[21,26]. An important feature of our behavioural results was that closed-loop stimulation does not simply reinforce stimulated movements, but rather produces a change in the mean velocity. Thus, changes in striatal activity should be apparent even for unstimulated movements. Specifically, the slope relating firing rate to velocity should be changed (Fig. 4b). Moreover, if photostimulation is necessary to alter plasticity in the recorded neuron then slope changes should correlate with photostimulation (Fig. 4b). To test this prediction we analysed a population of striatal units ($n=35$) during closed-loop stimulation of

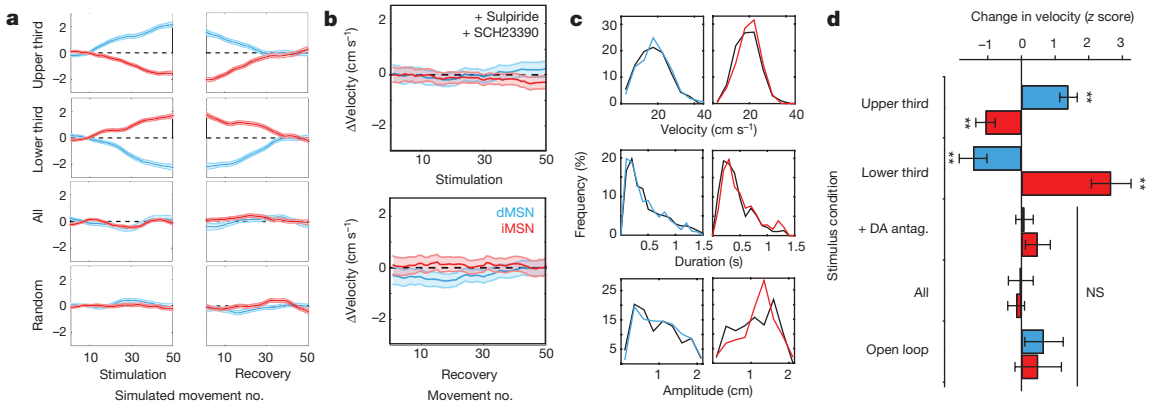


Figure 3 | Changes in velocity are consistent with dopamine-dependent reinforcement learning. a, Simulation of MeSH learning rule (see text for details). Change in average peak velocity (arbitrary units) as a function of trial number for dMSN-stimulation (blue) and iMSN-stimulation (red) simulations. **b**, ΔVelocity as a function of trial for stimulation of dMSN (blue) and iMSN (red) on the fastest third of 50 stimulation trials in the presence of dopamine receptor antagonists. Data from 14 stimulation and 11 sham dMSN sessions; 8 stimulation and 9 iMSN sham sessions. **c**, Movement parameter distributions for control sessions (black) and sessions following dopamine antagonist administration (coloured). **d**, Summary of the changes in velocity for experiments as indicated for dMSN (blue) and iMSN (red) stimulation sessions as defined in text. DA antag., dopamine receptor antagonist. Shaded area and error bars indicate standard error of the mean. $^{**}P < 0.005$, two tailed *t*-test; NS, not significant.

dMSNs on the fastest third of movements. Consistent with the model predictions, we observed an increase in the slope of the velocity tuning associated with an increase in the average velocity (Fig. 4d). Changes in tuning tended to be most prominent on units responsive to photostimulation (putative dMSNs; Fig. 4d) whereas neurons with weak stimulation responses (putative iMSNs or distant dMSNs) tended to show decreases in apparent tuning slope (Fig. 4d).

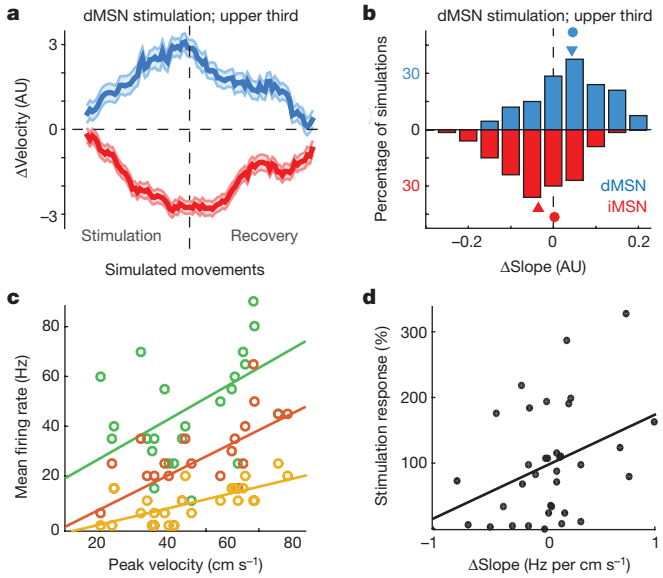


Figure 4 | Corticostriatal circuit model implements MeSH rule and experimental validation. a, Change in movement velocity (arbitrary units, AU) as a function of number of simulated trials (shaded area indicates standard deviation; $n = 100$) for simulations in which dMSN (blue) or iMSN (red) were stimulated during fastest ~30% of reaches. **b**, Histogram of the change in the average slope between dMSN (blue) or iMSN (red) activity and movement velocity for simulations with dMSN stimulation. Triangles indicate mean. Circles indicate mean change in weight of corticostriatal synapse (see text for details). **c**, Mean firing rate during movement (0–505 ms after onset) is plotted as a function of peak movement velocity with overlaid linear fit ('tuning slope') for 3 of 35 example striatal units. **d**, Stimulation response (per cent change in baseline firing rate during photostimulation) as a function of the change in the tuning slope (Δslope) for unstimulated trials (subthreshold velocity) of sessions in which dMSN were stimulated on the fastest third of movements. Linear fit overlaid ($r = 0.47$).

Here we provide the first demonstration, to our knowledge, that the innate biases in the direct and indirect pathway to increase or decrease the frequency of movement, respectively, do not extend to fixed biases in the control of movement parameters. The direct and indirect pathways engage opponent, activity-dependent plasticity mechanisms that can produce sustained biases in future behaviour. Each pathway is sufficient to produce bidirectional changes and, to some extent, is innervated by distinct cortical populations[27], suggesting that bidirectional control by each pathway could allow for adaptive control of goal-directed actions in different contexts or under different demands. These data argue that phasic activity in the striatum during specific movements is sufficient to selectively reinforce changes in a movement parameter independent of a generalized change in motivation consistent with a role for dopamine-dependent signalling in dorsal striatum in the control of movement vigour[21,26,28].

Our results reveal a bidirectional control of behaviour by MSNs that may be contrasted with the recent observation that self-stimulation supported by MSNs is unidirectional and dopamine-independent[13]. The differences between the findings may reflect the different experimental paradigms. Selectively biasing striatal activity in the context of a reward-based operant task could engage mechanisms distinct from the reinforcing properties of photostimulation itself. In the latter case strong stimulation may be sufficient to replace dopaminergic inputs or support self-stimulation in a dopamine-independent manner[29]. In addition, we observed a symmetric recovery to baseline following cessation of stimulation that is also distinct from the differential extinction of self-stimulation[13]. However, a recent modelling study argued that apparent differences in extinction are consistent with equivalent learning rates following dMSN and iMSN stimulation[12] consistent with our observation of opponent, but symmetric effects.

Here we proposed a circuit implementation by which a continuous parameter defining a purposive movement can be selectively reinforced by a stimulation-dependent enhancement of bidirectional synaptic plasticity. Importantly, it has been shown that striatal neurons are capable of bidirectional synaptic plasticity[22,23]; however, plasticity is mediated by distinct signalling events in the two populations[23]. Resolving the roles played by the intersection of these different cellular and circuit factors that govern bidirectional plasticity will be critical to understand the role of dopamine in instrumental learning. In addition to kinematic parameters of movement, other aspects of reinforcement learning

are governed by continuous parameters such as rates[30] or value[6]. The circuit implementation we propose could provide a general mechanism by which activity-dependent plasticity in striatum produces learned changes in continuous parameters with monotonic representations in neural activity.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 May 2015; accepted 3 March 2016.
Published online 2 May 2016.

1. Dudman, J. T. & Gerfen, C. R. in *The Rat Nervous System* (ed. G Paxinos) Ch. 17, (Elsevier, 2015).
2. Balleine, B. W., Liljeholm, M. & Ostlund, S. B. The integrative function of the basal ganglia in instrumental conditioning. *Behav. Brain Res.* **199,** 43–52 (2009).
3. Mink, J. W. The basal ganglia: focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* **50,** 381–425 (1996).
4. Frank, M. J. Computational models of motivated action selection in corticostriatal circuits. *Curr. Opin. Neurobiol.* **21,** 381–386 (2011).
5. Gurney, K. N., Humphries, M. D. & Redgrave, P. A new framework for cortico-striatal plasticity: behavioural theory meets *in vitro* data at the reinforcement-action interface. *PLoS Biol.* **13,** e1002034 (2015).
6. Schultz, W. Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* **57,** 87–115 (2006).
7. Desmurget, M. & Turner, R. S. Motor sequences and the basal ganglia: kinematics, not habits. *J. Neurosci.* **30,** 7685–7690 (2010).
8. Tumer, E. C. & Brainard, M. S. Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. *Nature* **450,** 1240–1244 (2007).
9. Andalman, A. S. & Fee, M. S. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc. Natl Acad. Sci. USA* **106,** 12518–12523 (2009).
10. Gerfen, C. R. *et al.* D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science* **250,** 1429–1432 (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. Collins, A. G. & Frank, M. J. Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychol. Rev.* **121,** 337–366 (2014).
13. Kravitz, A. V., Tye, L. D. & Kreitzer, A. C. Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nature Neurosci.* **15,** 816–818 (2012).
14. Tai, L. H., Lee, A. M., Benavidez, N., Bonci, A. & Wilbrecht, L. Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nature Neurosci.* **15,** 1281–1289 (2012).
15. Niv, Y., Daw, N. D., Joel, D. & Dayan, P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl.)* **191,** 507–520 (2007).
16. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: an Introduction.* (MIT Press, 1998).
17. Yin, H. H., Ostlund, S. B., Knowlton, B. J. & Balleine, B. W. in *Eur. J. Neurosci.* **22,** 513–523 (2005).
18. Paninski, L., Fellows, M. R., Hatsopoulos, N. G. & Donoghue, J. P. Spatiotemporal tuning of motor cortical neurons for hand position and velocity. *J. Neurophysiol.* **91,** 515–532 (2004).
19. Churchland, M. M. & Shenoy, K. V. Temporal complexity and heterogeneity of single-neuron activity in premotor and motor cortex. *J. Neurophysiol.* **97,** 4235–4257 (2007).
20. Moran, D. W. & Schwartz, A. B. Motor cortical representation of speed and direction during reaching. *J. Neurophysiol.* **82,** 2676–2692 (1999).
21. Panigrahi, B. *et al.* Dopamine is required for the neural representation and control of movement vigor. *Cell* **162,** 1418–1430 (2015).
22. Pawlak, V. & Kerr, J. N. Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. *J. Neurosci.* **28,** 2435–2446 (2008).
23. Shen, W., Flajolet, M., Greengard, P. & Surmeier, D. J. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* **321,** 848–851 (2008).
24. Cooper, L. N. & Bear, M. F. The BCM theory of synapse modification at 30: interaction of theory with experiment. *Nature Rev. Neurosci.* **13,** 798–810 (2012).
25. Izhikevich, E. M. & Desai, N. S. Relating STDP to BCM. *Neural Comput.* **15,** 1511–1523 (2003).
26. Turner, R. S. & Desmurget, M. Basal ganglia contributions to motor control: a vigorous tutor. *Curr. Opin. Neurobiol.* **20,** 704–716 (2010).
27. Wall, N. R., De La Parra, M., Callaway, E. M. & Kreitzer, A. C. Differential innervation of direct- and indirect-pathway striatal projection neurons. *Neuron* **79,** 347–360 (2013).
28. Mazzoni, P., Hristova, A. & Krakauer, J. W. Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. *J. Neurosci.* **27,** 7105–7116 (2007).
29. Phillips, A. G. & Fibiger, H. C. The role of dopamine in maintaining intracranial self-stimulation in the ventral tegmentum, nucleus accumbens, and medial prefrontal cortex. *Can. J. Psychol.* **32,** 58–66 (1978).
30. Gallistel, C. R. & Gibbon, J. Time, rate, and conditioning. *Psychol. Rev.* **107,** 289–344 (2000).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by funding from the Howard Hughes Medical Institute. J.T.D. is a Group Leader at Janelia Research Campus. We thank A. Lee, A. Karpova, N. Spruston, and members of our laboratory for critical reading and feedback on the manuscript. We also thank M. Frank for helpful discussions of the OpAL model.

Author Contributions E.A.Y. performed the experiments and analysed the data. E.A.Y. and J.T.D. designed the experiments, performed the modelling, and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.T.D. (dudmanj@janelia.hhmi.org).

METHODS

Subjects. Experimental subjects were eight adult (over 2 months old) male mice, four each of Drd1a-Cre (http://www.informatics.jax.org/allele/MGI:3836631) or Drd2-Cre (http://www.informatics.jax.org/allele/MGI:3836635) crossed with a mouse with an allele for Cre-dependent expression of channelrhodopsin-2 fused to enhanced yellow fluorescent protein (Ai32; https://www.jax.org/strain/012569). Mouse lines expressing Cre-recombinase were produced by the GENSAT project (GENSAT project, Rockefeller University, New York, USA)[31,32] and obtained from the MMRC (https://www.mmrrc.org). Ai32 mice were obtained from Jackson Laboratory and produced by the Allen Institute for Brain Science (https://www.alleninstitute.org)[33]. Number of animals and sessions were based upon previous studies using an intersession control model. Experimenters were not blinded to the condition or animal strain. All animals were handled in accordance with guidelines approved by the Institutional Animal Care and Use Committee (IACUC) of Janelia Research Campus which is IAAALAC accredited.

Animal care. Mice were individually housed in a temperature- and humidity-controlled room maintained on a reversed 12-h light/dark cycle. Following 1 week of recovery from surgery, the water consumption of the mice was limited to at least 1 ml per day. Mice underwent daily health checks, and water restriction was eased if mice fell below 70% of their body weight at the beginning of deprivation. Mice were acclimated to head fixation and trained to lick drops of water sweetened with saccharin.

Behavioural training. Mice spent 4–8 weeks adjusting to being head-fixed and learning to displace a side-mounted joystick, placed 2.7 cm away from their platform, to a threshold of roughly 0.5 cm. After this initial training, the threshold for a successful trial ('criterion threshold') was reduced 0.1 cm so that every joystick movement would be rewarded. Reaching movements were self-paced. Most movement amplitudes easily exceeded this amplitude threshold (Fig. 1c).

Mice were restricted to consume 1.5 ml of water per day to maintain motivation for task completion. Movement was measured by recording voltage changes applied across the variable potentiometer connected to the joystick and were found to be linearly proportional to displacement over the range of movement amplitudes used by mice. At the start of each trial, joystick position was centred to coordinates (0,0), and animals were trained to manoeuvre the joystick to certain displacement thresholds equal to a set resistance change of a potentiometer. Both movements away from and towards the body were allowed. Delivery of a sweetened water reward (~0.05–0.1 ml per trial; controlled by an audible solenoid valve) signalled a successful movement and advancement to the next trial. Water delivery was delayed by 1,000 ms after the joystick position crossed a specific distance threshold. The threshold was set at an arbitrary, low value such that false positives were not detected, but all trial-initiating movements in well-trained mice were rewarded. A force of 0.1 N was required to displace and hold the joystick at an eccentric position. For reference, this is at least fivefold less than a mouse can pull towards itself using its forelimbs for several seconds[34].

No other task-related stimulus was present and the behaviour was performed in a darkened behaviour box. Rewards were followed by a 4,000 ms inter-trial interval (ITI) in which no movements would be rewarded. The joystick position at the end of the ITI (almost always near the central default position) was used as the initial position for the subsequent trial. Mice performed at least 125 trials per session. The initial 25 trials were only used for daily acclimation of the animal to the behavioural setup. Blocks of 50 trials were performed with the stimulation block followed by the no stimulation block. Many blocks could be completed, but only the first block from each condition were used in these analyses. Sham stimulation sessions were identical to stimulation sessions, including the attachment of the optic fibre to the mouse's head, with the exception that the laser was not turned on. Stimulation sessions were conducted at random, and the data from the subsequent session were not used in analysis.

Fibre implantation and optical stimulation. Implantation surgery was performed under full anaesthesia (1.5% isoflurane). The skull was exposed and fibre optic probes were unilaterally inserted 2.2 mm into the brain at 0.5 mm anterior and 1.8 mm lateral to bregma. Fibre optic probes were made of glass fibres (100 μm core) fitted with zirconia LC connectors. Head fixation caps were implanted at the end of the procedure and all elements and remaining skull were covered with dental acrylic as described previously[35]. All surgical procedures were performed under aseptic conditions.

Fibre implants were targeted to the dorsomedial aspect of the striatum (DMS)[1]. Extended Data Fig. 1a shows the localization of the tips of optical fibres implanted for dMSN and iMSN stimulation. To characterize this specific location in more detail we performed bilateral injections of a retrograde tracer (Lumiflor beads) into the approximate DMS location of fibre implants. We found extensive retrograde labelling of neocortical neurons over a relatively extended rostro-caudal axis that was biased towards the medial aspect of neocortex (Extended Data Fig. 1b, c), consistent with previous results from our lab[36]. Based upon the anatomical atlas of the mouse brain[37] these cortical structures are annotated as M2 (secondary motor cortex) and Cg (cingulate cortex). However, we note that recent functional mapping of neocortex indicates that these sites are also within the boundaries of the rostral and caudal forelimb regions (Extended Data Fig. 1c)—areas that are sufficient to produce forelimb movements in response to microstimulation[38].

Closed-loop photostimulation on high- and low-velocity blocks was accomplished through online monitoring of instantaneous joystick velocity. Thresholds for triggering the laser were set for each animal such that approximately one-third of baseline movements would be suprathreshold. The velocity threshold within a session was fixed, but on occasion it was changed from one session to the next. Thresholds for all mice and all sessions were within 6% of each other. For stimulation of all movements, we set the velocity threshold low enough that all movements were suprathreshold. For low-velocity triggering (Fig. 3), we took advantage of the reliable, stereotyped nature of the reaches to predict peak velocity from early velocity. To trigger photostimulation, velocity needed to initially pass a low, 'onset' threshold while not exceeding a higher 'too fast' threshold for the next 20 ms. Using our real-time velocity triggering, we correctly stimulated 96% of upper-third (fast) reaches protocol and 84% of lower-third (slow) reaches protocol. Our false positive stimulation rate was 9% for both protocols. Photostimulation consisted 10-ms pulses at 16.7 Hz for 450 ms from a 473 nm blue laser set so that the power at the tip of the implanted optic probe was 3–6 mW. This was at a frequency below that which individual neurons could reliably follow (Extended Data Fig. 7). Upper-third stimulation data consist of 22 stimulation and 25 sham sessions in dMSN mice, 26 stimulation and 20 sham sessions in iMSN mice. Lower-third stimulation data consist of 16 stimulation and 18 sham sessions in dMSN mice, 20 stimulation, 16 sham sessions in iMSN mice.

We next sought to estimate the extent of light spread based upon the laser power, fibre diameter and duty cycle of our pulse train using a combination of simulation[39] and electrophysiology. The simulation result is shown in Extended Data Fig. 1d. In brief, the peak intensity of light stimulation was reduced to 1% maximum by ~1 mm below and 0.5 mm lateral to the optical fibre. To directly estimate the change in stimulation efficacy as a function of distance we performed recordings with a 4-shank silicon probe (NeuroNexusTech; 'Buzsaki32' site layout) on which 1 shank was affixed with an optical fibre. Consistent with the estimate of light scattering from the simulation we found that direct light activation was substantially weaker on the neighbouring shanks (Extended Data Fig. 1e). At the location of our fibres (Extended Data Fig. 1) the dorsal striatum extends for ~2 mm is all dimensions and the DMS roughly extends for 1 mm. Thus, these data indicate that direct photostimulation was restricted to the dorsal striatum.

Behavioural analysis. All behavioural events were recorded on separate channels at 1 kHz (BlackRock Microsystems; Salt Lake City, Utah). Data analysis was performed using written routines in MATLAB 2014 (MathWorks; Natick, Massachusetts) to extract individual forelimb movement trajectories (Extended Data Fig. 8). Quantification of individual movements considered only the outward component of the reach and quantified the peak amplitude and velocity. The beginning of the reach was assessed offline for each reach and was determined to be the first time point constituting the increasing velocity associated with that reach. The duration was computed as the full duration of the movement and tortuosity is a measure of the directness of the reach path, defined as the path length divided by the end point distance. Z scores were computed for each stimulation session, using the average sham session mean and standard deviation, within each animal, then combined across animals. Simulations of behavioural learning were implemented in MATLAB and are described in detail in the Extended Data. Unless otherwise noted, statistical significance refers to $P < 0.05$, two-tailed Student's t-test.

Electrophysiology. Extracellular electrophysiology was performed in the dorsal striatum of awake, behaving mice. 32 channel silicon probe arrays with attached integrated optical fibres (NeuroNexus Tech; 'Buzsaki32' site arrangement) were acutely implanted in the dorsal striatum (centre of array was positioned 0.5 mm anterior and 1.8 mm lateral to bregma and −2.0 mm to −3.0 mm depth from surface). Electrodes were prepared for recording by reducing the site impedance below 750 kOhm. Broadband continuous data (0.1 Hz–7.5 kHz) were recorded with simultaneous sampling of voltage from the joystick, the lick port, and digital signals from the behaviour control system (30 kHz sample rate on all channels, Blackrock Microsystems, Salt Lake City, Utah). Continuous voltage signals were highpass filtered (0.5–7 kHz) offline and events that exceeded four times the standard deviation of the continuous voltage signal were extracted (spikes). Spike sorting into individual units was performed in MATLAB using custom-written software. Spikes were isolated according to waveform amplitude distribution and principal components of the amplitude array across the eight electrodes (~25 μm spacing) of each shank ($n = 8$) of the silicon probe array. The event times for each individual single unit were then aligned to movement start as extracted from the continuous

voltage signal from the joystick. Velocity–firing-rate slopes were computed using the mean activity of each unit over the epoch spanning 0 to 400 ms after reach initiation. The evoked response for each stimulus (Fig. 1f) was averaged across all isolated neurons, with spikes placed into 2-ms-wide bins.

Pharmacology. We injected D1 and D2 receptor antagonists (SCH23390, 0.02 mg kg^{-1} and sulpiride, 25 mg kg^{-1}, co-injected intraperitoneally) before stimulation sessions. Sham sessions were ones in which the same animals received 0.9% saline injection instead of drug.

The MeSH learning rule. To determine whether the changes in reach velocity due to stimulation were consistent with a reinforcement learning rule, we developed a simple computational model:

$$M_{i+1} = M_i + \omega_r\,(m_i - \overline{M_i}) + \omega_S S - \omega_p\,(m_i - P)/|m_i - P|$$

where M is the Gaussian distribution of values m for a given movement parameter, from which m_i is chosen at random on trial i. Performance reinforcement in the form of reward, r, shifts the mean of M relative the reward r, to which is always given and therefore will be present in every trial. Additional stimulation-induced reinforcement also occurs, shifting the mean according to the type of stimulation, S (+1 or −1 for dMSN or iMSN, respectively; 0 for no stimulation) at a fixed proportion of the reward rate, that is, $\omega_S = C_s\omega_r$. Finally, we have added a restorative set point, P, which is based upon the original mean of the distribution.

Corticostriatal circuit model. We implemented a simple simulation of 500 D1 and 500 D2 striatal units. Activity was continuously varied between 0 and 1. The activity of a given unit was defined as: $r_i = w_i \times$ active, where active $\varepsilon[0,1]$.

Movement velocity was a product of the total cortical output summed with the net contribution of striatal activity (see schematic in Extended Data Fig. 6). This is an explicit model of the structure of the corticostriatal projection where striatal neurons receive collateral input from corticocortical and coorticofugal outputs from neocortex[1]. Thus, movement velocity was: mvmt_velocity $= \alpha\sum$active $+$ $\beta\sum$d1_msn $-\gamma\sum$d2_msn. Simulations were conducted with a variety of weightings, but for examples in the manuscript we used $\alpha = 0.5$, $\beta = 1$, $\gamma = 1$.

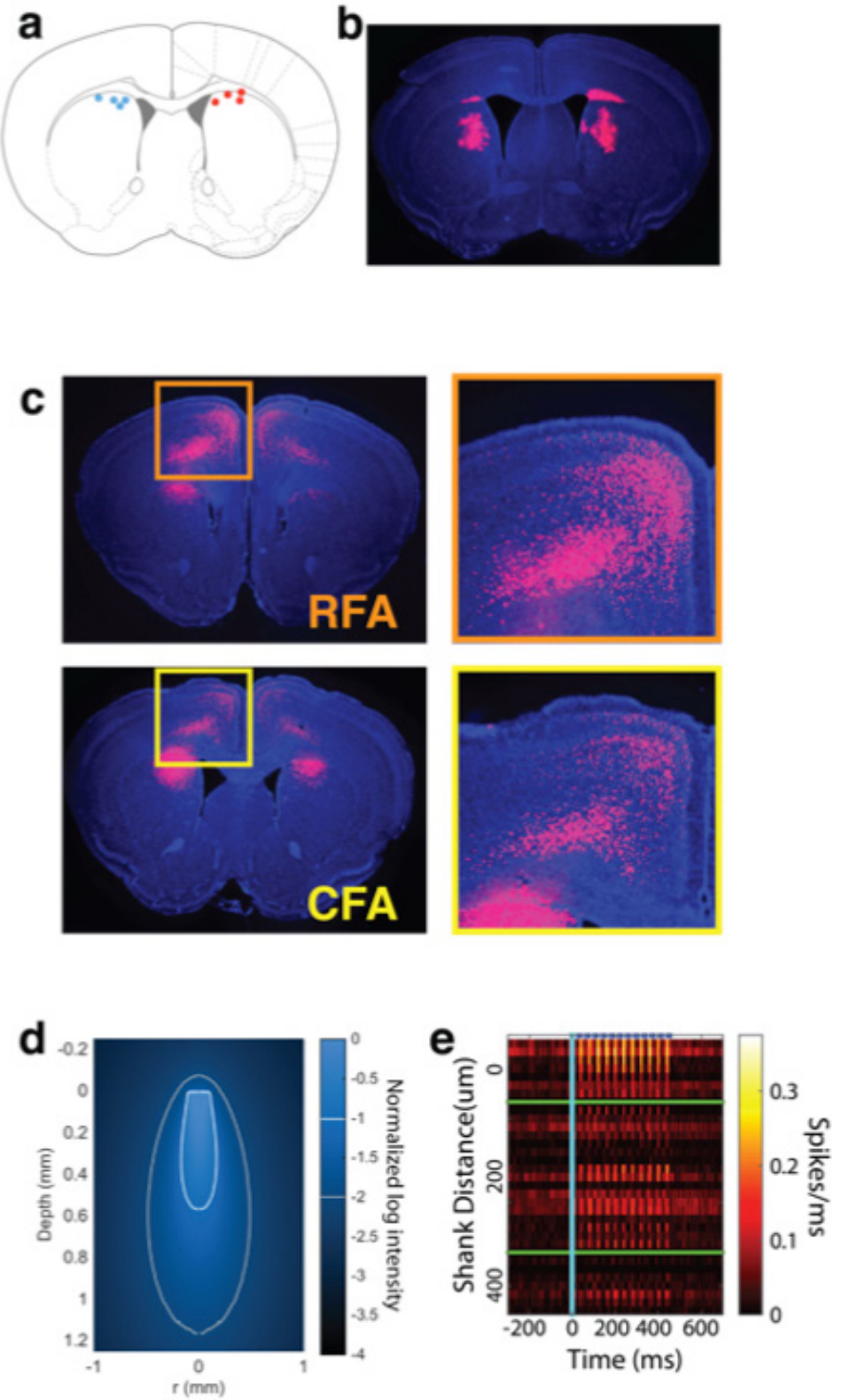
Synapse weights were updated incrementally according to a simple update equation: $w_i(t+1) = w_i(t) +$ active $\times\,\alpha_{\text{learn}} - (1 -$ active$)\beta_{\text{learn}}$.

Thus, if the unit was active its weight was increased by α_{learn} or otherwise decreased by β_{learn} if inactive. Various parameterizations of learning rates could be used, but we typically used $\alpha_{\text{learn}} = -\beta_{\text{learn}} = 0.05$ in the absence of photostimulation and $\alpha_{\text{learn}} = -\beta_{\text{learn}} = 0.09$ during photostimulation. Altered learning rates were only applied to the stimulated population.

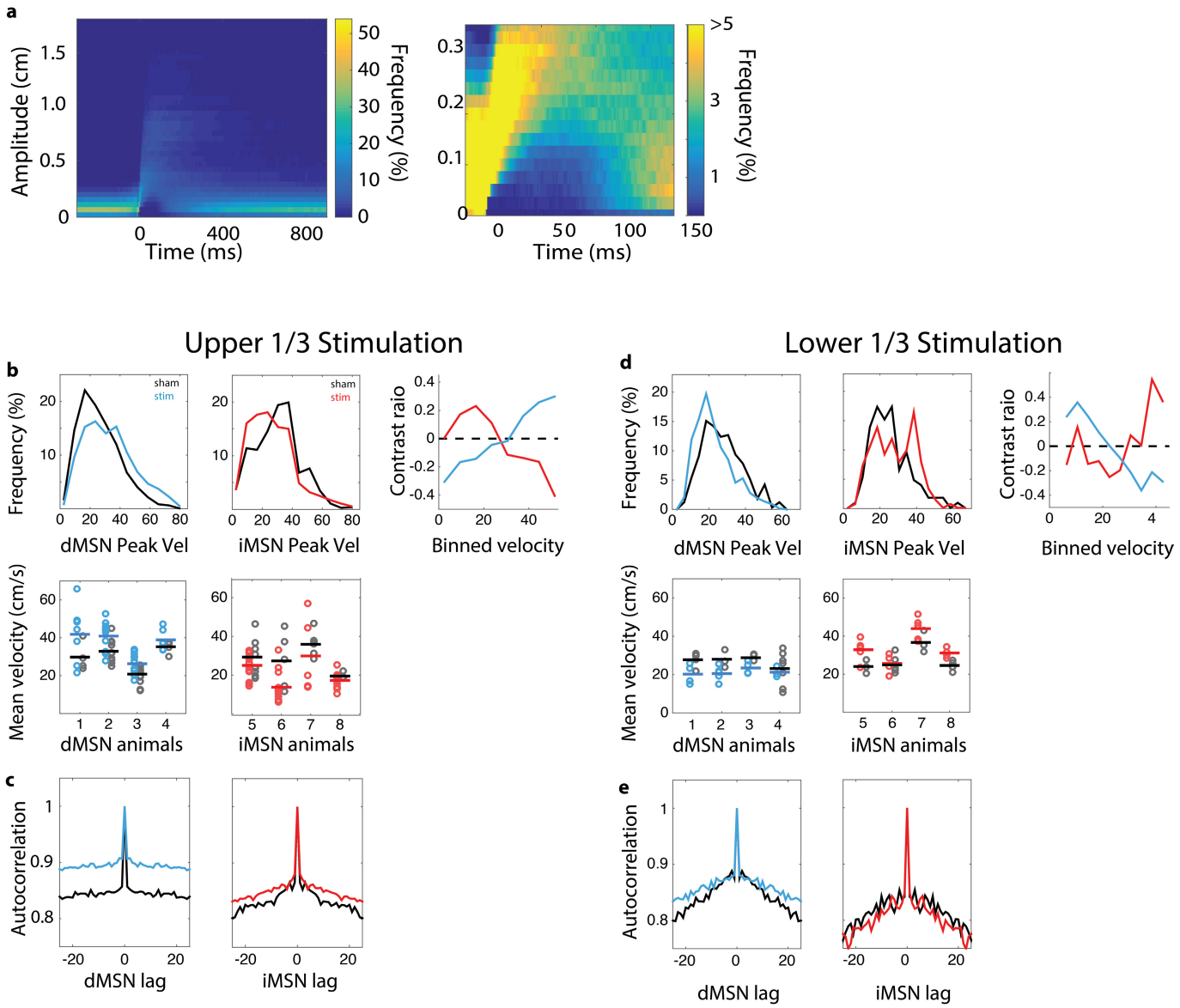
***In vitro* intracellular recordings.** Methods for Extended Data Fig. 7 were as described previously[40]. Briefly, for the preparation of *in vitro* brain slices, mice were deeply anaesthetized with isoflurane, decapitated, and the brain placed into ice-cold modified artificial cerebral spinal fluid (aCSF) (in mM: 52.5 NaCl, 100 sucrose, 26 $NaHCO_3$, 25 glucose, 2.5 KCl, 1.25 NaH_2PO_4, 1 $CaCl_2$, 5 $MgCl_2$ and in μM: 100 kynurenic acid) that had been saturated with 95%O_2/5%CO_2. 300-μm thick coronal slices were cut (Leica VT1200S; Leica Microsystems, Germany), transferred to a holding chamber and incubated at 35 °C for 30 min in modified aCSF (in mM: 119 NaCl, 25 $NaHCO_3$, 28 glucose, 2.5 KCl, 1.25 NaH_2PO_4, 1.4 $CaCl_2$, 1 $MgCl_2$, 3 Na pyruvate and in μM: 400 ascorbate and 100 kynurenic acid, saturated with 95%O_2/5%CO_2) and then stored at room temperature.

For recordings, slices were transferred to a recordings chamber and superfused with modified aCSF (in mM: 119 NaCl, 25 $NaHCO_3$, 18 glucose, 2.5 KCl, 1.25 NaH_2PO_4, 1.4 $CaCl_2$, 1 $MgCl_2$, 3 Na pyruvate and in μM: 400 ascorbate and saturated with 95%O_2/5%CO_2) maintained at 32–34 °C, at a flow rate of 2–3 ml per minute. Patch pipettes (resistance 5–8 MΩ) were pulled on a laser micropipette puller (Model P-2000, Sutter Instrument Co., Sunnyvale, California) and filled with a KGluconate-based intracellular solution (in mM: 137.5 KGluconate, 2.5 KCl, 10 HEPES, 4 NaCl, 3 GTP, 40 ATP, 10 phosphocreatine, pH 7.5). Intracellular recordings were made using a MultiClamp700B amplifier (Molecular Devices, Sunnyvale, California) interfaced to a computer using an analogue to digital converter (PCI-6259; National Instruments, Austin, Texas) controlled by custom written scripts in Igor Pro (Wavemetrics, Eugene, Oregon). Software is available at http://www.dudmanlab.org.

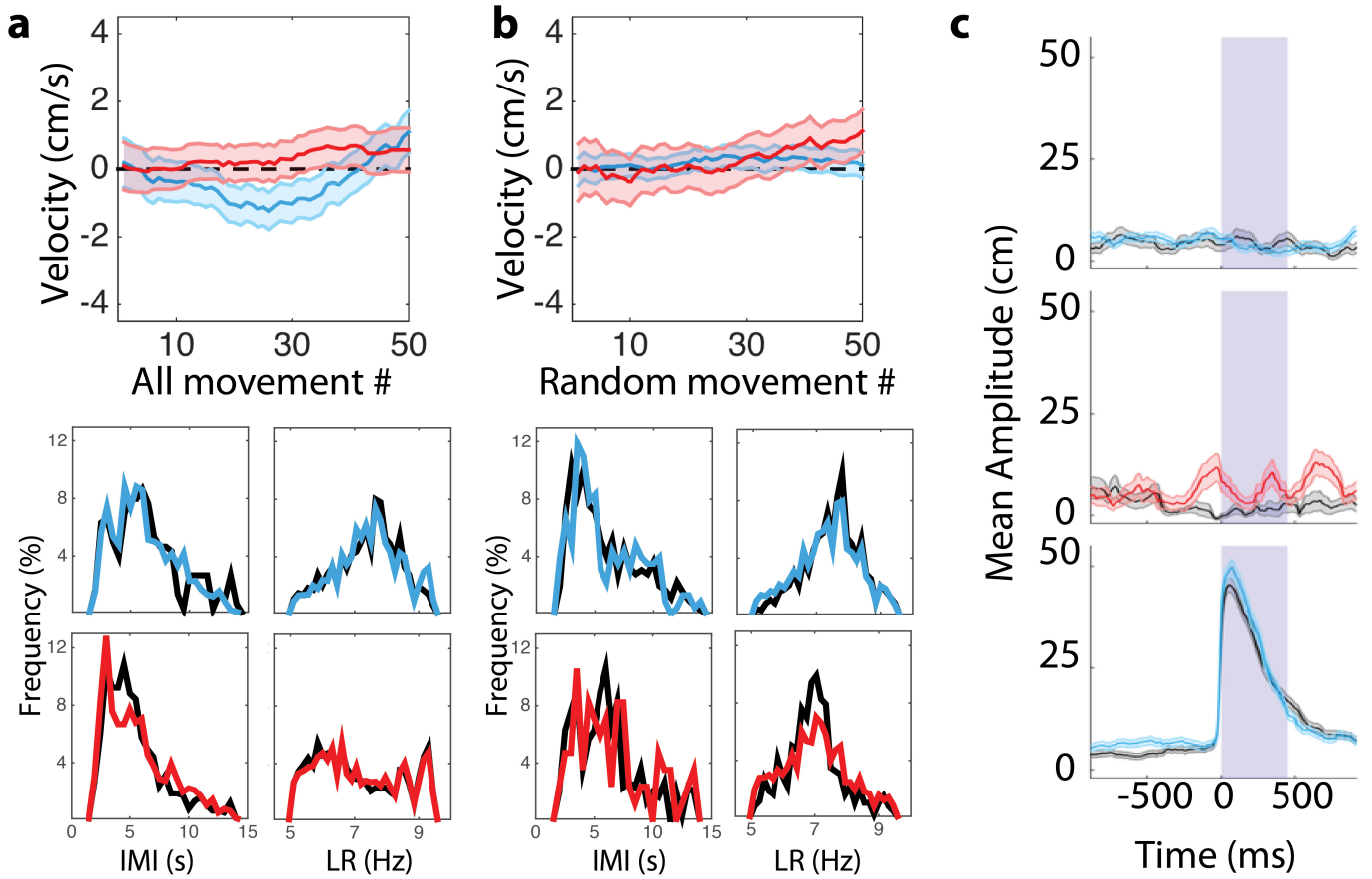
31. Gerfen, C. R., Paletzki, R. & Heintz, N. GENSAT BAC Cre-recombinase driver lines to study the functional organization of cerebral cortical and basal ganglia circuits. *Neuron* **80,** 1368–1383 (2013).
32. Gong, S. *et al.* Targeting Cre recombinase to specific neuron populations with bacterial artificial chromosome constructs. *J. Neurosci.* **27,** 9817–9823 (2007).
33. Madisen, L. *et al.* A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. *Nature Neurosci.* **15,** 793–802 (2012).
34. Deacon, R. M. Measuring the strength of mice. *J. Vis. Exp.* (2013).
35. Osborne, J. E. & Dudman, J. T. RIVETS: a mechanical system for *in vivo* and in vitro electrophysiology and imaging. *PLoS ONE* **9,** e89007 (2014).
36. Pan, W. X., Mao, T. & Dudman, J. T. Inputs to the dorsal striatum of the mouse reflect the parallel circuit architecture of the forebrain. *Front. Neuroanat.* **4,** 147 (2010).
37. Paxinos, G. & Franklin, K. *The Mouse Brain in Stereotaxic Coordinates.* (Gulf Professional Publishing, 2004).
38. Tennant, K. A. *et al.* The organization of the forelimb representation of the C57BL/6 mouse motor cortex as defined by intracortical microstimulation and cytoarchitecture. *Cereb. Cortex* **21,** 865–876 (2011).
39. Stujenske, J. M., Spellman, T. & Gordon, J. A. Modeling the spatiotemporal dynamics of light and heat propagation for *in vivo* optogenetics. *Cell Rep.* **12,** 525–534 (2015).
40. Brown, J., Pan, W. X. & Dudman, J. T. The inhibitory microcircuit of the substantia nigra provides feedback gain control of the basal ganglia output. *Elife* **3,** e02397 (2014).



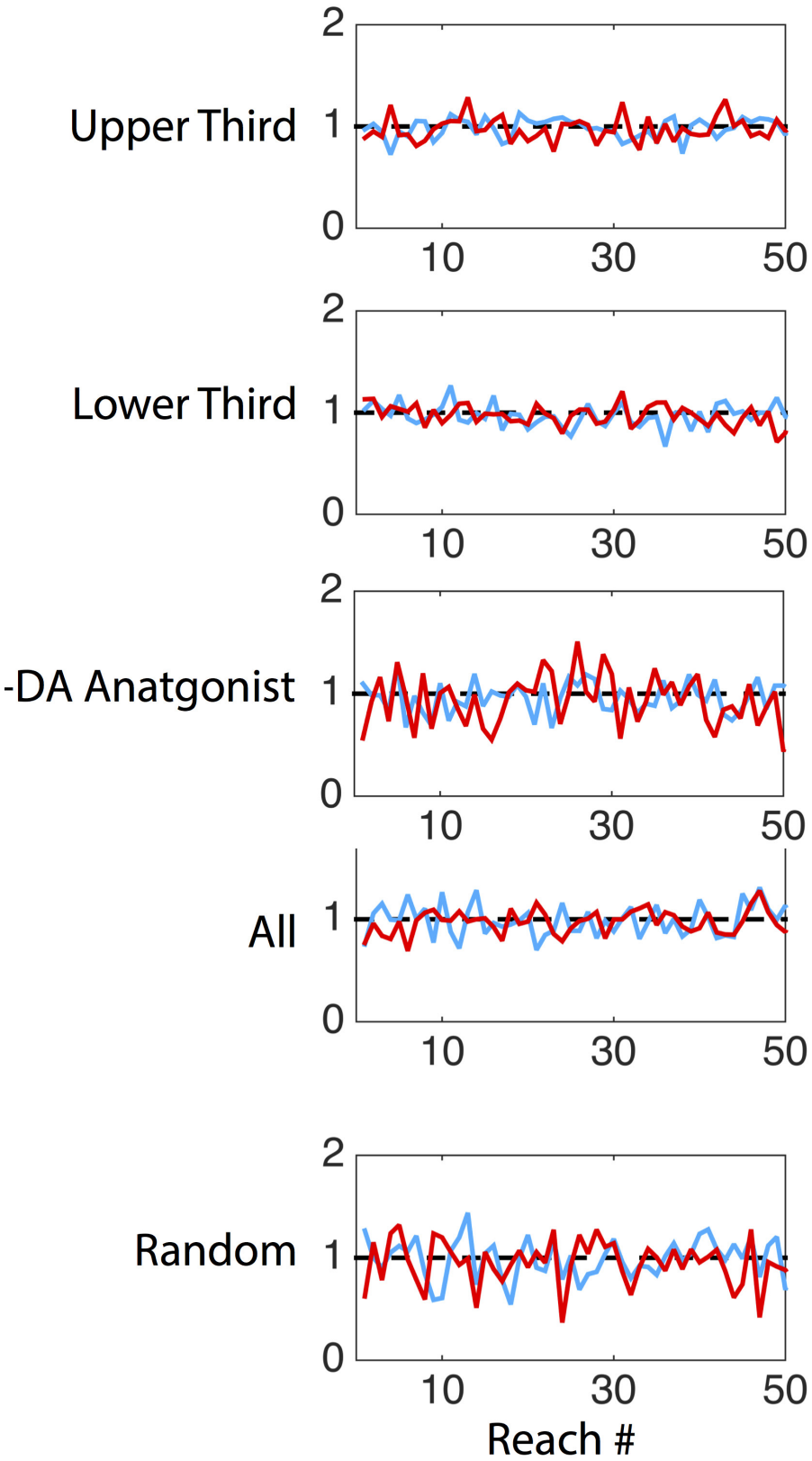
Extended Data Figure 1 | Anatomical localization of stimulated neurons and their corticostriatal inputs. a, Schematic showing the histological reconstruction of optic fibre endpoints in dMSN (cyan) and iMSN (red) mice. Hemispheric segregation for display purposes only; actual implants were randomized left and right. **b**, Location of retrobead tracer injection just below the coordinates of fibre implant terminus. **c**, Location of the labelled projection neurons innervating the retrobead injection site corresponding to the rostral forelimb area (RFA, orange) and caudal forelimb area (CFA, yellow). **d**, Results of simulation of light intensity and spread through the brain based upon the model of Stujenske *et al.*[39]. Light intensity drops below 10% peak intensity 285 μm below the fibre. At this depth the lateral fall-off is a drop to 1% peak intensity 250 μm. Contour lines for 10% intensity and 1% intensity are overlaid in white and grey, respectively. **e**, Average photostimulation-evoked activity of individual neurons (each row = 1 neuron) sorted by electrode shank. We observed minimal fidelity in the light-evoked responses of neurons on the farthest shank consistent with the predicted fall-off in light intensity from simulations. Blue dashes at top identify timing of laser pulses. Robust entrainment of spiking to photostimulation was also observed *in vitro* (Extended Data Fig. 7).



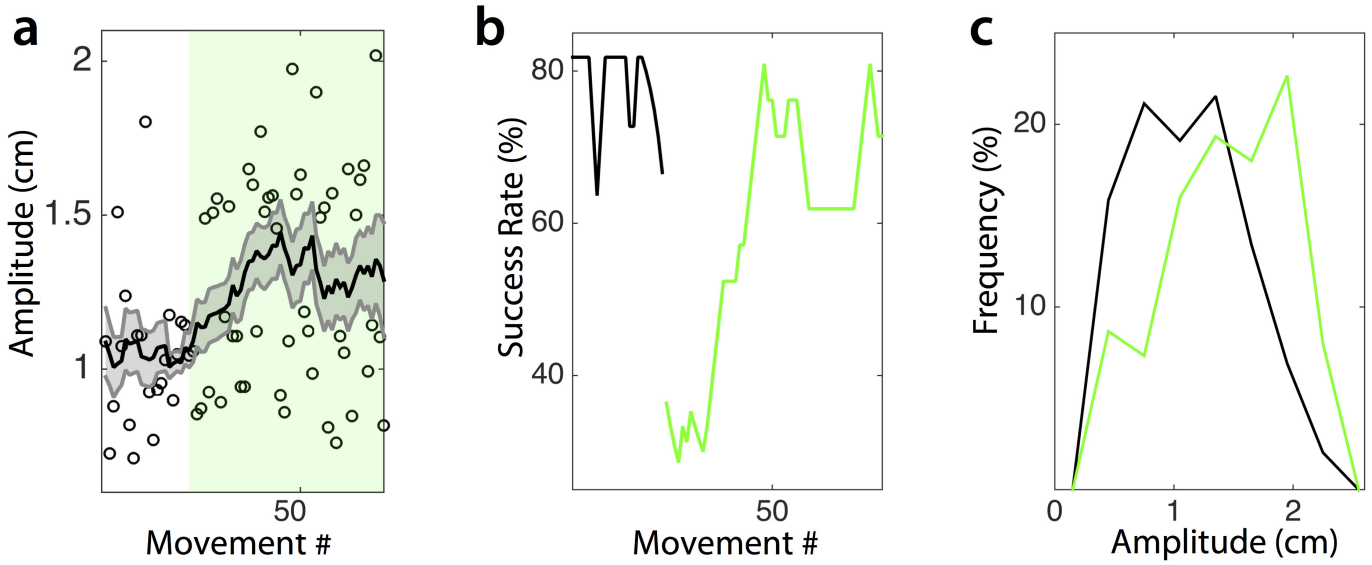
Extended Data Figure 2 | Selective stimulation of MSNs produces changes in peak velocity. a, Density plot of movement trajectories showing the percentage of movements that passed through a given amplitude throughout time, aligned to movement onset. Particularly in the magnified plot on the right, we can see that very few movement trajectories passed near the central 0.2 cm within the first 100 ms. **b**, Peak velocity distributions for sham and stimulation data sets for dMSN stimulation (blue, left column), and iMSN stimulation (red, middle column). Top row shows complete distributions across animals. Note that no particular part of the distribution is pronounced following stimulation. Bottom row shows means for all experimental sessions by animal. Animal number is for indexing purposes only. Most experiments were carried out concurrently across animals. The right column shows the contrast ratio (difference divided by sum) for dMSN and iMSN stimulation effects (blue and red, respectively). These data show a steady mean shift in the data; for instance, iMSN stimulation (red) is positive for velocities slower than the sham (control) mean (an increase in frequency), and is negative for velocity values greater than the sham mean (a decrease in frequency). We have curtailed the contrast ratio plot where too few values existed to achieve reliable estimates (55 cm s^{-1}). **c**, Autocorrelation of movement velocities for dMSN stim (blue), iMSN (red), and sham (black) data. **d**, **e**, Same analyses as **b**, **c**, but for those sessions in which stimulation occurred on the slower, lower third of movement velocities.



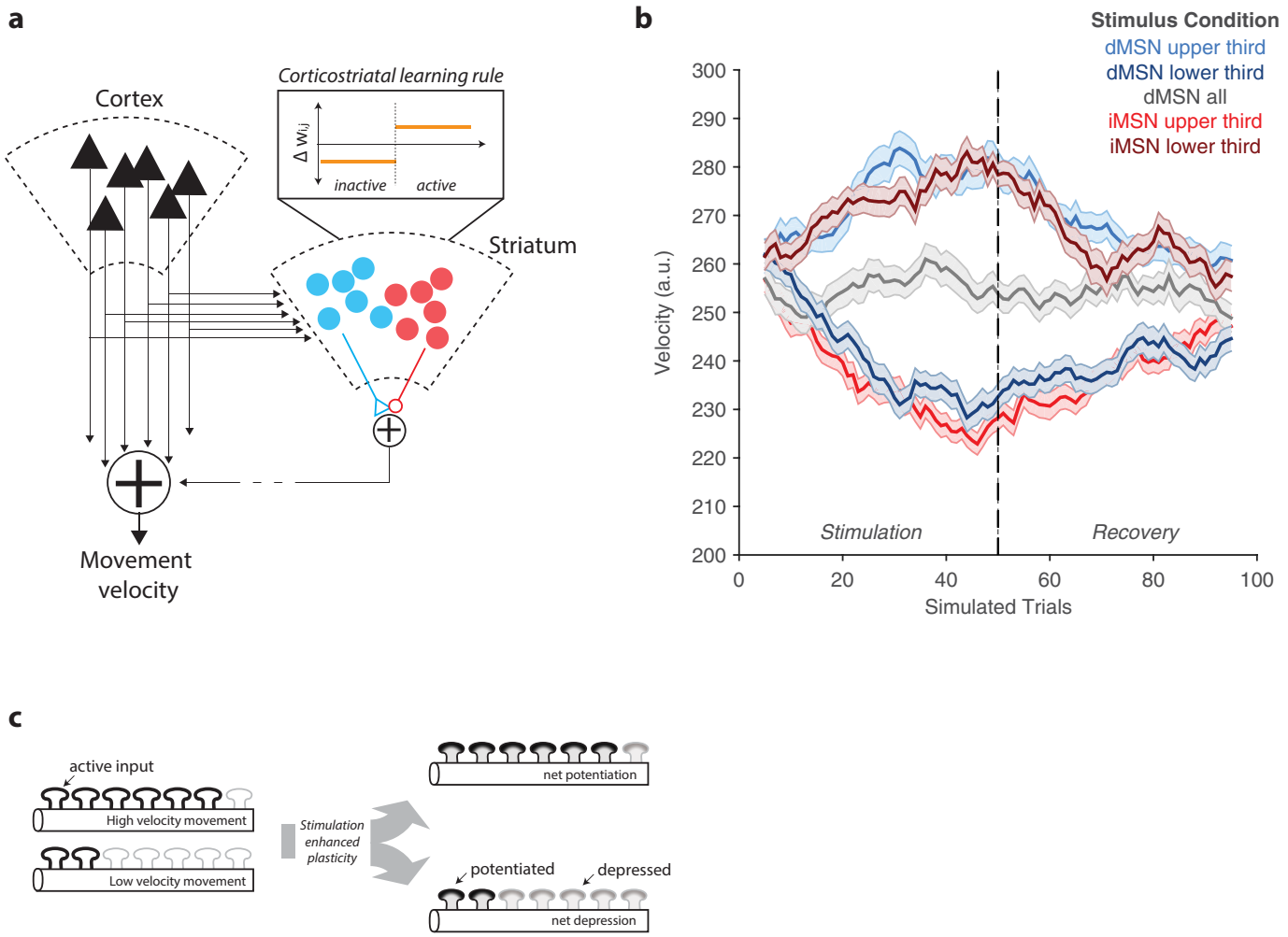
Extended Data Figure 3 | Trained animals can adjust amplitude to changing task requirements. a, Reach amplitudes from a sample of seven sessions in two mice in which the eccentricity of the threshold to receive reward was suddenly increased at random. Green field identifies reaches performed with the increased amplitude threshold. Shaded area represents standard error of the mean. **b**, Success rate before (black) and after (green) the jump in amplitude threshold. **c**, Distribution of reach amplitudes across sessions for pre- (green) and post- (black) amplitude threshold jump. These data indicate that the mean amplitude was not saturated and suggest that behaviour remains outcome-dependent (that is, goal-directed).



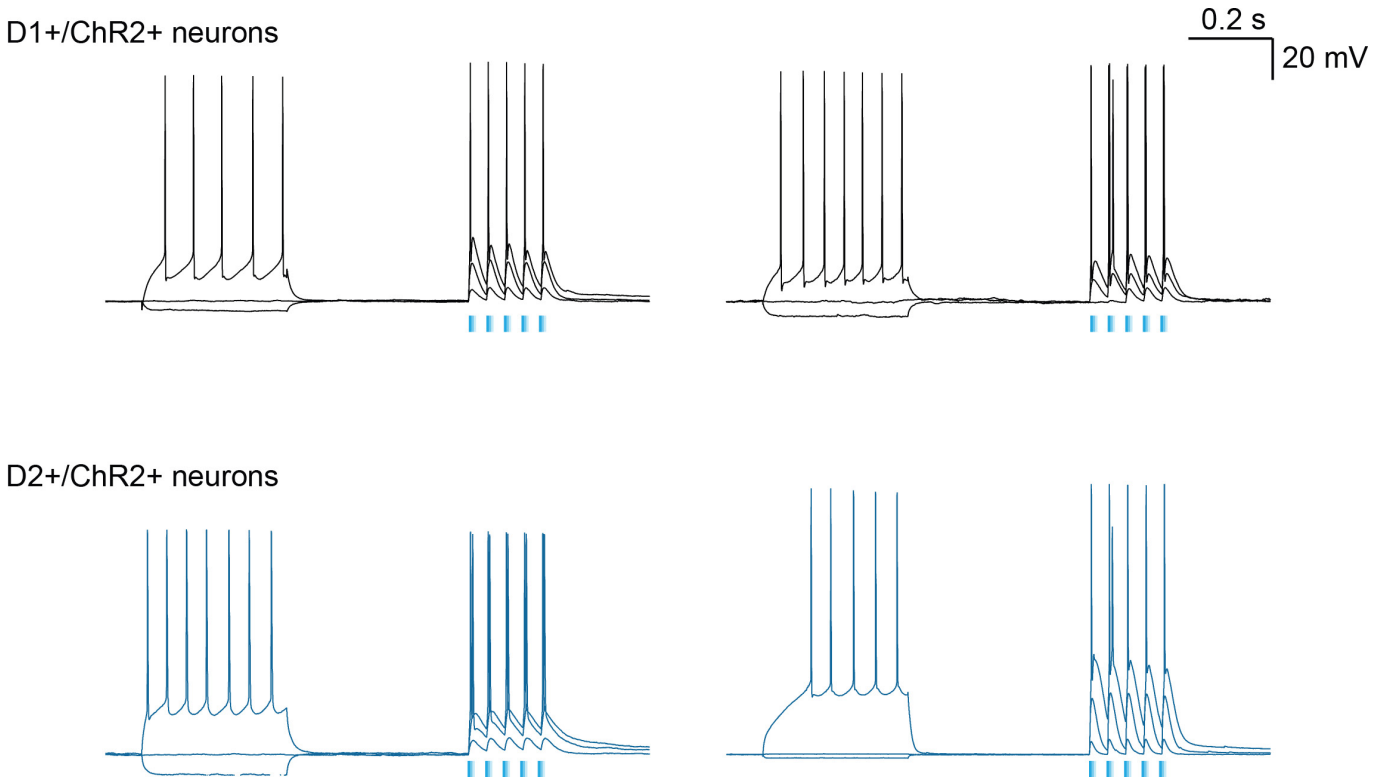
Extended Data Figure 4 | Variance of movement velocity does not change throughout a session. Each session was *z*-scored and the standard deviation for each movement number is plotted for dMSN (cyan) and iMSN (red) stimulation.



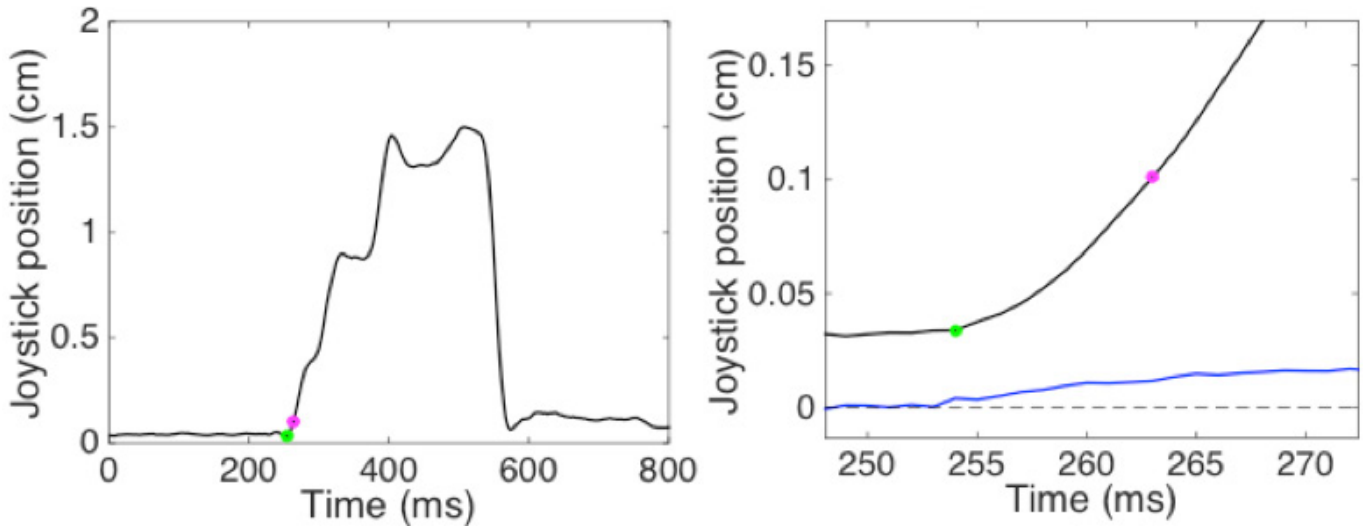
Extended Data Figure 5 | Non-selective stimulation does not affect motor control or initiation. **a**, **b**, 'All-stim' (**a**; 15 stimulation and 17 sham sessions from 4 dMSN mice; 17 stimulation and 20 sham sessions from 4 iMSN mice) and 'random stim' (**b**; 11 stimulation and 15 sham dMSN sessions from 3 dMSN mice, 8 stimulation and 12 sham sessions from 3 iMSN mice) summary data showing (top) the mean velocity as a function of movement number for dMSN (cyan) and iMSN (red) stimulation sessions and (bottom) histograms of the inter-move interval (IMI; left) and lick rate (LR; right). Shaded area indicates standard error of the mean. We found no differences between sham (black lines) and stimulation (coloured lines) sessions for either dMSN stimulation (cyan) or iMSN stimulation (red). **c**, Plot of average trajectory position aligned to stimulation onset for random dMSN (top) and iMSN stimulation (middle) for stimulation (coloured) and sham (black) sessions (in sham sessions, timing was randomly chosen, but no stimulation was given). Stimulation did not systematically induce forelimb movement. For reference, the bottom trace displays closed-loop dMSNs aligned to stimulation onset. Grey field represents the 450 ms stimulation period.



Extended Data Figure 6 | A corticostriatal circuit model that implements the MeSH learning rule. a, In the left panel, we present a schematic of the corticostriatal pathway consistent with known anatomical data[1]. Descending cortical outputs, largely from layer V of the neocortex, project subcortically and intracortically elaborating axon collaterals onto direct (blue) and indirect (red) MSNs in the dorsal striatum. By typical convention we assume that dMSNs have a net positive effect on behaviour (increase in velocity in this case) and iMSNs have a net inhibitory effect (decrease in velocity). These pathways are combined at the basal ganglia output nucleus (substantia nigra pars reticulata; not shown) and then combined with cortical drive to produce the net movement velocity. The model assumes that both dMSNs and iMSNs are positively correlated with cortical activity and with movement velocity. We assume a monotonic relationship between cortical activity and movement velocity[20]. The model is initialized at a presumptive steady-state in which weights between cortical inputs and dMSN and iMSN units are noisy, but distributed around 0.5 and bounded [0,1]. Most simulations were performed with 100 cortical units and 250 dMSNs and iMSNs each. Under all conditions weights are subject to updating according to a balanced plasticity rule (inset) in which inactive units are subject to depression and active units are subject to potentiation. All synapses drift back towards a mean of 0.5 thereby implementing a homeostatic set point to the weight distribution. Finally, photostimulation is assumed to enhance (90% increase) the magnitude of both depression and potentiation on stimulated trials in the stimulated population. Random sets of cortical inputs are assumed to be active on any given reach and are drawn from a gamma (shape parameters: 8, 63) distribution that gives a distribution of velocities similar to that observed experimentally. Further details of the model are provided in the Methods. **b**, Example simulations of 100 trials (first 50 receive stimulation according to conditions described in legend followed by 50 unstimulated recovery trials). Curves reflect averages and standard errors of 100 repetitions of the simulated condition. Other conventions as in main figures. **c**, Schematic of dendrite of MSN containing synapses active during limb movements. Synaptic plasticity enhanced by stimulation (inset in panel **a**) produces a net bias in synaptic weights when delivered in closed-loop. This bias can become uniform by permuting the active synapses on each simulated trial.



Extended Data Figure 7 | D1$^+$ and D2$^+$ MSNs can follow repetitive stimulus trains of ~20 Hz photostimulation *in vitro*. Example MSNs recorded *in vitro* in brain slices containing the DMS. Upper row are two example D1$^+$ dMSNs recorded from DMS of a Drd1a-Cre::Ai32 mouse. Lower row are two example D2$^+$ dMSN recorded from DMS of a Drd2a-Cre::Ai32 mouse. 3 example traces shown from each. Spiking is evoked by increasing current injection (traces selected for approximately similar evoked spiking rate) and ~500 ms later by a train of 5 pulses of blue light of increasing duty cycle. All cells recorded were able to follow rapid phasic stimulation and action potentials were reliably evoked on every stimulus of these 20 Hz trains (approximately similar to the stimulus trains used in stimulation experiments described in the text). Examples were selected from seven neurons from three D1$^+$ animals, and four neurons from three D2$^+$ animals that were recorded for this particular stimulation design.



Extended Data Figure 8 | Characterization of movement onset and reach initiation threshold crossing time. A sample reach from a sham block of 50 is shown, with eccentricity in black and velocity in blue (in right panel). The reach start is identified with the greed dot, the threshold crossing, 9 ms later when the reach eccentricity surpassed 0.1 cm, is identified with the magenta dot. The beginning of the reach (green) was assessed offline for each reach and was determined to be the first time point, sampled at 1 kHz, constituting the increasing velocity associated with that reach.

Extended Data Table 1 | Lick and move rates from each animal for sham upper third stimulation sessions

dMSN lick frequency (Hz):	Sham (sem)		Stim (sem)
Mouse 1	7.2 (0.3)		7.4 (0.3)
Mouse 2	7.0 (0.1)		6.9 (0.1)
Mouse 3	7.2 (0.1)		7.2 (0.2)
Mouse 4	7.3 (0.1)		7.2 (0.1)
iMSN lick frequency (Hz):			
Mouse 5	7.1 (0.6)		7.4 (0.3)
Mouse 6	6.9 (0.9)		6.6 (0.3)
Mouse 7	6.8 (0.1)		7.1 (0.2)
Mouse 8	6.7 (0.1)		6.8 (0.1)
dMSN inter-move interval (seconds):			
Mouse 1	5.9 (0.6)		5.6 (0.5)
Mouse 2	6.2 (0.3)		5.9 (0.3)
Mouse 3	6.1 (0.4)		6.8 (0.2)
Mouse 4	5.7 (0.3)		6.7 (0.4)
iMSN inter-move interval (seconds):			
Mouse 5	7.6 (0.4)		7.3 (0.4)
Mouse 6	5.9 (0.3)		6.2 (0.3)
Mouse 7	6.2 (0.4)		6.8 (1.1)
Mouse 8	6.0 (0.3)		6.3 (0.4)

Increases in the mean for lick frequency indicate a faster rate of licking (more motor output), while increases in the mean for inter-move interval indicate a slower rate of licking (less motor output). Lick rate changes per animal were not significant ($P > 0.4$ for all but one iMSN animal, which exhibited a paradoxically faster lick rate, $P = 0.18$). Inter-move interval changes were also not significant within each animal ($P > 0.3$ except one dMSN animal which paradoxically showed an increased inter-move interval, $P = 0.15$). Mouse number is only meant as an index and does not reflect timing or order of experiments, most of which were done in the same period of time.