

1 **Ramping activity in midbrain dopamine neurons signifies the use of**  
2 **a cognitive map**

3  
4 Akash Guru<sup>1,2</sup>, Changwoo Seo<sup>1,2</sup>, Ryan J. Post<sup>1,2</sup>, Durga S. Kullakanda<sup>1†</sup>, Julia A. Schaffer<sup>1‡</sup>, and  
5 Melissa R. Warden<sup>1,2\*</sup>

6  
7 <sup>1</sup>Department of Neurobiology and Behavior

8 <sup>2</sup>Cornell Neurotech

9 Cornell University, Ithaca NY 14853 USA

10  
11 \* Corresponding author: [mrwarden@cornell.edu](mailto:mrwarden@cornell.edu)

---

† Current address: NYU Grossman School of Medicine, New York University, 550 1st Ave, New York, NY 10016

‡ Current address: Laboratory of Behavioral and Genomic Neuroscience, National Institute on Alcohol Abuse and Alcoholism, Rockville, MD 20852, USA.

1        **Journeys to novel and familiar destinations employ different navigational strategies. The**  
2        **first drive to a new restaurant relies on map-based planning, but after repeated trips the**  
3        **drive is automatic and guided by local environmental cues<sup>1,2</sup>. Ventral striatal dopamine rises**  
4        **during navigation toward goals and reflects the spatial proximity and value of goals<sup>3</sup>, but the**  
5        **impact of experience, the neural mechanisms, and the functional significance of dopamine**  
6        **ramps are unknown<sup>4,5</sup>. Here, we used fiber photometry<sup>6-8</sup> to record the evolution of activity**  
7        **in midbrain dopamine neurons as mice learned a variety of reward-seeking tasks, starting**  
8        **recordings before training had commenced and continuing daily for weeks. When mice**  
9        **navigated through space toward a goal, robust ramping activity in dopamine neurons**  
10       **appeared immediately – after the first rewarded trial on the first training day in completely**  
11       **naïve animals. In this task spatial cues were available to guide behavior, and although ramps**  
12       **were strong at first, they gradually faded away as training progressed. If instead mice**  
13       **learned to run a fixed distance on a stationary wheel for reward, a task that required an**  
14       **internal model of progress toward the goal, strong dopamine ramps persisted indefinitely.**  
15       **In a passive task in which a visible cue and reward moved together toward the mouse, ramps**  
16       **appeared and then faded over several days, but in an otherwise identical task with a**  
17       **stationary cue and reward ramps never appeared. Our findings provide strong evidence that**  
18       **ramping activity in midbrain dopamine neurons reflects the use of a cognitive map<sup>9,10</sup> – an**  
19       **internal model of the distance already covered and the remaining distance until the goal is**  
20       **reached. We hypothesize that dopamine ramps may be used to reinforce locations on the way**  
21       **to newly-discovered rewards in order to build a graded ventral striatal value landscape for**  
22       **guiding routine spatial behavior.**

23

1 The decision to continue pursuing a goal or abandon the quest depends on how much progress has  
2 been made, how much remains to be done, and the value of the goal. For example, a climber will  
3 be more deterred by rain at a mountain's base than near the summit, and will be more reluctant to  
4 abandon a prized peak than a training hill. Information about progress toward goals and their value  
5 is essential for adaptively balancing time and energy between activities, and commitment to goals  
6 and the vigour of goal-directed actions are both regulated by goal progress and value<sup>11-16</sup>.

7  
8 Ventral striatal dopamine (DA) progressively rises as rodents navigate toward spatially distant  
9 rewards<sup>3</sup>, a surprising recent finding that was not anticipated by temporal difference learning  
10 models of DA function<sup>4</sup> but which has broadened our understanding of the role of ventral striatal  
11 DA in sustaining and invigorating goal-directed behavior<sup>17</sup>. DA ramps reflect the value and  
12 proximity of goals, scaling by the value of the reward and stretching or compressing in different  
13 environments to span the distance between start and goal locations<sup>3</sup>. Ventral striatal DA also rises  
14 during self-initiated actions to obtain rewards<sup>18-23</sup>, and ramping activity has been observed in  
15 ventral tegmental area (VTA) DA neurons via single-unit electrophysiology and calcium  
16 imaging<sup>24-30</sup>. These observations have inspired efforts to understand the functional significance of  
17 DA ramps, which have been hypothesized to signal reward expectation, reward prediction error,  
18 discounted reward, state value, or uncertainty<sup>3-5,23,30-36</sup>. Here, we sought to probe the neural  
19 mechanisms underlying DA ramps by examining how ramping activity in VTA DA neurons  
20 evolves with experience and how this evolution is affected by the presence or absence of  
21 informative sensory cues.

22

1 We began by asking whether activity in VTA DA neurons increases as animals navigate through  
2 space toward goals, as ventral striatal DA release can occur independently of VTA DA neural  
3 activity<sup>31,37–40</sup>. The VTA of DAT-Cre mice<sup>41</sup> was injected with a Cre-dependent genetically  
4 encoded calcium indicator, AAV-DJ-EF1 $\alpha$ -DIO-GCaMP6m<sup>42</sup>, and an optical fiber was implanted  
5 over the VTA (Fig. 1a) in order to record population activity in DA neurons<sup>6–</sup>  
6 <sup>8</sup>. Mice learned a simple self-paced reward-seeking task, in which they ran back and forth on a  
7 long linear track to obtain a ‘big’ reward (10  $\mu$ L of water) at one end and a ‘small’ reward (5  $\mu$ L)  
8 at the other (Fig. 1b). Activity in DA neurons increased as mice ran toward rewards (Fig. 1c-f,  
9 Extended Data Table 1; GFP control data, Extended Data Fig. 1), and continuously rising activity  
10 was evident on individual trials (Fig. 1c, d). DA neuron activity did not reflect instantaneous run  
11 speed, which peaked near the middle of the track (Extended Data Fig. 2a, b). Ramp slope depended  
12 on reward magnitude (Fig. 1e, f) and followed the locations of the big and small rewards as they  
13 switched positions (Extended Data Fig. 2c). Surprisingly, DA ramps on the linear track emerged  
14 immediately – on the first day of training in completely naïve animals – during the first run to a  
15 goal location following a single experience with reward (Fig. 1g-k). We observed this phenomenon  
16 for runs to both the big reward and runs to the small reward (Extended Data Fig. 3). This rapid  
17 appearance of DA ramps is unlike the timescale of cue-triggered DA reward prediction error (RPE)  
18 signals, which emerge slowly after hundreds of conditioning trials<sup>43,44</sup>, but is reminiscent of the  
19 timescale of fast hippocampal remapping<sup>45–48</sup>.

20

21 On the track, and in other instrumental tasks where DA ramps have been observed, animals move  
22 to obtain rewards. However, ramping activity in DA neurons has been observed during Pavlovian  
23 conditioning under some conditions<sup>24</sup>, suggesting that DA ramps may emerge even in the absence

1 of goal-directed movement. To test this idea, we designed a ‘passive approach’ task in which the  
2 mouse did not go to the reward – instead, the reward came to the mouse (Fig. 2a). Mice initiated  
3 trials by breaking an infrared beam positioned next to a transparent wall with a small opening. A  
4 cue light and a reward spout were co-mounted on an actuator positioned 10 cm from the opening,  
5 and upon trial initiation the cue light illuminated and the cue/reward began moving toward the  
6 mouse. When the cue/reward reached the mouse, the reward could be consumed through the  
7 opening. We found that VTA DA neural activity ramped up as the distance between the mouse  
8 and the cue/reward decreased, even though no action was required during cue/reward movement  
9 (Fig. 2b-e). If the cue/reward was instead stationary and permanently located on the chamber wall,  
10 DA ramps were absent (Fig. 2d, e). Thus, observing the progress of an approaching reward was  
11 sufficient to produce ramping activity in DA neurons – goal-directed movement was not required.

12  
13 In the tasks that we have examined thus far, mice could use systematically changing sensory cues  
14 to assess their progress toward the goal. We asked whether an internal representation of goal  
15 progress was sufficient to produce DA ramps in the absence of these cues. We addressed this  
16 question by training mice to run on a wheel for a fixed distance. Mice self-initiated a trial by  
17 starting to run, and if they ran without stopping for at least 5 turns of the wheel and stopped before  
18 9 turns, a cue light was illuminated and a water reward was delivered via a spout adjacent to the  
19 wheel (‘contingent’ task, Fig. 3a, Extended Data Fig. 4). We also ran a yoked control group in  
20 which running was permitted but not required, and in which rewards were delivered on a schedule  
21 determined by the contingent group (‘non-contingent’ task, Fig. 3b, Extended Data Fig. 4). During  
22 performance of the ‘contingent’ task, DA neural activity robustly increased as mice ran, despite  
23 the absence of informative sensory cues (Fig. 3c, e, g; Extended Data Fig. 5a), and ramps were

1 evident on single trials (Fig. 3c). This finding shares some features with internally-generated  
2 hippocampal episode/time fields, which have been observed during wheel running<sup>49-51</sup>. During the  
3 ‘non-contingent’ task, performed by a separate group of mice, DA neural activity did not increase  
4 during running bouts but instead rose steeply upon cue onset/reward delivery (Fig. 3d, f, g;  
5 Extended Data Fig. 5b). Thus, an internal model of progress toward the goal is sufficient for the  
6 expression of DA ramps – informative sensory cues are not required.

7  
8 Finally, we examined the evolution of DA ramps over extended training in all tasks. Ramps on the  
9 linear track were steepest on the first and second training days, and gradually faded as daily  
10 training continued (Fig. 4a, b). Switching the spatial positions of big and small rewards  
11 reinvigorated ramps (Fig. 4a, b), as observed previously<sup>21</sup>, but did not affect run time or path length  
12 (Extended Data Fig. 6). DA ramps during passive conditioning with an approaching cue/reward  
13 took several days to develop, and gradually transitioned to cue-driven activity as training  
14 progressed (Fig. 4c, d). Ramps were never observed during passive conditioning with a stationary  
15 cue/reward (Fig. 4c, d). Unlike DA ramps during track runs and passive conditioning, DA ramps  
16 on the wheel never faded, remaining robust even over extended training (Fig. 4e, f). The wheel is  
17 unique among the tasks that we tested because all systematically changing sensory cues have been  
18 eliminated. Thus, it is not possible to develop an association between a unique sensory state and  
19 reward, and mice must instead rely on an internal model of progress toward the goal. Comparing  
20 the evolution of DA ramps in different tasks reveals that DA ramps persist indefinitely when an  
21 internal representation of progress toward the goal is required, but fade when changing sensory  
22 cues are available to guide behavior.

1 Here, we have examined how ramping activity in VTA DA neurons evolves with experience, and  
2 how this evolution is affected by environmental cues. We have shown that DA ramps appear  
3 immediately when naïve mice engage in spatial navigation – on the first trial following the initial  
4 discovery of reward. We have also shown that DA ramps gradually fade when changing sensory  
5 cues are available to guide behavior, but never fade when mice must use an internal model of goal  
6 progress. Finally, we have shown that DA ramps don't require physical movement, but can instead  
7 be produced by cues/rewards moving toward passive mice.

8  
9 Our findings strongly support the hypothesis that ramping activity in dopamine neurons reflects  
10 the use of an internal model – a cognitive map of how much progress has been made toward a goal  
11 and how much remains to be done – and as such imply that DA ramps reflect goal-directed or  
12 model-based behavioral control rather than habitual or model-free<sup>1,52,53</sup>. Our findings cast doubt  
13 on the hypothesis that DA ramps reflect the reward prediction error term in temporal difference  
14 (TD) learning algorithms. First, there simply isn't any time for the required repeated sampling and  
15 updating. The phasic DA RPE signal<sup>54</sup> requires hundreds of trials to emerge<sup>43,44</sup>, but our findings  
16 reveal that DA ramps appear after a single rewarded trial. Second, TD learning models predict that  
17 DA ramps should flatten when systematically changing sensory cues are not available and be  
18 maintained when they are available<sup>36</sup>, but here we show the opposite – strong DA ramps persist  
19 indefinitely in the absence of informative sensory cues but gradually fade when they are available.  
20 These divergent findings and predictions suggest that distinct neural mechanisms likely underlie  
21 DA ramps and DA RPE signals.

22

1 We speculate that DA ramps may reflect the transmission of goal proximity information from  
2 cortical or cortical-like brain regions, such as the hippocampus/entorhinal cortex or orbitofrontal  
3 cortex, to VTA DA neurons. The hippocampus is thought to represent a cognitive map of spatial  
4 and non-spatial relationships<sup>9,10,55-57</sup>, sends a massive projection to the ventral striatum<sup>58-60</sup>, and  
5 regulates VTA DA neural activity via the ventral striatum<sup>61,62</sup>. In addition to dorsal hippocampal  
6 place cell sequences that may reflect prospection and retrospection<sup>63-68</sup>, the distance to goals and  
7 objects is explicitly encoded in the average firing rates of ventral hippocampus and entorhinal  
8 neurons<sup>69-73</sup>. Intriguingly, goal distance encoding in human hippocampus and entorhinal cortex is  
9 suppressed when subjects are instructed to use local sensory cues and avoid thinking about goals  
10 or directions to goals<sup>74</sup>. Frontal cortical regions including the orbitofrontal cortex (OFC) have also  
11 been implicated in encoding goal proximity, and the OFC has been hypothesized to represent a  
12 cognitive map of task space<sup>53</sup>. The OFC sends a direct projection to VTA DA neurons<sup>75</sup>, and OFC  
13 lesions degrade reward expectancy signals in VTA DA neurons<sup>26</sup>, suggesting a potential role for  
14 OFC circuits in the construction of DA ramps. Consistent with this idea, OFC value signals are  
15 amplified when animals direct their gaze closer to conditioned cues<sup>76</sup>.

16  
17 How might DA ramps be used to shape learning? In laboratory environments, stimulation of VTA  
18 DA neurons robustly reinforces both place occupancy and actions at specific places<sup>77-79</sup>, but it is  
19 unknown whether DA ramps are reinforcing, and if so what they reinforce. One possibility is  
20 spatial location, as suggested by the importance of DA in conditioned place preference and the  
21 strength of hippocampal inputs to the ventral striatum<sup>58-60</sup>. By reinforcing locations encountered  
22 on the way to newly discovered rewards, dopamine ramps could build a spatial value landscape in  
23 the ventral striatum, an idea that may generalize to more abstract relational domains<sup>9</sup>. Such a

1 representation could be used as a heuristic to guide routine navigation and other habitual spatial  
2 behavior<sup>80–82</sup>, freeing cortical and cortical-like circuits for planning and other effortful, deliberative  
3 cognitive operations<sup>83</sup>

## 1      **References**

- 2      1.      Graybiel, A. M. Habits, Rituals, and the Evaluative Brain. *Annu. Rev. Neurosci.* **31**, 359–  
3      387 (2008).
- 4      2.      Collett, M., Collett, T. S., Bisch, S. & Wehner, R. Local and global vectors in desert ant  
5      navigation. *Nature* **394**, 269–272 (1998).
- 6      3.      Howe, M. W., Tierney, P. L., Sandberg, S. G., Phillips, P. E. M. & Graybiel, A. M.  
7      Prolonged dopamine signalling in striatum signals proximity and value of distant rewards.  
8      *Nature* **500**, 575–9 (2013).
- 9      4.      Niv, Y. Neuroscience: Dopamine ramps up. *Nature* **500**, 533–5 (2013).
- 10     5.      Lloyd, K. & Dayan, P. Tamping Ramping: Algorithmic, Implementational, and  
11     Computational Explanations of Phasic Dopamine Signals in the Accumbens. *PLoS*  
12     *Comput. Biol.* **11**, e1004622 (2015).
- 13     6.      Cui, G. *et al.* Concurrent activation of striatal direct and indirect pathways during action  
14     initiation. *Nature* **494**, 238–42 (2013).
- 15     7.      Gunaydin, L. A. *et al.* Natural neural projection dynamics underlying social behavior. *Cell*  
16     **157**, 1535–51 (2014).
- 17     8.      Seo, C. *et al.* Intense threat switches dorsal raphe serotonin neurons to a paradoxical  
18     operational mode. *Science* **542**, aau8722 (2019).
- 19     9.      Tolman, E. C. Cognitive maps in rats and men. *Psychol. Rev.* **55**, 189–208 (1948).
- 20     10.     O’Keefe, J. & Nadel, L. *The hippocampus as a cognitive map.* (Clarendon Press, 1978).
- 21     11.     Hull, C. L. The goal-gradient hypothesis and maze learning. *Psychol. Rev.* **39**, 25–43  
22     (1932).
- 23     12.     Brown, J. S. Gradients of approach and avoidance responses and their relation to level of

- 1 motivation. *J. Comp. Physiol. Psychol.* **41**, 450–465 (1948).
- 2 13. Tversky, A. & Kahneman, D. The framing of decisions and the psychology of choice.  
3 *Science* **211**, 453–458 (1981).
- 4 14. Liberman, N. & Förster, J. Expectancy, value and psychological distance: A new look at  
5 goal gradients. *Soc. Cogn.* **26**, 515–533 (2008).
- 6 15. Koo, M. & Fishbach, A. The Small-Area Hypothesis: Effects of Progress Monitoring on  
7 Goal Adherence. *J. Consum. Res.* **39**, 493–509 (2012).
- 8 16. Hilborn, A., Petteorelli, N., Orme, C. D. L. & Durant, S. M. Stalk and chase: How hunt  
9 stages affect hunting success in Serengeti cheetah. *Anim. Behav.* **84**, 701–706 (2012).
- 10 17. Salamone, J. D. & Correa, M. The mysterious motivational functions of mesolimbic  
11 dopamine. *Neuron* **76**, 470–85 (2012).
- 12 18. Phillips, P. E. M., Stuber, G. D., Heien, M. L. A. V, Wightman, R. M. & Carelli, R. M.  
13 Subsecond dopamine release promotes cocaine seeking. *Nature* **422**, 614–8 (2003).
- 14 19. Roitman, M. F., Stuber, G. D., Phillips, P. E. M., Wightman, R. M. & Carelli, R. M.  
15 Dopamine operates as a subsecond modulator of food seeking. *J. Neurosci.* **24**, 1265–71  
16 (2004).
- 17 20. Wassum, K. M., Ostlund, S. B. & Maidment, N. T. Phasic mesolimbic dopamine signaling  
18 precedes and predicts performance of a self-initiated action sequence task. *Biol.*  
19 *Psychiatry* **71**, 846–854 (2012).
- 20 21. Collins, A. L. *et al.* Dynamic mesolimbic dopamine signaling during action sequence  
21 learning and expectation violation. *Sci. Rep.* **6**, 20231 (2016).
- 22 22. Syed, E. C. J. *et al.* Action initiation shapes mesolimbic dopamine encoding of future  
23 rewards. *Nat. Neurosci.* **19**, 34–6 (2016).

- 1      23.    Hamid, A. A. *et al.* Mesolimbic dopamine signals the value of work. *Nat. Neurosci.* **19**,
- 2                    117–26 (2016).
- 3      24.    Fiorillo, C. D., Tobler, P. N. & Schultz, W. Discrete coding of reward probability and
- 4                    uncertainty by dopamine neurons. *Science* **299**, 1898–902 (2003).
- 5      25.    Roesch, M. R., Calu, D. J. & Schoenbaum, G. Dopamine neurons encode the better option
- 6                    in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* **10**, 1615–
- 7                    1624 (2007).
- 8      26.    Takahashi, Y. K. *et al.* Expectancy-related changes in firing of dopamine neurons depend
- 9                    on orbitofrontal cortex. *Nat. Neurosci.* **14**, 1590–1597 (2011).
- 10     27.    Totah, N. K. B., Kim, Y. & Moghaddam, B. Distinct prestimulus and poststimulus
- 11                    activation of VTA neurons correlates with stimulus detection. *J. Neurophysiol.* **110**, 75–85
- 12                    (2013).
- 13     28.    Gomperts, S. N., Kloosterman, F. & Wilson, M. A. VTA neurons coordinate with the
- 14                    hippocampal reactivation of spatial experience. *Elife* **4**, 1–22 (2015).
- 15     29.    Engelhard, B. *et al.* Specialized coding of sensory, motor and cognitive variables in VTA
- 16                    dopamine neurons. *Nature* **570**, 509–513 (2019).
- 17     30.    Kim, H. R. *et al.* A unified framework for dopamine signals across timescales. *bioRxiv*
- 18                    803437 (2019). doi:10.1101/803437
- 19     31.    Mohebi, A. *et al.* Dissociable dopamine dynamics for learning and motivation. *Nature*
- 20                    **570**, 65–70 (2019).
- 21     32.    Schultz, W. *et al.* Explicit neural signals reflecting reward uncertainty. *Philos. Trans. R.*
- 22                    *Soc. Lond. B. Biol. Sci.* **363**, 3801–11 (2008).
- 23     33.    Gershman, S. J. Dopamine ramps are a consequence of reward prediction errors. *Neural*

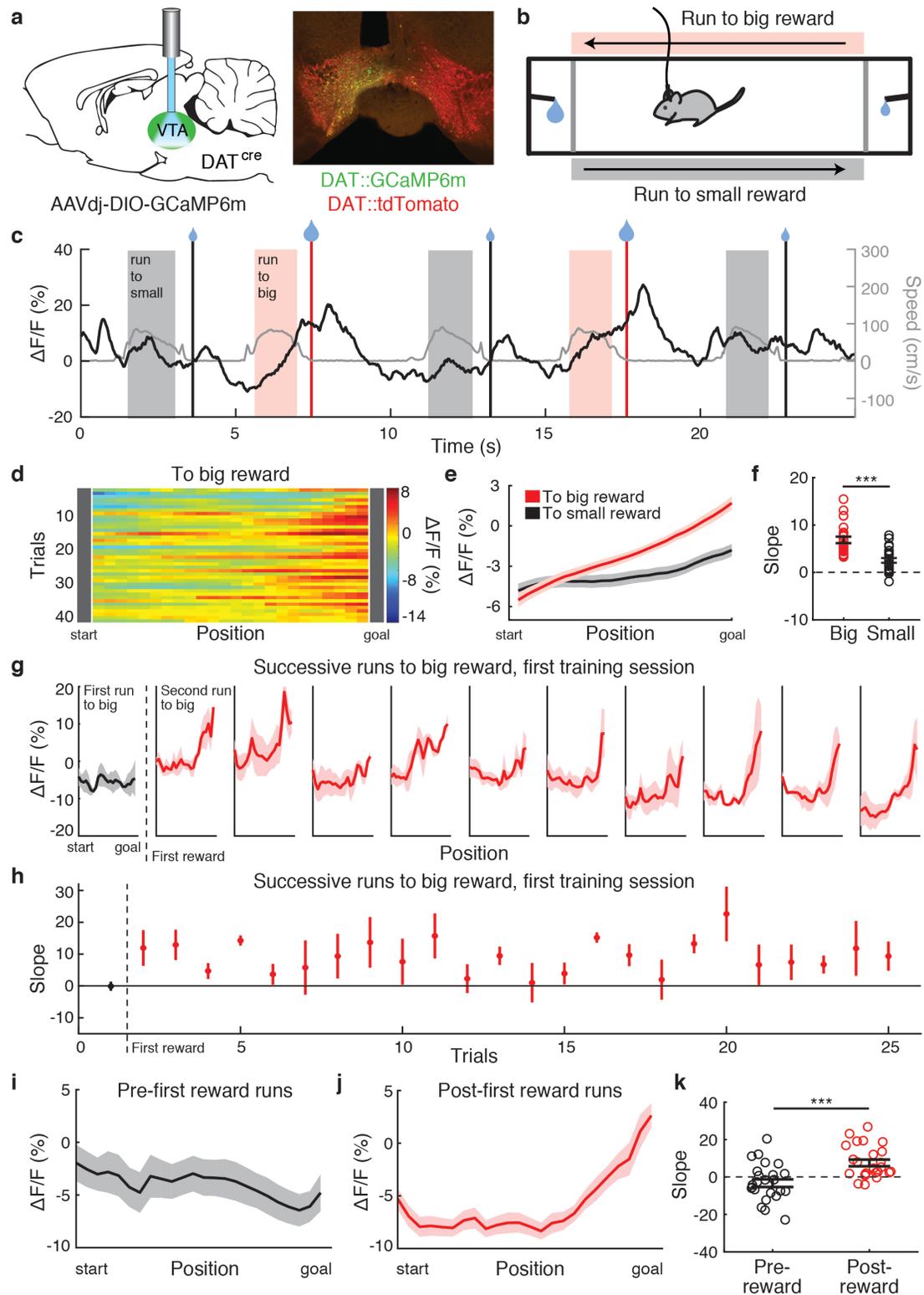
- 1            *Comput.* **26**, 467–71 (2014).
- 2        34. Morita, K. & Kato, A. Striatal dopamine ramping may indicate flexible reinforcement  
3            learning with forgetting in the cortico-basal ganglia circuits. *Front. Neural Circuits* **8**, 36  
4            (2014).
- 5        35. Berke, J. D. What does dopamine mean? *Nat. Neurosci.* **21**, 787–793 (2018).
- 6        36. Mikhael, J. G., Kim, H. R., Uchida, N. & Gershman, S. J. Ramping and State Uncertainty  
7            in the Dopamine Signal. *bioRxiv* 805366 (2019). doi:10.1101/805366
- 8        37. Floresco, S. B., Yang, C. R., Phillips, A. G. & Blaha, C. D. Basolateral amygdala  
9            stimulation evokes glutamate receptor-dependent dopamine efflux in the nucleus  
10            accumbens of the anaesthetized rat. *Eur. J. Neurosci.* **10**, 1241–51 (1998).
- 11        38. Jones, J. L. *et al.* Basolateral amygdala modulates terminal dopamine release in the  
12            nucleus accumbens and conditioned responding. *Biol. Psychiatry* **67**, 737–44 (2010).
- 13        39. Threlfell, S. *et al.* Striatal dopamine release is triggered by synchronized activity in  
14            cholinergic interneurons. *Neuron* **75**, 58–64 (2012).
- 15        40. Cachope, R. *et al.* Selective activation of cholinergic interneurons enhances accumbal  
16            phasic dopamine release: Setting the tone for reward processing. *Cell Rep.* **2**, 33–41  
17            (2012).
- 18        41. Zhuang, X., Masson, J., Gingrich, J. A., Rayport, S. & Hen, R. Targeted gene expression  
19            in dopamine and serotonin neurons of the mouse brain. *J. Neurosci.* **143**, 27–32 (2005).
- 20        42. Chen, T.-W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity.  
21            *Nature* **499**, 295–300 (2013).
- 22        43. Menegas, W., Babayan, B. M., Uchida, N. & Watabe-Uchida, M. Opposite initialization  
23            to novel cues in dopamine signaling in ventral and posterior striatum in mice. *Elife* **6**, 1–

- 1           26 (2017).
- 2       44.   Coddington, L. T. & Dudman, J. T. The timing of action determines reward prediction  
3           signals in identified midbrain dopamine neurons. *Nat. Neurosci.* **21**, 1563–1573 (2018).
- 4       45.   Muller, R. U. & Kubie, J. L. The effects of changes in the environment on the spatial  
5           firing of hippocampal complex-spike cells. *J. Neurosci.* **7**, 1951–68 (1987).
- 6       46.   Gothard, K. M., Skaggs, W. E. & McNaughton, B. L. Dynamics of mismatch correction in  
7           the hippocampal ensemble code for space: Interaction between path integration and  
8           environmental cues. *J. Neurosci.* **16**, 8027–8640 (1996).
- 9       47.   Wills, T. J., Lever, C., Cacucci, F., Burgess, N. & O’Keefe, J. Attractor dynamics in the  
10          hippocampal representation of the local environment. *Science* **308**, 873–876 (2005).
- 11      48.   Jezek, K., Henriksen, E. J., Treves, A., Moser, E. I. & Moser, M. B. Theta-paced  
12          flickering between place-cell maps in the hippocampus. *Nature* **478**, 246–249 (2011).
- 13      49.   Pastalkova, E., Itskov, V., Amarasingham, A. & Buzsáki, G. Internally generated cell  
14          assembly sequences in the rat hippocampus. *Science* **321**, 1322–7 (2008).
- 15      50.   MacDonald, C. J., Lepage, K. Q., Eden, U. T. & Eichenbaum, H. Hippocampal ‘time  
16          cells’ bridge the gap in memory for discontinuous events. *Neuron* **71**, 737–749 (2011).
- 17      51.   Kraus, B. J., Robinson, R. J., White, J. A., Eichenbaum, H. & Hasselmo, M. E.  
18          Hippocampal ‘time cells’: time versus path integration. *Neuron* **78**, 1090–101 (2013).
- 19      52.   Dolan, R. J. & Dayan, P. Goals and habits in the brain. *Neuron* **80**, 312–325 (2013).
- 20      53.   Wilson, R. C., Takahashi, Y. K., Schoenbaum, G. & Niv, Y. Orbitofrontal cortex as a  
21          cognitive map of task space. *Neuron* **81**, 267–278 (2014).
- 22      54.   Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward.  
23          *Science* **275**, 1593–9 (1997).

- 1 55. Gershman, S. J., Markman, A. B. & Otto, A. R. Retrospective revaluation in sequential  
2 decision making: A tale of two systems. *J. Exp. Psychol. Gen.* **143**, 182–194 (2014).
- 3 56. Stachenfeld, K. L., Botvinick, M. M. & Gershman, S. J. The hippocampus as a predictive  
4 map. *Nat. Neurosci.* **20**, 1643–1653 (2017).
- 5 57. Schuck, N. W. & Niv, Y. Sequential replay of nonspatial task states in the human  
6 hippocampus. *Science* **364**, (2019).
- 7 58. Kelley, A. E. & Domesick, V. B. The distribution of the projection from the hippocampal  
8 formation to the nucleus accumbens in the rat: An anterograde and retrograde-horseradish  
9 peroxidase study. *Neuroscience* **7**, 2321–2335 (1982).
- 10 59. Yuan, J. *et al.* Cell-Type-Specific Afferent Innervation of the Nucleus Accumbens Core  
11 and Shell. *Front. Neuroanat.* **12**, 1–16 (2018).
- 12 60. Van Der Meer, M. A. A., Ito, R., Lansink, C. S. & Pennartz, C. M. A. A. Hippocampal  
13 Projections to the Ventral Striatum: From Spatial Memory to Motivated Behavior. in  
14 *Space, Time and Memory in the Hippocampal Formation* (eds. Derdikman, D. & Knierim,  
15 J. J.) 497–516 (Springer Vienna, 2014).
- 16 61. Floresco, S. B., Todd, C. L. & Grace, A. A. Glutamatergic afferents from the  
17 hippocampus to the nucleus accumbens regulate activity of ventral tegmental area  
18 dopamine neurons. *J. Neurosci.* **21**, 4915–22 (2001).
- 19 62. Lisman, J. E. & Grace, A. A. The hippocampal-VTA loop: Controlling the entry of  
20 information into long-term memory. *Neuron* **46**, 703–713 (2005).
- 21 63. O’Keefe, J. & Dostrovsky, J. The hippocampus as a spatial map. Preliminary evidence  
22 from unit activity in the freely-moving rat. *Brain Res.* **34**, 171–5 (1971).
- 23 64. Foster, D. J. & Wilson, M. A. Reverse replay of behavioural sequences in hippocampal

- 1 place cells during the awake state. *Nature* **440**, 680–683 (2006).
- 2 65. Johnson, A. & Redish, A. D. Neural ensembles in CA3 transiently encode paths forward  
3 of the animal at a decision point. *J. Neurosci.* **27**, 12176–12189 (2007).
- 4 66. Dragoi, G. & Tonegawa, S. Preplay of future place cell sequences by hippocampal cellular  
5 assemblies. *Nature* **469**, 397–401 (2011).
- 6 67. Wikenheiser, A. M. & Redish, A. D. Hippocampal theta sequences reflect current goals.  
7 *Nat. Neurosci.* **18**, 289–294 (2015).
- 8 68. Ito, H. T., Zhang, S. J., Witter, M. P., Moser, E. I. & Moser, M. B. A prefrontal-thalamo-  
9 hippocampal circuit for goal-directed spatial navigation. *Nature* **522**, 50–55 (2015).
- 10 69. Royer, S., Sirota, A., Patel, J. & Buzsaki, G. Distinct Representations and Theta Dynamics  
11 in Dorsal and Ventral Hippocampus. *J. Neurosci.* **30**, 1777–1787 (2010).
- 12 70. Chrastil, E. R., Sherrill, K. R., Hasselmo, M. E. & Stern, C. E. There and back again:  
13 Hippocampus and retrosplenial cortex track homing distance during human path  
14 integration. *J. Neurosci.* **35**, 15442–15452 (2015).
- 15 71. Ciocchi, S., Passecker, J., Malagon-Vina, H., Mikus, N. & Klausberger, T. Selective  
16 information routing by ventral hippocampal CA1 projection neurons. *Science* **348**, 560–  
17 563 (2015).
- 18 72. Sarel, A., Finkelstein, A., Las, L. & Ulanovsky, N. Vectorial representation of spatial  
19 goals in the hippocampus of bats. *Science* **355**, 176–180 (2017).
- 20 73. Høydal, Ø. A., Skytøen, E. R., Andersson, S. O., Moser, M. B. & Moser, E. I. Object-  
21 vector coding in the medial entorhinal cortex. *Nature* **568**, 400–404 (2019).
- 22 74. Howard, L. R. *et al.* The hippocampus and entorhinal cortex encode the path and  
23 euclidean distances to goals during navigation. *Curr. Biol.* **24**, 1331–1340 (2014).

- 1       75.   Beier, K. T. *et al.* Circuit Architecture of VTA Dopamine Neurons Revealed by  
2           Systematic Input-Output Mapping. *Cell* **162**, 622–34 (2015).
- 3       76.   McGinty, V. B., Rangel, A. & Newsome, W. T. Orbitofrontal Cortex Value Signals  
4           Depend on Fixation Location during Free Viewing. *Neuron* **90**, 1299–1311 (2016).
- 5       77.   Tsai, H.-C. *et al.* Phasic firing in dopaminergic neurons is sufficient for behavioral  
6           conditioning. *Science* **324**, 1080–4 (2009).
- 7       78.   Witten, I. B. *et al.* Recombinase-driver rat lines: tools, techniques, and optogenetic  
8           application to dopamine-mediated reinforcement. *Neuron* **72**, 721–33 (2011).
- 9       79.   Steinberg, E. E. *et al.* A causal link between prediction errors, dopamine neurons and  
10          learning. *Nat. Neurosci.* **16**, 966–973 (2013).
- 11      80.   Graybiel, A. M. The basal ganglia and chunking of action repertoires. *Neurobiol. Learn.*  
12          *Mem.* **70**, 119–136 (1998).
- 13      81.   Packard, M. G. & Knowlton, B. J. Learning and Memory Functions of the Basal Ganglia.  
14          *Annu. Rev. Neurosci.* **25**, 563–593 (2002).
- 15      82.   Graybiel, A. M. & Grafton, S. T. The striatum: where skills and habits meet. *Cold Spring*  
16          *Harb. Perspect. Biol.* **7**, a021691 (2015).
- 17      83.   Kahneman, D. *Thinking, Fast and Slow*. (Macmillan, 2011).
- 18



1

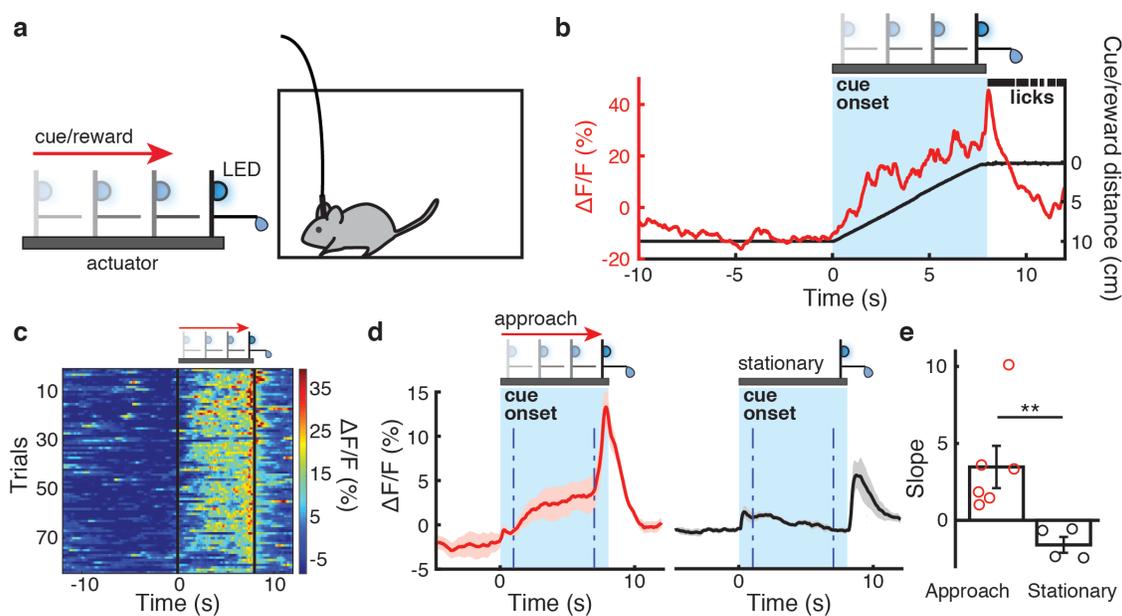
2

**Fig. 1 | Ramping activity in VTA DA neurons appears immediately after the first-ever**

3

**rewarded trial in naïve mice. a, Recording configuration, left. GCaMP6m expression in VTA**

1 DA neurons in a DAT-Cre/Ai14 mouse, right. **b**, Track schematic. **c**, Example track photometry  
2 data from a DAT::GCaMP6m mouse. GCaMP  $\Delta F/F$  in black, speed in grey. **d**, Example  $\Delta F/F$   
3 heatmap from a DAT::GCaMP6m mouse. **e**, Mean  $\Delta F/F$  with respect to position for big and small  
4 rewards (n=23; 6 animals, 3-4 sessions). **f**, Mean ramp slope during runs to big and small rewards.  
5 \*\*\* $P < 0.001$ , Wilcoxon signed-rank test **g**, Mean  $\Delta F/F$  for successive runs to the big reward  
6 during the first training session. Black – before the first-ever reward exposure; red – after the first-  
7 ever reward exposure (n=3). **h**, Ramp slope for successive runs to the big reward during the first  
8 track session. **i**, Mean  $\Delta F/F$  with respect to position for the 8 track runs to the big reward location  
9 immediately preceding the first-ever reward delivery (n=24; 3 animals, 8 trials). **j**, Mean  $\Delta F/F$  with  
10 respect to position for the 8 track runs to the big reward location immediately following the first-  
11 ever reward delivery (n=24; 3 animals, 8 trials). **k**, Mean ramp slopes for the 8 trials before and  
12 after the first-ever reward (n=24; 3 animals, 8 trials). \*\*\* $P < 0.001$ , Wilcoxon rank-sum test. Error  
13 bars indicate s.e.m.



1

2 **Fig. 2 | Approaching but not stationary conditioned cues produce VTA DA ramps. a**, Passive

3 conditioning with approaching cue/reward. **b**, Approaching cue/reward photometry example trial

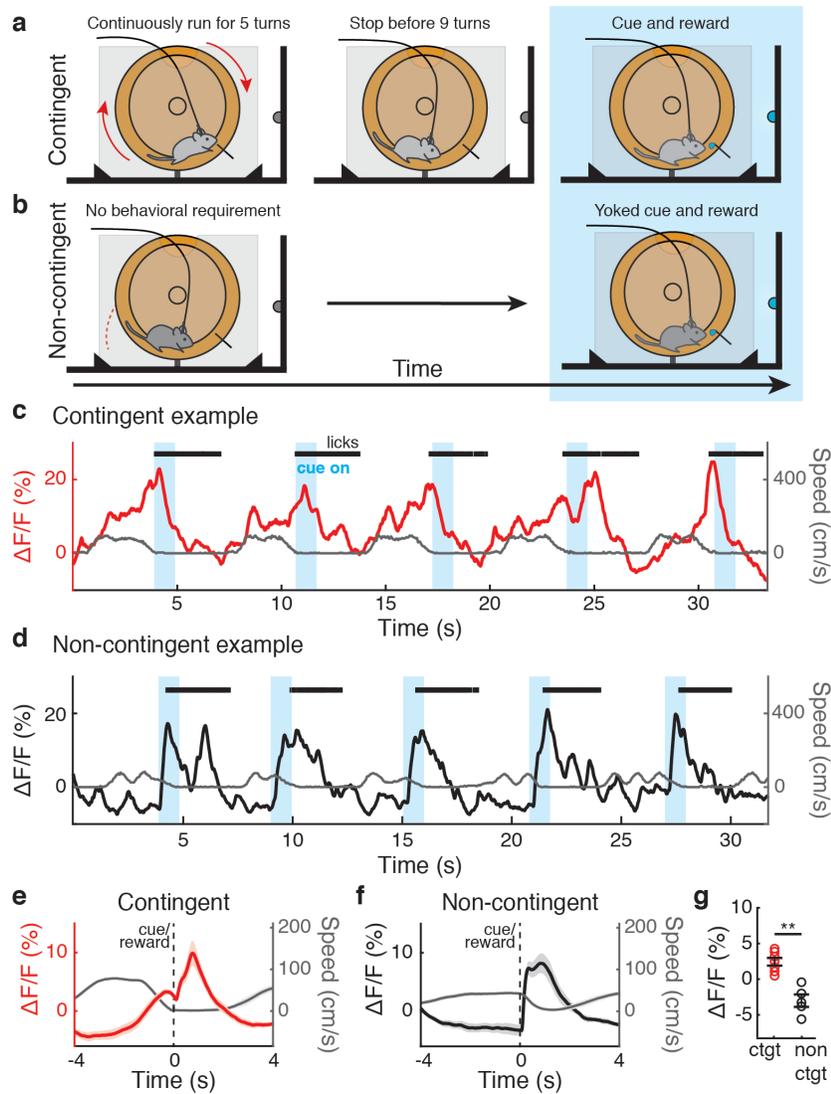
4 from a DAT::GCaMP6m mouse. GCaMP  $\Delta F/F$  in red, cue/reward distance in black, licks in black.

5 **c**, Example  $\Delta F/F$  heat map, approaching cue/reward. **d**, Mean  $\Delta F/F$  aligned to cue onset for

6 approaching cue/reward (n=6, left) and for stationary cue/reward (n=4, right). Dashed lines

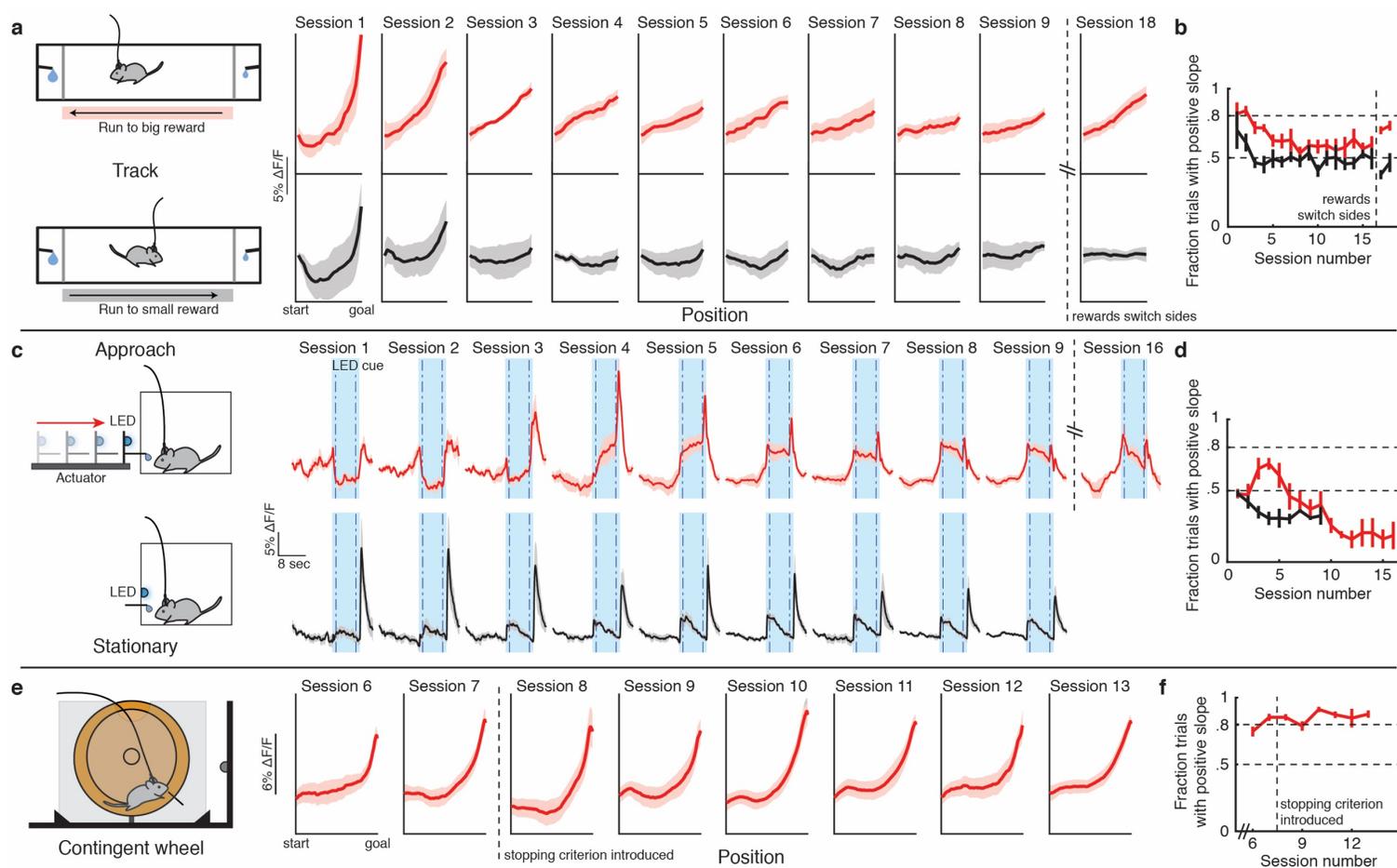
7 indicate slope analysis epoch. **e**, Mean ramp slopes for approaching and stationary cues/rewards.

8 **\*\*** $P < 0.01$ , Wilcoxon rank-sum test. Error bars indicate s.e.m.



**Fig. 3 | An internal model of progress toward a goal is sufficient to produce VTA DA ramps.**

**a**, Contingent wheel task schematic. **b**, Non-contingent wheel task schematic. **c**, Example wheel photometry data from a contingent DAT::GCaMP6m mouse. GCaMP  $\Delta F/F$  in red, speed in grey, licks in black. **d**, Example wheel photometry data from a non-contingent DAT::GCaMP6m mouse. GCaMP  $\Delta F/F$  in black, speed in grey, licks in black. **e**, Mean  $\Delta F/F$  (red) and mean speed (grey) aligned to stop-triggered cue/reward onset, contingent group (n=7). **f**, Mean  $\Delta F/F$  (black) and mean speed (grey) aligned to cue/reward onset, non-contingent group (n=5). **g**, Mean pre-cue  $\Delta F/F$ , contingent and non-contingent groups.  $**P < 0.01$ , Wilcoxon rank-sum test. Error bars indicate s.e.m.



2 **Fig. 4 | VTA DA ramps persist when an internal model of goal proximity is required but fade**  
 3 **when environmental cues are available. a**, Track schematic, left. Mean  $\Delta F/F$  with respect to  
 4 position for track runs to big (top) and small (bottom) rewards by session in DAT::GCaMP6m  
 5 mice, right (n=3 session 1, n=4 session 2, n=5 sessions 3-18). Rewards switch sides after session  
 6 16. **b**, Mean fraction of trials with positive ramp slope for track runs to big (red) and small (black)  
 7 rewards by session. **c**, Approaching and stationary conditioned cue/reward schematic, left. Mean  
 8  $\Delta F/F$  for approaching (top, n=6 session 1-10, n=3 session 11-16) and stationary (bottom, n=4  
 9 session 1-6,8,9, n=3 session 7) cues/rewards by session, right. **d**, Mean fraction of trials with  
 10 positive ramp slope for approaching (red) and stationary (black) cues/rewards by session. **e**, Wheel  
 11 schematic, left. Mean  $\Delta F/F$  with respect to position for wheel runs by session, right (n=6). **f**, Mean  
 12 fraction of trials with positive ramp slope for wheel runs by session. Error bars indicate s.e.m.

## 1 **Methods**

### 2 **Mice**

3 All procedures conformed to guidelines established by the National Institutes of Health and have  
4 been approved by the Cornell University Institutional Animal Care and Use Committee. DAT-Cre  
5 mice (The Jackson Laboratory, Bar Harbor, ME) were crossed with Ai14 Cre reporter mice (The  
6 Jackson Laboratory, Bar Harbor, ME) to generate DAT-Cre/Ai14 mice. DAT-Cre/Ai14 or DAT-  
7 Cre mice (postnatal 2-10 months) were used for DA-specific viral vector expression. All Cre driver  
8 lines were fully backcrossed to C57BL/6J mice. Both male and female mice were used. All mice  
9 were maintained on a 12-hour reverse light-dark cycle with *ad libitum* access to food and water,  
10 except during behavioral experiments.

### 11

### 12 **Viral vectors**

13 We used AAV-DJ-EF1 $\alpha$ -DIO-GCaMP6m (Stanford Vector Core, Stanford, CA) or AAV5-CAG-  
14 Flex-GFP (UNC Vector Core, Chapel Hill, NC) for photometry experiments.

### 15

### 16 **Surgery**

17 Mice were anaesthetized using isoflurane (5%) and prepared for stereotactic surgery. Fur was  
18 shaved and mice were placed in a stereotaxic frame (Kopf Instruments, Tujunga, CA). A heating  
19 pad was placed under the mice to prevent hypothermia. Isoflurane (1-2%) was delivered via a nose  
20 cone throughout the surgery. Ophthalmic ointment was used to protect the eyes. Buprenorphine  
21 (0.05 mg/kg, subcutaneous) was given before the start of surgery. A mixture of 0.5% lidocaine and  
22 0.25% bupivacaine (100  $\mu$ L) was injected intradermally along the incision line. The scalp was  
23 disinfected with betadine scrub and 70% ethanol. An incision was made using a scalpel along the  
24 scalp midline. The exposed skull was thoroughly cleaned, and a craniotomy was made above the  
25 VTA.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24

Virus was targeted to the VTA (-3.1 AP, 0.3-0.4 ML, 4.6 & 4.3 DV), and slowly pressure-injected (100 nl/min) using a 10  $\mu$ L Hamilton syringe (nanofil; WPI, Sarasota, FL), a 33-gauge bevelled needle, and a micro-syringe pump controller (Micro 4; WPI, Sarasota, FL). After each injection, the needle was left in place for 10 minutes and then slowly withdrawn. A total of 800 nl (500 nl at -4.6 DV and 300 nl at -4.3 DV) of vector was injected. Animals were implanted with a 400  $\mu$ m diameter, 0.48 NA optical fiber (Doric Lenses, Québec, Canada) above the VTA (-3.1 AP, 0.3-0.4 ML, 4.3-4.4 DV). A layer of metabond (Parkell, Inc., Edgewood, NY) and dental acrylic (Lang Dental Manufacturing, Wheeling, IL) was applied to firmly hold the fiber in place, and the surrounding skin was sutured. Post-operative buprenorphine (0.05 mg/kg), carprofen (5 mg/kg), and lactated ringers (500  $\mu$ l) were administered subcutaneously. Behavioral testing began after a minimum 3-week viral expression period.

### **Perfusion and Histological Verification**

Animals were deeply anaesthetized with Fatal-Plus at a dose of 90 mg/kg and transcardially perfused with 20 ml of PBS (phosphate-buffered saline), followed by 20 ml of 4% paraformaldehyde solution. Brains were quickly extracted and stored in 4% paraformaldehyde solution at 4°C for 24 hrs. Brains were then transferred to 30% sucrose in PBS solution and allowed to equilibrate for 2 days. Brains were then mounted on a freezing microtome using OCT and sectioned coronally (45  $\mu$ m). Sections were washed in PBS and mounted on slides with PVA-DABCO. Images were acquired using a Zeiss LSM 800 confocal scanning laser microscope with a 5X and 20X air objective.

### **Fiber Photometry**

1 Fiber photometry was performed as previously described<sup>8</sup>. 473 nm and 405 nm data were collected  
2 for all GCaMP and GFP photometry experiments using a custom-built fiber photometry system.  
3 473 nm and 405 nm diode lasers (Omicron Luxx, Rodgau-Dudenhofen, Germany) were modulated  
4 at 700 Hz and 500 Hz, respectively, using a dual optical chopper (New Focus Model 3502,  
5 Newport, Irvine, CA). The laser beams were combined using a mirror (KM100-E02, Thorlabs,  
6 Newton, NJ) and a dichroic filter (LM01-427, Semrock, Rochester, NY), and the power was  
7 adjusted using a neutral density filter to 15-80  $\mu$ W. The laser beams were then aligned to a  
8 fluorescence filter cube (DFMB, Thorlabs, Newton, NJ). Within the cube, the excitation beams  
9 were reflected with a dichroic filter (FF495, Semrock, Rochester, NY), collimated (F240FC-A  
10 with AD11F, Thorlabs, Newton, NJ), and coupled to an optical patch cord (400  $\mu$ m, Doric Lenses,  
11 Quebec, Canada). The optical patch cord was connected to the animal using a zirconia sleeve  
12 (ADAF1, Thorlabs, Newton, NJ). Emitted fluorescence received through the same optical patch  
13 cord was passed through a dichroic filter (FF495, Semrock, Rochester, NY) and a bandpass filter  
14 (FF03-525/50, Semrock, Rochester, NY), and was collected by a femtowatt photoreceiver (New  
15 Focus 2151, Newport, Irvine, CA). Voltage from the photoreceiver was channelled through two  
16 lock-in amplifiers (SR810 DSP, Stanford Research System, Sunnyvale, CA), digitized using a  
17 DAQ (U6 pro, LabJack, Lakewood, CL) at a sampling rate of 250 Hz, and recorded by LabJack  
18 software.

## 20 **Behavioral Testing**

21 Mice were water restricted prior to all behavioral testing. Body mass was measured daily to  
22 maintain at least 80% of baseline body weight.

23  
24 *Linear track:* A 122 x 12.5 cm custom-made plexiglass track contained reward ports on both ends.  
25 Licking was monitored using dual contact lickometers (ENV-250B, Med Associates, Fairfax, VT)

1 when mice completed the circuit between the reward spout and a copper plate placed beneath the  
2 reward port. Water rewards were delivered upon lick initiation using variable-rate infusion syringe  
3 pumps (PHM-107, Med Associates, Fairfax, VT) on both ends. Mice were required to alternate  
4 between the reward ports to trigger reward delivery. 10  $\mu$ l of water was delivered at one port and  
5 5  $\mu$ l of water was delivered at the other port. The location of the big reward (10  $\mu$ l) and the small  
6 reward (5  $\mu$ l) was held constant for a given session but was sometimes switched between sessions,  
7 as indicated. Behavioral testing was conducted under dark/red light and each session lasted 20-30  
8 min per day. Video was recorded using an infrared (IR) USB 3.0 camera (Chameleon3, Pointgrey,  
9 Richmond, BC, Canada), which provided TTL pulses for every recorded frame. Lick timestamps,  
10 reward delivery timestamps, and video TTL pulses were sent to the photometry DAQ for  
11 synchronized recording. Video data were processed offline using Ethovision-XT (Noldus  
12 Information Technology, VA) to collect position and velocity data. Reward port access was  
13 blocked at the beginning of each session for 5 minutes while mice were allowed to freely explore  
14 the track and the photometry signal was recorded. For the linear track experiments in Figure 4, a  
15 linear track (132 x 12.5 cm) was modified with IR sensors and controllers (ENV-253SD, ENV-  
16 253 Med Associates, Fairfax, VT). The IR sensors were placed 22 cm from the reward port on  
17 both sides and rewards were delivered when the mouse exited the reward port and crossed the  
18 closest IR sensor. The locations of big (8  $\mu$ l) and small (4  $\mu$ l) rewards were kept constant for 16  
19 sessions and switched for session 17 and 18. Photometry data was recorded during each session.

20  
21 *Running wheel:* The behavioral apparatus consisted of a running wheel (Comfort Exercise Wheel,  
22 Kaytee Products Inc., Chilton, WI) mounted on a breadboard (Thorlabs, Newton, NJ). The  
23 breadboard was kept flat or at an angle of 20 degrees to facilitate running. The wheel was  
24 connected to an optical encoder (E4T 200 cycles per revolution, US Digital, Vancouver, WA)  
25 using a custom-made shaft to record wheel movement. A blue LED visual cue was mounted in

1 front of the wheel. Clear plexiglass covered the front of the wheel. A metal lick spout was attached  
2 to the plexiglass and was connected to a custom capacitive lickometer. Water reward was delivered  
3 using a syringe pump (NE-500, New Era Pump Systems, Farmingdale, NY). All hardware  
4 operation and behavioral data acquisition (1 KHz sampling rate) were controlled by a data  
5 acquisition card (USB-6212, National Instruments, Austin, TX) with custom LabVIEW code.

6  
7 At the start of training, water-restricted mice were allowed to run on the wheel (Comfort Exercise  
8 Wheel, Kaytee Products Inc., Chilton, WI) in their home cage for 3-4 days. The next day they were  
9 introduced to the behavioral chamber and were allowed to run on the wheel without reward while  
10 photometry data was recorded.

11  
12 The next day, mice were randomly split into 2 groups – contingent and non-contingent. In the  
13 contingent group, reward delivery was contingent on correct behavioral performance, while  
14 rewards were independent from behavior for the non-contingent group. Contingent group mice  
15 were required to run a minimum of 5 wheel-turns (1000 encoder units) without stopping, and then  
16 stop (wheel speed zero) on their own without any external cues. A continuous run of at least this  
17 distance followed by a self-initiated stop resulted in 1 second of blue light via the LED and delivery  
18 of 7  $\mu$ l water. Mice collected the reward and then initiated the next trial. If mice stopped before 5  
19 turns, the encoder count was reset, and no reward was delivered. Velocity was calculated every 50  
20 ms from the encoder count. After 7 sessions, an upper bound run distance was introduced. Mice  
21 were required to run at least 5 wheel-turns (1000 encoder units) and stop before 9 wheel-turns  
22 (1800 encoder units) to get a reward. If they stopped after 9 wheel-turns, the encoder count was  
23 reset, and no reward was delivered. The animals then initiated the next trial.

1 Mice in the non-contingent group did not have any behavioral requirements. They were not  
2 required to run and stop to collect the reward, but they typically ran. Every day, the timing of the  
3 rewards/cues from a mouse in the contingent group was used to deliver yoked rewards/cues to the  
4 non-contingent group of animals. The reward schedules and exposure to the wheel were kept the  
5 same for both groups.

6  
7 *Passive Approaching Cue/Reward:* Animals were placed in a clear plexiglass chamber with a small  
8 opening for reward delivery on the left side. The chamber had 2 IR sensors (E18-D80NK,  
9 SMAKN, Shenzhen, GD, China), one on each side, and a blue LED on the right side. A linear  
10 actuator (L12-100-50-12-P, Actuonix, BC, Canada) was positioned 10 cm away from the clear  
11 plexiglass left wall. A custom 3D-printed holder was mounted on a micro linear slide rail (S9-100,  
12 Actuonix, BC, Canada), which was connected to the linear actuator. A blue LED and a metal lick  
13 spout connected to a custom lickometer were placed on the holder, and the actuator could be  
14 programmed to move the LED and the lick spout together to the opening in the left wall. The whole  
15 box was placed in a sound-attenuating cubicle (ENV-018MD, Med Associates, Fairfax, VT). The  
16 behavioral chamber was controlled using a data acquisition card (USB-6212, National Instruments,  
17 Austin, TX) with custom code written in LabVIEW, and behavioral data was acquired by the  
18 photometry data acquisition system to allow for synchronized behavioral and neural recording.

19  
20 Mice were habituated to the behavioral chamber for 1-2 days with a fixed spout and delivery of  
21 free reward at constant intervals. After habituation, mice were exposed to the full task. Trials were  
22 initiated by breaking the IR beam at the chamber wall closest to the actuator, on the left side. Upon  
23 trial initiation the actuator LED illuminated, and the actuator began moving the LED and the  
24 reward spout toward the wall opening. The mouse was not required to continuously block the IR  
25 sensor during actuator movement. The actuator speed was controlled using a voltage regulator,

1 and it took 8 seconds for the actuator to fully approach. The mouse had full visual access to actuator  
2 position at all times because the actuator LED was visible through the clear plexiglass walls. The  
3 actuator stopped and the LED extinguished when the reward spout reached the wall opening.  
4 Delivery of a 7  $\mu$ l water reward was triggered by the first lick. If there was no lick within 15  
5 seconds of spout availability, the actuator retracted, and the trial ended. After each trial there was  
6 an 8 second inter-trial interval for actuator retraction when a new trial could not be initiated. The  
7 end of the inter-trial interval was indicated by LED illumination on the right chamber wall. Mice  
8 were required to turn off this LED by blocking the right-side IR sensor before initiating a new trial.

9  
10 *Passive Stationary Cue/Reward:* Mice were placed in a plastic chamber with a blue LED and a  
11 reward spout connected to a custom lickometer. The chamber was located inside a sound-  
12 attenuating cubicle (Med Associates, Fairfax, VT). Each trial began with the illumination of a blue  
13 LED. After 8 seconds, the LED turned off and a 7  $\mu$ l water reward was delivered using a variable-  
14 rate infusion syringe pump (PHM-107, Med Associates, Fairfax, VT). Licks were detected using  
15 a custom lickometer. Inter-trial intervals were drawn from an exponential distribution with a scale  
16 parameter of 13 seconds, and a subset of inter-trial intervals between 9 and 40 seconds was used.  
17 The behavioral program was controlled using a data acquisition card (USB-6212, National  
18 Instruments, Austin, TX) with custom code written in LabVIEW.

19

## 20 **Data analysis**

21 All data analysis and statistical tests were performed using custom code in MATLAB  
22 (MathWorks, Natick, MA). The 473 nm and 405 nm fluorescent channels were low-pass filtered  
23 at 15 Hz. For the linear track task, the photometry data was then downsampled to 30 Hz using the  
24 video TTL pulse times. For all other experiments, the photometry data was not downsampled. The

1 whole-session 405 nm reference channel was fit to the 473 nm channel using linear least squares.  
2 Relative fluorescence changes, reported as  $\Delta F/F$ , were calculated using the following equation:

$$3 \frac{\Delta F}{F_0} = \frac{473 \text{ nm signal-fitted 405 nm signal}}{\text{mean 473 nm signal}} \times 100$$

4  
5  
6 For example traces, the  $\Delta F/F$  was detrended. For Fig. 1g, h and Extended Data Fig. 3a, b  $\Delta F/F$  was  
7 not detrended or lowpass filtered. For all other analyses, the  $\Delta F/F$  was lowpass filtered using a  
8 Butterworth filter with a cutoff frequency of 0.03 Hz, which was then subtracted from the  $\Delta F/F$  to  
9 remove very slow fluorescence fluctuations.

10  
11 *Linear track:* The mouse's position on the track was binned and the mean  $\Delta F/F$  for each bin was  
12 calculated in order to produce position-aligned photometry traces. Trials were averaged for each  
13 animal, and population plots were constructed from averages across animals. Speed was processed  
14 similarly. For slope calculation, the coefficients were calculated by linearly fitting each trial with  
15 polyfit (MATLAB). For fraction of trials with positive slope, the number of trials with positive  
16 slope was divided by the total number of trials. Normalized path length was calculated by summing  
17 the total distance moved in the x and y directions and normalizing by the length of the track.

18  
19 *Running wheel:* A run was defined as continuous wheel movement for at least 0.5 wheel-turns  
20 (100 encoder counts), and run onset was defined as either the moment this threshold was crossed  
21 (with respect to time) or 100 encoder counts (with respect to position). Wheel speed was defined  
22 as the distance moved per 50 ms bin measured by counting the number of encoder samples. Each  
23 wheel-turn corresponded to a distance of 42 cm. The total encoder count was reset when this  
24 number reached zero, which was defined as run offset. Only trials with at least 2.5 wheel-turns

1 (500 encoder counts) were considered for analysis. Trials were rewarded if the total distance was  
2 between 5 (1000 encoder count) and 9 (1800 encoder count) wheel-turns. For  $\Delta F/F$  by position  
3 plots, normalized position on each trial ranged from 0 to 1, starting from run onset and stopping at  
4 run offset as defined above.  $\Delta F/F$  traces were averaged for each animal and then averaged across  
5 animals for population plots. Slope was calculated for each trial using polyfit (MATLAB). The  
6 fraction of trials with positive slope was calculated using only correct trials and was defined as the  
7 number of trials with positive slope divided by the total number of trials. For all wheel analyses  
8 across sessions, 1 mouse was excluded because of file corruption. Extended Data Fig. 4d, the  
9 number of correct trials during the first 50 trials of a session was used to calculate percent correct  
10 trials.

11  
12 *Approaching and Stationary Conditioned Cue/Reward:* Slope was calculated for each trial using  
13 data between 1 and 7 seconds following cue/actuator onset. Trials were averaged for each animal  
14 and session averages were obtained by averaging across animals. For fraction trials with positive  
15 slope, the number of trials with positive slope was divided by the total number of trials.

16  
17 *Significance testing, fraction of trials with positive slope:* The hierarchical bootstrap method<sup>84</sup> was  
18 used (Fig. 4, Extended Data Table 1). For each session, the number of animals was sampled with  
19 replacement, and for each sampled animal trials were sampled with replacement. Then, the  
20 proportion of trials with positive slope was calculated. This procedure was repeated 10,000 times.  
21 Finally, the probability of the observed fraction of positive slopes given the null hypothesis of an  
22 equal fraction of positive and negative slopes was determined from the bootstrap population. This  
23 probability was tested for significance with an error rate of 0.05, and Bonferroni correction was  
24 used for multiple comparisons.

1     **Data Availability:** The data that support the findings of this study are available from the  
2     corresponding author upon reasonable request.

3

4     **Code Availability:** Code is available from the corresponding author upon reasonable request.

5

6     **Methods References:**

7     84.     Saravanan, V., Berman, G. J. & Sober, S. J. Application of the hierarchical bootstrap to  
8     multi-level data in neuroscience. *bioRxiv* 819334 (2019). doi:10.1101/819334

1     **Acknowledgments:** We thank I.T. Ellwood, R.R. Hoy, J.H. Goldberg, J.R. Fetcho, H.K. Reeve,  
2     A.K. Churchland, E.L. Troconis, B.J. Sleezer, Y.Y. Ho, W. Gu, Y. Baumel, D.A. Bulkin, R. Chen,  
3     A. Roeser, and T. Bollu for helpful discussions; I.T. Ellwood, R.R. Hoy, J.H. Goldberg, A.H. Bass,  
4     R.M. Harris-Warrick, C. Linster, N. Yapici, M.J. Sheehan, D.M. Smith, B.J. Sleezer, Y.Y. Ho, and  
5     R. Chen for comments on the manuscript; T. Bollu and B.M. Kardon for technical advice and  
6     lickometer design; and A.K. Recknagel for technical assistance. This work was supported by the  
7     Mong Family Foundation (A.G., C.S.), and by NIH DP2MH109982, the Alfred P. Sloan  
8     Foundation, the Whitehall Foundation, and the Brain and Behavior Research Foundation  
9     (M.R.W.). M.R.W. is a Robertson Neuroscience Investigator–New York Stem Cell Foundation.

10

11     **Author Contributions:** A.G. and M.R.W. conceived the project and designed the experiments;  
12     A.G. performed stereotaxic surgery; A.G., C.S., R.J.P., D.S.K., and J.A.S. developed behavioral  
13     data acquisition methods, conducted photometry experiments and performed histology; C.S. and  
14     A.G. developed photometry acquisition methods; A.G. and M.R.W. analysed the data and prepared  
15     the manuscript; M.R.W. supervised all aspects of the work.

16

17     **Competing Interests:** The authors declare no competing interests.

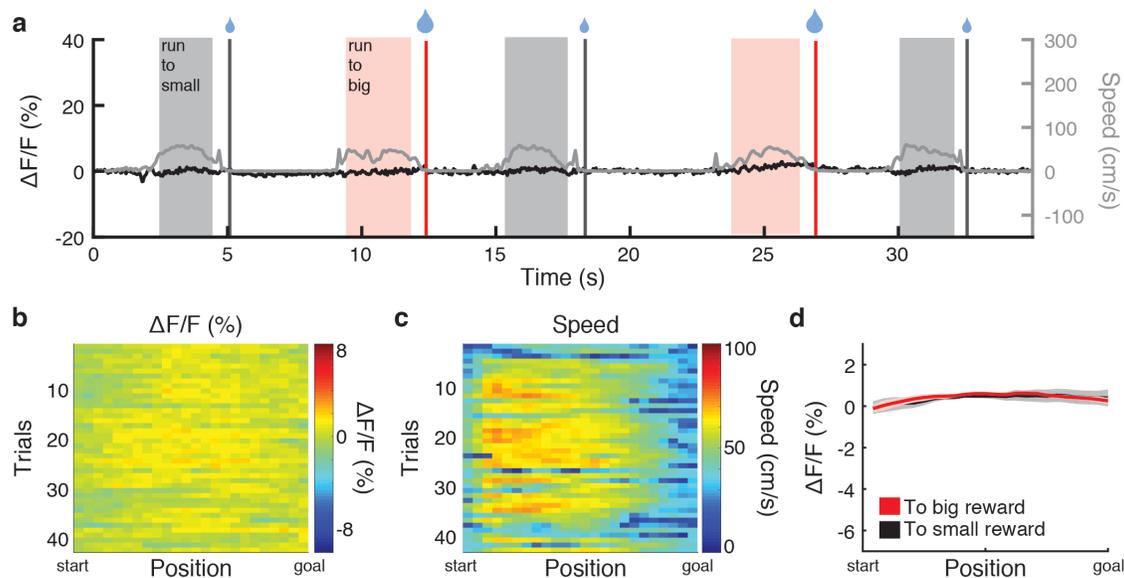
18

19     **Additional Information:**

20     **Extended Data** is available for this paper.

21     **Correspondence and requests for materials** should be addressed to M.R.W.

22



1

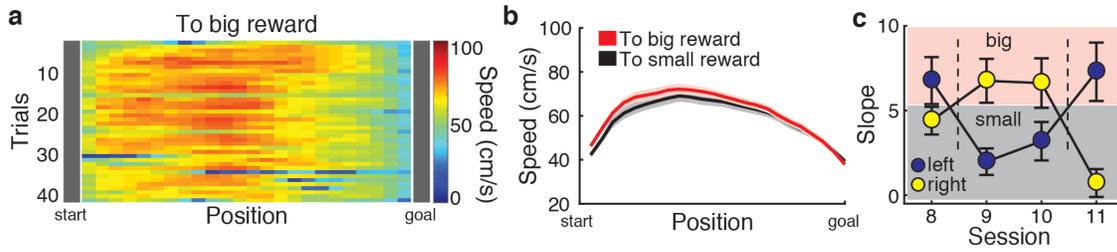
2

3

4

5

**Extended Data Fig. 1 | GFP control data.** **a**, Example track photometry data from a DAT::GFP mouse. GFP  $\Delta F/F$  in black, speed in gray. **b**, Example  $\Delta F/F$  and **c**, speed heatmaps from a DAT::GFP mouse. **d**, Mean GFP  $\Delta F/F$  with respect to position for big and small rewards ( $n=3$ ). Error bars indicate s.e.m.



1

2

**Extended Data Fig. 2 | Running speed and reward-switching VTA DA ramp slope data. a,**

3

Example track speed heatmap from a DAT::GCaMP mouse. **b,** Mean speed during runs to big and

4

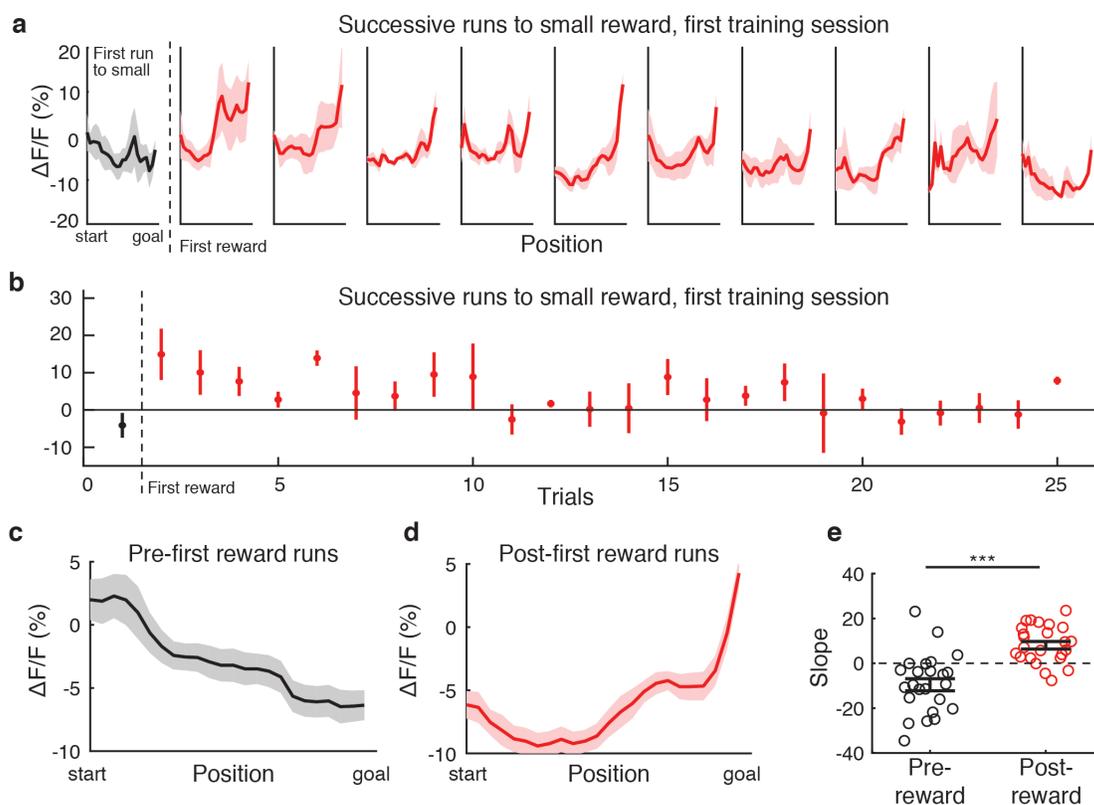
small rewards (n=23; 6 animals, 3-4 sessions). **c,** Mean ramp slope following switches in reward

5

position (n=6 for sessions 8, 9, and 11, and n=5 for session 10). Dashed lines indicate switches.

6

Error bars indicate s.e.m.



1

2

3

4

5

6

7

8

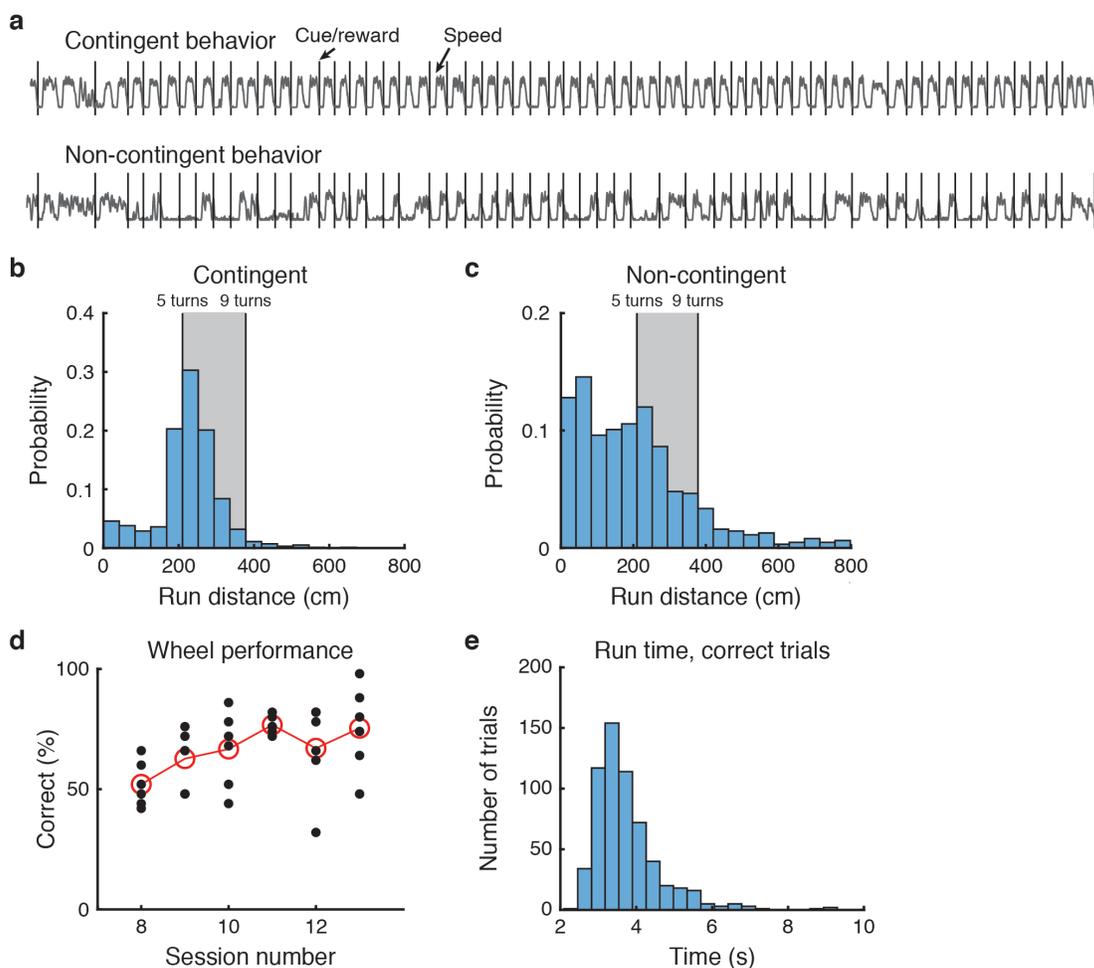
9

10

11

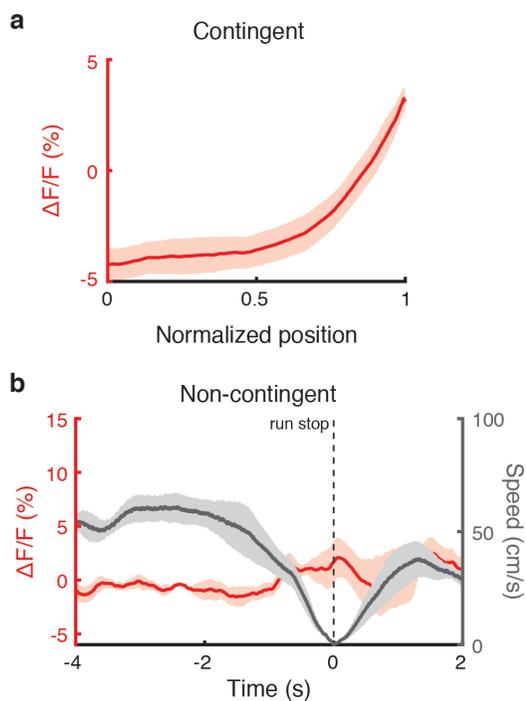
12

**Extended Data Fig. 3 | VTA DA ramps during runs to the small reward appear after the first rewarded trial at that location. a**, Mean  $\Delta F/F$  for successive runs to the small reward during the first training session. Black – before the first reward at the small goal location; red – after the first reward ( $n=3$ ). **b**, Ramp slope for successive runs to the small reward during the first track session. **c**, Mean  $\Delta F/F$  with respect to position for the 8 track runs to the small reward location immediately preceding the first reward delivery ( $n=24$ ; 3 animals, 8 trials). **d**, Mean  $\Delta F/F$  with respect to position for the 8 track runs to the small reward location immediately following the first reward delivery ( $n=24$ ; 3 animals, 8 trials). **e**, Mean ramp slopes for the 8 trials before and after the first reward ( $n=24$ ; 3 animals, 8 trials).  $***P < 0.001$ , Wilcoxon rank-sum test. Error bars indicate s.e.m.



1

2 **Extended Data Fig. 4 | Behavioral performance during wheel running.** **a**, Example wheel data  
3 from a contingent mouse (top) and a non-contingent mouse (bottom). Cue/reward delivery in  
4 black, speed in gray. **b**, Run distance histogram, contingent group. Rewarded distance in gray. **c**,  
5 Run distance histogram, non-contingent group. **d**, Wheel performance (percent correct) by session,  
6 first 50 trials per session. Mean in red. **e**, Run times for correct runs during the contingent wheel  
7 task.



1

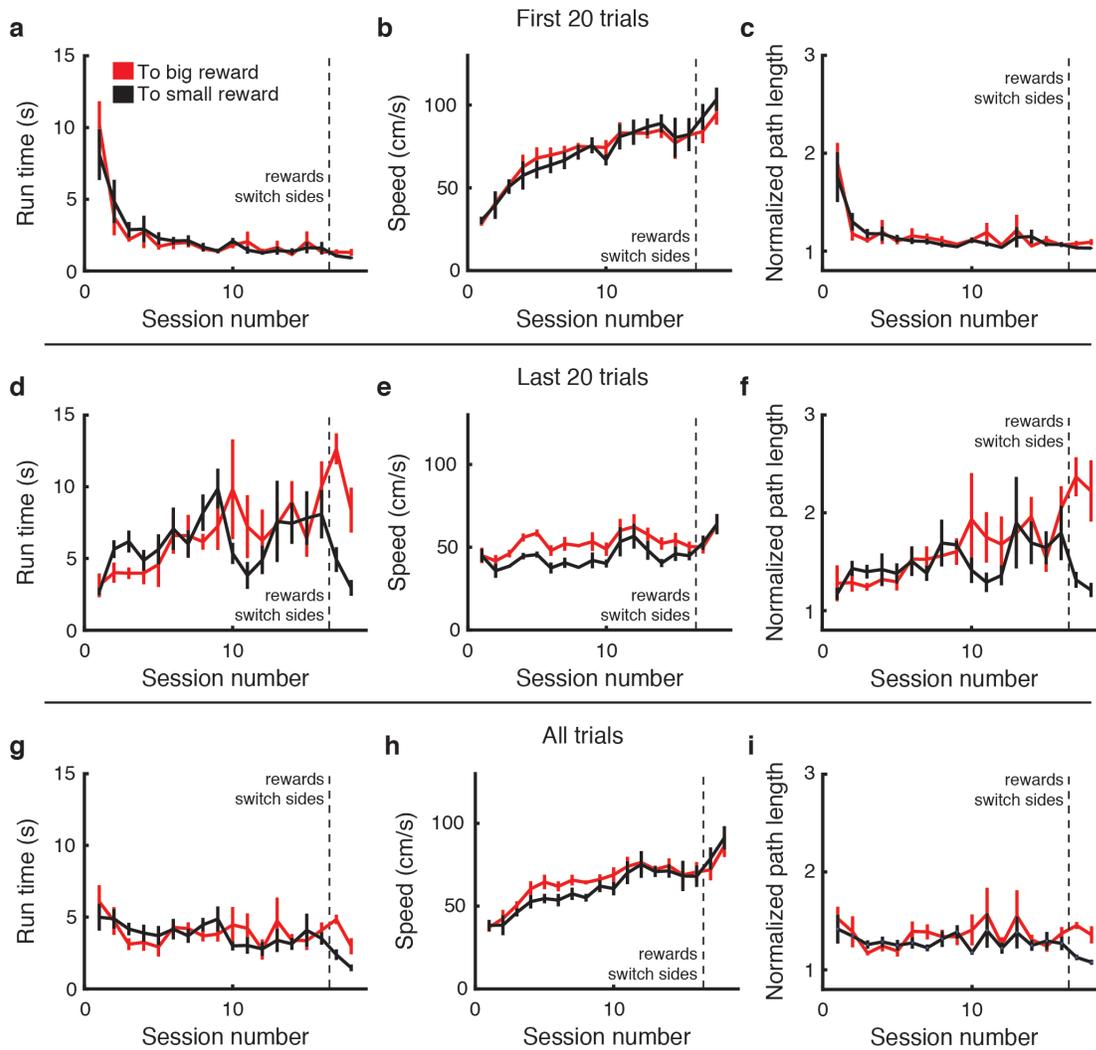
2

3

4

5

**Extended Data Fig. 5 | VTA DA neural activity during wheel running. a,** Mean  $\Delta F/F$  with respect to position, correct runs, contingent group. **b,** Mean  $\Delta F/F$  (red) and speed (gray) aligned to stop, runs between 5 and 9 wheel turns with no rewards, non-contingent group. Error bars indicate s.e.m.



1

2

**Extended Data Fig. 6 | Behavioral performance during navigation toward rewards, track. a,**

3

Mean run time for track runs to big (red) and small (black) rewards by session, first 20 trials per

4

session (n=3 session 1, n=4 session 2, n=5 sessions 3-18). **b,** Mean speed by session, first 20 trials.

5

**c,** Mean normalized path length by session, first 20 trials. **d, e, f,** Same as a, b, c but using last 20

6

trials per session. **g, h, i,** Same as a, b, c but using all trials per session. Error bars indicate s.e.m.

Figure	Number of animals or trials	Statistical test	Comparison	Significance	P-value
1 f	n =23 (6 animals 4 sessions)	Wilcoxon signed-rank test (two-tailed)	Big vs small reward	***	P<0.0001
1 k	n =24 (3 animals 8 trials)	Rank-sum test (two-tailed)	Pre vs Post reward	***	0.0003
2 e	N=6 Approach N=4 Classical	Rank-sum test (two-tailed)	Approach vs stationary	**	0.0095
3 g	n=7 Contingent n=5 Non-contingent	Rank-sum test (two-tailed)	Contingent vs Non-contingent	**	0.0025
4 a, b	N=3, sess. 1 N= 4, sess. 2 N=5, sess. 3-18	Hierarchical Bootstrap, Bonferroni correction	Probability that the null hypothesis is correct. (one-tailed)		

Session	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Big	p<0.05	p<0.05	p<0.05	p<0.05	N.S	p<0.05	p<0.05											
small	N.S	N.S	N.S	N.S	N.S	N.S	N.S	N.S	N.S	N.S	N.S	N.S	N.S	N.S	N.S	N.S	N.S	N.S

4 c, d	Approach N=6 sess. 1-10 N=3 sess. 11-16 Stationary N=4 N=3, sess. 7	Hierarchical Bootstrap, Bonferroni correction	Probability that the null hypothesis is correct. (one-tailed)		
--------	---------------------------------------------------------------------------------	-----------------------------------------------	---------------------------------------------------------------	--	--

Session	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Approach	N.S	N.S	N.S	p<0.05	N.S											
Classical	N.S	N.S	N.S	N.S	N.S	N.S	N.S	N.S	N.S							

4 f	N=6	Hierarchical Bootstrap, Bonferroni correction	Probability that the null hypothesis is correct. (one-tailed)		
-----	-----	-----------------------------------------------	---------------------------------------------------------------	--	--

Session	6	7	8	9	10	11	12	13
Wheel	p<0.05							

Extended Figure 3 e	n =24 (3 animals 8 trials)	Rank-sum test	Pre vs Post reward	***	P<0.0001
---------------------	-------------------------------	---------------	--------------------	-----	----------

1

2 **Extended Data Table 1 | Statistical Results.**