Published in final edited form as:

Nature. 2013 August 29; 500(7464): 575–579. doi:10.1038/nature12475.

Prolonged Dopamine Signalling in Striatum Signals Proximity and Value of Distant Rewards

Mark W. Howe[1], **Patrick L. Tierney**[1], **Stefan G. Sandberg**[2], **Paul E.M. Phillips**[2], and **Ann M. Graybiel**[1]

[1]McGovern Institute for Brain Research and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[2]Department of Psychiatry & Behavioral Sciences and Department of Pharmacology, University of Washington, Seattle, WA 98195, USA

Abstract

Predictions about future rewarding events have a powerful influence on behaviour. The phasic spike activity of dopamine-containing neurons, and corresponding dopamine transients in the striatum, are thought to underlie these predictions, encoding positive and negative reward prediction errors[1–5]. Many behaviours, however, are directed toward distant goals, for which transient signals might fail to provide sustained drive. Here we report a novel, extended mode of reward-predictive dopamine signalling in the striatum that emerged as rats moved toward distant goals. These dopamine signals, which were detected with fast-scan cyclic voltammetry (FSCV), gradually increased or--in rare instances--decreased as the animals navigated mazes to reach remote rewards, rather than having phasic or steady tonic profiles. These dopamine increases (ramps) scaled flexibly with both the distance and size of the rewards. During learning, these dopamine signals exhibited spatial preferences for goals in different locations and readily changed in magnitude to reflect changing values of the distant rewards. Such prolonged dopamine signalling could provide sustained motivational drive, a control mechanism that may be important for normal behaviour and that can be impaired in a range of neurologic and neuropsychiatric disorders.

The spike activity patterns of midbrain dopamine-containing neurons signal unexpected and salient cues and outcomes[1–4,6,7], and the dynamics of these phasic neural signals have been found to follow closely the principles of reinforcement learning theory[3–6]. In accord with this view, selective genetic manipulation of the phasic firing of dopamine neurons alters some forms of learning and cue-guided movements[8,9]. Episodes of transient dopamine release in the ventral striatum have been detected with FSCV, and these also occur in response to primary rewards and, after learning, to cues predicting upcoming rewards[10–13]. Thus dopamine transients in the striatum share many features of the phasic spike activity of midbrain dopamine neurons.

Correspondence and requests for materials should be addressed to A.M.G. (graybiel@mit.edu).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature

Author Contributions. M.W.H. and A.M.G. designed the experiments, analyzed the data, and wrote the manuscript. M.W.H. conducted the experiments. P.L.T. helped to develop the FSCV chronic recording setup. S.G.S. and P.E.M.P provided technical training on FSCV and advice on data analysis. All authors contributed to discussion and interpretation of the findings.

The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper.

Classic studies of such dopamine transients have focused on Pavlovian and instrumental lever-press tasks, in which rewards were within arm's reach[1–3,10–13]. In many real-life situations, however, animals must move over large distances to reach their goals. These behaviours require that ongoing motivational levels be flexibly adjusted according to changing environmental conditions. The importance of such control of ongoing motivation is reflected in the severe impairments suffered in dopamine deficiency disorders, including Parkinson's disease. In addition, in pioneering experimental studies, dopamine signalling has been implicated in controlling levels of effort, vigour and motivation during the pursuit of goals in maze tasks[14–17]. It has been unclear how phasic dopamine signalling alone could account for persistent motivational states[18]. We adapted chronic FSCV to enable prolonged measurement of real-time striatal dopamine release as animals learned to navigate toward spatially distant rewards.

We measured dopamine levels in the dorsolateral striatum (DLS) and ventromedial striatum (VMS) (Extended Data Figs. 1 and 2, Methods) as rats navigated mazes of different size and shape to retrieve reward (Figs. 1–4, Methods). The rats were trained first on an associative T-maze task to run and to turn right or left as instructed by tones to receive chocolate milk reward at the indicated end-arms[19] (n = 9, Figs. 1, 2, and 4). To our surprise, instead of mainly finding isolated dopamine transients at the initial cue or at goal-reaching, we primarily found gradual increases in the dopamine signals that began at the onset of the trial and ended after goal-reaching (Fig. 1a,b). These ramping dopamine responses, identified in session averages by linear regression (Pearson's $R > 0.5$, $P < 0.01$), were most common in the VMS (75% of sessions) but were also present at DLS recording sites (42% of sessions). They were evident both in single trials (Fig. 1a–c) and in population averages (Fig. 1e,f, Extended Data Figs. 2g,h and 3), bore no clear relationship to run speed within or across trials (Fig. 1d), and matched, in electrochemical profile, dopamine release evoked by tonic electrical stimulation *in vivo (*Extended Data Fig. 2i,j). Before goal-reaching, the ramps had similar amplitudes in correct (65% overall) and incorrect trials (Fig. 1e,f). After goal-reaching, the signals were significantly larger in correct trials, especially in the VMS (paired t-test, $P = 0.01$, Fig. 1e,f). Notably, a subset of the session-averaged signals in the DLS (22%, 58/262, 7 probes in 5 rats) exhibited sustained inhibition up to goal-reaching (Extended Data Fig. 3). Such negative signals were rare in the VMS (5%, 15/300 recordings), suggesting that ramping dopamine signals in the DLS, but not VMS, exhibit heterogeneity in polarity.

We identified isolated phasic transients at warning click indicating trial start and after goal-reaching. These were clearly distinct from the slower ramping responses in ~10% of single trials (Extended Data Fig. 4) but were often superimposed on the ramping signals, indicating that the signals recorded could include combinations of transient increases after warning click, slower ramps to goal-reaching, and transient increases after goal-reaching (Extended Data Fig. 4d). The peak magnitudes of the dopamine ramps were comparable to, or slightly smaller than, those of isolated phasic dopamine signals recorded here (Extended Data Fig. 4) and in other studies[11,12], and they were correlated with the peak magnitudes of free-reward evoked dopamine measured on the same probes (Pearson's $R = 0.45$, $P < 0.001$, Extended Data Fig. 5), indicating that the ramping signals could be subject to similar regulatory mechanisms and display similar anatomic heterogeneity as classical phasic reward-evoked dopamine signals.

We took advantage of the trial-to-trial variability in the rats' run-times (Fig. 2a) to determine whether ramping dopamine release reflected elapsed time or reward proximity, or whether the ramps reflected sums of multiple, accumulated transients to fixed maze cues[20]. If the dopamine ramps tracked elapsed time, peak dopamine values should have scaled directly with trial-time (same slope, different peak height; Fig. 2b,f). If the ramping reflected

distance or spatial location relative to goal-reaching (proximity), peak dopamine levels should have been equivalent for shorter and longer trials (different slope, same peak height; Fig. 2c,f). If the ramps were generated by summation of multiple transients, then for characteristic transient dynamics, the signals should have tended to peak at lower values for longer runs than for shorter runs (different slope, different peak height; Extended Data Fig 6a,b, Supplementary Discussion). The measured peak dopamine values at goal-reaching were nearly equivalent for short and long trials (Fig. 2e), and were not correlated with trial length (Fig. 2d–f), or with run velocity or acceleration (Extended Data Fig. 6e,f). Moreover, on trials in which rats paused mid-run, the signals remained sustained (or dipped slightly) and resembled the actual proximity to reward (Extended Data Fig. 7). These observations indicated that the ramping signals could represent a novel form of dopamine signalling that provides a continuous estimate of the animal's spatial proximity to distant rewards (Fig. 2, Extended Data Fig. 6, Supplementary Discussion).

Given that phasic responses of dopamine-containing neurons can reflect the relative value of stimuli[21], we asked, in a subset of rats, whether the ramping dopamine signals could also be modulated by the size of the delivered rewards (Methods). We used mazes with T, M or S configurations and different total lengths (Fig. 3, Extended Data Fig. 8). We required the animals to run toward one or the other maze-end and varied the rewards available at the alternate goal-regions. With all three mazes, dopamine ramping became strongly biased toward the goal with the larger reward (Fig. 3, Extended Data Fig. 8). Run speed was slightly higher for the high-reward maze arms (Fig. 3i,k), but these small differences were unlikely to account fully for the large differences in the dopamine signals recorded. When we then reversed the locations of the small and large rewards, the ramping signals also shifted, across sessions or just a few trials, so as to favour the new high-value maze-arm (Fig. 3, Extended Data Fig. 8). These bias effects were statistically significant for each experimental paradigm (Extended Data Fig. 8h–j, Mann-Whitney U-test, $P < 0.05$) and across all rats (Fig. 3d, $n = 4$, Mann-Whitney U-test, $P = 0.02$).

In the M-maze, the ramps became extended to cover the longer end-arm distances to goal-reaching, and critically, peaked at nearly the same level prior to goal-reaching as did the ramping signals recorded in the T-maze, despite the longer distance travelled (Fig. 3e). This result suggested that the ramping dopamine signals do not signal reward proximity in absolute terms, but, instead, scale with the path distance to a fixed level that depends on the relative reward value.

To determine whether such value-related differences in the ramping dopamine signals would occur when the actions to reach the distant goal-sites were equivalent, we used the "S"-shaped maze. The ramping signals were larger for the run-trajectories leading to the larger rewards (Fig. 3c,j and Extended Data Fig. 9), despite the fact that the sequence of turns and the lengths of the runs needed to reach the larger and smaller rewards were equivalent for both trajectories ($n = 2$ rats, 4 and 5 sessions/rat, Fig. 3c,j,k, Extended Data Figs. 8 and 9).

In rats performing the free-choice associative version of the T-maze task, robust dopamine signal biases existed in about 20% of sessions (Mann-Whitney U-test, $P < 0.05$) and significantly more often than chance overall (z-test, $P < 0.00001$ vs. bootstrapped variances; Methods, Fig. 4a,b,d). These biases were maintained across consecutive training sessions for individual animals (Fig. 4b), did not relate to run-speed biases (Fig. 4c, Extended Data Fig. 10c) or recording hemisphere (Fig. 4d, Extended Data Fig. 10a,b,f), and, notably, emerged gradually over days as performance improved and training progressed (Fig. 4e,f). Though not obviously related to imbalances in maze cues or differences in left-right performance, they displayed a weak association with right end-arm choice biases that developed late in training (Extended Data Fig. 10d,e,g). Thus end-arm biases in the ramping dopamine signals

could develop even in the absence of experimentally imposed discrepancies in value, possibly reflecting developing internal value estimates (Supplementary Discussion).

Ramping spike-firing has been recorded for putative midbrain dopamine neurons in head-fixed primates under conditions of reward uncertainty[22] and for nigral non-dopamine-containing neurons[1]. We asked whether the magnitudes of the ramping dopamine signals that we recorded in the striatum changed as performance improved on the free-choice associative T-maze task (Fig. 4e). They did not (Pearson's $R = -0.08$, $P = 0.19$; Extended Data Fig. 10h–j), suggesting that uncertainty about reward probability was unlikely to have controlled the magnitude of the ramping signals[22] (Supplementary Discussion).

Classic studies of dopamine neuron firing and striatal dopamine release have largely focused on transient responses associated with unpredicted rewards and reward-predictive cues. Here we demonstrate that, in addition to such transient dopamine responses, prolonged dopamine release in the striatum can occur, changing slowly as animals approach distant rewards during spatial navigation. These dopamine signals appear to represent the relative spatial proximity of valued goals, perhaps reflecting reward expectation[23]. It remains unclear whether these signals represent goal proximity on the basis of environmental cues, effort, or internally scaled estimates of distance. However, the brain possesses mechanisms for representing both allocentric spatial context and relative distance from landmarks[24], which could, in principle, be integrated with dopaminergic signalling to produce such extended dopamine signals.

Transient dopaminergic responses to learned reward-predictive cues have been proposed to initiate motivated behaviours[25,26], but with this mode of signalling alone, it is difficult to account for how dopamine acts to maintain and direct motivational resources during prolonged behaviours (Supplementary Discussion). The ramping dopamine signals that we describe here, providing continuous estimates of how close rewards are to being reached, and weighted by the relative values of the rewards when options are available, seem ideally suited to maintain and direct such extended energy and motivation.

Methods

All experimental procedures were approved by the Committee on Animal Care at the Massachusetts Institute of Technology and were in accordance with the US National Research Council Guide for the Care and Use of Laboratory Animals. Sample sizes were chosen based on signal variability estimates from other published studies using FSCV.

Implant procedures

Deeply anaesthetized male Long Evans rats ($n = 9$) were implanted under sterile precaution according to approved surgical procedures[19] with headstages carrying 1–3 independently movable voltammetry microsensor probes targeting the DLS (AP +0.5 mm, ML ±3.5 mm, DV 3.5–4.0 mm), of the right ($n = 3$) or left ($n = 5$) hemisphere, or the DLS bilaterally ($n = 1$), with 1–3 probes targeting the VMS of the same hemispheres (AP +1.5 mm, ML ±2.1 mm, DV 6–7 mm), and with a unilateral Ag/AgCl reference electrode in the posterior cortex (AP −2.3 mm, ML ±3.5mm, DV ~0.5mm). Five rats that underwent maze training and three additional rats for acute stimulation experiments were implanted with tungsten bipolar stimulation electrodes (FHC Inc.) straddling the ipsilateral medial forebrain bundle (MFB; AP −4.6 mm, ML ±1.3 mm, DV 7–8 mm) to verify striatal dopamine release (see below).

Behavioural training

All behavioural training was conducted on a custom built "grid maze" with fully reconfigurable tracks and walls. Approximately 4 weeks after implantation, training began on an associative T-maze task with auditory instruction cues (Figs. 1, 2, and 4)[19]. Voltammetric recordings began when animals learned to smoothly run down the track to retrieve reward. Early sessions with sporadic maze behaviour, such as wall rearing and sluggish initiation of maze running, were discarded. Daily behavioural sessions consisted of 40 trials. Each trial began with a brief warning click, followed 0.5 s later by the lowering of a swinging gate, allowing the rat to run down the maze. Half-way down the long-arm, a tone was triggered (1 or 8 kHz), indicating which end-arm to visit in order to retrieve chocolate milk reward (0.3 ml) delivered through automated syringe pumps (Pump Systems Inc.) upon the rat's arrival. The spatial position of each rat was continually monitored via video tracking (Neuralynx Inc.). Tone delivery and syringe pumps were controlled by in-house behavioural software written in MATLAB (Mathworks Inc.). After 15–35 T-maze sessions per rat, a subset of rats (n = 3) received 17 training sessions (4–6 sessions each) on the M-maze task in which the end-arms of the T-maze were extended (Fig. 3). These rats received a larger amount of reward (0.4 ml) at one goal site than at the other (0.1 ml for 2 rats and 0.2 ml for 1 rat). After 2–3 sessions with a given set of spatial reward contingencies, the reward amounts at the two goals were reversed. One rat (M31) was required to make turn choices in response to tones as in the previous T-maze task, whereas the other two rats (M36 and M47) were directed pseudo-randomly to one end-arm of the maze on each trial by removing the track to the opposite arm (20 trials to each arm) without tone presentation. Two rats were trained on the S-maze task (Fig. 3). These rats were required simply to run back and forth to retrieve a large volume of chocolate milk (0.4 ml) at one goal and a small volume (0.1 ml) at the other goal. Consecutive visits to the same reward site did not trigger the reward pumps.

Voltammetry data acquisition and analysis

Waveform generation and data acquisition for voltammetry recordings were done with two PCI data acquisition cards and software written in LabVIEW (National Instruments). Triangular voltage waveforms were applied to chronically implanted carbon fibre electrodes, relative to the reference electrode, at 10 Hz. Electrodes were held at −0.4 V between scans, and were ramped to 1.3 V and back to −0.4 V during each scan[27]. Current produced by redox reactions was recorded during voltage scans.

We compiled a library of current vs. applied voltage (C/V) templates for dopamine and pH changes of varying magnitudes by stimulating the MFB (60 Hz, 24 pulses, 300 μA) to induce dopamine release in the striatum in 5 rats maintained under isoflurane anaesthesia. We used these templates from all 5 rats as a training set to perform chemometric analysis[28] on voltammetry measurements obtained during behaviour with in-house MATLAB software. This procedure allowed us to distinguish changes in current due to dopamine release from changes due to pH or to other electroactive substances[28]. In a separate set of rats, we stimulated the MFB (10 Hz, 60 pulses, 100–120 μA) under urethane anaesthesia to mimic the slower, low amplitude ramping signals that we observed in behaving animals (Extended Data Fig. 2). Current changes were converted to estimated dopamine concentration by using calibration factors obtained from *in-vitro* measurements of fixed dopamine concentrations. Behavioural video tracking was synchronized with voltammetry recordings by marker TTL signals sent to the voltammetry data acquisition system.

For each trial, voltammetry data were normalized by subtracting average background current at each potential measured during the 1-s baseline period before warning click. Session averaged traces (Figs. 1 and 3 and Extended Figs. 3 and 10) were computed by averaging the dopamine signals recorded in a single session across 40 trials, and then averaging these

traces to obtain global averages across all rats and electrodes. Each session-averaged trace (1 from each probe from each session) was considered as an independent measurement for computing s.e.m. Concatenation of dopamine and proximity signals (see below) was performed by scaling the peri-event windows using the median inter-event intervals between consecutive events across all trials (Figs. 1–3 and Extended Data Figs. 3, 4 and 10). Traces between two consecutive events were plotted by taking data from each event to half of the median inter-event interval. Maze arm selectivity (Fig. 4 and Extended Data Fig. 10) was computed by the following equation:

$$Dopamine\ Selectivity\ Index = ([\mathrm{DA}]_{left} - [\mathrm{DA}]_{right})/([\mathrm{DA}]_{left} + [\mathrm{DA}]_{right})$$

where $[DA]_{left}$ and $[DA]_{right}$ represent dopamine concentration during trials to the left and right arms of the maze, respectively.

Session-averaged dopamine traces were identified as having positive or negative ramping characteristics (Extended Data Fig. 3) if they exhibited a significantly positive or negative linear regression coefficient (Pearson's, $R > 0.5$ or $R < -0.5$ and $P < 0.01$) over the entire trial period. Trials with phasic responses around trial start (Extended Data Fig. 4) were identified by calculating the relative difference between consecutive time points (100 ms/sample) in a 1-s window with its centre sliding in 0.1-s steps for a 1-s period from 0.5 s to 1.5 s after warning click. For a given window position, if the differences were all positive values across the first half of the window (0.5 s) and negative across the second half, we determined that a significant inflection point was present in that trial. Comparison of dopamine signals on long and short trials (Fig. 2) was done by selecting trials that fell in the bottom third (short trials) and in the top third (long trials) of the trial-time distribution for each rat that displayed a dopamine peak within 0.5 s of goal-reaching. Trials with noisy video tracking data were discarded from this analysis. The simulations for the time-elapsed model (Fig. 2b,f) were made by calculating the average slope of the ramping signals across all trials on each session and by using linear extrapolation to predict the peak dopamine values on each trial within that session. Peak values were normalized to the median peak value for each session individually and averaged for short and long trials to generate the predictions in Fig. 2f. The multi-transient model (Extended Data Fig. 6) was implemented in MATLAB and tested using a range of physiologically realistic estimates for the slope and decay times for previously observed transient signals *in vivo*[11]. For the model results shown in Extended Data Fig. 6a and c, simulations were run 100 times, and temporal offset times for the transients were determined by drawing randomly (normrnd function in MATLAB) from a normal distribution with means of 0.8 s (with standard deviation of 0.5 s) and 1.4 s (with standard deviation of 7 s) for short and long trials, respectively (Extended Data Fig. 6a, c). Spatial proximity to goal (Fig. 2c and Extended Data Fig. 7b,e) was calculated by summing the pixel differences in the "x" and "y" spatial dimensions for each recorded rat's position. These traces were averaged across all short and long trials separately to generate the traces shown in Fig 2c. Session-by-session estimations of peak dopamine concentration were made by randomly generating peak trial values using the mean and standard deviation of peak values present in the experimental data. All peak values for short and long trials (Fig. 2f and Extended Data Fig. 6b,d) were normalized to the mean dopamine concentration for each session for both simulated and experimental data.

The presence of population selectivity in the dopamine signals on the associative T-maze task (Fig. 4) was determined by comparing the variance of selectivity indices from the T-maze recordings to the distribution of variances obtained by shuffling the dopamine concentrations on the two end-arms and bootstrapping 10,000 times (Z-test comparing variance of the data to the variances of the bootstrapped distributions, $P < 0.00001$). To

identify changes in selectivity and ramping magnitude across training, Z-scores of selectivity indices (Fig. 4f) were computed for each rat by taking the absolute values of all selectivity indices, then by normalizing across sessions before combining all values across rats. Choice selectivity (Extended Data Fig. 10) was computed similarly to the selectivity score for dopamine:

$$Behavioural\ Selectivity\ Index = (Left\ Arm\ Choices - Right\ Arm\ Choices)/(Left\ Arm\ Choices + Right\ Arm\ Choices)$$

Run time biases and percent correct biases (Extended Data Fig. 10) between the two arms were also computed in this way. Correlations between peak dopamine magnitude and percent correct performance were calculated by normalizing the average peak dopamine values on each trial to the average peak value across all trials within that session.

Histology

Probe positions were verified histologically[19]. Brains were fixed by transcardial perfusion with 4% paraformaldehyde in 0.1 M $NaKPO_4$ buffer, post-fixed, washed in the buffer solution, and cut transversely at 30 μm on a freezing microtome, and they were stained with cresylecht violet to allow reconstruction of the recording sites (Extended Data Fig. 1). For a subset of the probes, a constant current (20 mA, 20 s) was passed through the probe prior to fixation to make micro-lesions at probe-tip locations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Gregory Telian, Letitia Li, Tshiamo Lechina, and Nick Hollon for help, and Drs. Min Jung Kim, Kyle S. Smith, Terrence J. Sejnowski, Leif G. Gibb and Yasuo Kubota for their comments. This work was supported by NIH grant R01 MH060379 (A.M.G.), National Parkinson Foundation (A.M.G.), CHDI Foundation grant A-5552, the Stanley H. and Sheila G. Sydney Fund (A.M.G.), a Mark Gorenberg fellowship (M.W.H.), and NIH grants R01 DA027858 and R01 MH079292 (P.E.M.P.).

References

1. Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature. 2012; 482:85–88. [PubMed: 22258508]
2. Matsumoto M, Hikosaka O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature. 2009; 459:837–841. [PubMed: 19448610]
3. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997; 275:1593–1599. [PubMed: 9054347]
4. Waelti P, Dickinson A, Schultz W. Dopamine responses comply with basic assumptions of formal learning theory. Nature. 2001; 412:43–48. [PubMed: 11452299]
5. Bayer HM, Glimcher PW. Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron. 2005; 47:129–141. [PubMed: 15996553]
6. Schultz W. Getting formal with dopamine and reward. Neuron. 2002; 36:241–263. [PubMed: 12383780]
7. Redgrave P, Gurney K. The short-latency dopamine signal: a role in discovering novel actions? Nat Rev Neurosci. 2006; 7:967–975. [PubMed: 17115078]
8. Zweifel LS, et al. Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior. Proc Natl Acad Sci U S A. 2009; 106:7281–7288. [PubMed: 19342487]

9. Wang LP, et al. NMDA receptors in dopaminergic neurons are crucial for habit learning. Neuron. 2011; 72:1055–1066. [PubMed: 22196339]
10. Day JJ, Jones JL, Wightman RM, Carelli RM. Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs. Biol Psychiatry. 2010; 68:306–309. [PubMed: 20452572]
11. Day JJ, Roitman MF, Wightman RM, Carelli RM. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. Nat Neurosci. 2007; 10:1020–1028. [PubMed: 17603481]
12. Gan JO, Walton ME, Phillips PE. Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. Nat Neurosci. 2010; 13:25–27. [PubMed: 19904261]
13. Phillips PE, Stuber GD, Heien ML, Wightman RM, Carelli RM. Subsecond dopamine release promotes cocaine seeking. Nature. 2003; 422:614–618. [PubMed: 12687000]
14. Braun AA, Graham DL, Schaefer TL, Vorhees CV, Williams MT. Dorsal striatal dopamine depletion impairs both allocentric and egocentric navigation in rats. Neurobiol Learn Mem. 2012; 97:402–408. [PubMed: 22465436]
15. Salamone JD, Correa M, Farrar A, Mingote SM. Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. Psychopharmacology (Berl). 2007; 191:461–482. [PubMed: 17225164]
16. Whishaw IQ, Dunnett SB. Dopamine depletion, stimulation or blockade in the rat disrupts spatial navigation and locomotion dependent upon beacon or distal cues. Behav Brain Res. 1985; 18:11–29. [PubMed: 3911980]
17. Salamone JD, Correa M. The mysterious motivational functions of mesolimbic dopamine. Neuron. 2012; 76:470–485. [PubMed: 23141060]
18. Niv Y, Daw ND, Joel D, Dayan P. Tonic dopamine: opportunity costs and the control of response vigor. Psychopharmacology (Berl). 2007; 191:507–520. [PubMed: 17031711]
19. Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature. 2005; 437:1158–1161. [PubMed: 16237445]
20. Niv Y, Duff MO, Dayan P. Dopamine, uncertainty and TD learning. Behav Brain Funct. 2005; 1:6. [PubMed: 15953384]
21. Tobler PN, Fiorillo CD, Schultz W. Adaptive coding of reward value by dopamine neurons. Science. 2005; 307:1642–1645. [PubMed: 15761155]
22. Fiorillo CD, Tobler PN, Schultz W. Discrete coding of reward probability and uncertainty by dopamine neurons. Science. 2003; 299:1898–1902. [PubMed: 12649484]
23. Hikosaka O, Sakamoto M, Usui S. Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. J Neurophysiol. 1989; 61:814–832. [PubMed: 2723722]
24. Derdikman D, Moser EI. A manifold of spatial maps in the brain. Trends Cogn Sci. 2010; 14:561–569. [PubMed: 20951631]
25. Flagel SB, et al. A selective role for dopamine in stimulus-reward learning. Nature. 2011; 469:53–57. [PubMed: 21150898]
26. Berridge KC. The debate over dopamine's role in reward: the case for incentive salience. Psychopharmacology (Berl). 2007; 191:391–431. [PubMed: 17072591]
27. Clark JJ, et al. Chronic microsensors for longitudinal, subsecond dopamine detection in behaving animals. Nat Methods. 2010; 7:126–129. [PubMed: 20037591]
28. Keithley RB, Heien ML, Wightman RM. Multivariate concentration determination using principal component regression with residual analysis. Trends Analyt Chem. 2009; 28:1127–1136.

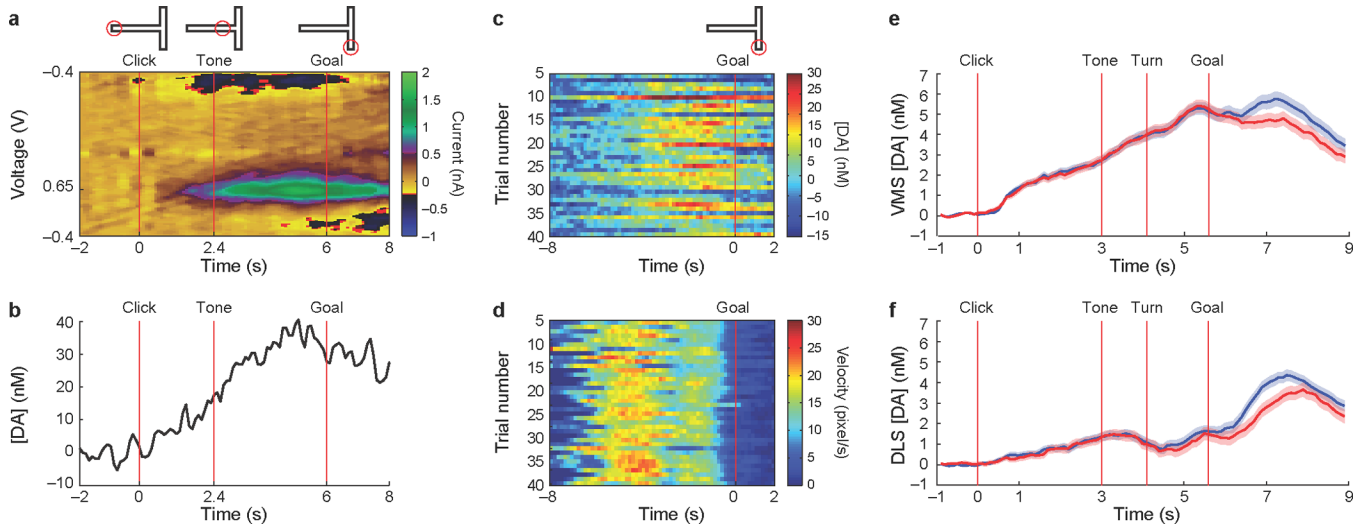

Figure 1. Ramping striatal dopamine signals occur during maze runs

a, b, Baseline subtracted current (**a**) and dopamine concentration ([DA], **b**) measured by FSCV in VMS during a single T-maze trial. **c, d** Trial-by-trial changes in dopamine concentration (**c**) and velocity (**d**) relative to goal-reaching. **e, f,** Dopamine concentration (mean ± s.e.m.) for VMS (**e**, *n* = 300 session-averaged recordings from 18 probes across 214 sessions and for DLS (**f**, *n* = 262, 13 probes) for correct (blue) and incorrect (red) trials, averaged over all 40-trial sessions.

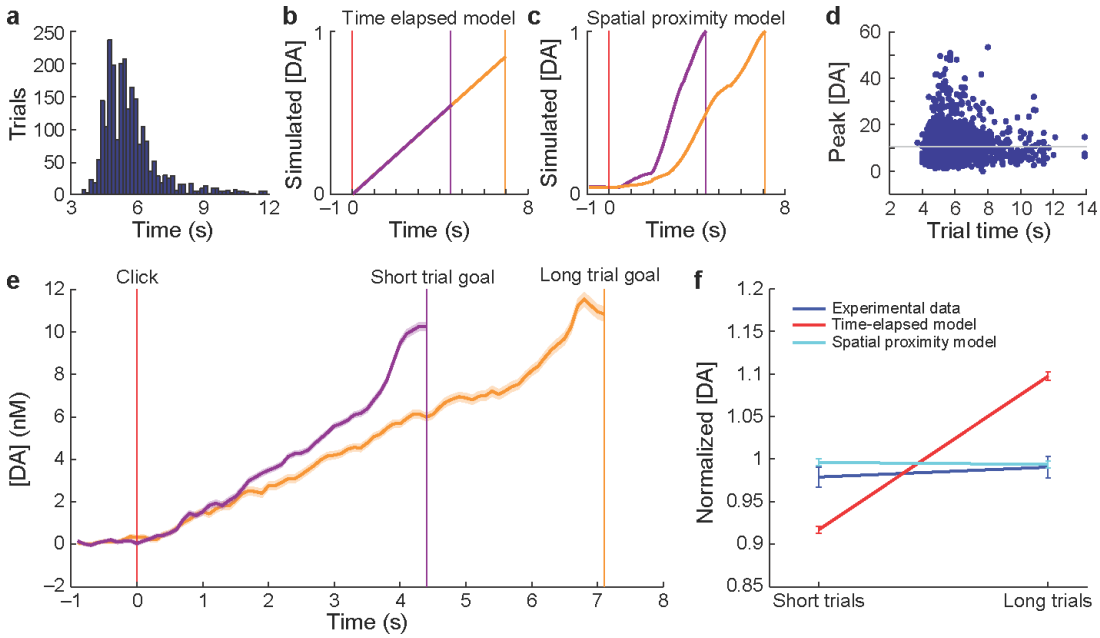

Figure 2. Ramping dopamine signals proximity to distant rewards

a, Distribution of trial times (from warning click to goal-reaching, n = 3933 trials). **b, c,** Dopamine release modelled as a function of time elapsed since maze-running onset (**b**) and as a function of spatial proximity to visited goal (**c**) for short (purple) and long (orange) trials (see Methods). Vertical lines indicate trial start (red) and end (purple and orange) times. **d,** Peak dopamine concentration vs. trial time for all ramping trials (n = 2273, Pearson's R = 0.0004, P = 0.98). **e,** Experimentally recorded dopamine release (mean ± s.e.m.) in short (n = 327, purple) and long (n = 423, orange) trials. Dopamine peaks at equivalent levels, as in proximity model in **c. f,** Normalized peak dopamine levels (mean ± s.e.m.) predicted by time-elapsed (red) and proximity (light blue) models, and measured experimental data (dark blue).

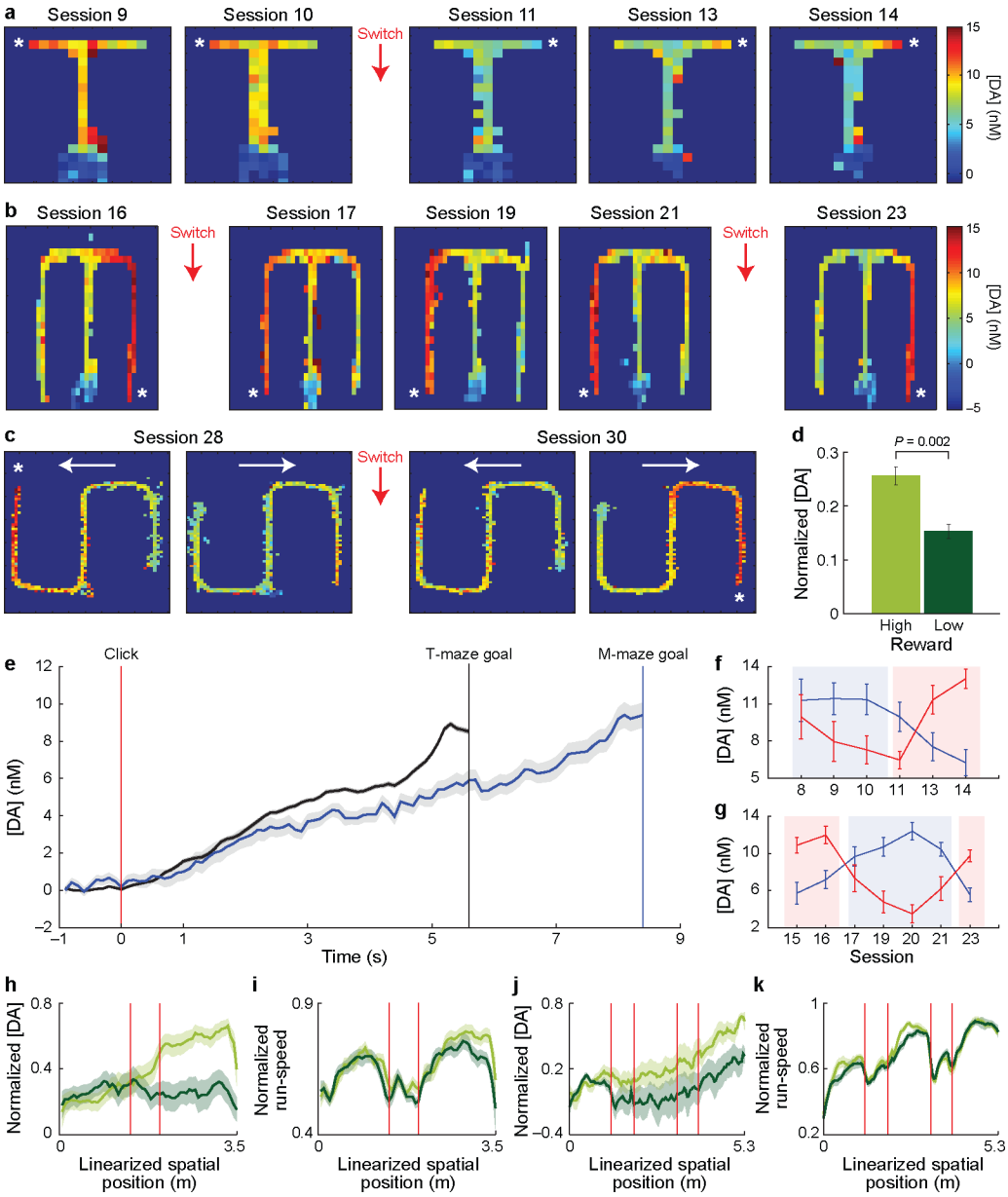

Figure 3. Dopamine ramping is sensitive to reward magnitude

a, b, Average dopamine signals from a VMS probe, for consecutive T-maze (**a**) and M-maze (**b**) sessions with asymmetric rewards. Asterisks indicate the goal with larger reward. Red arrows (and Switch) indicate reversal of reward amounts. **c,** Dopamine signals from a different rat running in the S-maze. White arrows indicate run direction. **d,** Average (± s.e.m.) peak dopamine across all value experiments (n = 4 rats). **e,** Average (± s.e.m.) VMS dopamine during T-maze (n = 44 sessions in 3 rats, black) and M-maze (n = 17, blue) sessions in same rats. **f, g,** Average (± s.e.m.) peak dopamine signals for the sessions plotted in **a** (**f**) and **b** (**g**) for trials to left (blue) and right (red) goals. Shading indicates arm with larger reward. **h, i,** Average normalized dopamine (**h**) and running speed (**i**) for runs to high (light green) and low (dark green) reward goals in the M-maze. Vertical lines indicate turns. **j, k,** Average normalized dopamine (**j**) and running speed (**k**) in the S-maze (n = 9 sessions in 2 rats), plotted as in **h** and **i**.

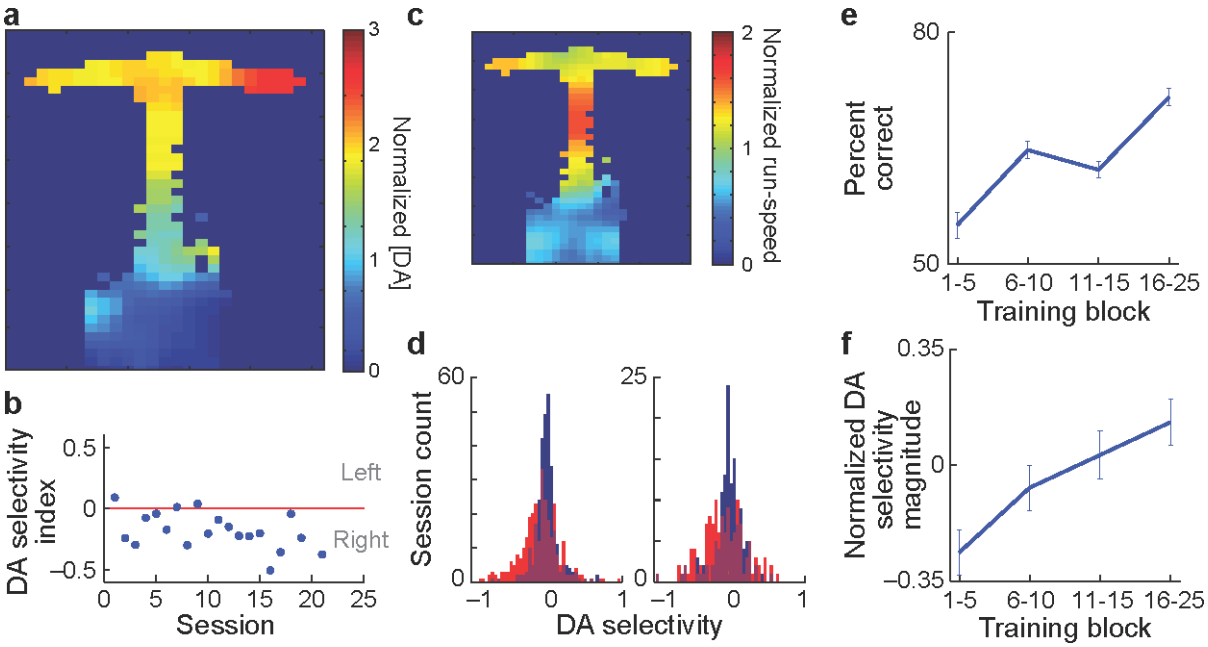

Figure 4. Ramping dopamine selectivity can emerge with training without experimentally imposed reward discrepancies

a, Average normalized dopamine at a VMS site as a function of maze location (n = 19 sessions). **b,** Dopamine selectivity indices (Methods) for all individual sessions averaged in **a. c,** Average running speed for sessions in **a. d,** Selectivity indices for all VMS (left) and DLS (right) recordings (red) compared to shuffled data (blue) for all rats (n = 9). **e, f,** Average percent correct performance (**e**) and average Z-score normalized dopamine selectivity (**f**) across training blocks. Error bars, s.e.m.