Article

Natural behaviour is learned through dopamine-mediated reinforcement

https://doi.org/10.1038/s41586-025-08729-1

Received: 17 June 2024

Accepted: 3 February 2025

Published online: 12 March 2025

Check for updates

Jonathan Kasdin[1,4], Alison Duffy[2,4], Nathan Nadler[1], Arnav Raha[1], Adrienne L. Fairhall[2], Kimberly L. Stachenfeld[1,3] & Vikram Gadagkar[1]✉

Many natural motor skills, such as speaking or locomotion, are acquired through a process of trial-and-error learning over the course of development. It has long been hypothesized, motivated by observations in artificial learning experiments, that dopamine has a crucial role in this process. Dopamine in the basal ganglia is thought to guide reward-based trial-and-error learning by encoding reward prediction errors[1], decreasing after worse-than-predicted reward outcomes and increasing after better-than-predicted ones. Our previous work in adult zebra finches—in which we changed the perceived song quality with distorted auditory feedback—showed that dopamine in Area X, the singing-related basal ganglia, encodes performance prediction error: dopamine is suppressed after worse-than-predicted (distorted syllables) and activated after better-than-predicted (undistorted syllables) performance[2]. However, it remains unknown whether the learning of natural behaviours, such as developmental vocal learning, occurs through dopamine-based reinforcement. Here we tracked song learning trajectories in juvenile zebra finches and used fibre photometry[3] to monitor concurrent dopamine activity in Area X. We found that dopamine was activated after syllable renditions that were closer to the eventual adult version of the song, compared with recent renditions, and suppressed after renditions that were further away. Furthermore, the relationship between dopamine and song fluctuations revealed that dopamine predicted the future evolution of song, suggesting that dopamine drives behaviour. Finally, dopamine activity was explained by the contrast between the quality of the current rendition and the recent history of renditions—consistent with dopamine's hypothesized role in encoding prediction errors in an actor–critic reinforcement-learning model[4,5]. Reinforcement-learning algorithms[6] have emerged as a powerful class of model to explain learning in reward-based laboratory tasks, as well as for driving autonomous learning in artificial intelligence[7]. Our results suggest that complex natural behaviours in biological systems can also be acquired through dopamine-mediated reinforcement learning.

Complex behaviours, such as speaking or producing music, are often learned by comparing ongoing performance to internal goals or templates without explicit external reinforcement in the form of rewards or punishments. The neural mechanisms that underlie internally guided trial-and-error learning of motor skills are poorly understood. This is because much of our knowledge of the brain mechanisms of trial-and-error learning comes from studies of animals learning simple tasks motivated by external rewards such as food or juice[8,9]. Basal-ganglia-projecting midbrain dopamine neurons in the ventral tegmental area (VTA) are thought to provide a reinforcement signal for reward-based trial-and-error learning by encoding reward prediction errors (RPEs)[1]; they increase their firing on unexpected rewards or reward-predicting cues, and decrease their firing when expected rewards are withheld. Whether developmental learning of complex natural behaviours also occurs through dopamine-mediated reinforcement is not known.

Vocal learning, an essential substrate for spoken language, is a rare trait exhibited by only a few animal groups, including humans and songbirds[10–12]. Similar to speech acquisition, song learning in zebra finches requires auditory feedback and involves a developmental critical period during which juveniles learn to make a copy of an adult tutor's song through internally guided trial and error[13–15]. The zebra finch brain contains a specialized song system[16] dedicated to learning and producing song, including a dopaminergic projection from the VTA to the song system nucleus Area X[17,18], the singing-related basal ganglia. Previous work, by changing the perceived quality of specific song syllables with

[1]Department of Neuroscience, Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA. [2]Department of Neurobiology and Biophysics and Computational Neuroscience Center, University of Washington, Seattle, WA, USA. [3]Google DeepMind, New York, NY, USA. [4]These authors contributed equally: Jonathan Kasdin, Alison Duffy. ✉e-mail: vikram.gadagkar@columbia.edu

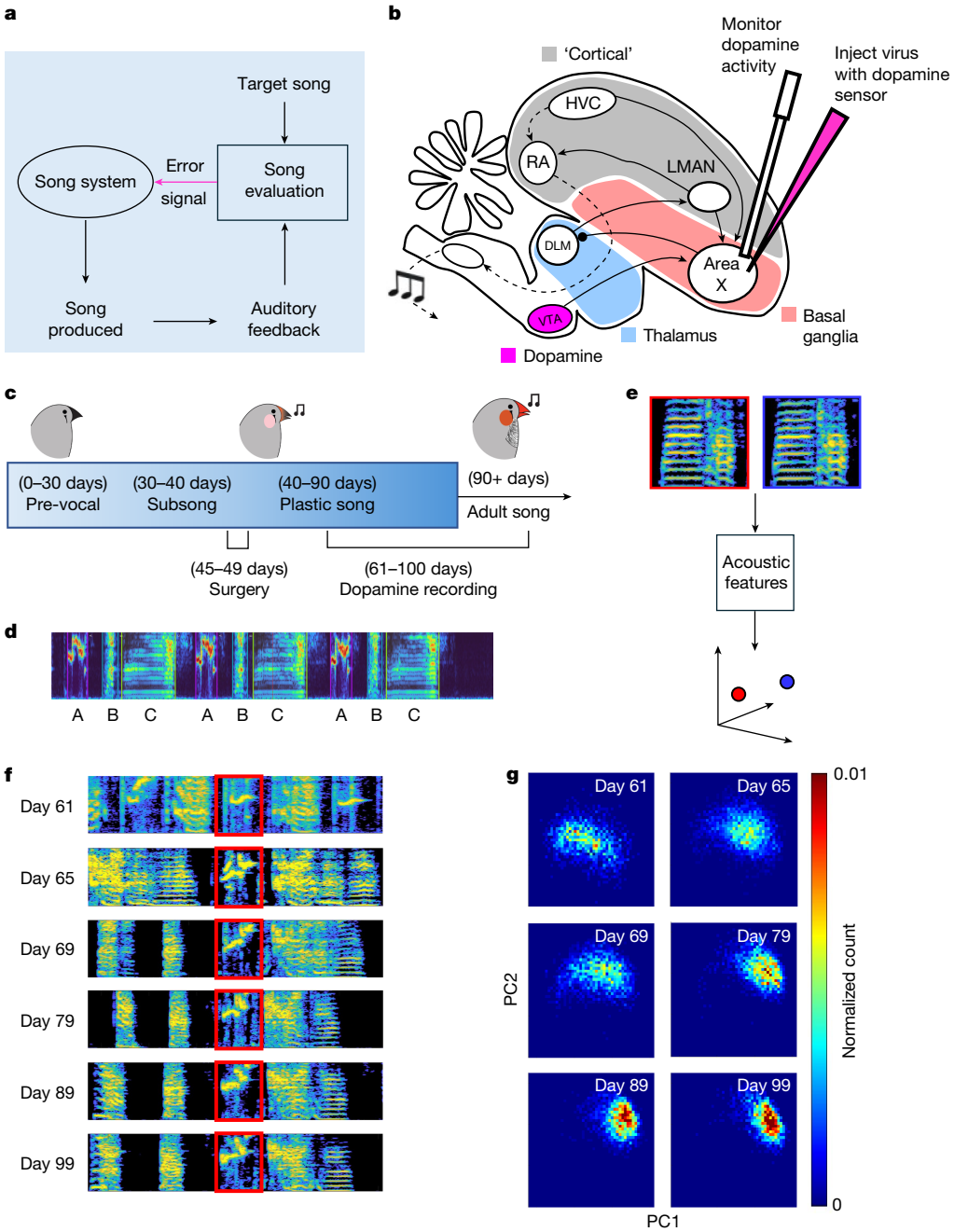


Fig. 1 | Synchronous recording of developmental song learning and dopamine. a, Dopamine is hypothesized to provide an error signal to the song system that guides developmental vocal learning. **b**, Dopamine release in Area X, the singing-related basal ganglia, was recorded using fibre photometry during song learning. DLM, medial nucleus of the dorsolateral thalamus; HVC, high vocal centre; LMAN, lateral magnocellular nucleus of the anterior nidopallium; RA, robust nucleus of the arcopallium. **c**, Stages of zebra finch song learning along with surgery and recording timelines. **d**, Song syllables were labelled using a deep-neural-network-based annotation tool (example syllables denoted by A, B, C)[32]. **e**, Acoustic features were extracted for each labelled song syllable using MATLAB's Audio Toolbox and renditions were placed in acoustic feature space. **f**, An example syllable (inside red box) on six different days as it is refined across development. **g**, For the same syllable and example days shown in **f**, each rendition is projected onto the first two acoustic principal components (PC1 and PC2), showing refinement across development.

distorted auditory feedback[19,20] in adult birds, showed that these VTA_X neurons provide an evaluative signal by encoding performance prediction error (PPE)[2,21,22]; they were suppressed after distorted syllables and activated after undistorted syllables. In addition, optogenetics experiments showed that activating VTA_X terminals in Area X on high-pitch syllable renditions induced the birds to increase the pitch of these syllables[23,24], and that disrupting dopamine signalling to Area X impaired pitch modification[23,25]. Thus, it has been shown that dopamine can drive changes to adult song and that dopamine activity is consistent with encoding PPEs in artificial settings. However, it remains unknown whether dopamine guides song learning through reinforcement[26–28] during the developmental critical period, which we set out to test.

Tracking dopamine during song learning

To test whether dopamine activity resembles a reinforcement signal during song learning (Fig. 1a), we recorded developmental song trajectories[29,30] in juvenile male zebra finches and concurrently measured

dopamine release in Area X using dopamine fibre photometry[3] (Fig. 1b). Male zebra finch hatchlings ($n = 6$) were raised naturally by their parents in our breeding colony until the age of 47 ± 2 days, after which they were isolated; memorization of the tutor's song typically occurs by 35 days of age[31]. Injection of a virus encoding the dopamine sensor[3] $GRAB_{DA3h}$ into Area X was followed by implantation of an optical cannula at age 47 ± 2 days (see Extended Data Fig. 1). Song learning trajectories and dopamine activity were recorded between 61 and 100 days of age—a period that includes both the plastic song phase of learning, when syllables can be identified but are still variable both in sequence and in spectral content, as well as ten days of early adulthood (Fig. 1c; see Methods). Zebra finch song typically crystallizes when birds reach adulthood at around 90 days of age. Song syllables ($n = 25$ syllables across 6 birds) were annotated by training a deep neural network (Deep Audio Segmenter; DAS[32]) on the entire recorded song trajectory, and all subsequent analyses were performed for each song syllable independently (Fig. 1d and Methods). Acoustic features were extracted for each labelled song syllable using MATLAB's Audio Toolbox and renditions were parametrized in a 124-dimensional acoustic feature space (Fig. 1e and Methods). Figure 1f,g shows the evolution of an example syllable across development, which gets progressively refined as the song crystallizes. Analysis of synchronously recorded song trajectories and associated dopamine activity allowed us to ask whether and how dopamine relates to song learning.

Dopamine encodes relative syllable quality

If dopamine guides song learning, we would expect high-quality syllables (renditions that are closer to their adult version than renditions in recent history) to be associated with high dopamine, and low-quality syllables (renditions that are further away from their final form than recent renditions) to be associated with low dopamine. For each syllable rendition, we calculated the proximity (Euclidean distance) to the median of the adult renditions of the syllable (age > 90 days) in feature space (Fig. 2a,b). Because dopamine is hypothesized to encode the relative value of performance[2], we defined the 'relative quality' of a syllable as proximity to the adult version of the syllable relative to the mean of the previous 11 renditions (Methods).

To test for online error responses, for each day we compared the dopamine activity associated with the 10% of syllable renditions with the highest relative quality with that of the 10% of renditions with the lowest relative quality (Fig. 2c–g). As previously reported, the baseline level of dopamine in Area X is known to increase during singing[22]; to isolate the effect of syllable quality from these singing-related dopamine increases, we subtracted the mean dopamine response across all renditions of a syllable on each day from the dopamine response for each rendition. The mean-subtracted data are presented in Fig. 2 (see also Methods and Extended Data Fig. 2). Figure 2c shows the dopamine response for an example syllable, with each row of the heat map showing the average across a day. For a quantification of the rendition-to-rendition variability in dopamine responses, see Extended Data Fig. 3. Figure 2d shows dopamine responses averaged across all $n = 25$ syllables. Figure 2e shows the dopamine responses averaged across all days for each syllable sorted by the magnitude of the dopamine response. The variability in dopamine responses between syllables could not be explained by variability in distance to adult syllable, duration or change in distance to adult syllable within or across days (Extended Data Fig. 4). Figure 2f shows that the dopamine response systematically tracks syllable quality from the bottom 10% to the top 10% across all ten deciles of relative distance. Dopamine in Area X was significantly increased after renditions of high relative quality in 24 out of 25 syllables and significantly decreased after renditions of low relative quality in 25 out of 25 syllables in our dataset (Methods). Activations followed syllable onset with a latency of 273 ± 67 ms; suppressions mirrored the activations, following syllable onset with a latency of 242 ± 60 ms (Fig. 2h). We defined the error response as the difference between the average traces for the top and bottom 10% relative quality of syllable renditions; 25 of 25 syllables showed a significant error response (Methods). Although song is known to exhibit circadian oscillations, with song quality improving during the course of a day, these oscillations had only a subtle influence on the dopamine responses (Extended Data Fig. 5).

In addition to error signalling, dopamine is known to correlate with movement vigour[33], which might be reflected in syllable amplitude during singing. We observed no significant change in the dopamine error response when syllable amplitude was normalized to its median value (Extended Data Fig. 6). Together, the observed precisely timed phasic activations and suppressions suggest that dopamine provides an evaluative signal during song learning, with syllable renditions that are closer to the final version relative to recent renditions resulting in high dopamine and renditions that are further away from the adult version relative to recent renditions resulting in low dopamine.

Dopamine predicts future song evolution

We next asked whether this dopaminergic evaluative signal guides song learning by influencing future song. If dopamine activity in Area X is not just correlated with syllable quality but drives song learning, fluctuations in dopamine should predict song movement along the learning trajectory.

To test whether dopamine fluctuations predict changes during song learning, we used canonical correlation analysis (Methods) to find the direction in acoustic feature space along which song is maximally correlated with dopamine within individual blocks of 150 consecutive song renditions (Fig. 3a). This method allows us to identify a direction in song space along which dopamine is strongly co-varying with song, and thus potentially reinforcing, without having to select an a priori performance quality or similarity measure. We label this direction the 'dopamine vector', which is defined locally for each song block and varies across blocks and development. We then projected future song renditions onto the dopamine vector and asked whether future song moved in the direction of high dopamine along this acoustic dimension (Methods). Figure 3b shows one example song block (150 consecutive renditions). The coloured circles show individual renditions, where colour represents a scalar measurement of dopamine response on each rendition, and the black circle is the median location along the first and second acoustic principal component (PC) dimensions of renditions within this block (Methods). The grey crosses represent individual renditions in a future song block, starting 500 renditions later within the same day, and the black cross is the median location of this future song block. The red arrow depicts the dopamine vector computed for the original song block using the first 23 PC dimensions (see Methods). The black arrow depicts the direction of song movement, measured as the difference between the median song locations of the future and present song blocks.

In Fig. 3c, we plot the movement of song (in sliding blocks of 150 consecutive renditions) along the original dopamine vector over 600 renditions into the future (on a single day). In this song-block example, future song moves in the direction of high dopamine. We then repeated this procedure over many days of development with the local dopamine vectors defined by the local 'current' song block (current song blocks were measured every 25 renditions). This analysis allowed us to ask whether, on average, future song moved along the dopamine vector defined by the current song. Figure 3d shows the average movement of the median of future song blocks along the dopamine vector for the same example syllable, with the average computed across all local dopamine vectors in the recorded song trajectory across development (see Extended Data Fig. 7 for more example syllables).

We then assessed the significance of this effect across syllables. Movement along the dopamine vector averaged across all 25 syllables

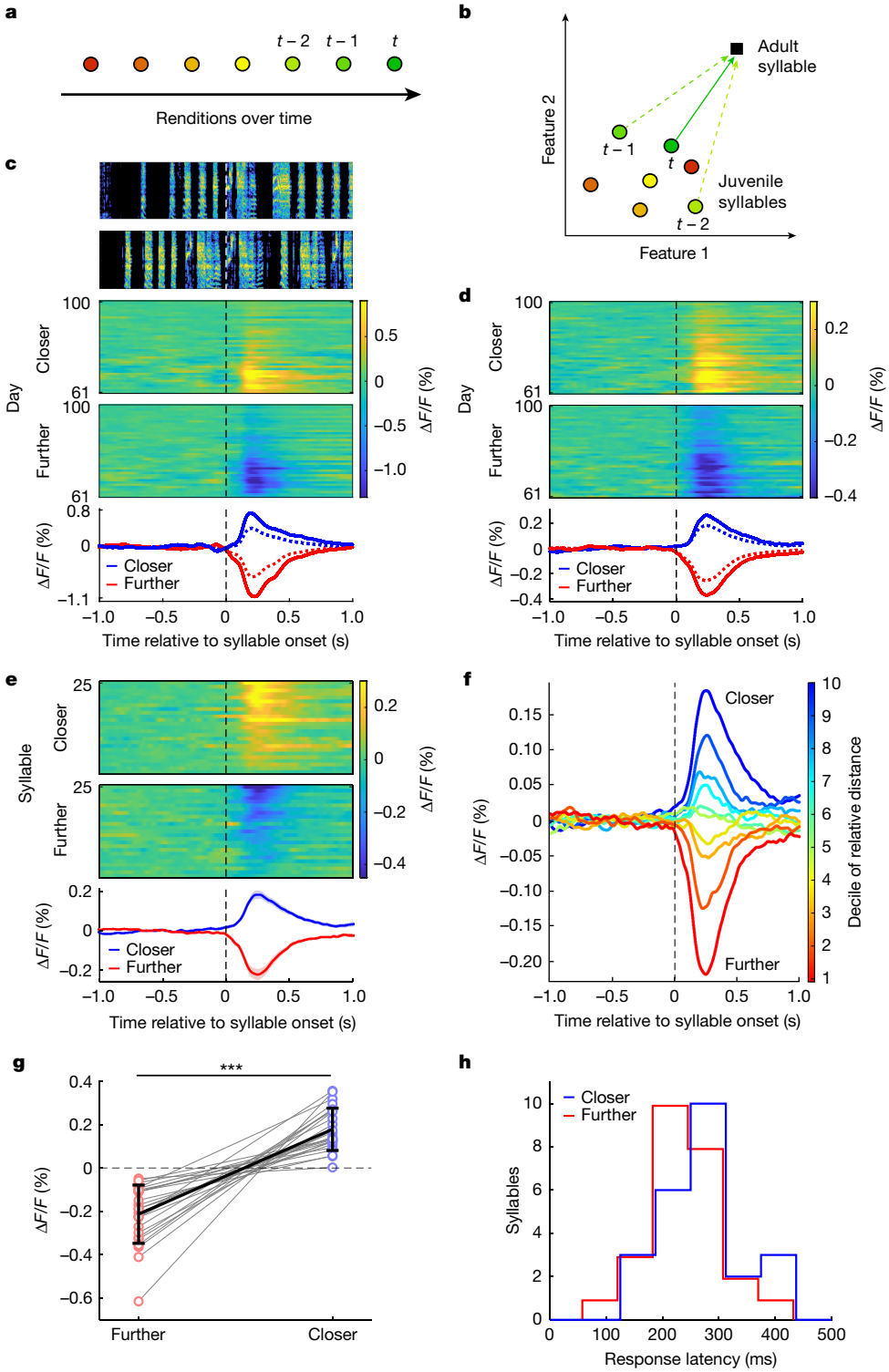


Fig. 2 | Dopamine encodes relative syllable quality during song learning. **a**, Schematic showing a series of consecutive renditions of a given song syllable (*t*, most recent rendition; *t* − 1 and *t* − 2, two previous renditions). **b**, Schematic showing computation of relative distance to adult syllable (circles, juvenile syllable renditions; black square, median location of adult (day > 90) renditions; arrows, distance to adult syllable). Relative distance was calculated as the difference between the current distance to adult syllable and the mean distance of the previous 11 renditions (see Methods). **c**, Dopamine responses for an example syllable after mean subtraction (see Methods and Extended Data Fig. 2). Example spectrograms and day-averaged dopamine responses across development for syllable renditions with the closest 10% relative distance (top) and furthest 10% relative distance (bottom) recorded in Area X during song learning, plotted above average Δ*F*/*F* signals (all plots aligned to syllable onset; blue, closer renditions; red, further renditions; solid line, day 61–75; dashed line, day 61–100). **d**, Δ*F*/*F* signals during developmental song learning averaged across *n* = 25 syllables in *n* = 6 birds, plotted similarly to **c**. **e**, Δ*F*/*F* signals are plotted similarly to **d**, but each row represents the average across days 61–100 for one syllable (shading, ±s.e.m). **f**, Δ*F*/*F* signals averaged across all 25 syllables and days 61–100 are plotted similarly to **e**, but for each decile (10% of syllable renditions) of relative distance. **g**, Scatter plot of averaged Δ*F*/*F* signals for all syllables (Methods) for closer (blue) and further (red) renditions (***$P < 0.001$, paired *t*-test; black bars, mean ± s.d.). **h**, Distribution of response latency for significant responses (*n* = 24 significant increases; *n* = 25 significant decreases; see Methods).

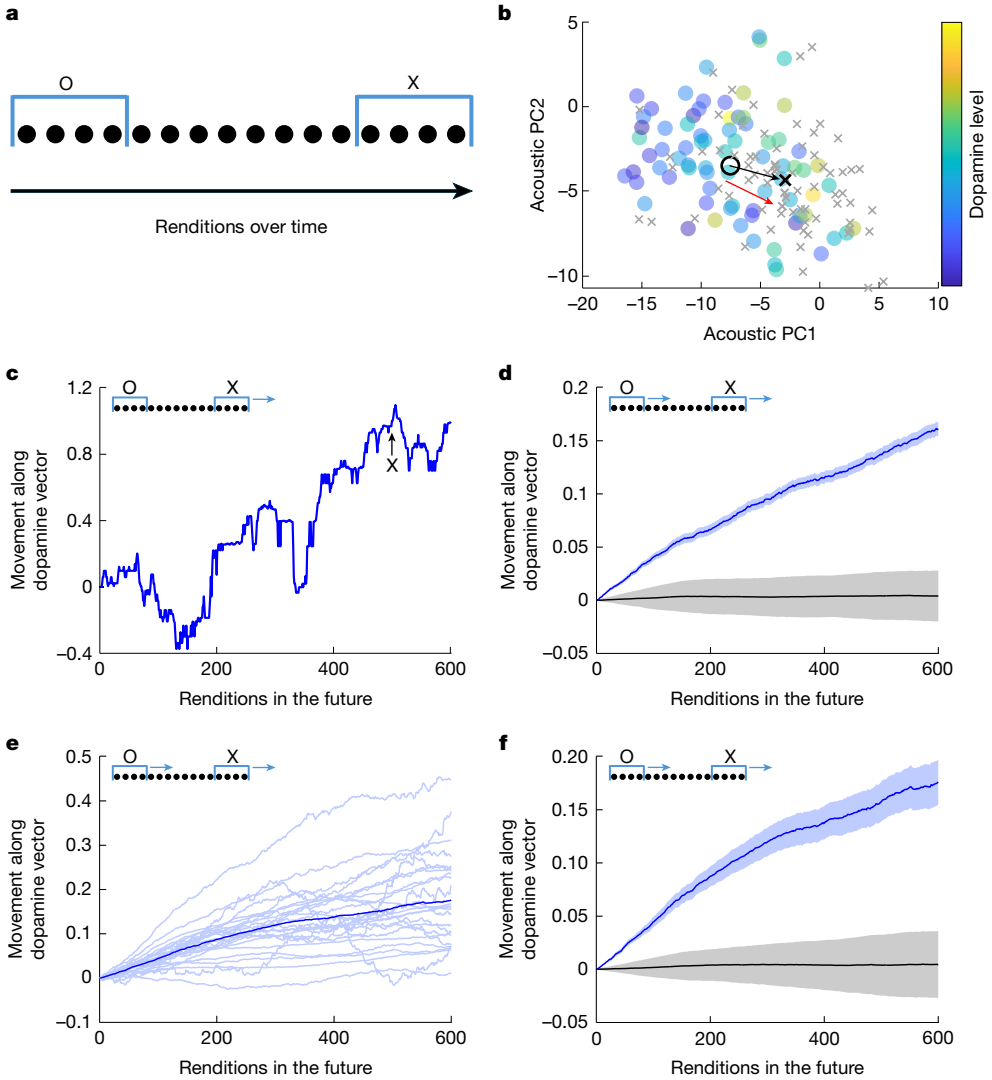


Fig. 3 | Dopamine predicts future song evolution during development. **a**, Schematic showing renditions of a syllable over time. Two groups of 150 consecutive renditions from earlier (O) and later (X) parts of the day. **b**, Dopamine predicts song movement for an example syllable. A 'dopamine vector' (red arrow), calculated from 150 consecutive renditions earlier in the day (coloured circles, coloured by associated dopamine release; black circle, median location), is defined as the direction in acoustic PC space of maximal correlation with dopamine (direction computed in the first 23 PC dimensions). This direction closely aligns with the direction of song movement (black arrow) at a later time (grey crosses, 150 consecutive renditions later in the day; black cross, median location; 75/150 renditions selected for visualization). **c**, Song movement along the dopamine vector during the course of a day for the example syllable shown in **b**. Each point along the *x* axis represents a group of 150 renditions shifted a rendition forward relative to the circles (see inset schematic) shown in **b** (black X, location of crosses plotted in **b**). **d**, Average across all dopamine vectors for the example syllable shown in **b**, calculated for blocks of 150 syllable renditions every 25 renditions (blue trace and shading, mean ± s.e.m.; black line with grey shading, mean ± 1.96 s.d. of shuffled data; see Methods). **e**, Movement along the dopamine vector (as in **d**) for each of the $n = 25$ syllables (thick blue line, mean across syllables). **f**, Movement along the dopamine vector averaged across all $n = 25$ syllables, plotted as in **d** ($P < 0.02$, computed from shuffled distribution, see Methods).

in our dataset was in the direction of increasing dopamine and was significantly different from shuffled data (Fig. 3e,f) (black line and grey shading represent mean ± 1.96 s.d. of 100 shuffles; Methods). The amount of movement along the dopamine vector was significantly correlated with the magnitude of the error response ($r = 0.46$, $P < 0.05$; see Methods). Thus, our data show that dopamine predicts future song evolution across our population of syllables and birds, providing evidence consistent with the hypothesis that dopamine drives future song. We reperformed this analysis across different song-block sizes and found significant movement along the dopamine vector at a song-block size as low as 10, showing that even very few dopamine transients contain information about future song trajectory (Extended Data Fig. 7).

We next sought to ascertain whether the relationship between song and dopamine is purely prospective. To address this, we extended our analysis to past renditions and found that past song trajectories move along the dopamine vector in the direction of lower dopamine, but by a significantly smaller amount (Extended Data Fig. 7). To investigate how this temporal asymmetry depends on local change in song, we performed an additional analysis in which we divided local song trajectories into two geometries: trajectories that undergo sharp turns in acoustic feature space and trajectories that remain aligned to their current course (see Methods). We found that when the song trajectory undergoes a sharp turn, current dopamine transients are significantly and positively correlated with song along the future direction of song movement, but uncorrelated along the direction of past song (Extended Data Fig. 8a–e). Furthermore, we found that when an acoustic element present in the tutor's syllable is being incorporated into the juvenile's syllable, renditions that contain the element are associated with higher dopamine, compared with renditions that do

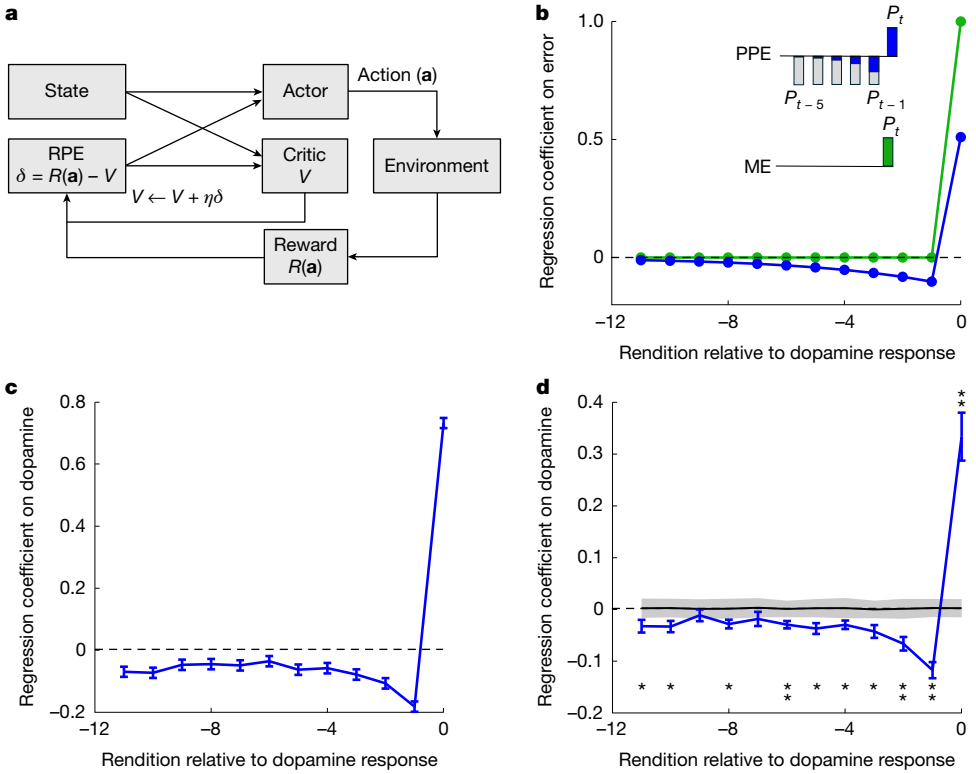


Fig. 4 | Dopamine tracks performance history to generate prediction errors. **a**, Schematic of an actor–critic reinforcement-learning algorithm[6] (see Methods). **b**, Coefficients of a linear regression fit of the error history to the hypothesized dopamine signal in a synthetic agent[35] (blue trace: coefficients if dopamine encodes current PPE, the difference between the current performance and a running average of previous performances; green trace: coefficients if dopamine encodes current ME). Inset, schematic of expected coefficients for a PPE signal (blue) and an ME signal (green). **c**, Example syllable fit of song error history to dopamine (blue trace; error bars, ±s.e.m.). **d**, Average over syllables ($n = 25$) (blue line, average regression coefficients over syllables; error bars, ±s.e.m.; grey line, average coefficient value across shuffles; grey shading, 95% percentile of the shuffled distribution; $^{**}P < 0.001$; $^{*}P < 0.05$, two-sided t-test).

not (Extended Data Fig. 8f–k). These results provide further evidence consistent with dopamine driving song learning and not simply correlating with song performance.

Dopamine tracks performance history

We next investigated the exact nature of the dopamine signal that guides learning. Specifically, we sought to determine whether dopamine encodes a simple mismatch error (ME) between ongoing performance and a fixed internal goal, or whether it represents a PPE (ref. 2), taking into account expectations based on recent performance. Birdsong learning has been proposed to follow an actor–critic[4,5,34] model of reinforcement learning. In such a learning framework, the PPE is defined as the difference between the current performance on a given trial and an iteratively updated prediction of performance, the value function, which is a weighted average of past performances (Fig. 4a).

We asked whether dopamine's dependence on performance history more closely resembles PPE or ME. In a modelled actor–critic reinforcement-learning agent, the current PPE is equal to the discounted sum of past rewards, in which earlier performances are summed with decaying weights, and subtracted from the current reward[35] (Fig. 4b). Figure 4b shows the coefficients for this subtraction. This is in contrast with ME, which is simply a function of the current reward (Fig. 4b). We similarly fit the linear relationship between performance and dopamine in our data to test whether the dopamine reward function was consistent with it being a PPE or an ME signal (see Methods). Figure 4c shows the linear regression fit of performance history and current performance onto dopamine across all days of recording for a single example syllable. Figure 4d shows the average linear regression coefficients onto dopamine over all syllables in our dataset (see Extended Data Fig. 9 for further individual syllable examples). Dopamine fluctuation is predicted by the difference between the current performance ($t = 0$) and a discounted sum of the previous performances ($t = -1$ to $t = -11$). This is consistent with dopamine activity in Area X acting as a PPE signal rather than as a pure ME or 'reward' signal. We measured how well the PPE model performs as we increase the number of history steps (see Methods and Extended Data Fig. 10). We found that model performance generally increased as we increased this history to a point, then started to decrease for increasingly longer histories. The number of previous renditions at which performance peaked varied across syllables, peaking on average at 11 renditions. Thus, we use this value for the results shown here (Fig. 4c,d) as well as Fig. 2. Together, our data indicate that dopamine tracks the quality of both current and past performance to provide a prediction error signal that guides the learning of natural behaviour.

Discussion

Our results support a model in which dopamine guides developmental song learning by providing a trial-by-trial PPE signal to the singing-related basal ganglia. This finding connects an extensive literature on the role of dopamine as an RPE signal in laboratory tasks to natural, internally guided developmental learning of motor skills such as vocal learning. Speech and language learning in humans are also thought to be mediated by basal-ganglia-dependent reinforcement mechanisms[12]. Birdsong is one of a small number of examples of vocal learning and shares notable parallels with speech acquisition[10–15]. More broadly, songbirds have emerged as an important model system to elucidate the neural mechanisms underlying internally guided learning of fine motor skills that require extensive practice[36]. Such

motor-skill learning requires evaluation of ongoing performance relative to an internal goal and subsequent modification of future motor output on the basis of these evaluation or error signals. We found that during developmental song learning in zebra finches, song syllable renditions that were closer to the adult version of the song than recent renditions were associated with an increase of dopamine in Area X, the singing-related basal ganglia. Correspondingly, syllable renditions that were further from the eventual song than recent renditions were associated with a decrease of dopamine. These results provide evidence that dopamine in the basal ganglia serves as a trial-by-trial evaluative signal for internally guided motor learning. Although syllable durations are short compared with the observed dopamine transients, we note that numerical simulations of gradient-based reinforcement-learning models have been found to converge in reasonable timescales despite a delayed and temporally imprecise reinforcement signal[37].

Our longitudinal recordings of song learning trajectories and concurrent dopamine activity allowed us to ask whether dopamine predicts future evolution of song. We found that song moves in the direction predicted by increased dopamine, supporting a model in which dopamine guides song evolution through reinforcement on a trial-by-trial basis. Our data allowed us to go one step further and investigate the exact nature of this reinforcement signal, by asking whether dopamine transients encode a simple mismatch error between ongoing song and the ultimate goal or whether they represent a prediction error[1,2,35] that takes into account recent performance history. By comparing our data with an actor–critic[4,5] reinforcement-learning model, we showed that dopamine activity exhibits the characteristics of prediction error signalling, comparing ongoing performance with predictions based on recent history.

Our data are consistent with a causal role for dopamine in guiding behaviour; however, future experiments such as syllable-specific optogenetic suppression or activation[23,24] of Area X dopamine during development will be required to demonstrate causality. Although the variability in dopamine responses between different syllables could not be simply explained by variability in syllable quality, we note that syllables with the strongest dopamine responses tended to be more complex compared with syllables with the weakest responses (Extended Data Fig. 4). Future work will investigate the possibility of differential contributions of dopamine to the learning of different syllable types.

The identification of VTA dopamine as the evaluation signal that mediates the developmental learning of internally guided complex motor skills also opens up future avenues of work studying the causes and consequences of this error signal[26,27]. How neural circuits encode the song template or internal goal[38,39], and how circuits upstream of the VTA[40,41] compare ongoing performance to the internal goal[4,42,43] to generate an error signal are not known. Several studies have begun to address these questions by characterizing the inputs to the VTA that differentially modulate the error signal[4,42–45]. Similarly to mammals, the downstream effect of the dopaminergic evaluation signal is thought to occur through dopamine-modulated cortico-striatal plasticity[26], which can link each time step in the song to the specific vocalization that produced a good match to the internal goal at that time step.

Reinforcement learning has contributed to some of the biggest recent advances in artificial intelligence and machine learning[7], from mastering chess and Go[46], to the automated discovery of novel algorithms[47]. It has long been speculated that internally guided reinforcement[48] underpins much of complex learned behaviour, from language and dance to sport and music. Here we provide direct evidence that dopamine in the song-related basal ganglia guides birdsong learning through reinforcement. VTA–basal ganglia circuits are evolutionarily conserved across birds and mammals[49,50] which suggests that a similar mechanism has a role in various other forms of internally guided motor-skill learning, including music, speech and language learning in humans.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-025-08729-1.

1. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
2. Gadagkar, V. et al. Dopamine neurons encode performance error in singing birds. *Science* **354**, 1278–1282 (2016).
3. Zhuo, Y. et al. Improved green and red GRAB sensors for monitoring dopaminergic activity in vivo. *Nat. Methods* **21**, 680–691 (2024).
4. Chen, R. & Goldberg, J. H. Actor–critic reinforcement learning in the songbird. *Curr. Opin. Neurobiol.* **65**, 1–9 (2020).
5. Joel, D., Niv, Y. & Ruppin, E. Actor–critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Netw.* **15**, 535–547 (2002).
6. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (MIT Press, 1998).
7. Botvinick, M. et al. Reinforcement learning, fast and slow. *Trends Cogn. Sci.* **23**, 408–422 (2019).
8. Wickens, J. R., Reynolds, J. N. & Hyland, B. I. Neural mechanisms of reward-related motor learning. *Curr. Opin. Neurobiol.* **13**, 685–690 (2003).
9. Costa, R. M. Plastic corticostriatal circuits for action learning: what's dopamine got to do with it? *Ann. N. Y. Acad. Sci.* **1104**, 172–191 (2007).
10. Jarvis, E. Vocal learning and spoken language. *Science* **366**, 50–54 (2019).
11. Davenport, M. H. & Jarvis, E. D. Birdsong neuroscience and the evolutionary substrates of learned vocalization. *Trends Neurosci.* **46**, 97–99 (2023).
12. Konopka, G. & Roberts, T. F. Insights into the neural and genetic basis of vocal communication. *Cell* **164**, 1269–1276 (2016).
13. Prather, J., Okanoya, K. & Bolhuis, J. J. Brains for birds and babies: neural parallels between birdsong and speech acquisition. *Neurosci. Biobehav. Rev.* **81**, 225–237 (2017).
14. Doupe, A. J. & Kuhl, P. K. Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* **22**, 567–631 (1999).
15. Brainard, M. S. & Doupe, A. J. Translating birdsong: songbirds as a model for basic and applied medical research. *Annu. Rev. Neurosci.* **36**, 489–517 (2013).
16. Burke, J. E. & Schmidt, M. F. Neural control of birdsong. *eLS* **1**, 345–355 (2020).
17. Person, A. L., Gale, S. D., Farries, M. A. & Perkel, D. J. Organization of the songbird basal ganglia, including area X. *J. Comp. Neurol.* **508**, 840–866 (2008).
18. Lovell, P. V. et al. ZEBrA: Zebra finch Expression Brain Atlas—a resource for comparative molecular neuroanatomy and brain evolution studies. *J. Comp. Neurol.* **528**, 2099–2131 (2020).
19. Tumer, E. C. & Brainard, M. S. Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. *Nature* **450**, 1240–1244 (2007).
20. Andalman, A. S. & Fee, M. S. A basal ganglia–forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc. Natl Acad. Sci. USA* **106**, 12518–12523 (2009).
21. Duffy, A., Latimer, K. W., Goldberg, J. H., Fairhall, A. L. & Gadagkar, V. Dopamine neurons evaluate natural fluctuations in performance quality. *Cell Rep.* **38**, 110574 (2022).
22. Roeser, A. et al. Dopaminergic error signals retune to social feedback during courtship. *Nature* **623**, 375–380 (2023).
23. Hisey, E., Kearney, M. G. & Mooney, R. A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. *Nat. Neurosci.* **21**, 589–597 (2018).
24. Xiao, L. et al. A basal ganglia circuit sufficient to guide birdsong learning. *Neuron* **98**, 208–221 (2018).
25. Hoffmann, L. A., Saravanan, V., Wood, A. N., He, L. & Sober, S. J. Dopaminergic contributions to vocal learning. *J. Neurosci.* **36**, 2176–2189 (2016).
26. Fee, M. S. & Goldberg, J. H. A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience* **198**, 152–170 (2011).
27. Mackevicius, E. L. & Fee, M. S. Building a state space for song learning. *Curr. Opin. Neurobiol.* **49**, 59–68 (2018).
28. Doya, K. & Sejnowski, T. A novel reinforcement model of birdsong vocalization learning. In *Adv. Neural Information Processing Systems 7 (NIPS 7)* (eds Tesauro, G. et al.) 101–108 (MIT Press, 1995).
29. Tchernichovski, O., Mitra, P. P., Lints, T. & Nottebohm, F. Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* **291**, 2564–2569 (2001).
30. Kollmorgen, S., Hahnloser, R. H. R. & Mante, V. Nearest neighbours reveal fast and slow components of motor learning. *Nature* **577**, 526–530 (2020).
31. Funabiki, Y. & Konishi, M. Long memory in song learning by zebra finches. *J. Neurosci.* **23**, 6928–6935 (2003).
32. Steinfath, E., Palacios-Munoz, A., Rottschafer, J. R., Yuezak, D. & Clemens, J. Fast and accurate annotation of acoustic signals with deep neural networks. *eLife* **10**, e68837 (2021).
33. Lerner, T. N., Holloway, A. L. & Seiler, J. L. Dopamine, updated: reward prediction error and beyond. *Curr. Opin. Neurobiol.* **67**, 123–130 (2021).
34. Toutounji, H., Zai, A. T., Tchernichovski, O., Hahnloser, R. H. R. & Lipkind, D. Learning the sound inventory of a complex vocal skill via an intrinsic reward. *Sci. Adv.* **10**, eadj3824 (2024).
35. Bayer, H. M. & Glimcher, P. W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
36. Adam, I. et al. Daily vocal exercise is necessary for peak performance singing in a songbird. *Nat. Commun.* **14**, 7787 (2023).
37. Fiete, I. R., Fee, M. S. & Seung, H. S. Model of birdsong learning based on gradient estimation by dynamic perturbation of neural conductances. *J. Neurophysiol.* **98**, 2038–2057 (2007).

38. Ikeda, M. Z., Trusel, M. & Roberts, T. F. Memory circuits for vocal imitation. *Curr. Opin. Neurobiol.* **60**, 37–46 (2019).
39. Louder, M. I. M. et al. Transient sensorimotor projections in the developmental song learning period. *Cell Rep.* **43**, 114196 (2024).
40. Tian, J. et al. Distributed and mixed information in monosynaptic inputs to dopamine neurons. *Neuron* **91**, 1374–1389 (2016).
41. Watabe-Uchida, M., Eshel, N. & Uchida, N. Neural circuitry of reward prediction error. *Annu. Rev. Neurosci.* **40**, 373–394 (2017).
42. Chen, R. et al. Songbird ventral pallidum sends diverse performance error signals to dopaminergic midbrain. *Neuron* **103**, 266–276 (2019).
43. Kearney, M. G., Warren, T. L., Hisey, E., Qi, J. & Mooney, R. Discrete evaluative and premotor circuits enable vocal learning in songbirds. *Neuron* **104**, 559–575 (2019).
44. Bottjer, S. W., Brady, J. D. & Cribbs, B. Connections of a motor cortical region in zebra finches: relation to pathways for vocal learning. *J. Comp. Neurol.* **420**, 244–260 (2000).
45. Mandelblat-Cerf, Y., Las, L., Denisenko, N. & Fee, M. S. A role for descending auditory cortical projections in songbird vocal learning. *eLife* **3**, e02152 (2014).
46. Schrittwieser, J. et al. Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature* **588**, 604–609 (2020).
47. Fawzi, A. et al. Discovering faster matrix multiplication algorithms with reinforcement learning. *Nature* **610**, 47–53 (2022).
48. Markowitz, J. E. et al. Spontaneous behaviour is structured by reinforcement without explicit reward. *Nature* **614**, 108–117 (2023).
49. Colquitt, B. M., Merullo, D. P., Konopka, G., Roberts, T. F. & Brainard, M. S. Cellular transcriptomics reveals evolutionary identities of songbird vocal circuits. *Science* **371**, eabd9704 (2021).
50. Pfenning, A. R. et al. Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science* **346**, 1256846 (2014).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2025

Methods

Zebra finches

Subjects were six male zebra finches (*Taeniopygia guttata*) ≥45 days old, reared by both parents in a breeding colony and randomly selected before experiments. All experiments were performed in accordance with NIH guidelines and were approved by the Columbia Institutional Animal Care and Use Committee.

Surgery

At 47 ± 2 days old, the subject zebra finches were separated from the colony for surgery. During implant surgeries, each bird was anaesthetized with isoflurane (1–3%) and 1 µl of the viral construct AAV2/9-hSyn-GRAB$_{DA3h}$ (ref. 3) was injected into Area X at two sets of coordinates (600 nl at +5.9 mm anterior, +1.5 mm lateral relative to lambda and −2.8 mm ventral relative to pial surface; head angle 20°, 400 nl at +5.9 mm, +1.5 mm, −2.5 mm, 20°) using a Nanoject III (Drummond Scientific) injector. During the same surgery, an optical cannula attached to a metal ferrule (Doric, 400-µm core) was implanted above the injection site (+5.9 mm, +1.5 mm, −2.4 mm, 20°).

Histology

For histology to verify cannula placement, birds were transcardially perfused, and brains were fixed in 4% paraformaldehyde solution. Brains were sliced into 100-µm-thick sagittal sections and imaged using a Nikon AZ100 Multizoom microscope.

Dopamine photometry and song recordings

After surgery, birds were placed in a sound isolation chamber for around two weeks to allow for viral expression. At around 60 days old, birds were moved to the experimental rig equipped with a Doric photometry system and a microphone. Custom MATLAB code detected singing onset, which triggered song recording as well as the LEDs for photometry. Two LEDs (470 nm at 208.6 Hz, dopamine dependent; and 405 nm at 572.2 Hz, dopamine independent) were used for excitation. Demodulation in Doric Neuroscience Studio produced separate emission signals generated by the 470-nm and 405-nm excitation, which were then used to compute the percentage change of the fractional fluorescence signal, defined as $\Delta F/F$ (%) = 100 × [470_signal − 405_signal × mean (470_signal_baseline)/mean (405_signal_baseline)]/mean (470_signal_baseline). Baseline was defined as the 1 s preceding syllable onset. Song and dopamine data were recorded in this set-up during plastic song and into adulthood, between 61 and 100 days of age. Dopamine data corresponding to song syllables were included if photometry data were present from −1 to +1 s relative to syllable onset. Statistical methods were not used to predetermine sample sizes.

Syllable labelling and acoustic parametrization

Song syllable identities were based on crystallized adult song, determined at more than 90 days old, and were labelled across development using a deep-neural-network-based song annotation tool (Deep Audio Segmenter; DAS)[32]. Ten per cent of total syllables were hand-labelled throughout the course of development and were used to train unique classification models for each syllable in each bird. After automated classification, putative syllables were further filtered by excluding outliers on the basis of syllable duration and absolute distance to the adult syllable. Five hundred randomly selected renditions, evenly distributed across development for each syllable, were hand checked in this phase, and false positives were discarded. Each syllable used for further analysis had an F_1 score (harmonic mean of precision [1 − false positive rate] and recall [1 − false negative rate]) of at least 90% based on this check. Unique syllable types in zebra finch song vary in acoustic structure and are best characterized by different acoustic features[51]. For example, harmonic stacks are well characterized by pitch, but syllables with broad band frequency spectrums have unstable pitch estimates. Thus, acoustic features were selected to build a generic, high-dimensional, acoustic parameterization of syllables that could be applied without curation across acoustic profiles. Consistent with this, we found that the correlation between dopamine and individual song features varied across individual syllable types. The MATLAB function audioFeatureExtractor (Audio Toolbox) was used to extract 31 time-varying acoustic features across each syllable. Each syllable was divided into four segments by duration, and each feature was averaged within each syllable segment to create a 124-dimensional, smoothed, time-varying acoustic parameterization of each syllable. Principal component analysis was then used to reduce the dimensionality individually for each syllable for further analysis. Days were excluded from future analyses if there were not enough ($n > 11$) labelled renditions on that day to compute relative distance (number of days included per syllable: mean, 28 and range, 4–40, across 25 syllables). Experimenters were blind to dopamine data while relative proximity of syllables was being calculated.

Data analysis

Syllable quality and dopamine responses. The top 23 PCs (explaining more than 75% of variance) were extracted from the 124-dimensional feature space containing all instances of a syllable. We tested the robustness of our analyses in Figs. 2 and 3 across 5–50 PC dimensions and found the results to be qualitatively unchanged. The 'adult syllable' was defined as the median location in this feature space of all renditions produced by the bird after crystallization (age > 90 days)[29,52]. High-relative-quality, or 'closer', renditions were defined as those that were closer to the adult syllable than the mean distance of the previous 11 renditions (see Methods section 'Selecting the length of history filter in the PPE representation of dopamine' for choice of 11). Similarly, low-relative-quality, or 'further', renditions were defined as those that were further from the adult syllable than the mean distance of the preceding 11 renditions. The 10% highest- and 10% lowest-relative-quality renditions, as defined above, were then identified for each syllable on each day for further analysis. For Fig. 2c, $\Delta F/F$ (%) is shown for 1 s before and after syllable onset, with each row corresponding to an average of either the 10% highest-quality or 10% lowest-quality renditions during the day. For Fig. 2d, each row corresponds to an average across all syllables that have at least 20 renditions in both the furthest and closest 10% with dopamine data on that day. For Fig. 2e, each row represents a single syllable averaged across all renditions throughout development. For Fig. 2f, each trace represents one decile of data as arranged by relative distance averaged across all syllables and across development, with the top and bottom traces equivalent to those in Fig. 2e. For Fig. 2g, the $\Delta F/F$ (%) values for each syllable (red and blue circles) were obtained by averaging the $\Delta F/F$ (%) traces in a 100-ms window between 200 ms and 300 ms from syllable onset, and then averaging across development. The difference between populations was significant ($P = 2.75 \times 10^{-9}$; paired t-test). For Fig. 2h, response latency was defined as the time to peak or trough in the z-scored $\Delta F/F$ (%) between 0 ms and 500 ms from syllable onset only if the peak or trough was significant (defined as greater than 1.96 for peaks and less than −1.96 for troughs, comprising a two-tailed z test). z-scored $\Delta F/F$ was calculated as ($\Delta F/F$ (%) − mean [baseline $\Delta F/F$ (%)])/s.d. [baseline $\Delta F/F$ (%)]. Significances for each syllable were calculated individually; for multiple-comparisons correction, see Extended Data Fig. 3. Throughout Fig. 2, results are presented in mean-subtracted format; that is, for each day, the average dopamine trace for that syllable across all renditions is subtracted out (see Extended Data Fig. 2).

Singing-related dopamine responses. In Extended Data Fig. 2a,b, the dopamine response to the onset and offset of a bout of singing was analysed. Bouts were defined as sequences of vocalizations separated by less than 150 ms, starting with an introductory note or song syllable and ending with a song syllable. Bouts were identified algorithmically using

an amplitude threshold, with a subset of bouts checked for accuracy. Only bout onsets preceded by 1 s and offsets followed by 1 s of silence were included. For this analysis, the baseline for $\Delta F/F$ (%) generation was computed using the 0.5 s preceding bout onset or offset. Extended Data Fig. 2c–g illustrates the subtraction of mean singing-related dopamine responses to isolate the error response. For each syllable on each day, a mean dopamine trace was created by averaging the dopamine response across all renditions, not just the top and bottom 10%. In Fig. 2, as well as Extended Data Figs. 2g, 3b–d, 4, 5d–f, 6 and 8f–k, dopamine data are analysed after first subtracting that syllable's mean trace on each day. In Extended Data Fig. 2e, the blue and red traces are identical to the ones in Extended Data Fig. 2c, but the 'equidistant' decile is shown as well. To find the closest and furthest 10% on each day, the renditions in a day are sorted by relative distance, ranging from negative numbers (closer syllables) to positive numbers (further syllables). In this sorted list of renditions, the set of 10% consecutive renditions that have their mean relative distance equal to 0 are defined as 'equidistant' (no change in distance compared with recent renditions). In Extended Data Fig. 2g, the responses to the closest, furthest and equidistant 10% across development are quantified following mean subtraction. The closer and further responses differed significantly from 0 ($P = 2.67 \times 10^{-9}$, $P = 3.73 \times 10^{-8}$; one-sided t-test), whereas the equidistant responses did not ($P = 0.96$; one-sided t-test). (As in Fig. 2, the response is defined as the mean of the trace in the 200–300-ms window after syllable onset).

Rendition-to-rendition variability in dopamine responses. In Extended Data Fig. 3a, an example spectrogram of song (top) is shown along with concurrent dopamine activity (bottom). For this dopamine trace, the baseline for $\Delta F/F$ (%) generation was computed over the full 8-s time period. In Extended Data Fig. 3b,c, data are shown for an example syllable on one example day (day 67). In Extended Data Fig. 3c,d, bars show the mean dopamine response over the 200–300 ms after syllable onset.

Variability in dopamine responses across syllables. In Extended Data Fig. 4c, the average Euclidean distance in the 23-PC space for each rendition was averaged for each day for each of the 25 syllables. Then, for each day between 61 and 100, an average was taken across all the syllables that had at least 100 renditions on that day. Shading represents s.e.m. across syllables, with the sample size for each point being the number of syllables with data on that day. In Extended Data Fig. 4d–g, the magnitude of the dopamine response is plotted against various syllable metrics. For each of the 25 syllables, the dopamine response is defined as the difference between the mean response for the closest 10% across development and the mean response for the furthest 10% between 200 ms and 300 ms after syllable onset. These values correspond to the differences between the pairs of points presented in Fig. 2g. In Extended Data Fig. 4d, each syllable's location on the x axis corresponds to the mean duration across all renditions of the syllable across development. In Extended Data Fig. 4e, each syllable's location on the x axis corresponds to the mean of the standard deviation (s.d.) of the absolute distance from the adult syllable on each day of development with more than 100 renditions. In Extended Data Fig. 4f, each day with more than 100 renditions for each syllable was examined separately, and the change per hour within a day was found by fitting a slope to the absolute distance relative to the adult version. All days across development were then averaged to give each syllable its position on the x axis. In Extended Data Fig. 4g, the change in distance per day across development for each syllable was found by computing the mean absolute distance to the adult version for each day with more than 100 renditions in development, then fitting a slope across days. Correlation and significance values in Extended Data Fig. 4d–g were computed using MATLAB's corr function, returning the Pearson linear correlation coefficient (panel d, $r = 0.22$, $P = 0.29$; panel e, $r = 0.04$, $P = 0.86$; panel f, $r = 0.17$, $P = 0.43$; panel g, $r = 0.07$, $P = 0.76$).

Effects of circadian oscillations. In Extended Data Fig. 5a, the average absolute distance from the adult syllable over the course of the day is shown. Here, the distance of all renditions falling within each hour of the day across development was averaged for each of the $n = 25$ syllables. Each point along the y axis represents an average across all of the $n = 25$ syllables with at least 20 renditions during that hour across all days of development. In Extended Data Fig. 5b, the dopamine response (averaged across 200–300 ms after syllable onset) was calculated for all renditions across development for each of the 12 hours of the day for each of the $n = 25$ syllables. Then, for each hour of the day, the average across all of the $n = 25$ syllables with at least 20 dopamine-aligned renditions across development is shown. For Extended Data Fig. 5c, for each of the $n = 25$ syllables, all renditions in each of the 12 hours of the day were examined, and the percentage of renditions falling in the closest 10% and furthest 10% in terms of relative distance during the day was found; the average percentage of the closest 10% of trials (blue line) and the furthest 10% (red line) for each hour averaged across all of the $n = 25$ syllables with more than 200 renditions during that hour across development is shown. For Extended Data Fig. 5d–f, dopamine responses were computed separately for the renditions occurring in the first three hours of the day ('morning'), and the latter nine hours of the day ('afternoon'). This temporal division cut the data roughly in half ($n = 51{,}560$ morning renditions and $n = 51{,}465$ afternoon renditions across all $n = 25$ syllables). For each rendition, relative distance was still computed using the median adult (day > 90) rendition across all hours of the day as the target, but the dopamine responses of renditions were analysed separately depending on the time of day. The dopamine response to 'closer' trials was not significantly different in the morning compared with the afternoon ($P = 0.12$; paired t-test), whereas the response to 'further' trials was significantly greater in the morning ($P = 0.0046$; paired t-test).

Contribution of syllable amplitude to the dopamine response. In Extended Data Fig. 6a, the normalized amplitude of an example syllable across $n = 40$ days of development is shown. For each syllable rendition, amplitude was computed as the root mean square of the raw audio trace. Then, amplitude for all renditions on a day was averaged. Finally, each day's mean amplitude was divided by the mean amplitude of all days > 90, which is the syllable's 'adult' amplitude. In Extended Data Fig. 6b, the process described above was performed for each syllable, then, on each day, a syllable was included if there were at least 20 renditions of that syllable on that day. In Extended Data Fig. 6c, all renditions with accompanying dopamine data are shown for an example syllable. Each rendition's point along the x axis corresponds to the rendition's amplitude, calculated as the root mean square of the raw audio trace, and each rendition's point along the y axis is the average dopamine response from 200–300 ms after syllable onset. The Pearson's correlation coefficient was generated with MATLAB's corr function ($r = 0.01$, $P = 0.08$). In Extended Data Fig. 6d, the correlation and significance of the relationship described in Extended Data Fig. 6c is shown for all syllables. To generate the amplitude-normalized plots in Extended Data Fig. 6f–h, each quarter of the syllable was first normalized to its median amplitude before feature extraction. Amplitude normalization did not significantly change the response to closer ($P = 0.43$, paired t-test) or further ($P = 0.92$, paired t-test) trials.

Linking dopamine activity to future and past song. For Fig. 3 and Extended Data Fig. 7, canonical correlation analysis (CCA), computed using the MATLAB function canoncorr, was used to find the linear combination of acoustic features that maximally correlated with rendition-to-rendition dopamine fluctuations in local blocks of consecutive syllables ($n = 150$ renditions)[53,54]. To compare a scalar dopamine value to syllable acoustic structure, the dopamine ($\Delta F/F$) trace was averaged in 12.5-ms windows. The timing onset of the dopamine window used to compute the direction of maximal correlation with

syllable acoustic variations was selected to be the peak, summed CCA correlation between song fluctuations and dopamine variation across all song-block–dopamine comparisons (direction computed in the first 23 PC dimensions). Across syllables, peaks were selected from within 0 to 500 ms after syllable onset. Timing latency between syllable onset and dopamine was held fixed for individual syllables across all analyses (mean latency = 207 ms). Song renditions were projected onto the acoustic dimension of maximum correlation with dopamine (the 'dopamine vector'). The median of song along this dimension was computed within blocks of 150 song renditions in sliding steps of one rendition, for 600 renditions forward in time. The distance between future song blocks and the original song block was defined as the difference in song-block medians. This procedure was performed across all days and all song blocks (every 25 renditions) for each syllable. Song blocks were not used to estimate a dopamine vector if less than 40% of renditions within the block had associated dopamine data. The movement along the dopamine vector was then averaged across all song blocks across development. The significance of average movement along the dopamine vector was measured by shuffling dopamine traces relative to syllable renditions within individual days and reperforming the same analysis 100 times. The shuffled distribution of song movement was used to compute a two-sided P value at each step forward in time. Average movement along the dopamine vector over the full population of syllables was significant ($P < 0.02$) in future rendition steps 10–600. A second shuffling procedure was also tested by selecting consecutive sequences of dopamine responses randomly offset from the true sequence location in rendition number. This preserves possible underlying temporal correlations in the dopamine activity across renditions while randomizing the relationship to song. Significance remained unchanged. To test sensitivity to the song-block size in which the dopamine vector was computed, song-block size was varied from $n = 5$ to $n = 300$. Movement along the dopamine vector remained significant ($P < 0.05$) in block sizes $n = 10$ to $n = 300$ (Extended Data Fig. 7a).

As an additional comparison, we measured the movement along the 'adult song vector', which is the direction defined as the difference between the adult song median location and the current song-block median location in acoustic space. This is the acoustic direction the song is ultimately going during development, and thus is, on average, the linear direction in which the song will move the most. When we do this comparison, the average movement along the dopamine vector is 15.0% of the average movement along the adult song direction at 600 renditions into the future (averages taken across syllables).

To similarly measure the dopamine vector's orientation towards the past song trajectory (Extended Data Fig. 7c), song movement along the dopamine vector was computed in past rendition steps (−1 to −600 renditions relative to the onset of the current song block). We found that the past song moves along the negative direction of the dopamine vector (towards lower dopamine) but much less than the positive movement in the future direction. The significance of the difference in magnitudes of future versus past song movement along the dopamine vector was calculated using a paired t-test at each rendition step away from current song. We find that song moves significantly further along the dopamine vector in the future than in the past from rendition steps 51 to 600 ($P < 0.05$).

Movement along the dopamine vector for each syllable, defined as the amount of movement 600 renditions in the future divided by the s.d. of the shuffled data at that point, was significantly correlated with the magnitude of the error signal for that syllable (defined as the difference between the response to the closest 10% and furthest 10% renditions in relative distance averaged over 200–300 ms after syllable onset; $r = 0.46$, $P = 0.02$, Pearson's linear correlation coefficient, calculated using MATLAB's corr function).

Dopamine's relationship with the local geometry of the song learning trajectory. Future song moves in the direction of high dopamine and past song moves in the direction of low dopamine, but to a lesser degree. To further test dopamine's relationship with the immediate song trajectory (Extended Data Fig. 8), we characterized the local geometry of the song learning trajectory and measured dopamine's correlation with future and past song movement. To measure this, we locally defined three consecutive, non-overlapping blocks of song: the past song block, the current song block and the future song block. These three blocks of song are computed every 25 renditions across all days of recording. We defined the past song vector as the difference between the current and the past song blocks' median locations in acoustic feature space (the first 23 PCs of each syllable, as described earlier). We defined the future song vector as the difference between the future and current song blocks' median locations. We then projected the song renditions in the current song block onto the future and past vectors and measured the correlation with dopamine along these two dimensions in acoustic feature space. These measurements compare current dopamine's correlation with future and past song change. We categorized local song trajectories into two types: (1) trajectories in which the angle between the future and the negative of the past song vector is greater than π/2 (labelled 'arrows' because they represent relatively straight trajectories through song space); and (2) trajectories in which the angle between the future and the negative of the past song vectors is less than π/2 (labelled 'hinges' because they represent trajectories in which the song has taken sharp turns). We found that in local 'hinge' trajectories, dopamine was significantly positively correlated with the future song vector but not with the past song vector. This is consistent with dopamine driving song learning and not simply correlating with song. As expected in 'arrow' trajectories, in which the past and future vectors are relatively aligned, dopamine significantly and positively correlated with both the past and future vectors, but more strongly with the future vector. The significance of the difference between dopamine correlations along the future and past song vectors was calculated using a paired t-test. The significance of the individual correlations with dopamine along the future and past vectors for both hinge and arrow song trajectories were computed using two separate methods: (1) a two-sided t-test and (2) comparing to a shuffled distribution ($n = 100$ shuffles) in which dopamine trials were randomized in relation to song trials. Both methods yielded significant correlations with future song vectors in hinge trajectories ((1) $P < 0.00001$; (2) $P < 0.02$), but not with past song vectors in hinge trajectories. This analysis was repeated in song-block sizes 5–300 as in the analysis for Fig. 3, and was found to be robust to block size.

In addition, to test dopamine's relationship with imitative learning, we found three syllables that contained an element from the tutor's syllable that was also present in the final, adult, version, but was missing from renditions early in development. For each of these three syllables, we identified the day (days 61, 63 and 61, respectively) when roughly half of the renditions contained the tutor syllable's element ('mature' renditions), and half did not ('immature' renditions). For each syllable, all renditions on that day were hand-labelled as either mature or immature. By taking the mean of the dopamine response for each trial between 200 ms and 300 ms after syllable onset, the mature and immature populations could be compared, and all three syllables had a significant difference in dopamine response between mature and immature renditions ($P = 1.6 \times 10^{-8}$, $P = 4.9 \times 10^{-4}$, $P = 0.012$, two-sided t-test).

Data modelling

Actor–critic model and dopamine's dependence on performance history. The critical encoding of a prediction error that we test in the dopamine data is generic to a variety of formulations of reinforcement-learning models[6]. We focus here on the actor–critic implementation because this is the dominant model that has been posited for the birdsong system[4,28,37], but our evaluation of the relationship between dopamine and error versus prediction error is more general than this

formulation. For Fig. 4 and Extended Data Fig. 9, an actor–critic reinforcement-learning model was defined as follows. A randomly generated (drawn from a normal distribution with mean 0 and variance 1) n-dimensional vector $\mathbf{t}$ was defined as the target song. The actor agent's policy was initialized to an n-dimensional unit vector $\boldsymbol{\pi}$ (each value drawn from a normal distribution with mean 0 and variance 1) and the critic agent's value V was initialized to 0. At each time point, the actor agent generates an action vector $\mathbf{a}$ equal to the current policy with a random step $\mathbf{e}$ added to it, scaled by $\varepsilon = 0.1$:

$$e_i = N(0, \varepsilon^2)$$

$$\mathbf{a} = \boldsymbol{\pi} + \mathbf{e}$$

The reward for this action R is then calculated as the mean squared error (MSE) between the action and the target:

$$R(\mathbf{a}) = -\frac{1}{n}\sum_{i=1}^{n}(t_i - a_i)^2$$

$$R(\mathbf{a}) = -\frac{1}{n}\,\|\mathbf{t} - \mathbf{a}\|^2$$

The temporal difference error (or PPE), δ, was calculated as the difference between this reward and the critic agent's current expected value:

$$\delta = R(\mathbf{a}) - V$$

This is then used to update the actor's policy along with learning rate $\eta_{\text{actor}} = 0.15$,

$$\boldsymbol{\pi} \leftarrow \boldsymbol{\pi} + \eta_{\text{actor}} \times \delta \times \mathbf{e},$$

and critic's value along with learning rate $\eta_{\text{critic}} = 0.2$,

$$V \leftarrow V + \eta_{\text{critic}} \times \delta.$$

To evaluate dopamine's dependence on performance history, the MATLAB function fitlm was used to fit a linear regression model to the performance quality history (renditions(i), $i = -N_{\text{history filter}}$ to 0) and current scalar dopamine response ($\Delta F/F$ (%)) ($i = 0$) computed as described above. In brief, this model fits a linear model that predicts the dopamine response on the previous trial as the weighted sum of the quality of the current syllable as well as those in its recent history, identifying a coefficient for each lag distance that captures the relationship between syllables at that lag and dopamine on the current trial[35]. Positive coefficients for a given syllable lag indicate that high performance at that point in the history predict high dopamine on the current rendition, whereas negative coefficients indicate that low performance at that lag predicts higher dopamine. Performance quality was defined as the negative L_1 distance between the current song rendition and the median location of adult song in the first 23 PC dimensions of the full acoustic feature space. PCs were computed for each syllable individually across all days of recording. Linear fits were performed on all days of recording across all syllables and birds and separately for early development (age ≤ 75 days) and late development (age > 75 days).

Distinguishing PPE from ME. For Extended Data Fig. 10, to compare dopamine's representation of performance ME (analogous to reward error) and PPE (analogous to RPE), linear regression fits to current performance ME and current performance plus performance history were compared over a range of performance history steps (N_{history} = [1 2 3 4 5 6 10 15 20 25 30 35 40 50]). Linear regression is the appropriate model to test the PPE hypothesis because the functional form of the RPE (PPE here) equation from a reinforcement-learning theory framework is a linear combination of current reward minus a weighted sum of past rewards, in which the weight decays to zero back in history:

$$\text{RPE}_t = R_t - \eta\sum_{i=1}^{\infty}(1-\eta)^{i-1}R_{t-i},$$

with $0 < \eta < 1$. Model quality was compared using two measures: (1) the difference in the Akaike information criterion (ΔAIC)[55,56] and (2) the difference in MSE (ΔMSE) between the ME model and the PPE model, estimated from ten cross-validation folds. Syllables were excluded from the ME–PPE comparison if neither the ME nor the PPE linear regression onto dopamine had a significant R^2 value (computed using the fitlm function in MATLAB). All syllables were included in the ME/PPE comparison and had at least one ME or PPE linear regression onto dopamine with a significant R^2 value. We find that the PPE model is better than a pure performance mismatch model at predicting dopamine transients in the data across many different numbers of sample history terms: under the ΔAIC metric, at least 20/25 syllables were better fit by a PPE model for all history filters n = 1:10; under the ΔMSE metric, at least 19/25 syllables were better fit by a PPE model than by a ME model for all history filters n = 1:10.

Selecting the length of history filter in the PPE representation of dopamine. For Extended Data Fig. 10, across the PPE history filter parameter sweep, the minimum values of both the ΔAIC and the ΔMSE metrics indicate the largest improvement in quality of fit to dopamine over performance history fits. From reinforcement-learning theory, we expect the coefficients on performance history terms to asymptotically decay to 0 further back in the past. The expectation of performance quality is iteratively updated on each performance rendition, and the filter theoretically includes all past performance. However, the amount of data, level of noise and extent of underlying PPE representation in dopamine will all affect the number of history coefficients that can be found significant. As the coefficients approach 0, noise will overtake the benefit of including additional history terms in the dopamine function. Consistent with this, we found a positive correlation between the number of optimal history terms for each syllable and the amount of data for each syllable (ΔAIC: $r = 0.69$, $P < 0.001$; ΔMSE: $r = 0.72$, $P < 0.001$). The optimal number of history terms was averaged over all included syllables (22/25) to select the length of the history filter used in Figs. 2 and 4. Under the ΔAIC and the ΔMSE model comparison metrics, the average best history filter length was found to be 11 and 12 respectively. The more conservative model average, n = 11, was used in Figs. 2 and 4. Coefficients from fits across all recording days were averaged across syllables and plotted in Fig. 4d. In Fig. 4d, P values for coefficient averages over syllables were computed individually for each performance history step using a two-sided t-test. From 0 to −11 song history steps, P = [1.89×10^{-7}, 8.67×10^{-8}, 3.58×10^{-5}, 0.002, 0.001, 0.002, 3.61×10^{-4}, 0.19, 0.003, 0.323, 0.006, 0.014]. At a history filter of 11, the average r^2 across syllables = 0.01 ± 0.01 (s.d.), and 20/25 of these r^2 values are significant (P value of F-test on regression model, $P < 0.05$, Holm–Bonferroni method used to correct for multiple comparisons[57]). The significance of coefficient averages across the full syllable population was further assessed by randomizing the relationship between song rendition and dopamine. Dopamine responses were kept temporally ordered to account for the effect of possible underlying temporal correlations in the dopamine response across renditions. Population shuffles were performed 100 times to generate randomized population coefficient averages.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Further data are available upon request. Source data are provided with this paper.

Code availability

The code written for this study is available upon request.

51. Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B. & Mitra, P. P. A procedure for an automated measurement of song similarity. *Anim. Behav.* **59**, 1167–1176 (2000).
52. Immelman, K. in *Bird Vocalizations* (ed. Hinde, R. A.) 64–74 (Cambridge Univ. Press, 1969).
53. Krzanowski, W. J. *Principles of Multivariate Analysis: A User's Perspective* (Oxford Univ. Press, 1988).
54. Seber, G. A. F. *Multivariate Observations* (Wiley, 1984).
55. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **19**, 716–723 (1974).
56. Ljung, L. *System Identification: Theory for the User* (Prentice Hall, 1999).
57. Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **6**, 65–70 (1979).

Acknowledgements We thank L. Abbott, D. Aronov, M. Churchland, A. Litwin-Kumar, N. Sawtell, S. Siegelbaum and members of the V.G. laboratory for comments and suggestions; M. Eswaran and A. Sahilu for technical assistance; and K. J. Miller for analysis suggestions and feedback on the manuscript. Imaging was performed with support from the Zuckerman Institute's Cellular Imaging platform for instrument use and technical advice. We thank the Zebra Finch Expression Brain Atlas website (http://www.zebrafinchatlas.org) for histological reference images. Funding support was provided to A.L.F. and A.D. by the Simons Collaboration on the Global Brain, and to V.G. by the NIH (R00NS102520 and DP2AT012347) and the Searle, Klingenstein–Simons and McKnight scholars programs.

Author contributions J.K. and V.G. conceived the study, and designed and performed the experiments. J.K., A.D., N.N., A.R., A.L.F., K.L.S. and V.G. performed data analysis. A.D., A.R., A.L.F. and K.L.S. performed data modelling. J.K., A.D., N.N., A.R. and V.G. wrote the original draft of the manuscript. J.K., A.D., N.N., A.R., A.L.F., K.L.S. and V.G. edited and reviewed the final manuscript. A.L.F. and V.G. acquired funding. A.L.F., K.L.S. and V.G. supervised the project.

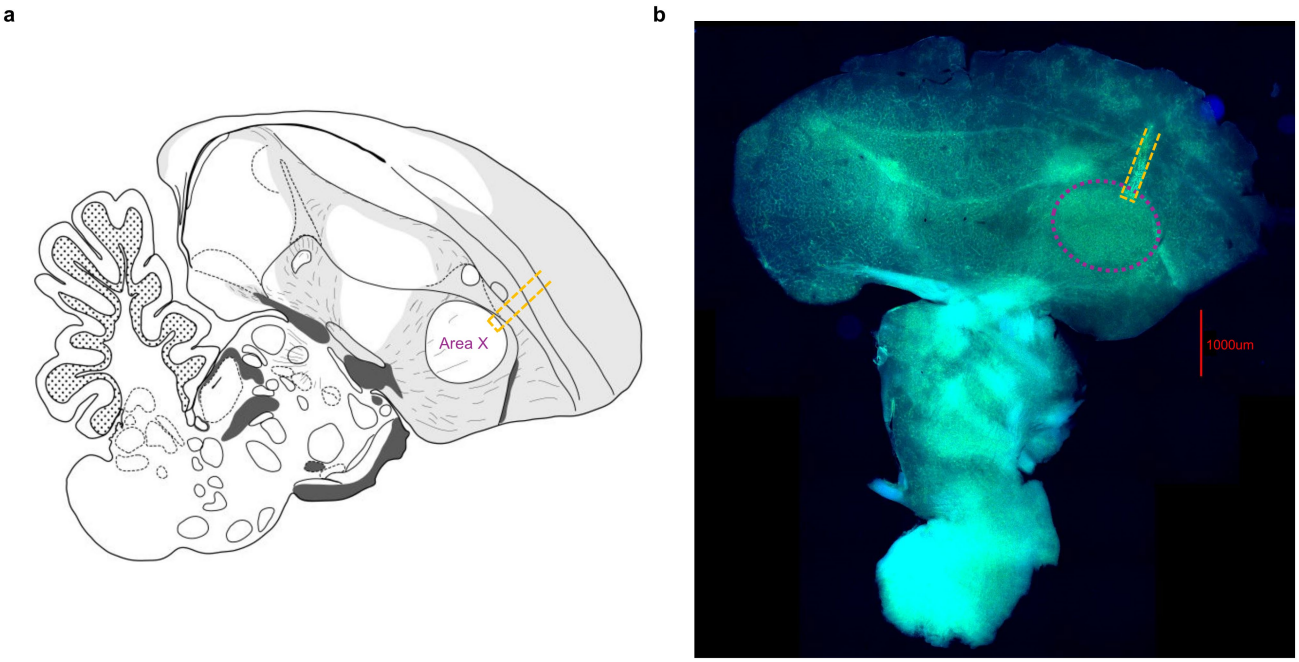
Competing interests The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-025-08729-1.

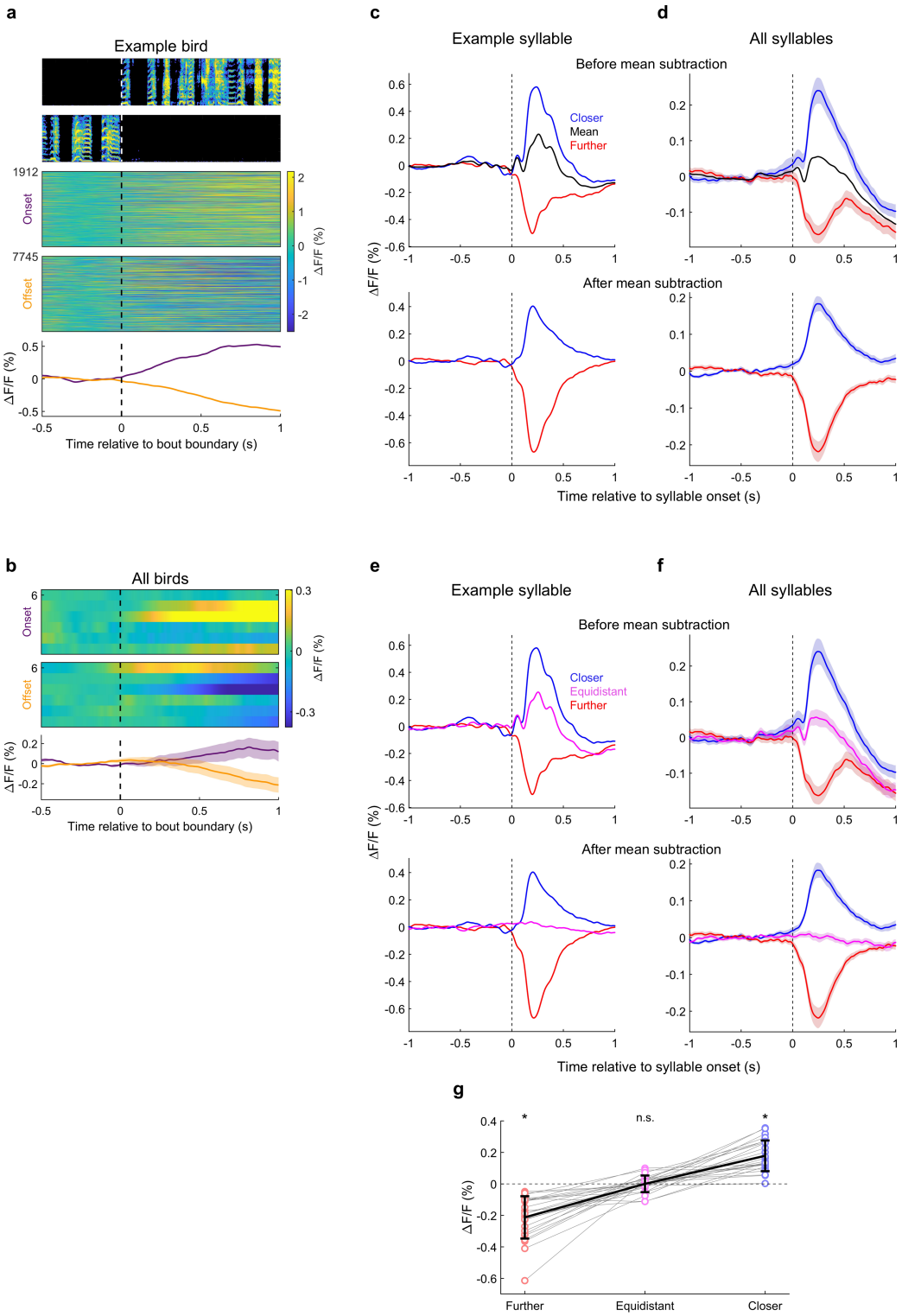
Correspondence and requests for materials should be addressed to Vikram Gadagkar.

Peer review information *Nature* thanks Ofer Tchernichovski and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

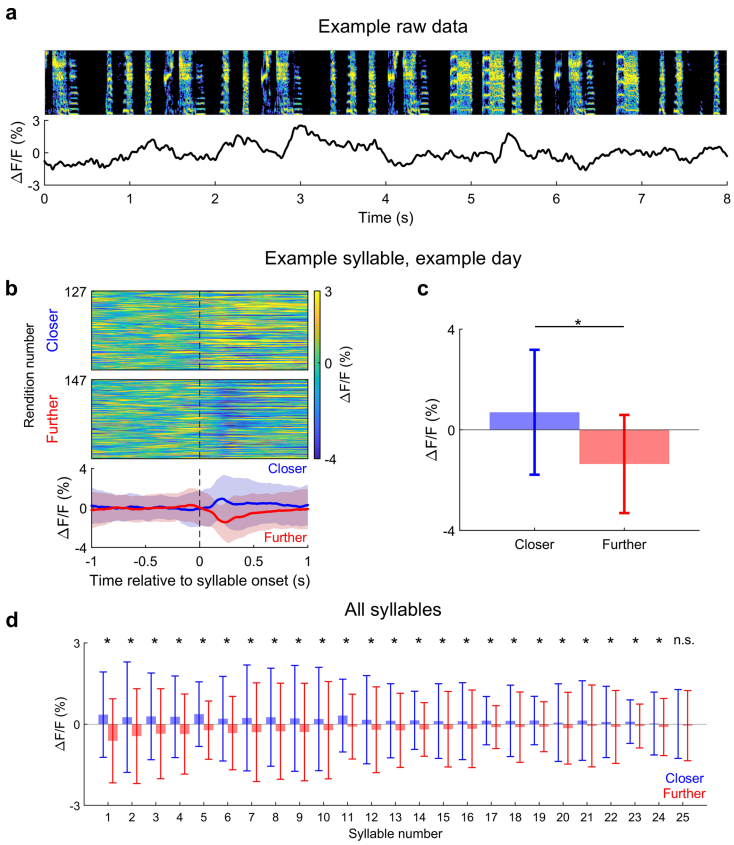
Reprints and permissions information is available at http://www.nature.com/reprints.



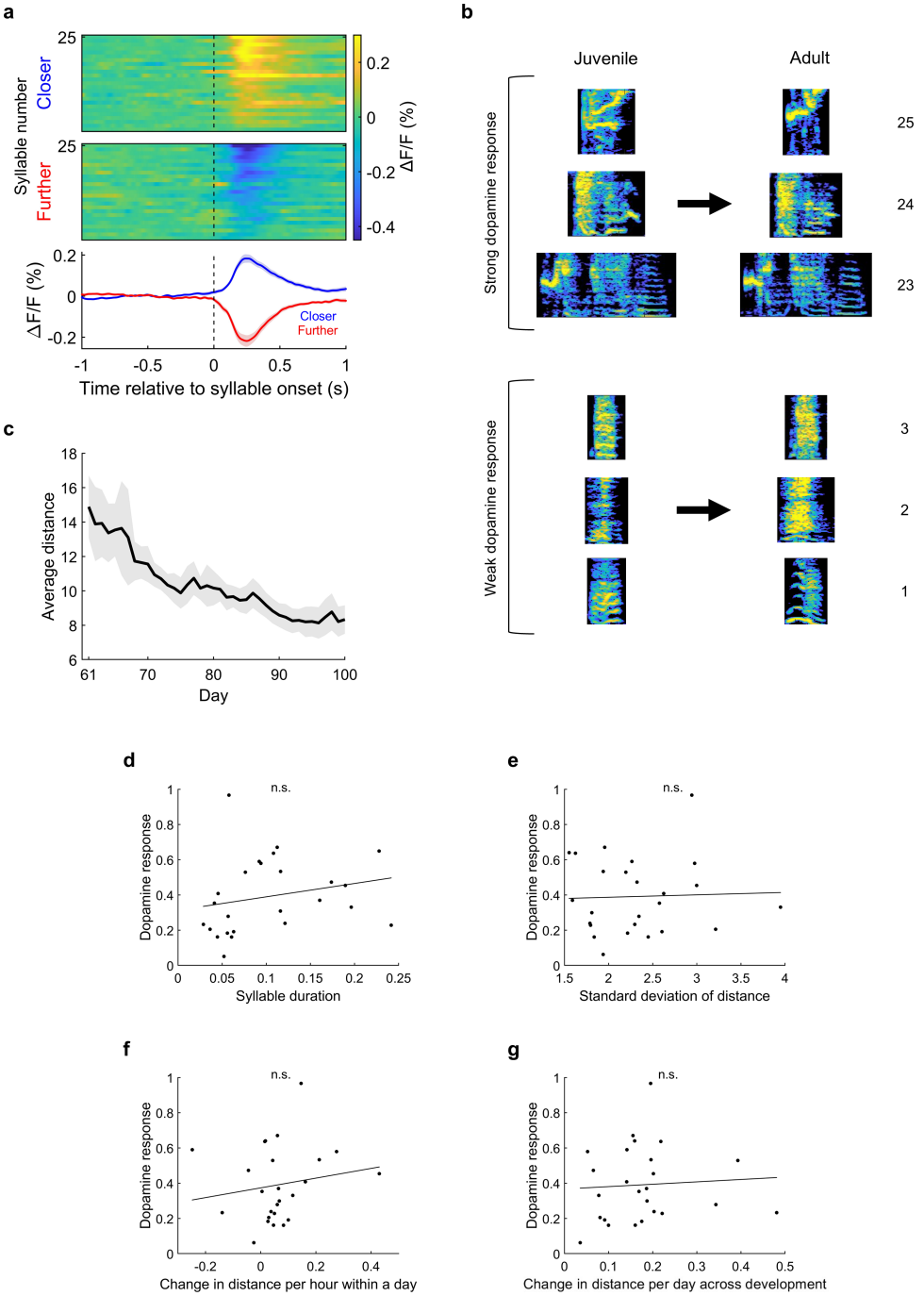
Extended Data Fig. 1 | Histological verification of fibre placement in Area X. **a**, Schematic from the zebra finch atlas. Adapted from ref. 18 (Wiley). Orange dashed lines indicate the intended placement of the cannula in Area X. **b**, Example brain slice showing fibre placement (purple dashed circle, Area X; orange dashed lines, fibre placement).



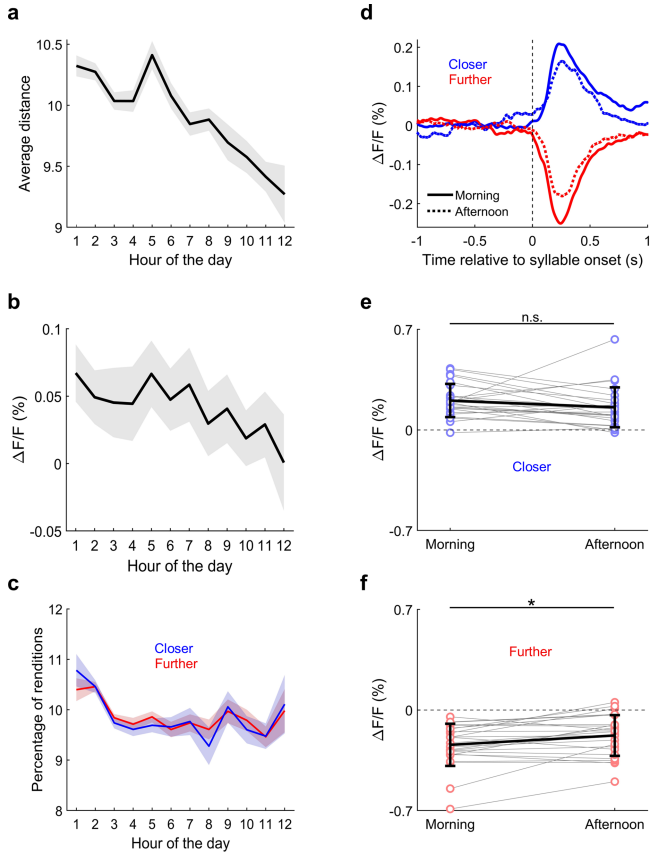
Extended Data Fig. 2 | Singing-related dopamine activity. a, Dopamine responses aligned to singing bout boundaries (see Methods) for an example bird across days 61–100 of development. Example spectrograms and dopamine responses aligned to bout onsets (n = 1912, top) and bout offsets (n = 7745, bottom), plotted above average ΔF/F signals (purple, bout onsets; orange, bout offsets). **b**, ΔF/F signals plotted similarly to **a** but each row represents the average across one of n = 6 birds (shading, ±s.e.m). **c**,**d**, Mean subtraction of singing-related dopamine activity to isolate the error response (see Methods). Top, ΔF/F signals for an example syllable (**c**) and averaged across all (n = 25) syllables (**d**) for the closest 10% relative distance (blue), furthest 10% relative distance (red) and mean (black) across all renditions (shading, ±s.e.m). Bottom, ΔF/F signals for the furthest and closest relative distance plotted as on top but after mean subtraction. **e**,**f**, ΔF/F signals plotted as in **c**,**d** with the dopamine response to the middle 10% of trials with 0 average relative distance (purple, equidistant - neither closer nor further). **g**, Scatter plot of averaged ΔF/F signals for all syllables (Methods) for closer (blue), further (red), and equidistant (purple) renditions (*$P < 0.05$, n.s. $P > 0.05$, not significantly different from zero, one sample t-test; black bars, mean ± s.d.).



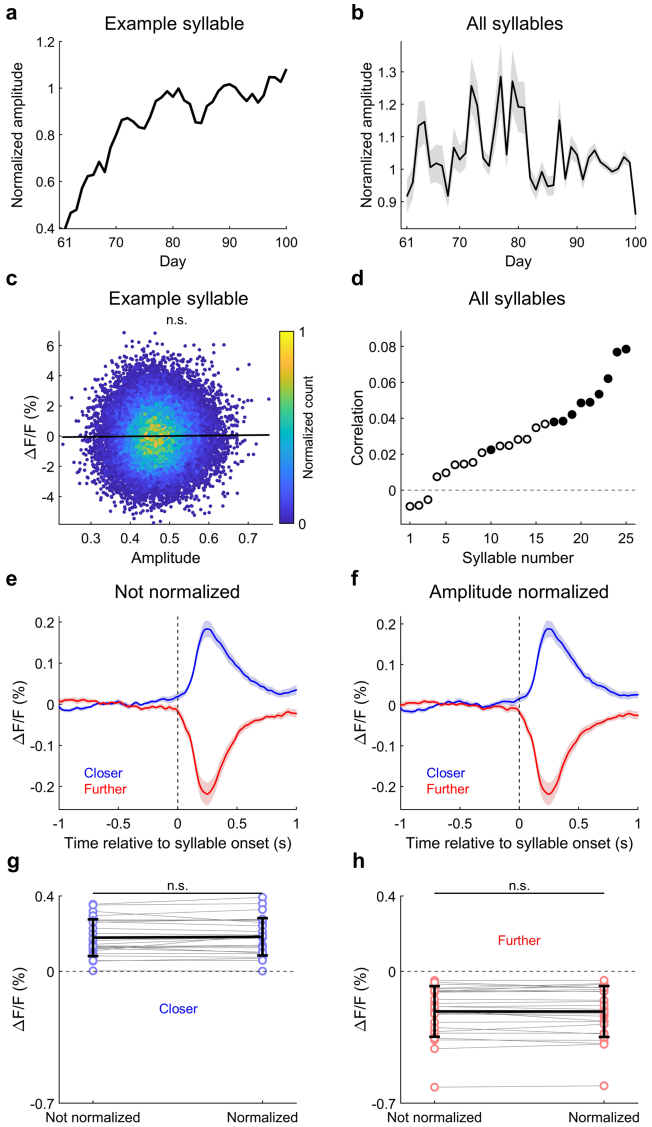
Extended Data Fig. 3 | Rendition-to-rendition variability in dopamine responses. a, Spectrogram of an example eight-second section of song plotted above the corresponding ΔF/F signal. **b**, Single rendition dopamine responses for an example syllable on a single day (day 67) with the closest 10% (top) and furthest 10% (bottom) relative distance plotted above average ΔF/F signals (blue, closer renditions; red, further renditions; shading, ±s.d.). **c**, Dopamine responses (see Methods) for the same data shown in **b** (bar and error bar, mean ± s.d.; $*P < 0.05$, two-sided t-test). **d**, Dopamine responses plotted similarly to **c** but averaged across development (day 61–100) for all n = 25 syllables, ordered by the difference between closer and further response ($*P < 0.05$, n.s. $P > 0.05$, not significant, 2-sided t-test with Holm–Bonferroni correction[57]).



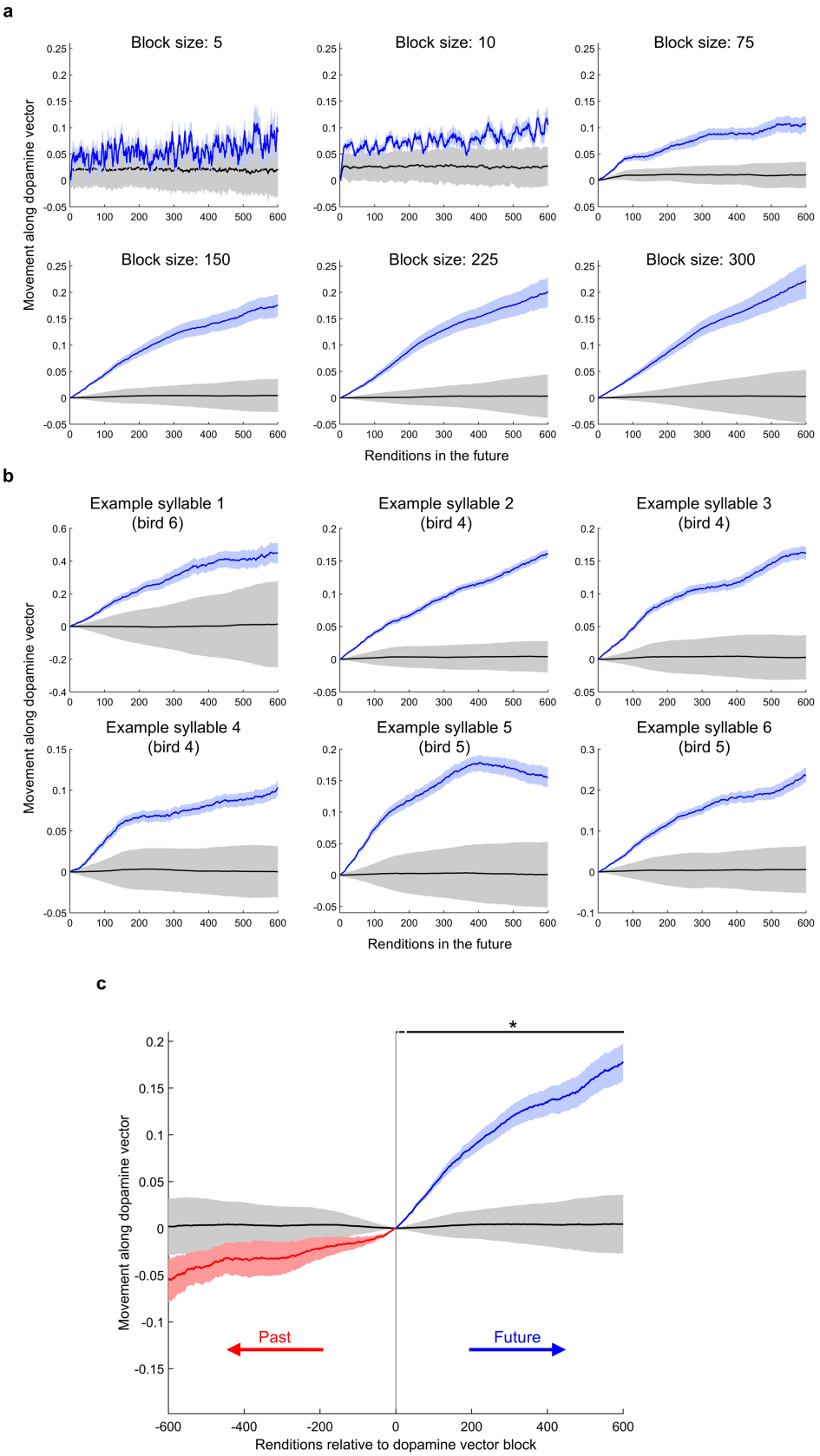
Extended Data Fig. 4 | Variability in dopamine responses across syllables. **a**, Syllable-averaged dopamine responses (reproduced from Fig. 2e, see Methods) across development (day 61–100) with the closest 10% relative distance (top) and furthest 10% relative distance (bottom) plotted above average ΔF/F signals (blue, closer renditions; red, further renditions; shading, ±s.e.m) sorted from weakest to strongest dopamine response. **b**, Example spectrograms for the three syllables with the strongest dopamine response (top) from early in development (day ≤ 76) and during adulthood (day > 90) (top) and the three syllables with the weakest dopamine response (bottom). Number indicates ranking (1, weakest; 25, strongest). **c**, Average over all syllables (n = 25) of the mean distance from all renditions on a given day to the adult (age > 90 days) median (grey shading, ±s.e.m.). **d**, Correlation between magnitude of dopamine response and syllable duration (black dots, 25 syllables; black line, best fit; n.s. $P > 0.05$, no significant correlation, Pearson linear correlation coefficient). **e**–**g**, Plotted as in **d** but for average s.d. per day of distance to adult version across development (**e**), per hour change in distance averaged across each day (**f**) and mean change in distance per day across development (**g**).



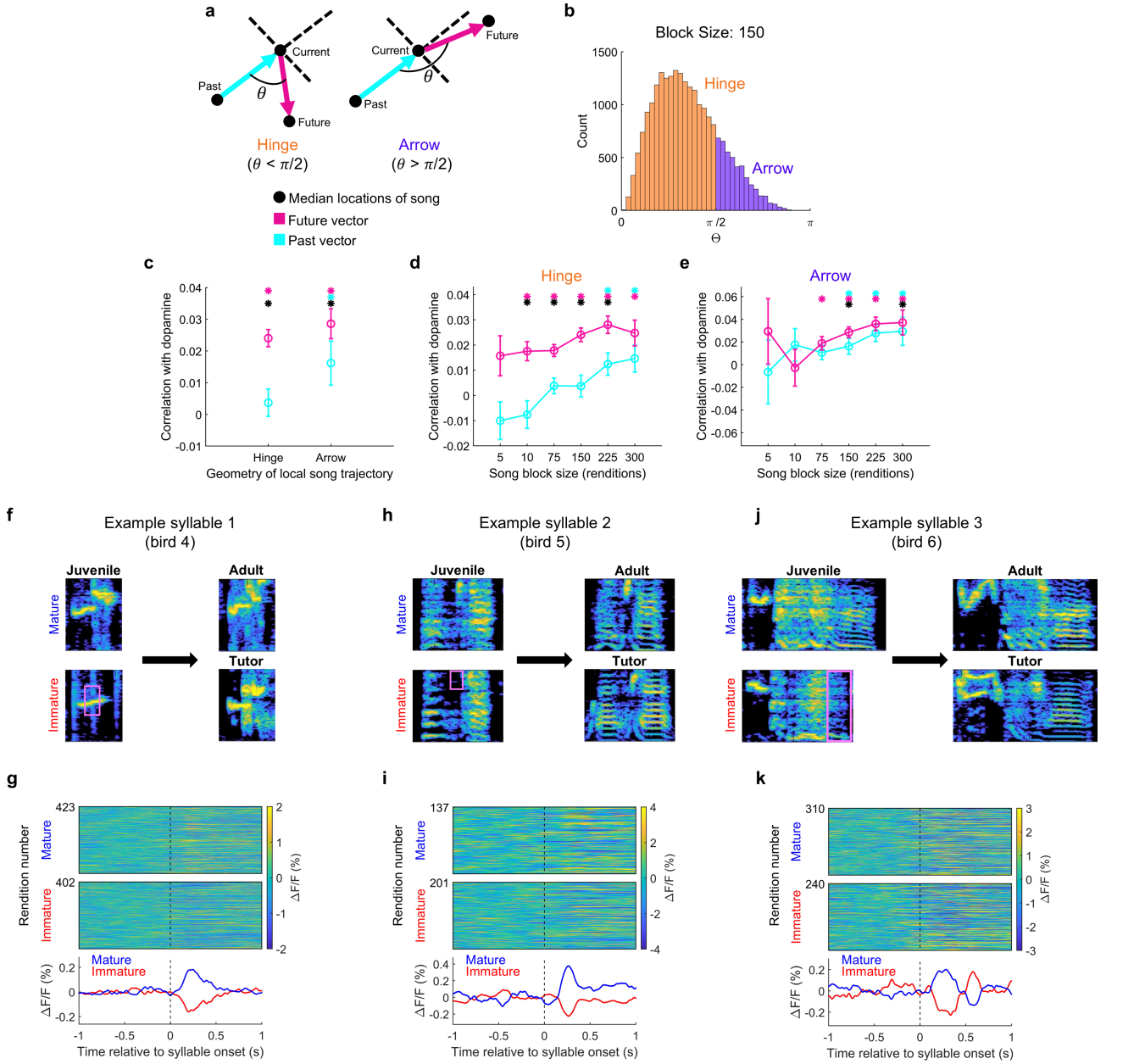
Extended Data Fig. 5 | Circadian oscillations have a subtle influence on the magnitude of the dopamine response. a, Distance to adult median averaged across all syllables (n = 25) and over development (day 61 – 100) decreases over the course of a day (line and shading, mean across syllables ± s.e.m.). **b**, Average dopamine response across syllable renditions (see Methods), plotted similarly to **a**. **c**, The percentage of renditions with closest 10% (blue) and furthest 10% (red) relative distance during a day that were sung in each hour of the day (line and shading, mean across syllables ± s.e.m.). **d**, Dopamine responses to the closest 10% (blue) and furthest 10% (red) relative distance renditions sung in the morning (first 3 hours of the day, solid lines, n = 51560 total renditions across birds) and afternoon (latter 9 hours of the day, dotted lines, n = 51465 total renditions across birds) averaged across n = 25 syllables across development (Day 61–100). **e**, Scatter plot of averaged ΔF/F signals for the closest 10% relative distance sung in the morning (left) and afternoon (right) for all syllables (n.s. $P > 0.05$, not significant, paired t-test; black bars, mean ± s.d.). **f**, Plotted similarly to **e** but for furthest 10% relative distance (*$P < 0.05$, significant, paired t-test).



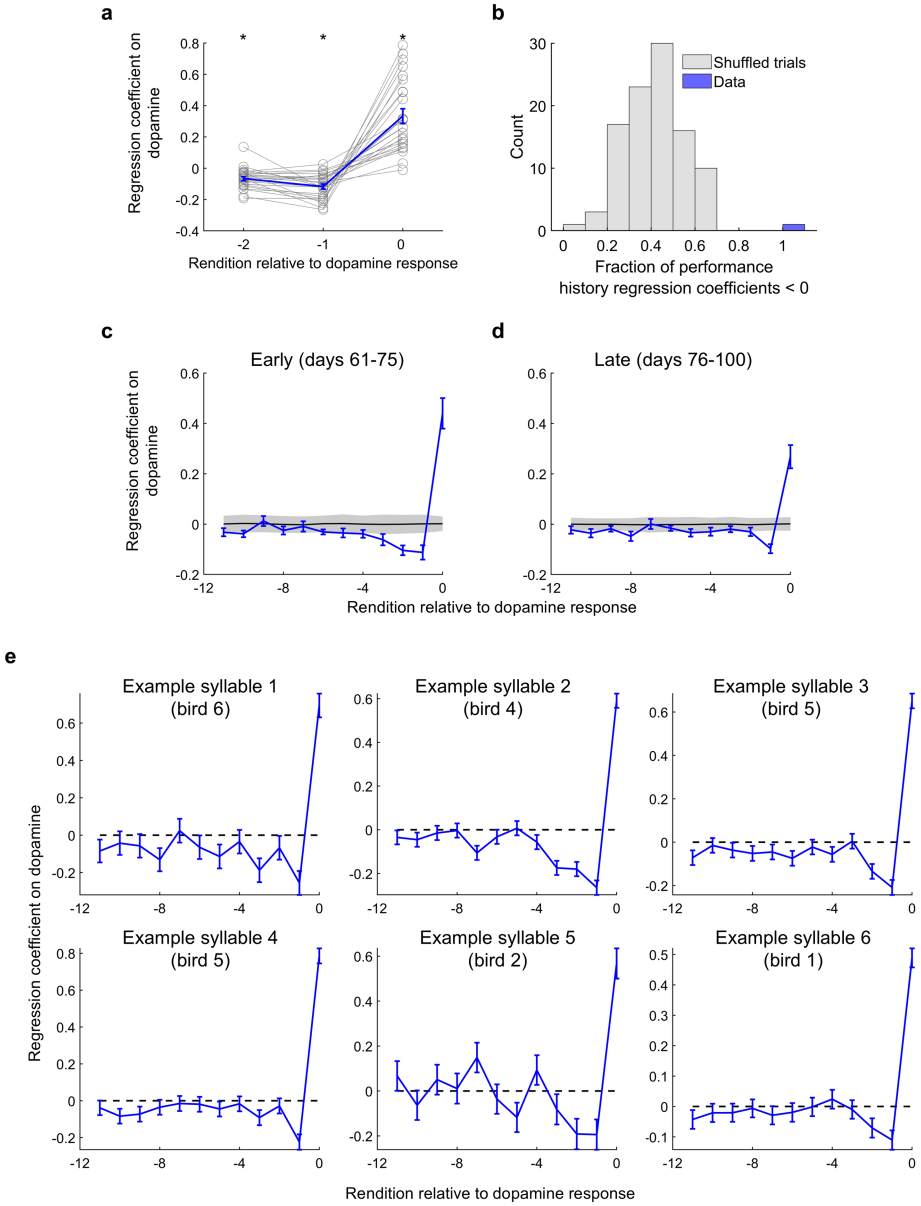
Extended Data Fig. 6 | Variation in syllable amplitude does not account for the dopamine response. a, Mean amplitude across all renditions on each day of development (day 61–100) for an example syllable, normalized to the average adult (day > 90) amplitude of that syllable. **b**, Plotted similarly to **a** but averaged across all (n = 25) syllables (shading, ±s.e.m.). **c**, Correlation of the ΔF/F signal with syllable amplitude (see Methods) across development (day 61–100) for an example syllable (black line, best fit line; n.s., $p > 0.05$, not significant, Pearson linear correlation coefficient). **d**, Pearson linear correlation coefficient between syllable amplitude and the ΔF/F signal for all syllables (open circles, $p > 0.05$, not significant; filled circles, $p < 0.05$, Pearson linear correlation coefficient with Holm–Bonferroni correction[57]). **e**,**f**, Syllable-averaged (n = 25) ΔF/F signals for 10% closest (blue) and 10% furthest (red) relative distance before (**e**, reproduced from Fig. 2e) and after (**f**) normalization (see Methods) by syllable amplitude (shading, ± s.e.m). **g**,**h**, Scatter plot of averaged ΔF/F signals for all n = 25 syllables for closer (**g**) and further (**h**) relative distance comparing not normalized (left) and amplitude-normalized (right) renditions (n.s. $P > 0.05$, not significant, paired t-test; black bars, mean ± s.d.).



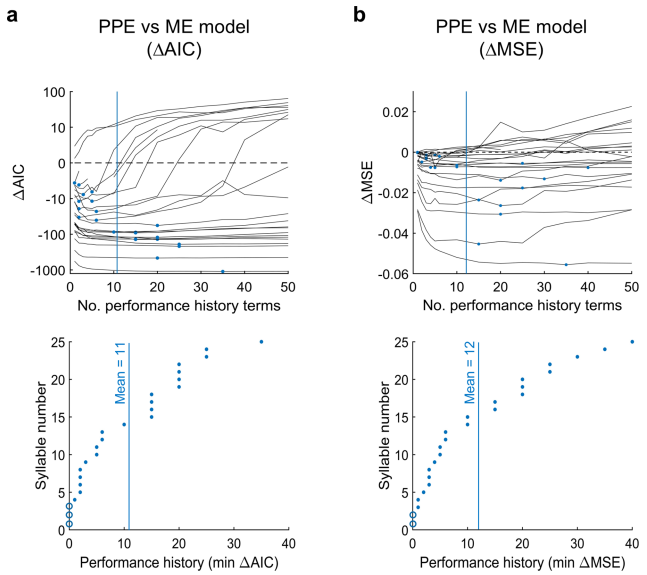
Extended Data Fig. 7 | Dopamine predicts future song evolution. a, Movement along the dopamine vector is not sensitive to song-block size (see Methods) and the effect is consistently significant at a block size of 10 or more renditions. Replication of analysis shown in Fig. 3f over a range of song-block sizes. Movement along the dopamine vector averaged across all n = 25 syllables (blue trace and shading, mean ± s.e.m.; black line and grey shading, mean ± 1.96 s.d. of shuffled data). **b**, Six example syllables plotted as in Fig. 3d. Average across all dopamine vectors for the example syllable across all focal groups of 150 syllable renditions. **c**, Extension of analysis shown in Fig. 3f (see Methods) for both forward (blue) and backward (red) in rendition steps relative to the focal song block, plotted similarly to **a** (black bar and star, $P < 0.05$, paired t-test, region in which the difference in absolute magnitudes between the movement along the dopamine vector is significantly greater in the future than in the past).



Extended Data Fig. 8 | Dopamine correlates more with the future than with the past at points of abrupt change in song. a, Schematic of analysis. Three consecutive, non-overlapping song blocks are selected: past, current, and future. Past and future vectors are computed from the difference in median song locations in each song block. Song renditions in the current song block are projected onto the past and future vectors and the correlation with dopamine in the current song block are computed along these dimensions. This is repeated at steps of 25 renditions across all song development (see Methods). Local song-block triplets are divided into two categories depending on the local geometry of the song trajectory: 'hinge' ($\theta < \pi/2$) trajectories and 'arrow' ($\theta > \pi/2$) trajectories (θ is defined between 0 and π). **b**, Distribution of angles between the future vector and the negative of the past vector across all sampled triplets of song blocks along song development across n = 25 syllables (block size, 150 renditions). **c**, Average correlation coefficient between current song renditions projected onto the future and past song vectors and current dopamine transients (points and error bars, mean ± s.e.m over syllables; blue/pink stars, $p < 0.05$, two-sided t-test, correlations between dopamine and past/future song vectors are significantly different from zero; black circle, $p < 0.05$, paired t-test, correlations between dopamine and past/future song vectors are significantly different). **d**,**e**, Plotted as in **c**, but across multiple song-block sizes for 'hinge' (**d**) and 'arrow' (**e**) trajectories. **f**, Example syllable spectrograms showing the emergence of an element present in the tutor song but not in early renditions during development. Left, syllable renditions produced on a single day during development (day = 61) were divided into 'Immature' versions (without the tutor element, purple box) and 'Mature' versions (with the tutor element). Right, example spectrograms of the adult version (day = 100) and corresponding tutor syllable. **g**, Dopamine responses for the example syllable shown in **a** on a single day during development (day = 61) for the 'Mature' (n = 423, top) and 'Immature' (n = 402, bottom) versions plotted above average ΔF/F signals (all plots aligned to syllable onset; blue, 'Mature' renditions; red, 'Immature' renditions; $P < 0.05$, two-sided t-test, significant difference between dopamine responses for 'Immature' and 'Mature' renditions). **h**,**i**, A second example syllable plotted as in **f**,**g** using day = 63 during development ($P < 0.05$, two-sided t-test, significant difference between dopamine responses for 'Immature' (n = 201) and 'Mature' (n = 137) renditions). **j**,**k**, A third example syllable plotted as in **f**,**g** using day = 61 during development ($P < 0.05$, two-sided t-test, significant difference between dopamine responses for 'Immature' (n = 240) and 'Mature' (n = 310) renditions).



Extended Data Fig. 9 | Dopamine tracks performance history in individual syllables. a, Magnified view of the regression coefficients over all development for the current rendition (0) and the previous two renditions (−1, −2) for all n = 25 syllables plotted as in Fig. 4d (error bars, mean ± s.e.m; *$P < 0.001$, two-sided t-test; see Methods). **b**, Fraction of performance history coefficients in linear regression fit to dopamine in Fig. 4b that are less than 0 compared to the shuffled distribution (see Methods). Coefficients are consistently negative above chance ($P < 0.05$, computed from shuffled distribution, see Methods). **c**,**d**, Linear regression fit of performance to dopamine for early (**c**) and late (**d**) development, plotted as in Fig. 4d. **e**, Plotted as in Fig. 4c. Coefficients of a linear regression fit of the song history to the dopamine signal for example syllables fit across development (blue trace, average over days; error bars, ±s.e.m.).



Extended Data Fig. 10 | Dopamine transients represent prediction errors better than mismatch errors. a, Top, comparison of the PPE and ME models with different numbers of performance history terms (see Methods). Each trace is a single syllable model comparison (blue dot, the minimum, negative Akaike information criterion, ΔAIC, the PPE model's best fit relative to the ME model). The magnitude and sign of ΔAIC indicate the relative superiority of one model over another. 25/25 syllables were included in the ME/PPE comparison because all syllables had at least one ME or PPE linear regression onto dopamine with a significant R^2 value (computed using the fitlm function in MATLAB). 22/25 syllable traces have a negative minimum (blue dot excluded from the three syllable traces which are always above 0; that is, the ME model always outperforms the PPE model). Blue line indicates the average best number of history terms across syllables (including the three syllables in which no. history terms = 0). Bottom, summary plot of number of history terms in best selected model from the top traces. 3/25 models had no PPE models which improved fit to dopamine over the ME model (shown with open circle). **b**, Top, as in **a**, but applying ΔMSE as a secondary model comparison metric (see Methods). As in **a**, the sign and magnitude of ΔMSE indicates the relative superiority of the ME versus PPE model. Negative values indicate that the PPE model outperforms the ME model. The ΔAIC and ΔMSE metrics found the average best number of history terms across syllables to be $n = 11$ and $n = 12$, respectively. Bottom, plotted similarly to **a** for ΔMSE.

nature portfolio

Corresponding author(s): Vikram Gadagkar

Last updated by author(s): Jan 19, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Photometry data was collected using Doric Neuroscience Studio versions 5.4.1.12 and 6.1.5.0. Song data was recorded using MATLAB 2014b.
Data analysis	Data were analyzed using custom code written in MATLAB 2022a.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The data used for this study are available from the corresponding author on request.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes used in this study are typical to the field, eg. Hisey et al., 2018; Xiao et al., 2018.
Data exclusions	Syllables were excluded from further analysis if labeling did not exceed a 90% F1 score.
Replication	Results were replicated in 6 birds.
Randomization	Juvenile birds were randomly selected within age groups.
Blinding	Data collection was fully automated to exclude experimenter bias. Experimenters were blind to dopamine data while relative proximity of syllables was being calculated.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	Juvenile male zebra finches (Taeniopygia guttata)
Wild animals	No wild animals were used in this study.
Reporting on sex	Only male zebra finches were used since females do not produce courtship song.
Field-collected samples	No field-collected samples were used.
Ethics oversight	All experiments were approved by the Columbia University Institutional Animal Care and Use Committee and were performed in accordance with NIH guidelines.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*