NEURAL BASIS OF PERFORMANCE EVALUATION IN BIRDS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Ruidong Chen

August 2020

NEURAL BASIS OF PERFORMANCE EVALUATION IN BIRDS

Ruidong Chen, Ph.D.

Cornell University 2020

Reinforcement learning (RL) is one of the most successful theories in the intersection between animal behavior, neural basis of learning, and artificial intelligence. RL-like computation in the brain is evidenced by representation of a key variable in the dopaminergic system, the reward prediction error. Yet most studies of the dopaminergic system have been in situations where an animal is motivated by external rewards, such as food or juice. Does RL apply to learning of complex behavior without immediate external reward? What are the neural circuits responsible for learning a motor sequence such as birdsong?

Here I introduce work in the songbird zebra finch demonstrating RL applies to performance learning based on a flexible internal benchmark. Dopaminergic neurons in the midbrain were shown to report a performance prediction error.

I expand the song system involved in song learning to neural circuits responsible for RL. I found ventral pallidum (VP), a major input area to the dopaminergic midbrain, to be involved in song learning. I discovered a rich representation in VP including performance prediction related activity that is signaled to the dopaminergic midbrain. I contextualize my findings in the songbird system and hypothesize a consolidator-actor-critic architecture for song learning.

I discovered neural representations in VP and the dopaminergic midbrain that switch between body movement and song timing depending on the animal's behavioral states of singing and non-singing.

I also report the first neural recordings from a parrot. Surprisingly, representations of both vocalization and body movements were found in the parrot primary vocal motor cortex. This finding sets up the foundation for future neurophysiological studies of social behavior in the parrot.

BIOGRAPHICAL SKETCH

Ruidong Chen graduated Guangzhou No. 47 High School in 2010. He spent a year in preparatory studies in Zhejiang University, where he was introduced to Dr. Strangelove in a class. In 2011 he started at University of Hong Kong for a bachelor's degree in computer science. On a study abroad program he went to University of Illinois at Urbana-Champaign, where he first learned about neuroscience in introduction to neurobiology. In the summer of 2013, he worked with the professor of that class Dr. Mark Nelson on programming a robot driven by curiosity. Ruidong graduated HKU in 2014 with First Class Honours and joined the department of Neurobiology and Behavior at Cornell University that fall, where he has been studying the neural basis of performance evaluation with Dr. Jesse Goldberg.

ACKNOWLEDGEMENTS

My experience of the last six years in graduate school has been rewarding thanks to the enthusiastic support at every stage from my advisor Jesse Goldberg. Jesse taught me how to be a scientist, both in the technical sense of how to do science, and in the broader sense of recognizing what matters most in science is supporting the people who do it. With every interaction and every decision, Jesse has exemplified this philosophy and I can only aspire to live by it.

My committee members Joe Fetcho, Melissa Warden, and Josh Dudman have been a source of invaluable feedback and advice whenever I asked. I want to especially thank Joe for his insights on how to broadly think about neuroscience, Melissa for the opportunity to work on a rotation project in her lab, and Josh for making important connections between my experimental findings and relevant theoretical work.

This dissertation depended on the efforts of many collaborators: Vikram Gadagkar, Pavel Puzerey, Eliza Baird-Daniel, Alex Farhang, Andrea Roeser, Tori Riccelli, Archana Podury, Kamal Maher, Don Murdoch, Eugene Kim, and Treasure Nwokeleme. I'd also like to thank excellent technicians Ariana Enzerink and Brian Kardon for their contributions to my work in the lab.

I'm grateful to people who have made my time in the lab fun and stimulating: Teja Bollu, Brendan Ito, Anindita Das, Mei Liu, and Heidi Huang. Special thanks to Teja, Andrea, Brendan, and Heidi for being delightful office mates until Quarantine forced us apart.

The department of neurobiology and behavior has become a second home to me. I've always felt welcome and supported by the current and former staff of NBB, including Kathie Ely, Saundra Anderson, Adrienne Mason, Lori Maine, Jaclene Weston, Julie Hughes, Terri

Natoli, and Sara Eddleman. I'm also grateful to Rich Moore, Stephanie Pierson, Jessie Gorges and other members of CARE team for taking care of our animals, and Brian Mlodzinski for maintaining computing equipment critical for our experiments.

Ithaca is an amazing place to live but wouldn't be half as fun without friends. I am grateful for the companionship from Akash Guru, Ryan Post, Changwoo Seo, Laura Manella, Joel Tripp, Hanna Chapman, Irene Ballagh, Rose Tatarsky, Sam Whitehead, Hayden Waller, John Ronalter, Chieh-Ren Hsia, and many others.

I'd like to thank Ron Harris-Warrick for his leadership in running the neurobiology course when I was a TA. And thanks to the students in my section, I learned how rewarding it is to be a teacher.

I'd like to thank Mark Nelson for giving me an opportunity in research in the summer of 2013. For someone who never knew any scientist outside of the classroom, it was simply inconceivable that I could be a scientist until then.

Finally, I am most thankful to my parents. I often wondered whether it is ethical and worthwhile to spend my time eight thousand miles away from home, studying what at times feel insignificant compared to staying close to and caring for my family. I felt this most acutely in the beginning of 2020, when I came back to the US to finish my PhD, while China was still the center of what would soon become the Pandemic. Yet my parents never lost faith in me and have supported and encouraged me to pursue whatever I wish to achieve. Ultimately they are the source of my strength.

TABLE OF CONTENTS

Biographical sketch iii
Acknowledgements iv
Table of contents vi
List of figures viii

Chapter 1: Introduction
Chapter text *1*
References *6*

Chapter 2: Songbird ventral pallidum sends diverse performance error signals to dopaminergic midbrain
Abstract *8*
Introduction *9*
Results *11*
Discussion *27*
Methods *32*
Acknowledgements *41*
References *42*
Appendix *49*

Chapter 3: Movement signaling in ventral pallidum and dopaminergic midbrain is gated by behavioral state in singing birds
Abstract *59*
Introduction *60*
Results *61*
Discussion *70*
Methods *72*
References *76*

Chapter 4: Actor-critic reinforcement learning in the songbird

Abstract *78*

Main Text *79*

Summary *94*

References *95*

Chapter 5: Neural activity related to vocal and nonvocal social behavior in the parrot

Abstract *99*

Introduction *100*

Results *102*

Discussion *105*

References *107*

LIST OF FIGURES

Figure 2.1 Auditory cortical and ventral pallidal stimulation drive diverse changes in VTAx neuron firing.

Figure 2.2 VP lesions impair song learning.

Figure 2.3 VP neurons exhibit performance and prediction error signals during singing.

Figure 2.4 VP neurons exhibit temporally precise song-locked activity during singing.

Figure 2.5 VP sends diverse error- and prediction-related signals to VTA.

Figure 2.6 An actor-critic circuit motif for computing performance prediction error.

Figure 2.S1 AIV and VP projects to VTA.

Figure 2.S2 Auditory cortical and ventral pallidal stimulation drive diverse changes in VTAx neuron firing.

Figure 2.S3 Firing patterns of VP neurons.

Figure 2.S4 Latency and duration of VP error responses.

Figure 2.S5 VP neurons can exhibit error signals that occur specifically during singing and was not attributable to animal movement.

Figure 2.S6 RA, Uva, AIV, DLM, Ov, and VTAx project to VP.

Figure 3.1 Measuring neural activity and movement as birds transition into and out of singing states.

Figure 3.2 VP and VTA neurons can exhibit movement-locked activity.

Figure 3.3 Example neurons that exhibit movement selectivity in the non-singing state only.

Figure 3.4 Example neurons that switch their tuning between song timing and movement at singing state boundaries.

Figure 4.1 Dopaminergic error signals in singing birds.

Figure 4.2 Actor-critic RL for both reward and song learning.

Figure 4.3 Variability- and timing- generating pathways in the song system.

Figure 5.1 Example call- and warble-locked AAC neurons.

Figure 5.2 AAC neural responses to bird's own calls, warble song, movements, and calls and movements generated by a partner bird.

CHAPTER 1
INTRODUCTION

Animals learn complex motor sequences through experience. Human language, music, sports are all familiar examples. In order to understand this aspect of nature, a theory for how motor learning is realized in a biological network is desired. This remains elusive, partially due to a lack of successful algorithms that can reproduce naturalistic learning, and partially due to a lack of relevant data against which to test competing theories. Correspondingly, a full understanding of the neural circuitry responsible for motor learning is lacking.

A promising framework for understanding motor learning is reinforcement learning (RL). At the root of RL is the intuitive idea that any action that leads regularly to preferred outcome should be reinforced over others. This idea was used to explain behavioral observations of cats learning to escape from a box by pulling and pushing various contraptions (Thorndike, 1911). A key observation was that the cat would start with apparently random actions, and only accidentally open the door. With experience, the time it took to open the same door reduced, suggesting an association has formed between the box and the correct action to open the door.

Expanding on this idea, theorists have formalized a family of algorithms that start with random actions, and gradually modify a policy through trials to improve the reward outcome (Sutton and Barto, 1998). A generic RL problem is structured as a Markov Decision Process, with a set of states that the agent can be in, actions available in each state, a probabilistic transition function that maps state and action to the next state, and reward available at goal states. In this mathematical formulation reward can be arbitrarily mapped onto states. A family of algorithms can take such a problem and efficiently solve the problem of navigating the MDP space to maximize total received reward over time. In order to test which, if any, RL algorithm

is used by neural circuits in guiding animal behavior, neuroscientists have studied choice behavior in association tasks. In these studies, reward has been mostly construed as an external feedback such as a drop of juice (Bayer and Glimcher, 2005; Schultz et al., 1997). This makes sense because it simulates the naturalistic situation where an animal encounters food or water during foraging. And applying RL algorithms to the behavior of animals in tasks and to neural activity has enjoyed enormous success. Neural circuits responsible for implementing RL has generated much interest ever since the landmark discovery of a reward prediction error (RPE) signal exhibited by dopaminergic cells in the midbrain (Schultz et al., 1997). The RPE signal was exactly as expected if dopamine is used as the error signal in the temporal difference learning algorithm (Schultz et al., 1997). Because it's a clear signal it also provided an access point into the RL circuitry of the brain.

The nature of reward in learning a complex sequence such as language is less clear. When a young child learns to speak, not every improvement in pronunciation is rewarded by either sweets or praise. Indeed, even children in cultures where parents do not often talk to them learn to speak (Cristia et al., 2019). Could there be an internally generated reward signal that's based on how 'good' one's own voice sounds?

The songbird zebra finch provides an ideal model for studying the neural basis of internally motivated skill learning. Zebra finches learn to produce a song motif composed of a sequence of syllables, which is copied from a tutor (Marler, 1997). After auditory exposure to the tutor song, not unlike babbling children, young birds spontaneously sing thousands of syllables per day in practice, and gradually improve the similarity to tutor song.

The zebra finch brain has a set of specialized structures known as the song system. In particular, a basal ganglia nucleus Area X receives dense dopaminergic innervation and is

required for song learning (Hoffmann et al., 2016; Person et al., 2008; Scharff and Nottebohm, 1991). By recording from Area X-projecting VTA dopamine neurons (VTAx) while perturbing the auditory feedback of a singing bird, we found that VTAx neurons paused their firing when a target syllable was distorted, consistent with a negative performance error; and increased their firing in a burst when a distorted syllable was left undistorted, consistent with a positive performance error (Gadagkar et al., 2016). The burst was especially informative, because it demonstrates an internal expectation that the target syllable had lower performance. In other words, the dopamine system has access to an internal estimate of performance learned through experience. What are the circuit mechanisms for maintaining and updating expected performance? In Chapter 2, I identify the upstream region ventral pallidum (VP) as a critical hub for the construction of performance prediction.

A more general form of RL is multi-objective reinforcement learning (MORL), where a single agent pursues multiple independent objectives by training sub-agents with their own specific reward functions (Liu et al., 2015). Animals may have multiple objectives that are connected to multiple effectors. It is not known how neural circuits performing RL can multiplex distinct objectives into distinct behavioral modules at the same time.

Here again songbirds provide a great opportunity to ask questions about how animals pursue multiple objectives at the same time. Zebra finches often sing and hop around the cage at the same time. In a collaboration with Don Murdoch, I tested the idea that different reinforcers may have restricted access to certain kinds of behavior. We found that bright strobe light could reinforce the choice of perch but not the pitch of syllables, whereas broadband noise could reinforce both modalities. Importantly the valence of broadband noise was inversed: while it's well known that DAF is negatively reinforcing to vocal variants (Andalman and Fee, 2009;

Tumer and Brainard, 2007), it was positively reinforcing to the navigation system and produced a perch preference (Murdoch et al., 2018). This finding is reminiscent of the saliency reversal of salt water observed in sodium-deprived rats, where the same homeostatic objective results in variable value of the same stimulus, depending on the animal's internal state of salt appetite (Zhang et al., 2009). In the case of MORL, opposing saliencies can coexist within different sub-agents. Whether reinforcement signals for multiple sub-agents coexist in animals remains to be seen. One possible test is to record from dopamine neurons with identified projection targets in a behavioral paradigm that invokes learning with multiple objectives (Murdoch et al., 2018).

In Chapter 3, I report switching of neural representation in songbird VTA and VP. I found many neurons had movement-locked firing when the bird was not singing, and lost that movement tuning during singing, even though birds were not still and instead moved their head quite often during singing. This may provide a hint to the possible neural mechanism of evaluating multiple objectives.

In Chapter 4, I hypothesize an actor-critic architecture for song learning based on findings from VP and VTA. Previous work in the study of mammalian basal ganglia provides a framework to identify each subregion as a component in an actor-critic RL agent (Houk et al., 1995; Joel et al., 2002; O'Doherty et al., 2004). Here I raise the possibility that it's a unifying principle that applies to the songbird as well. I further integrate this hypothesis with the idea that the vocal motor cortex RA serves as a consolidator.

From a comparative perspective, parrots have similarities in different aspects to songbirds and humans, and are a great model to ask what are general principles in vocal learning and performance, in relation to species specific neuroanatomy and ethology. To that end I started recording from interacting budgerigars last year. In Chapter 5, I introduce neural recordings

from the vocal motor cortex of a pair of budgerigars that demonstrate both similarity to the songbird and novel characteristics.

REFERENCES

Andalman, A.S., and Fee, M.S. (2009). A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. Proc Natl Acad Sci U S A 106, 12518-12523.

Bayer, H.M., and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 47, 129-141.

Cristia, A., Dupoux, E., Gurven, M., and Stieglitz, J. (2019). Child-Directed Speech Is Infrequent in a Forager-Farmer Population: A Time Allocation Study. Child Dev 90, 759-773.

Gadagkar, V., Puzerey, P.A., Chen, R., Baird-Daniel, E., Farhang, A.R., and Goldberg, J.H. (2016). Dopamine neurons encode performance error in singing birds. Science 354, 1278-1282.

Hoffmann, L.A., Saravanan, V., Wood, A.N., He, L., and Sober, S.J. (2016). Dopaminergic Contributions to Vocal Learning. J Neurosci 36, 2176-2189.

Houk, J., Adams, J., and Barto, A. (1995). A Model of How the Basal Ganglia Generate and Use Neural Signals that Predict Reinforcement, Models of Information Processing in the Basal Ganglia (eds. JC Houk, JL Davis and DG Beiser), 249/270. (MIT Press).

Joel, D., Niv, Y., and Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives. Neural Netw 15, 535-547.

Liu, C., Xu, X., and Hu, D. (2015). Multiobjective reinforcement learning: A comprehensive overview. IEEE Transactions on Systems, Man, and Cybernetics: Systems 45, 385-398.

Marler, P. (1997). Three models of song learning: evidence from behavior. J Neurobiol 33, 501-516.

Murdoch, D., Chen, R., and Goldberg, J.H. (2018). Place preference and vocal learning rely on distinct reinforcers in songbirds. Sci Rep 8, 6766.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R.J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 304, 452-454.

Person, A.L., Gale, S.D., Farries, M.A., and Perkel, D.J. (2008). Organization of the songbird basal ganglia, including area X. J Comp Neurol 508, 840-866.

Scharff, C., and Nottebohm, F. (1991). A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. J Neurosci 11, 2896-2913.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science 275, 1593-1599.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement learning: an introduction (Cambridge, MA: MIT Press).

Thorndike, E.L. (1911). Animal Intelligence (Darien, CT: Hafner).

Tumer, E.C., and Brainard, M.S. (2007). Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. Nature 450, 1240-1244.

Zhang, J., Berridge, K.C., Tindell, A.J., Smith, K.S., and Aldridge, J.W. (2009). A neural computational model of incentive salience. PLoS Comput Biol 5, e1000437.

CHAPTER 2

SONGBIRD VENTRAL PALLIDUM SENDS DIVERSE PERFORMANCE ERROR SIGNALS TO DOPAMINERGIC MIDBRAIN

Abstract

Motor skills improve with practice, requiring outcomes to be evaluated against ever-changing performance benchmarks. Yet it remains unclear how performance error signals are computed. Here we show that the songbird ventral pallidum (VP) is required for song learning and sends diverse song timing and performance error signals to ventral tegmental area (VTA). Viral tracing revealed inputs to VP from auditory and vocal motor thalamus, auditory and vocal motor cortex, and VTA. Our findings show that VP circuits, commonly associated with hedonic functions, signal performance error during motor sequence learning.

*Previously published as Chen, R., Puzerey, P.A., Roeser, A.C., Riccelli, T.E., Podury, A., Maher, K., Farhang, A.R., and Goldberg, J.H. (2019). Songbird Ventral Pallidum Sends Diverse Performance Error Signals to Dopaminergic Midbrain. Neuron 103, 266-276 e264

Introduction

When practicing a piano concerto, you could evaluate your performance relative to a fixed template, such as an auditory memory of Glenn Gould playing Chopin's Prelude No. 4. Yet reinforcement learning (RL) theory suggests improved learning if you instead learn from prediction errors, in which a note is reinforced only if it sounds better (closer to the template) than predicted based on your past performance (Sutton and Barto, 1998). Reinforcement learning proceeds via incremental improvement in performance quality, requiring performance to be evaluated against benchmarks that change with practice (Schmidt et al., 2018; Thelen, 1995). Neural mechanisms of performance benchmarking and evaluation remain poorly understood.

Songbirds provide a tractable model system to study performance evaluation. First, songbirds have a specialized circuit, the 'song system', that includes a projection from VTA dopamine (DA) neurons to the striatopallidal nucleus Area X (Gale et al., 2008; Reiner et al., 2004). Second, zebra finches gradually learn to imitate a sequence of song notes, or syllables, acquired from a single tutor, suggesting they have a 'fixed template' they aspire to learn (Marler, 1997). Yet consistent with RL theory, song syllables are not evaluated exclusively against this fixed template but are additionally evaluated against syllable-specific performance benchmarks that change with practice. Specifically, Area X projecting VTA (VTA_X) DA neurons recorded during singing are phasically suppressed by distorted auditory feedback (DAF) during singing (Gadagkar et al., 2016). DAF, though not generally aversive (Murdoch et al., 2018), induces a perceived vocal error on distorted renditions such that undistorted renditions are reinforced (Ali et al., 2013; Andalman and Fee, 2009; Hamaguchi et al., 2014; Keller and Hahnloser, 2009; Tumer and Brainard, 2007). Over day timescales, DAF also reduces the predicted quality (i.e. proximity to template) of DAF-targeted syllables. When a reliably distorted target syllable is left undistorted, VTA_X neurons exhibit phasic bursts at the precise moment of the song when DAF is

predicted to occur but does not occur; and the magnitude of this burst depends on past error probability (Gadagkar et al., 2016). Thus, VTA_X neurons signal errors in predicted song quality, i.e. the difference between how good (close to the template) a syllable sounded and how good it was predicted to be based on recent practice.

To signal performance prediction error, songbird DA neurons must compute the difference between actual 'just heard' error and the error that was predicted at that specific time-step of the song. The roles of upstream projections to VTA in these computations remain unclear. One projection to VTA comes from a high-order auditory cortical area that forms a 'cup' around the robust vocal motor nucleus of the arcopallium (RAcup) (Bottjer et al., 2000; Kelley and Nottebohm, 1979; Mello et al., 1998; Vates et al., 1996). RAcup, which is located in the ventral intermediate arcopallium (AIV), is necessary for song learning and sends error signals to VTA (Gale et al., 2008; Mandelblat-Cerf et al., 2014), but it remains unknown how AIV influences VTA firing.

The second major forebrain input to VTA in songbirds comes from a ventral pallidal (VP) region outside the classic song system, ventral and medial to Area X (Gale and Perkel, 2010; Gale et al., 2008). Yet it remains unknown if VP is important for song learning and what, if any, singing-related signals it exhibits.

Here we combine lesion, electrophysiology, and viral tracing studies to demonstrate for the first time that songbird VP is required for learning, exhibits performance error signals, and receives previously unknown inputs from nuclei of the 'classic' song system. More generally, our results directly implicate VP in learning a purely motor sequence task like birdsong.

Results

Auditory cortical stimulation causes diverse responses in VTA_X neurons

To identify circuits important for performance evaluation, we injected retrograde tracer into the Area X-projecting part of VTA, where performance error encoding DA neurons important for song learning reside (Gadagkar et al., 2016; Hisey et al., 2018; Hoffmann et al., 2016; Xiao et al., 2018). Consistent with past work, retrogradely labeled neurons were observed in AIV (Bottjer et al., 2000; Gale et al., 2008; Kelley and Nottebohm, 1979; Mandelblat-Cerf et al., 2014; Mello et al., 1998; Vates et al., 1996)(Figure 2.S1A-B). Previous studies showed that during singing, DAF causes activation of VTA-projecting AIV neurons and, at a slightly longer latency, pauses in VTA_X neurons (Gadagkar et al., 2016; Mandelblat-Cerf et al., 2014).

To test if AIV stimulation affects VTA_X firing, we recorded VTA neurons while electrically stimulating AIV in anesthetized birds (n = 7 birds, Figure 2.1A). AIV stimulation induced phasic rate changes in all wide spiking, antidromically identified VTA_X neurons (n=8 antidromic neurons, spike width 0.38±0.03ms, Figure 2.S2K,L) (Schultz and Romo, 1987). Responses included suppression followed by activation (n=3, Figure 2.S2B-D, latency 75+/-19ms), activations (n=3, Figure 2.S2F,H-I, latency 112+/-54ms), and suppressions (n=2, Figure 2.1F and S2E,J, latency 15-35ms). In putative VTA interneurons with thin spikes (n=13, spike width 0.26±0.02ms, Methods), AIV stimulation caused low latency activations (n=7, latency 29±12ms) or suppressions (n=5, latency 20±7.8ms). When we recorded simultaneously from a VTA_X neuron and a thin spiking interneuron at the end of the same electrode, we found that AIV stimulation activated the interneuron, which in turn could suppress the VTA_X neuron (Figure 2.1B-F).

Together, these findings suggest that VTA contains complex local circuitry, including one that implements feedforward inhibition to invert excitatory signals from AIV, consistent

with the idea that performance error-induced activations in AIV can drive pauses in VTA_X firing during singing. Notably, the AIV-VTA projection resembles cortical projections to VTA in mammals that also can target local GABAergic interneurons and inhibit dopaminergic firing (Beier et al., 2015; Carr and Sesack, 1999; Creed et al., 2014; Moreines et al., 2017; Patton et al., 2013).

Ventral pallidal stimulation causes diverse responses in VTA_X neurons

Following tracer injection into the Area X projecting part of VTA, retrogradely labeled neurons were also observed in a ventromedial basal ganglia region termed VP in previous studies (Figure 2.S1C) (Gale and Perkel, 2010; Gale et al., 2008; Reiner et al., 2004). To test if VP can influence VTA_X activity, we electrically stimulated VP while recording antidromically identified VTA_X neurons in anesthetized birds (n=4 birds, Figure 2.1A). Response to VP stimulation included suppression followed by activation (n=3/5 neurons tested, Figure 2.1G, S2N-P), suppression (n=1/5, Figure 2.S2Q), and activation (n=1/5, Figure 2.S2R). Responses to VP stimulation had lower latency than AIV (VP: 15+/-2.7 ms, AIV: 22.8+/-7.8 ms, $p<0.01$, WRS test, Figure 2.S2L). The observation that VP can influence VTA_X activity in complex ways is consistent with diverse VTA-projecting cell types in VP (Person et al., 2008), as well as the presence of feedforward inhibition in VTA that can invert incoming signals.

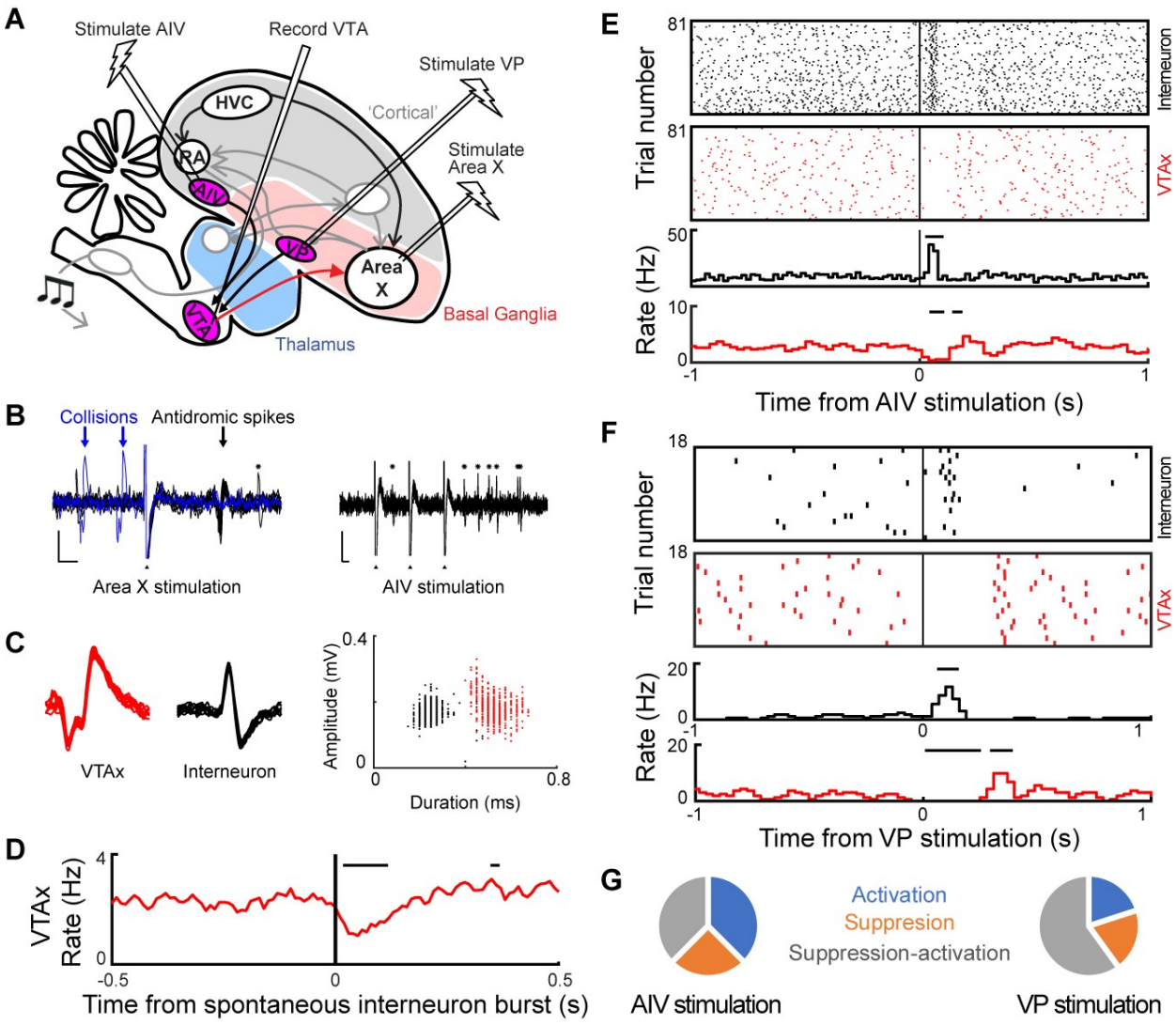

Figure 2.1. Auditory cortical and ventral pallidal stimulation drive diverse changes in VTA$_X$ neuron firing. See also Figure 2.S1 and 2.S2.

(**A**) Schematic of experimental strategy for recording VTA$_X$ neuronal response to AIV and VP stimulation. (**B** to **G**) Experiment conducted on simultaneously recorded wide-spiking VTA$_X$ neuron and thin-spiking VTA interneuron. (**B**) Antidromic identification (left), and AIV stimulation (right). Arrows, VTA$_X$ antidromic spiking (black) and collisions (blue). Asterisks: interneuron spikes. Filled triangles: stimulation artifacts. Scale bars: 0.1mV (vertical), 2ms (horizontal). (**C**) Left, overlay of raw trace from 10 VTA$_X$ spikes (red). Middle, 10 VTA interneuron spikes (black). Right, amplitude and duration of all recorded VTA$_X$ (red) and interneuron (black) spikes. (**D**) Cross-correlogram of spontaneous firing between the two units.

Horizontal bars, significant response ($p<0.05$, z-test) (Methods). (**E**) Raster plots (top) and rate histograms (bottom) of interneuron (black) and VTA_X neuron (red), aligned to AIV stimulation. Horizontal bars indicate significant response ($p<0.05$, z-test) (Methods). (**F**) Same as (**E**) for VP stimulation. (**G**) Summary of VTA_X response types (n=8 neurons with AIV stimulation, n=5 neurons with VP stimulation).

Ventral pallidal lesions impair song learning

To test if VP is important for song learning, we conducted lesion experiments in juvenile birds. To specifically lesion the part of VP that is part of the Area X - VP - VTA - Area X loop previously hypothesized to play a role in learning (Gale and Perkel, 2010), we carried out electrophysiologically-guided excitotoxic lesions of VP in juvenile birds and evaluated their adult songs (Methods). During lesion surgery, stimulation electrodes were implanted into Area X and the boundaries of orthodromic stimulation-evoked responses were mapped with recording electrodes in VP. This mapping specified the locations of excitotoxin injections (Methods). VP lesions significantly impaired song learning (Song imitation score, WRS test, $p=0.014$, n=6 lesion birds and 7 controls, Methods) (Figure 2.2).

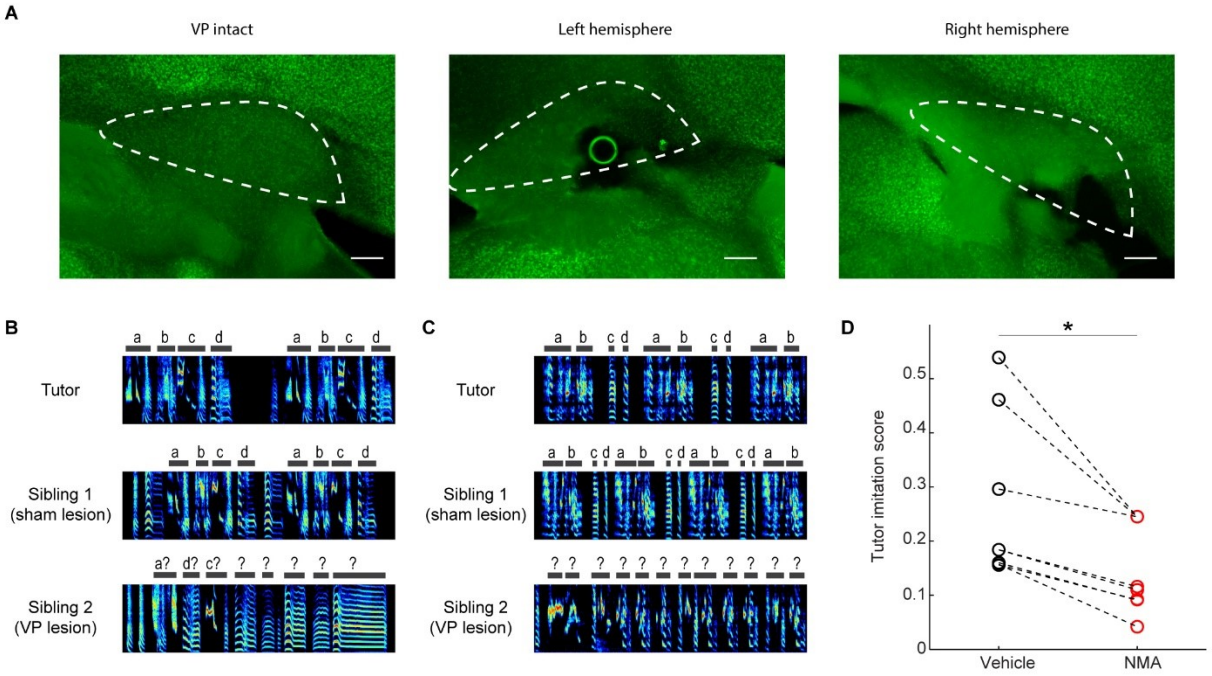

Figure 2.2. VP lesions impair song learning.

(**A**) Lesions were confirmed in neuronal nuclear stained (anti-NeuN) slices as extensive cell death. Scale bars, 200 microns. (**B**) Tutor song (top), adult song of sham lesioned (middle), and VP lesioned (bottom) siblings. Each spectrogram is 2 seconds long. (**C**) Same as (**B**) for another pair. (**D**) Adult song of vBG lesioned birds had lower similarity to their tutor compared to controls (rank-sum test between all controls and all lesioned birds, p=0.014. n=7 sham lesion, n=6 VP lesion).

Firing patterns of VP neurons in singing birds

To test how VP may guide learning, we recorded VP neurons in singing birds. With some exceptions detailed below, neurons could not easily be categorized into distinct cell classes because they exhibited a continuum of mean firing rates, spike widths, and discharge patterns (firing rate in singing, 43±3.3 Hz, range 0.38-310 Hz; spike width, 0.25±0.008 ms, range 0.05-0.59 ms; CV_{ISI} in singing: 1.10±0.03, range: 0.4-2.9; n=162 neurons; Figure 2.S3A-G). This heterogeneity is consistent with the diverse striatal and pallidal cell types intermingled inside songbird VP (Person et al., 2008).

Notably, neuronal discharge heterogeneity is also observed in mammalian VP, leading several studies to categorize neurons on the basis of their responses to primary rewards and cues that predict them (Ahrens et al., 2016; Ito and Doya, 2009; Ottenheimer et al., 2018; Richard et al., 2016; Richard et al., 2018; Smith et al., 2009; Tachibana and Hikosaka, 2012; Tindell et al., 2004). Below we will proceed similarly by categorizing neurons on the basis of their responses to song error and syllable timing.

VP neurons exhibit performance and prediction error signals during singing

To specifically test the role of VP neurons in learning, we used syllable-targeted distorted auditory feedback (DAF) to control perceived error during our recordings (Andalman and Fee, 2009; Tumer and Brainard, 2007)(Methods). Beginning days prior to recordings, a specific 'target' song syllable was either distorted with DAF or, on randomly interleaved renditions, left undistorted altogether (distortion rate at target syllable 48.0±1.4%, mean ± s.e.m., n=39 birds). Days of pre-training with syllable-targeted DAF reduces the predicted quality of target syllables such that undistorted renditions are signaled as better-than-predicted by VTA_X DA neurons (Gadagkar et al., 2016).

To test if VP neurons exhibited online error responses, we compared the activity between randomly interleaved renditions of distorted and undistorted songs. We defined an error response as a significant difference of firing in distorted and undistorted renditions. Surprisingly, although there are no known inputs to VP that carry auditory information in awake singing birds, significant error responses were observed in 31/128 VP neurons tested (Figure 2.3) (assessed by WRS test on number of spikes in the 125 ms window following DAF onset time (Keller and Hahnloser, 2009), 34/162 neurons were not tested due to low number of trials, Methods).

Neurons were either activated (n=16) or suppressed (n=15) by DAF during singing (quantified as z-scored rate difference between undistorted and distorted renditions, Methods). Biphasic responses were observed in 6/31 error responsive neurons: initial suppressions (or activations) were followed by significant activations (or suppressions) (Figure 2.3D, Methods).

A subset of error responsive neurons exhibited a significant maximum or minimum of firing rate following undistorted but not distorted target times (n=10/31 error responsive neurons, $p<0.05$, bootstrapping, Methods). We term these 'prediction error' responses because they cannot be explained by the external DAF sound and occur only following better-than-predicted song outcomes (Gadagkar et al., 2016). Prediction error responses could be rate peaks (n=8 error suppressed neurons, Figure 2.3B) or nadirs (n=2 error activated neurons, Figure 2.5D).

Overall, the latencies and durations of error responses were similar to those observed in downstream VTA_X DA neurons (latency: 55.0±3.8 ms; duration: 76.4±7.0 ms; n=31 cells, Figure 2.S4A-B) (Methods).

To test if error responses were performance-related (defined as error responses that only occur during singing), and not nonspecific auditory responses, we played back distorted and undistorted renditions of bird's own song (BOS) to non-singing birds in a subset of recorded neurons. Auditory error responses were rarely observed following passive playback of distorted versus undistorted BOS to awake, non-singing birds ($p>0.05$ in 33/35 neurons tested, including 14/16 error responsive neurons, WRS test)(Figure 2.S5A-D). To test if error responses were attributable to different movement patterns following distortions, we used custom head-mounted accelerometers to quantify movement during recordings. Movement did not differ between distorted and undistorted renditions (Figure 2.S5E), confirming that error responses were not

attributable to body movement (Gadagkar et al., 2016). Together these data demonstrate that VP neurons can exhibit error signals specifically during singing, consistent with performance error.

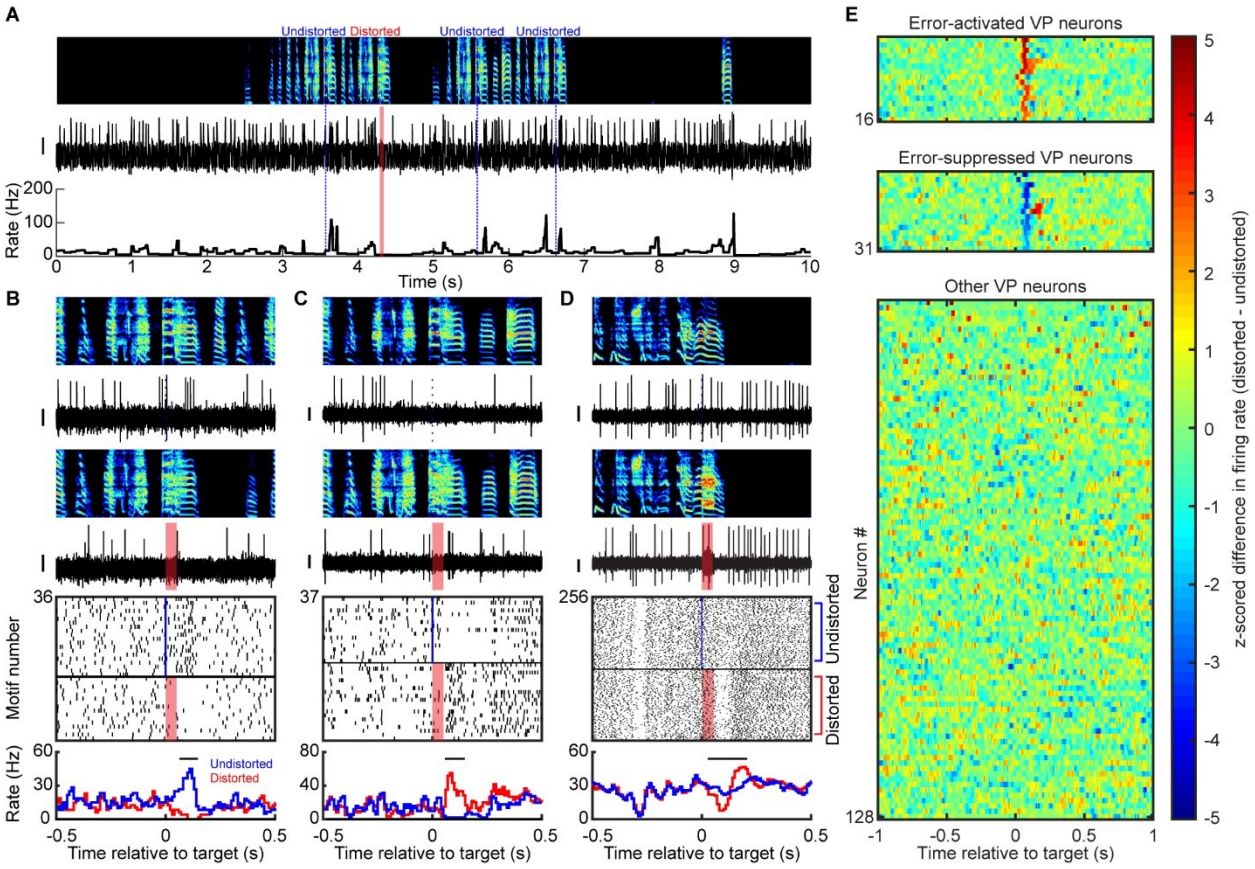

Figure 2.3. VP neurons exhibit performance and prediction error signals during singing. See also Figure 2.S3 , 2.S4, and 2.S5.

(**A**) Spectrogram (top), spike discharge (middle) and instantaneous firing rate (bottom) of a prediction error encoding VP neuron (DAF, red shading; undistorted targets, blue lines). (**B**) Expanded view of neuron from (**A**). Top to bottom: spectrograms, spiking activity during undistorted and distorted trials, corresponding spike raster plots and rate histograms (all aligned to target onset). Horizontal bars in histograms indicate significant different responses to distorted and undistorted renditions, and lack of horizontal bars indicate no significant different response was detected (error response, $p < 0.05$, WRS test) (Methods). (**C** and **D**) Additional examples of error-activated and error-suppressed neurons, same format as (**B**). (**E**) Each row plots the z-scored difference between distorted and undistorted target-aligned rate histograms. Error-activated neurons (top, n=16), error-suppressed neurons (middle, n=15), and non-error

responsive neurons (bottom, n=97) are independently sorted by maximal z-score. Scale bar for spiking activity in (**A**-**D**) is 0.15mV.

VP neurons exhibit temporally precise song-locked activity during singing

Many VP neurons also exhibited activity patterns temporally aligned to song. To quantify the precision of song locked discharge, we calculated the pairwise inter-motif correlation coefficients (IMCC) of instantaneous firing rates for all neurons recorded for 10 or more motifs (Goldberg et al., 2010; Goldberg and Fee, 2010; Kao et al., 2008; Olveczky et al., 2005)(Methods). A neuron that reliably discharges at the same time-steps across song renditions will have an IMCC=1; a neuron with random discharge unrelated to singing will have an IMCC=0. Most VP neurons exhibited significant song-locked discharge with a precision that varied among the population ($p<0.05$ in 96/162 neurons tested; IMCC=0.13±0.016, Figure 2.4A). Most neurons did not show time-locked responses to BOS playback ($p>0.05$ in 33/35 neurons tested, including 21/23 song-locked neurons, Methods).

Three cell 'types' were distinguished by extremely precise song-locked firing (IMCC>0.3) (Figure 2.4, n=10). A first type exhibited ultra-sparse discharge aligned to specific song syllables (n=2, sparseness index>0.5, Methods)(Figure 2.4A-B). These sparse neurons' discharge resembled striatal medium spiny neurons (MSNs) previously recorded in Area X (Goldberg and Fee, 2010; Woolley et al., 2014) and support the previous finding that striatal and pallidal cell types can be intermingled in songbird VP (Person et al., 2008). A second type exhibited stereotyped, rhythmic firing patterns visible as high frequency bursts aligned to specific song time-steps with millisecond precision (Figure 2.4A and 4C, n=6). In contrast to the first two types which did not exhibit error responses, a third type exhibited error responses as

well as time-locked response to BOS playback, consistent with a neuron that receives strong auditory input (n=2, IMCC>0.3 during BOS playback; Figure 2.4D, S5C-D).

Most song-locked neurons also exhibited significant rate modulations at various time-steps of the song (n=93/96 with significant rate maximum or minimum, Methods)(Figure 2.4E-F). Finally, yet other neurons were distinguished by dramatic increase or decrease in firing rate at the transition between non-singing and singing states (Figure 2.4G-I, n=12, rate difference >85%). All of these diverse cell types were spatially intermingled (Figure 2.S3H).

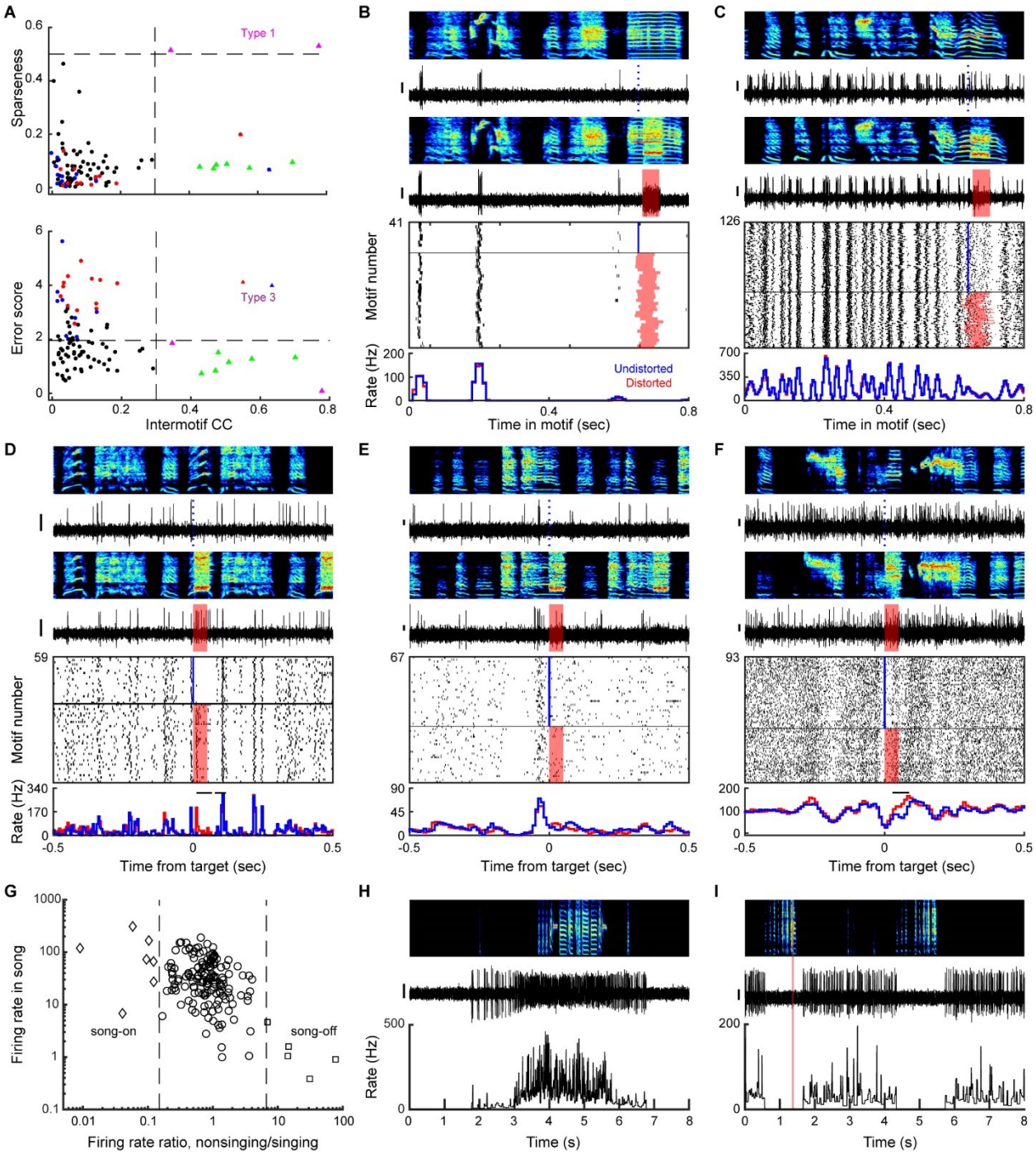

Figure 2.4. VP neurons exhibit temporally precise song-locked activity during singing. See also Figure 2.S3.

(**A**) Top: sparseness of song-locked VP neurons plotted against the intermotif correlation coefficient (IMCC). Dashed lines: IMCC=0.3 (vertical) and sparseness=0.5 (horizontal). Bottom: Maximum error score (absolute z-scored difference between distorted and undistorted trials) plotted against IMCC. Dashed lines: IMCC=0.3 and error score=1.96. (**B** to **F**) Top to bottom:

spectrograms, spiking activity during undistorted and distorted trials, corresponding spike raster plots and rate histograms for a VP neuron with sparse, temporally precise discharge (**B**), one with time-locked bursts that tile the song (**C**), one with time-locked bursts and error response (**D**), one with significant rate maximum immediately before target time (**E**), and one with significant rate minimum at target time (**F**). Horizontal bars on top of rate histograms indicate significant difference between distorted and undistorted firing ($p<0.05$, WRS test). (**G**) A scatter plot of the mean firing rate during singing plotted against the ratio between mean firing rates during non-singing and singing periods identified VP neurons gated by singing state. (**H**-**I**) Top to bottom: spectrograms, spiking activity, and corresponding instantaneous firing rate (IFR) for a song-on neuron which fired at high rate during singing, but silent outside song (**H**) and a song-off neuron which abruptly stopped firing during singing (**I**). Scale bar for spiking activity in (**B**-**F**,**H**,**I**) is 0.15 mV.

VP sends diverse error- and prediction-related signals to VTA

To test which VP signals are sent to VTA, we used antidromic and collision testing methods to identify VTA-projecting VP (VP_{VTA}, n=10, Figure 2.5A-B) and putative non VTA-projecting (VP_{Other}, n=92) neurons (n=60 not tested). Like the VP population more generally, VP_{VTA} neurons exhibited a range of firing rates and discharge patterns (Figure 2.S3). However, VP_{VTA} neurons were significantly more likely to exhibit error responses compared to non-projectors (n=5/10 VP_{VTA} neurons, 16/92 VP_{Other} neurons. $p<0.05$, WRS test, Methods).

VP_{VTA} neurons were significantly more likely to exhibit a minimum in firing rate immediately prior to the target time in the song (n=5/10 VP_{VTA} neurons, 12/92 VP_{Other} neurons, $p<0.01$, WRS test, Methods). These pre-target pauses are consistent with a predicted quality signal. One VP_{VTA} neuron exhibited a robust pre-target burst, but such pre-target activations were not enriched in the VP_{VTA} neurons relative to VP_{Other} neurons ($p>0.05$, WRS test,

Methods). Together these findings show that VP sends diverse error- and prediction-related signals to VTA.

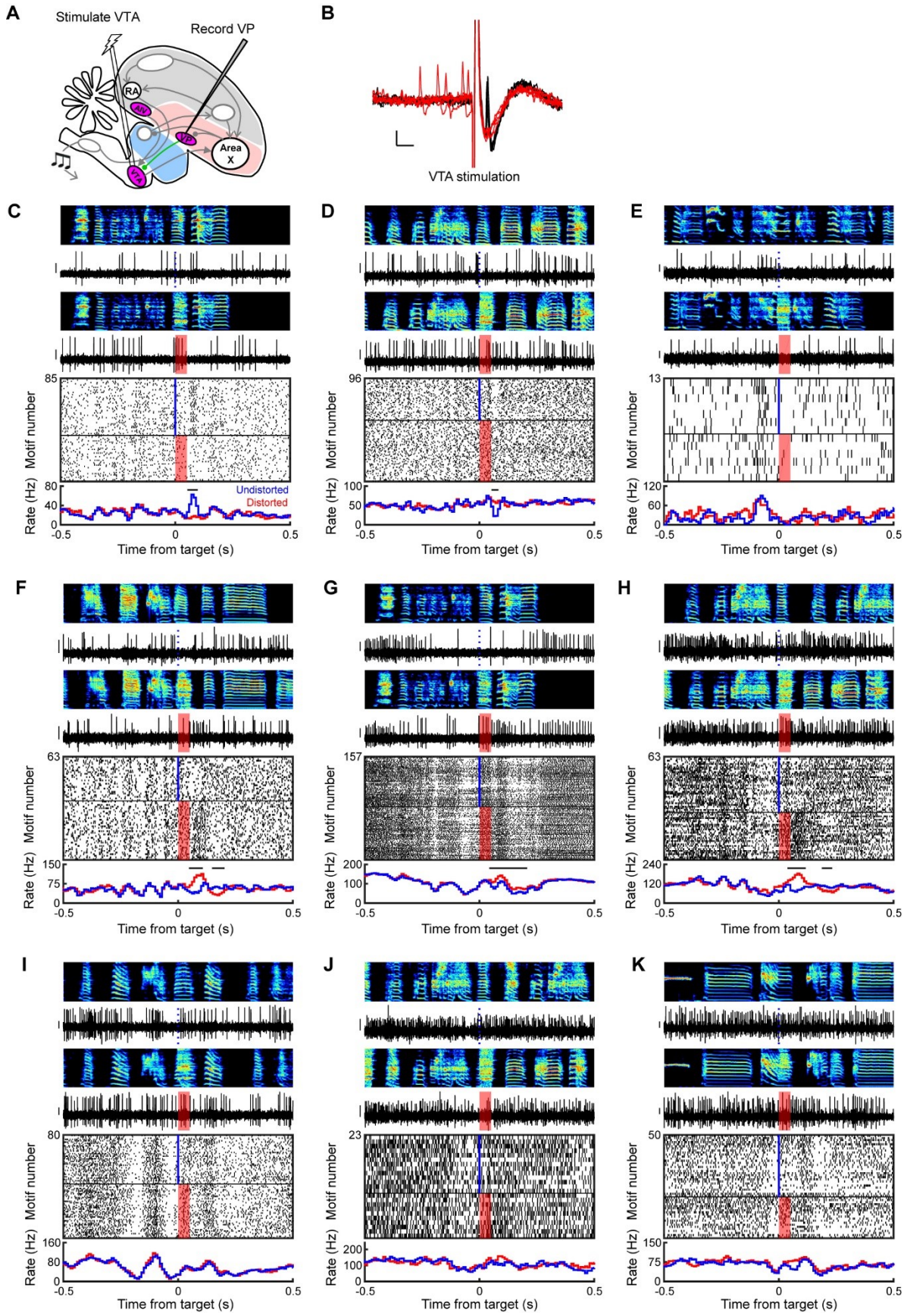
A
Stimulate VTA
Record VP
RA
4N
VTA
Area X
LMAN
B
VTA stimulation
C
D
E
F
G
H
I
J
K
Motif number
Rate (Hz)
Undistorted
Distorted
Time from target (s)
-0.5
0
0.5

Figure 2.5. VP sends diverse error- and prediction-related signals to VTA. See also Figure 2.S3.

(**A**) Stimulation and recording electrodes were chronically implanted into VTA and VP, respectively, for antidromic identification of VTA-projecting VP neurons (VP_{VTA}). (**B**) Antidromic (black) and collision (red) testing of the neuron shown in (**C**). Scale bars: 1ms (horizontal) and 0.1mV (vertical). (**C** to **K**) Song-locked firing patterns of 9 VP_{VTA} neurons, plotted as in Figure 2.3B, reveal diverse responses including prediction errors (**C** and **D**), pre-target burst (**E**), error-induced activation (**F**-**H**), and pre-target pauses (**F** to **J**). Y scale bar for spiking activity is 0.15 mV.

Viral tracing identifies inputs to VP from dopaminergic, vocal motor, and auditory regions

To test what inputs to VP could account for this diversity of singing and error-related firing, we combined retrograde and anterograde viral tracing strategies. Abundant fiber tracts course through VP, complicating the interpretation of results obtained with dextran and cholera toxin (CTB) tracers, which can be taken up by fibers of passage. We thus used a sparse GFP-expressing retrograde virus, self-complementary AAV9 (scAAV9-CBh-GFP) that is taken up by axon terminals (Xiao et al., 2018)(Methods). Following injection of viral tracer into VP, retrogradely labeled neurons were observed in several singing-related cortical, thalamic and midbrain structures including: (1) RA, a vocal motor cortex-like nucleus known to send precise motor command signals to brainstem motor neurons (Leonardo and Fee, 2005; Sober et al., 2008; Yu and Margoliash, 1996) (Figure 2.S6A-B, retrograde labelling observed in 4/11 hemispheres); (2) Uva, a motor thalamic nucleus known to send precise song timing information to HVC (5/19 hemispheres, Figure 2.S6E-G) (Danish et al., 2017; Hamaguchi et al., 2016); (3) DLM, the Area X-recipient thalamic nucleus known to send premotor signals to cortical nucleus LMAN (3/12 hemispheres, Figure 2.S6M-O) (Goldberg and Fee, 2012); (4) AIV, an auditory cortical area known to send error signals to VTA (4/11 hemispheres, Figure 2.S6I-K)

(Mandelblat-Cerf et al., 2014); (5) Ovoidalis, the primary auditory thalamus (2/11 hemispheres, Figure 2.S6M,N,P) (Lei and Mooney, 2010; Vates et al., 1996); and (6) VTA (9/21 hemispheres, Figure 2.S6Q-S).

We note that using this retrograde viral tracing strategy we only observed sparsely labeled cells in all identified input regions to VP (Figure 2.S9). This may indicate sparse connections between input regions and VP, but could also be due to sparse uptake by axon terminals and/or sparse viral expression (Methods). Indeed, in all VP injected birds we observed sparse input from the main known input to VP, Area X (Gale et al., 2008) (not shown). We additionally used dual tracer strategies and further confirmed that anterogradely labeled RA axons co-mingled with VTA-projecting VP neurons (n=1/2 hemispheres, Figure 2.S6C-D); retrograde tracers in VP and VTA could co-label AIV neurons (3/5 cells co-labeled in 3/3 hemispheres, Figure 2.S6I-L); retrograde tracers in VP and Area X can co-label VTA neurons (n=4/6 cells co-labeled in 2/2 hemispheres, Figure 2.S6Q-T); and that retrograde tracers in VP and HVC can co-label Uva neurons (3/9 cells co-labeled in 2/3 hemispheres, Figure 2.S6E-H). Together, these data show that VP receives inputs from RA, DLM, Ovoidalis, HVC-projecting Uva neurons, Area X-projecting VTA neurons, and VTA-projecting AIV neurons.

Discussion

By combining lesions, viral tracing, and electrophysiology we discovered that songbird ventral pallidum is required for song learning, receives information about song syllable timing and error, and sends diverse performance prediction and error related signals to VTA. These findings demonstrate that ventral pallidal circuits can play an essential role in performance evaluation during a purely motor sequence learning task like birdsong.

Despite the importance of error signals for motor sequence learning, the identification of online error signals remained elusive in singing birds (Achiro et al., 2017; Derégnaucourt et al., 2004; Ganguli and Hahnloser, 2011). HVC and LMAN do not exhibit responses to DAF during singing (Hamaguchi et al., 2014; Kozhevnikov and Fee, 2007; Leonardo, 2004; Vallentin and Long, 2015), ruling against hypothesized roles of the classic song system in online evaluation of auditory feedback (Doupe and Konishi, 1991; Nottebohm et al., 1990; Troyer and Doupe, 2000). Recent studies instead support the idea that auditory cortical areas extract error signals and send them to VTA, which in turn sends dopaminergic prediction error signals to Area X (Fee and Goldberg, 2011; Woolley, 2019). Specifically, early stages of the auditory cortical hierarchy exhibit singing-related auditory responses that include DAF-driven modulations (Keller and Hahnloser, 2009), and AIV, a higher order auditory cortical region, contains VTA-projecting neurons that are not simply activated by singing and instead are specifically activated by DAF-induced errors (Mandelblat-Cerf et al., 2014). VTA then sends performance error signals to Area X that can modify future performance (Gadagkar et al., 2016; Hisey et al., 2018; Xiao et al., 2018). Our finding that both AIV and VP can modulate VTA_X firing supports the idea that song evaluation signals can reach the classic song system through the VTA projection to Area X (Gale and Perkel, 2010).

VP is classically viewed as an output of the limbic system involved in seeking primary reinforcers such as food, drugs, or courtship (McAlonan et al., 1993; Mogenson et al., 1980; Smith et al., 2009). VP lesions in mammals cause anhedonia, reduce drug- and food-seeking, and impair reward-based place preference, implicating VP with both motivational 'wanting' and hedonic 'liking'. Consistent with this idea, VP neuronal activity is modulated by reward omissions, rewards, the cues that predict them, and their hedonic value (Ahrens et al., 2016; Ito and Doya, 2009; Ottenheimer et al., 2018; Richard et al., 2016; Richard et al., 2018; Tachibana and Hikosaka, 2012; Tindell et al., 2004).

Computational models of basal ganglia (BG) dependent reinforcement learning may provide insight into how VP's established hedonic functions relate to performance evaluation in singing birds. In classic actor-critic (AC) models the BG has two functional subdivisions: a ventral 'critic' with outputs to VTA and a dorsal 'actor' with outputs to the motor system (Figure 2.6A) (Daw et al., 2006; Joel et al., 2002; Takahashi et al., 2008). Both subdivisions receive dopaminergic reward prediction error signals and implement DA-modulated plasticity to weigh cortical inputs (which encode 'state', Figure 2.6A) according to their reward value. DA modulated plasticity in the critic therefore computes predicted state value, for example a cue-dependent reward prediction. Predicted state-value coding by VP, manifest as VP responses to conditioned stimuli in Pavlovian tasks (Ahrens et al., 2016; Ito and Doya, 2009; Richard et al., 2018; Tindell et al., 2004), can provide VTA with 'prediction' information necessary to compute reward prediction error (Tian et al., 2016). DA prediction error signals project back to the critic (to update predicted state-value) and also to an 'actor.' The actor also implements DA-modulated plasticity on 'state'-encoding inputs. But because the actor has topographically organized outputs to the motor system, this plasticity links a state representation to a reward-maximizing action

(Daw et al., 2006; Joel et al., 2002; Takahashi et al., 2008), manifest as action-value coding in dorsal BG structures (Samejima and Doya, 2007).

Our findings suggest that songbird VP may implement some functions analogous to the critic in the classic AC architecture, including the computation of predicted state value (which in birdsong is analogous to predicted syllable quality). Specifically, thalamic (Uva) or cortical (RA) inputs to VP could provide state representations in the form of 'time-step' in song that could explain the observed VP timing responses (Figure 2.4B-C)(Danish et al., 2017; Hamaguchi et al., 2016; Leonardo and Fee, 2005; Sober et al., 2008; Yu and Margoliash, 1996). Next, DA inputs to VP from VTA_X neurons could enable DA-modulated plasticity to weigh Uva inputs according to past error. For example, consider a three syllable song *a*-*b*-*c*. If syllable *b* is reliably distorted, then DA pauses (driven by DAF) would be coincident with those Uva inputs active at syllable *b*. Then DA-modulated plasticity of Uva inputs would re-weigh these synapses, resulting in a representation in VP of error-weighted timing or, equivalently, predicted error (Figure 2.6B). With an eligibility trace (Sutton and Barto, 1998; Yagishita et al., 2014), this process would explain the pre-target pauses observed in VP that are enriched in VP_{VTA} neurons (Figure 2.5, F-K). Finally, AIV neurons are DAF-responsive during singing and could explain error responses in VP (Figure 2.3) (Mandelblat-Cerf et al., 2014). Altogether, VP thus contains information necessary to signal the difference between predicted and actual error, manifest in the observed prediction error responses (Figure 2.3, A and B) that can be sent to VTA (Figure 2.5C and D). A key prediction of this model is the existence of DA-modulated plasticity of Uva or RA inputs to VP.

In biologically inspired AC models, the DA error signal reaches both ventral and dorsal BG domains (Daw et al., 2006; Joel et al., 2002; Takahashi et al., 2008). Notably, VP-projecting

VTA neurons also project to Area X, which is located more dorsolaterally and which has topographically organized outputs to the song motor system (Johnson et al., 1995; Luo et al., 2001) . DA modulated plasticity in Area X (Ding and Perkel, 2004) reinforces the way a target syllable is produced (Hisey et al., 2018; Hoffmann et al., 2016; Xiao et al., 2018), much like manipulation of striatal DA in mammals can reinforce place preference or action selection (Tsai et al., 2009; Wise and Schwartz, 1981). Thus we propose that Area X has anatomical and functional similarities to the 'actor' in classic AC architecture (Charlesworth et al., 2012; Fee and Goldberg, 2011)(Figure 2.6B), forming the counterpart to the ventral critic.

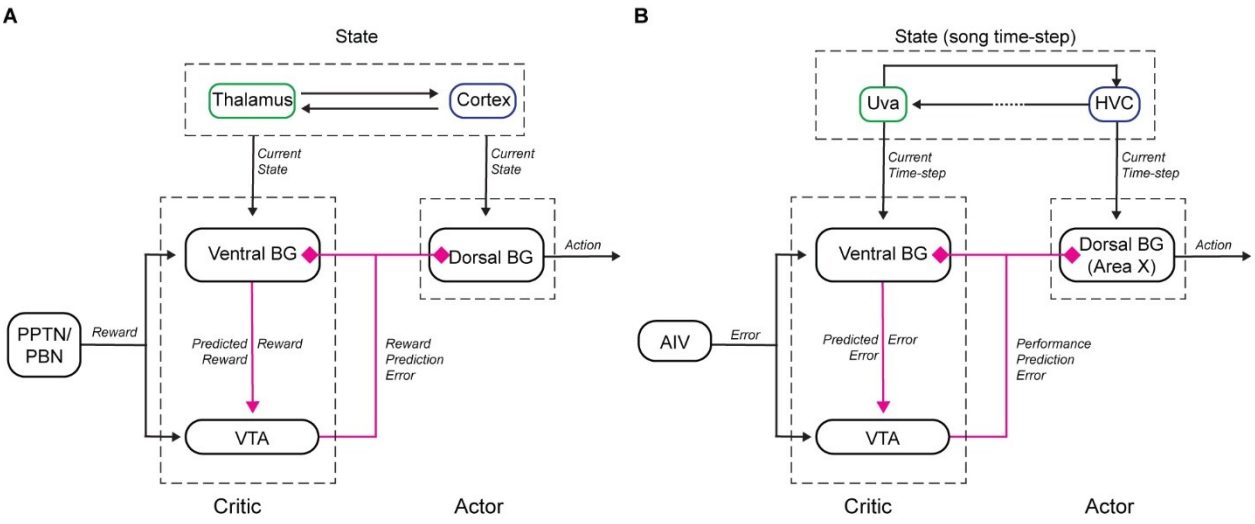

Figure 2.6. An actor-critic circuit motif for computing performance prediction error. See also Figure 2.S6.

(**A**) Actor-critic circuit motif in mammalian BG adapted from (Daw et al., 2006; Joel et al., 2002; Takahashi et al., 2008). (**B**) Anatomy and signaling in songbird VP reveals a similar motif.

The AC model helps unify mechanisms of reward seeking and performance evaluation into a common framework, but it fails to explain the diversity of VP signals and likely oversimplifies processes underlying the construction of the dopaminergic error signal. First, VP stimulation

could cause either activation or suppression of VTA_X firing (Figure 2.1E-F and S2N-R), which could be due to a mixture of glutamatergic and GABAergic VP_{VTA} neurons and/or differential engagement with local inhibition in VTA that can invert afferent signals (Yang et al., 2018). Second, VP_{VTA} neurons could be activated by distorted renditions, could be activated or suppressed by undistorted renditions, and could exhibit pre-target bursts or pauses that may encode predicted song error or predicted quality, respectively (Figure 2.5). Thus VP sends virtually every conceivable error-related signal to VTA. While it is possible to linearly combine these VP signals in specific ways to construct the known VTA_X signal, it remains unclear how these mixed inputs are transformed by the VTA microcircuit into a remarkably homogeneous dopaminergic prediction error signal in the VTA_X pathway. Notably, mixed responses to reward and reward-predicting cues are observed in VP (and other) inputs to mammalian VTA. even though VTA output is a similarly homogenous reward prediction error signal (Hong and Hikosaka, 2014; Tian et al., 2016). How relatively homogenous DA error signals are constructed from mixed inputs is similarly elusive in birds and mammals, and may involve complex, cell-type specific engagement with VTA microcircuitry (Yang et al., 2018). Future recordings of VTA_X neurons in lesioned animals could clarify how DA pauses and bursts depend on specifically on inputs from AIV and VP (Takahashi et al., 2016; Takahashi et al., 2011; Tian and Uchida, 2015).

In summary, we report that the ventral pallidum, a limbic structure associated with reward seeking in mammals, is necessary for birdsong learning, receives information from VTA and auditory and vocal motor areas, and sends diverse performance and prediction error signals to VTA.

Methods

Subjects

Subjects were 91 male zebra finches (at least 39 days post hatch, dph). Animal care and experiments were carried out in accordance with NIH guidelines and were approved by the Cornell Institutional Animal Care and Use Committee.

Surgery and histology

All surgeries were performed with isoflurane anesthetization. For functional mapping experiments (8 birds, 90 dph or older, Figure. 1), bipolar stimulation electrodes were implanted into AIV and Area X (Gadagkar et al., 2016; Mandelblat-Cerf et al., 2014). In 4/8 birds, additional stimulation electrodes were implanted into VP. AIV coordinate was determined by its anterior and ventral position to RA, and Area X coordinate was +5.6A, +1.5L relative to lambda and 2.65 ventral relative to pial surface, at a head angle of 20 degrees. VP coordinate was +4.9A, +1.3L relative to lambda and 3.9 ventral to pial surface at a head angle of 20 degrees. Recordings were made in VTA using a carbon fiber electrode (1 MOhm, Kation Scientific). VTA was identified by anatomical landmarks. Specifically, the boundaries of DLM and Ovoidalis were determined by spontaneous firing and auditory responses. Recordings were then made at the same AP position, +0.6L relative to lambda and 6.5 ventral relative to pial surface, at a head angle of 55 degrees. This AP position corresponds to the anterior part of VTA enriched in Area X projecting neurons (VTA_X) (Gadagkar et al., 2016; Mandelblat-Cerf et al., 2014). VTA_X neurons were further confirmed by antidromic response and collision testing. At the end of the experiment, small electrolytic lesions (30 uA for 60s) were made at each stimulation site. Location of the stimulating electrodes was verified histologically.

For VP lesion (13 birds, 39-52 dph), a bipolar stimulation electrode was implanted into Area X and the boundaries of VP was electrophysiologically mapped by finding units suppressed by Area X stimulation. 115nl of 2% N-methyl-DLaspartic acid (NMA; Sigma, St Louis, MO) or saline (for control birds) was injected into VP bilaterally. Lesions were histologically confirmed by labeling neurons with anti-NeuN (full bilateral VP lesions in 6 birds).

For awake-behaving electrophysiology (39 birds, 87-355 dph), custom made microdrives carrying an accelerometer (Analog Devices AD22301), linear actuator (Faulhaber 0206 series micromotor) and homemade electrode arrays (5 electrodes, 3-5 MOhms, microprobes.com) were implanted into VP by coordinates (4.4-5.4A, 1.1-1.5L, 3.5V, head angle 20 degrees). In 20/39 birds, a bipolar stimulation electrode was implanted into VTA using anatomical landmarks as described above. After each experiment, small electrolytic lesions (30 μA for 60 s) were made with one of the recording electrodes. Brains were then fixed, cut into 100 μm thick sagittal sections for histological confirmation of stimulation electrode tracks and reference lesions.

For VP tracing experiments (31 birds, 90 dph or older), 40nl of self-complementary adeno-associated virus carrying green fluorescent protein (scAAV9-CBh-GFP, UNC vector core) was injected into VP in two coordinates (4.6/4.9A, 1.3L, 4V). Upstream neurons retrogradely infected and expressing GFP could be observed in RA, AIV, Uva, Ov, DLM, and VTA (each input was checked in a subset of birds as indicated in the main text). To determine if VP shares common inputs with HVC, in addition to scAAV9 in VP, fluorescently labeled cholera toxin subunit B (CTB, Molecular Probes) was injected into HVC in 3 birds. To determine if VP share common inputs with Area X, CTB was injected into Area X for 2 birds. To test if RA axons co-mingle with VTA projecting VP neurons, CTB was injected into VTA and anterograde HSV-mCherry (MGH viral core) was injected into RA.

Recording VTA responses to stimulation of AIV/VP

Neurons were classified as Area X-projecting (VTA_X) based on antidromic stimulation and collision testing (200 μs pulses, 100-300 μA). Spike duration was determined by the interval between onset and offset time of spikes (Figure 2.1C, S2K). VTA neurons that did not respond to Area X stimulation were classified as putative interneurons. We cannot rule out the possibility that a subset of these neurons project to the basal ganglia outside the field of influence of stimulation. A burst of AIV (or VP) stimulation consisting three 200 μs pulses with 10ms inter-pulse-interval was delivered every 1.5-2 s, with 300 μA current amplitude. Putative interneurons were also tested for response to AIV stimulation. VTA_X neurons were found in an anterior part of VTA, intermingled with non-projecting local neurons (Gadagkar et al., 2016; Mandelblat-Cerf et al., 2014).

All VTA_X neurons and those putative interneurons with rate influenced by AIV stimulation were further analyzed. To determine if VTA neurons respond to AIV (VP) stimulation, spiking activity within ±1 second relative to stimulation onset was binned in a moving window of 30ms (for VTA_X neurons) or 10ms (for VTA interneurons) with a step size of 5ms. Each bin after stimulation onset was tested against all the bins in the previous 1 second (the prior) using a z-test. Windows where at least 2 consecutive bins with $p<0.05$ were considered significant. The response onset and offset were required to bracket lowest (for phasic decrease) or highest (for phasic increase) firing rate after stimulation onset. Response was quantified by normalized firing rate in the first significant window using the 1 sec before stimulation onset as baseline (Figure 2.S2L).

To determine if the simultaneously recorded putative interneuron (PIN) and VTA_X neuron were correlated, we constructed rate histogram of VTA_X neuron spiking events aligned to spontaneous spiking events of PIN with preceding inter-spike interval (ISI) > 10ms, and assessed significance of rate changes of VTA_X neuron using z-test (Figure 2.1D).

Song imitation score

Song learning in VP lesioned and control birds was assessed by song similarity between pupil (at 90 dph) and their tutors. We computed imitation scores using an automated procedure based on Sound Analysis Pro (SAP) algorithm (Mandelblat-Cerf and Fee, 2014; Tchernichovski et al., 2000). Briefly, the tutor motif was segmented into syllables by hand. Syllables in the pupil song were determined by finding the section of pupil song with highest SAP similarity to each tutor syllable. Additionally, a sequencing score was computed as the similarity of the next syllable in tutor song and the next section in the pupil song. Imitation score was the product of song similarity and sequence similarity(Mandelblat-Cerf et al., 2014).

Syllable-targeted distorted auditory feedback

Postoperative birds with microdrive implant were placed in a sound isolation chamber and given at least a day to habituate to distorted auditory feedback (DAF) as described previously (Gadagkar et al., 2016). Briefly, ongoing singing was analyzed by Labview software to target specific syllables, and two speakers inside the chamber played a 50ms DAF sound on top of bird's singing on 50% of randomly selected target renditions. DAF was either a broadband sound band passed at 1.5-8 kHz, the same spectral range of zebra finch song, or a segment of one of the bird's own non-target syllables displaced in time.

Passive playback of the bird's own song

For passive playback of the bird's own song (BOS), we played back randomly interleaved renditions of the undistorted and distorted motifs of the bird's own song during awake, non-singing periods. The loudness of playback was adjusted to match the average peak loudness of zebra finch song (Gadagkar et al., 2016; Mandelblat-Cerf et al., 2014).

Analysis of neural activity

Neural signals were band-passed filtered (0.25-15 kHz) in homemade analog circuits and acquired at 40 kHz using custom Matlab software. Single units were identified as VTA-projecting by antidromic identification and antidromic collision testing (Figure 2.5A-B). Spike sorting was performed offline using custom Matlab software. Instantaneous firing rates (IFR) were defined at each time point as the inverse of the enclosing ISI. Firing rate histograms were constructed with 10 ms bins and smoothed with a 3-bin moving average, except for Figure 2.4C and 4D, where the histograms had 4ms and 5ms bins. All data were acquired during undirected song, except for the neuron in Figure 2.5K, which was recorded during female-directed song.

Performance error response

To identify performance-error related neurons, we assessed the difference in firing rate between distorted and undistorted singing renditions as previously described (Keller and Hahnloser, 2009; Mandelblat-Cerf et al., 2014). Neurons with less than 10 trials of either distorted or undistorted renditions of the target syllable were excluded from this analysis (n=34/162 excluded). Briefly, we performed a WRS test on the number of spikes in distorted vs.

undistorted renditions in 30 ms windows. Windows were shifted in 5 ms steps and considered significant when at least 4 consecutive windows had $p<0.05$. Error-related neurons were classified as error-activated if the firing rate is higher in distorted trials in window of significance, and error suppressed if the firing rate is higher in undistorted trials.

To visualize error response, we calculated z-score of the difference between distorted and undistorted rate histograms (Figure 2.3E). We defined 'error score' for each neuron to be the maximum of absolute z-scored difference in the 125ms after target onset (Figure 2.4A).

To identify prediction-error related neurons, we quantified phasic rate changes following undistorted target time. 1000 surrogate rate histograms were generated by randomly time-shifting each trial of undistorted target aligned data over the duration of the motif. Response was considered significant when firing rate dropped below 5th percentile or exceeded 95th percentile of the surrogate data. Neurons with significant rate peak or nadir in the window of significance for error response were identified as prediction error neurons.

To test if error responses were attributable to purely auditory responses to a different sound, we performed the same analysis for distorted and undistorted renditions during passive playback of bird's own song (BOS) playback in 16/31 error neurons and 19/97 non-error neurons. Only two neurons exhibited an error response during passive playback. One such neuron exhibited similar song-locked firing during both singing and listening (Figure 2.S5C). One other error responsive neuron also had song time-locked response in BOS playback, although the part of playback that contained target syllable was consistently masked by calls (Figure 2.S5D).

We compared the latency and duration of error response to those of VTA_X neurons from a previous dataset (Gadagkar et al., 2016). Latency and duration were defined by the onset and onset-offset interval of significant windows in WRS test as described above.

To test if VP_{VTA} neurons were more likely to exhibit error response, we assigned a value of 1 (if error responsive) or 0 (if not error responsive) to each neuron tested for VTA antidromic stimulation. VP_{VTA} neurons were tested against the group of VP neurons tested but not antidromic using WRS test (p=0.05).

Song timing related activity

Sparseness index was used to identify putative song-related MSNs. This distinguishes MSNs from other striatal cell types in the dorsal basal ganglia nucleus Area X (Goldberg and Fee, 2010). For each neuron, we calculated rate histograms aligned to syllable onset for all syllables. Then we normalized these histograms over all syllables to generate a probability density function p_i over N bins. An entropy-based sparseness index was computed as follows:

$$Sparseness\ Index = 1 + \frac{\sum_{i=1}^{N} p_i \log(p_i)}{\log(N)}$$

Intermotif pairwise correlation coefficient (IMCC) was used to identify neurons that had highly time-locked firing to song motifs (timing neurons), as previously described (Goldberg et al., 2010; Goldberg and Fee, 2010; Kao et al., 2008; Olveczky et al., 2005; Woolley et al., 2014). Motif aligned IFR was time warped to the median duration of all motifs, mean-subtracted, and smoothed with a Gaussian kernel of 20ms SD, resulting in $\boldsymbol{r_i}$ for each motif. IMCC was defined as the mean value of all pairwise CC between $\boldsymbol{r_i}$ as follows:

$$IMCC = \frac{1}{N_{pairs}} \sum_{j>i}^{N_{pairs}} CC_{ij}$$

$$CC_{ij} = \frac{\boldsymbol{r_i} \cdot \boldsymbol{r_j}}{\sqrt{\boldsymbol{r_i}^2 \boldsymbol{r_j}^2}}$$

To assess the significance of IMCC values, we computed new IMCC for each neuron by adding random, circular time shifts to each spiketrain. This was repeated 1000 times. IMCC was considered significant when the real value was greater than the 95th percentile of the shuffled data.

To quantify significant song-locked rate modulations, we compared the highest rate peak and lowest nadir in target-aligned rate histogram to 1000 surrogate rate histograms generated by randomly time-shifting spike trains. Rate peaks exceeding the 95th percentile of surrogate rate maximum and rate nadirs below the 5th percentile of surrogate rate minimum were considered significant.

To test if VPvta neurons were more likely to exhibit rate maxima/minima immediately prior to target time, we assigned a value of 1 (if a significant rate maximum/minimum was present in the 100ms before target time) or 0 (if significant peak/nadir was not present) for each neuron tested for VTA antidromic stimulation. VPvta neurons were tested against the group of VP neurons tested but not antidromic using WRS test (p=0.05).

Quantification of movement

An accelerometer (Analog Devices AD22301) was mounted on microdrives to quantify gross body movements as described previously(Gadagkar et al., 2016). Briefly, movement onsets and offsets were determined by threshold crossings of the band-passed, rectified accelerometer signal. To test if error responses could be explained by a difference in movement rate following DAF, for each bird we calculated onset times of movements relative to song target time. Then we performed a WRS test (p=0.05) on the number of movement onsets in distorted vs. undistorted

renditions in 30 ms windows. Windows were shifted in 5 ms steps and considered significant when at least 4 consecutive windows had $p<0.05$.

Imaging

Imaging data was acquired with a Leica DM4000 B microscope and a Zeiss LSM 710 Confocal microscope. Image processing was done with ImageJ.

Acknowledgements

We thank Todd Roberts for providing virus, Joe Fetcho, Teja Bollu, and Josh Dudman for comments, and Vikram Gadagkar, Aaron Andalman, and Dmitriy Aronov for comments and analysis code. Funding to JHG was provided by the NIH (R01NS094667), the Pew Charitable Trust, and the Klingenstein Neuroscience Foundations. Funding to PAP was provided by the NIH (F32NS098634). Imaging data was acquired in the Cornell BRC-Imaging Facility using the shared, NIH-funded (S10RR025502) Zeiss LSM 710 Confocal.

REFERENCES

Achiro, J.M., Shen, J., and Bottjer, S.W. (2017). Neural activity in cortico-basal ganglia circuits of juvenile songbirds encodes performance during goal-directed learning. Elife 6.

Ahrens, A.M., Meyer, P.J., Ferguson, L.M., Robinson, T.E., and Aldridge, J.W. (2016). Neural activity in the ventral pallidum encodes variation in the incentive value of a reward cue. Journal of Neuroscience 36, 7957-7970.

Ali, F., Otchy, T.M., Pehlevan, C., Fantana, A.L., Burak, Y., and Olveczky, B.P. (2013). The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. Neuron 80, 494-506.

Andalman, A.S., and Fee, M.S. (2009). A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. Proc Natl Acad Sci U S A 106, 12518-12523.

Beier, K.T., Steinberg, E.E., DeLoach, K.E., Xie, S., Miyamichi, K., Schwarz, L., Gao, X.J., Kremer, E.J., Malenka, R.C., and Luo, L. (2015). Circuit Architecture of VTA Dopamine Neurons Revealed by Systematic Input-Output Mapping. Cell 162, 622-634.

Bottjer, S.W., Brady, J.D., and Cribbs, B. (2000). Connections of a motor cortical region in zebra finches: relation to pathways for vocal learning. J Comp Neurol 420, 244-260.

Carr, D.B., and Sesack, S.R. (1999). Terminals from the rat prefrontal cortex synapse on mesoaccumbens VTA neurons. Ann N Y Acad Sci 877, 676-678.

Charlesworth, J.D., Warren, T.L., and Brainard, M.S. (2012). Covert skill learning in a cortical-basal ganglia circuit. Nature 486, 251-255.

Creed, M.C., Ntamati, N.R., and Tan, K.R. (2014). VTA GABA neurons modulate specific learning behaviors through the control of dopamine and cholinergic systems. Frontiers in behavioral neuroscience 8, 8.

Danish, H.H., Aronov, D., and Fee, M.S. (2017). Rhythmic syllable-related activity in a songbird motor thalamic nucleus necessary for learned vocalizations. PloS one 12, e0169568.

Daw, N.D., Niv, Y., and Dayan, P. (2006). Actions, policies, values and the basal ganglia. Recent breakthroughs in basal ganglia research 10, 1214-1221.

Derégnaucourt, S., Mitra, P., Fehér, O., Maul, K., Lints, T., and Tchernichovski, O. (2004). Song development: In search of the error-signal. Annals of the New York Academy of Sciences 1016, 364-376.

Ding, L., and Perkel, D.J. (2004). Long-term potentiation in an avian basal ganglia nucleus essential for vocal learning. J Neurosci 24, 488-494.

Doupe, A.J., and Konishi, M. (1991). Song-selective auditory circuits in the vocal control system of the zebra finch. Proc Natl Acad Sci U S A 88, 11339-11343.

Fee, M.S., and Goldberg, J.H. (2011). A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. Neuroscience 198, 152-170.

Gadagkar, V., Puzerey, P.A., Chen, R., Baird-Daniel, E., Farhang, A.R., and Goldberg, J.H. (2016). Dopamine neurons encode performance error in singing birds. Science 354, 1278-1282.

Gale, S.D., and Perkel, D.J. (2010). A basal ganglia pathway drives selective auditory responses in songbird dopaminergic neurons via disinhibition. The Journal of neuroscience 30, 1027-1037.

Gale, S.D., Person, A.L., and Perkel, D.J. (2008). A novel basal ganglia pathway forms a loop linking a vocal learning circuit with its dopaminergic input. J Comp Neurol 508, 824-839.

Ganguli, S., and Hahnloser, R. (2011). Bird song learning without reinforcement: the Hebbian self-organization of sensorimotor circuits. In CoSyne (Salt Lake City, UT).

Goldberg, J.H., Adler, A., Bergman, H., and Fee, M.S. (2010). Singing-related neural activity distinguishes two putative pallidal cell types in the songbird basal ganglia: comparison to the primate internal and external pallidal segments. J Neurosci 30, 7088-7098.

Goldberg, J.H., and Fee, M.S. (2010). Singing-related neural activity distinguishes four classes of putative striatal neurons in the songbird basal ganglia. Journal of Neurophysiology 103, 2002-2014.

Goldberg, J.H., and Fee, M.S. (2012). A cortical motor nucleus drives the basal ganglia-recipient thalamus in singing birds. Nature Neuroscience 15, 620-627.

Hamaguchi, K., Tanaka, M., and Mooney, R. (2016). A distributed recurrent network contributes to temporally precise vocalizations. Neuron 91, 680-693.

Hamaguchi, K., Tschida, K.A., Yoon, I., Donald, B.R., and Mooney, R. (2014). Auditory synapses to song premotor neurons are gated off during vocalization in zebra finches. Elife 3, e01833.

Hisey, E., Kearney, M.G., and Mooney, R. (2018). A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. Nature neuroscience, 1.

Hoffmann, L.A., Saravanan, V., Wood, A.N., He, L., and Sober, S.J. (2016). Dopaminergic Contributions to Vocal Learning. J Neurosci 36, 2176-2189.

Hong, S., and Hikosaka, O. (2014). Pedunculopontine tegmental nucleus neurons provide reward, sensorimotor, and alerting signals to midbrain dopamine neurons. Neuroscience 282, 139-155.

Ito, M., and Doya, K. (2009). Validation of decision-making models and analysis of decision variables in the rat basal ganglia. Journal of Neuroscience 29, 9861-9874.

Joel, D., Niv, Y., and Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives. Neural Netw 15, 535-547.

Johnson, F., Sablan, M.M., and Bottjer, S.W. (1995). Topographic organization of a forebrain pathway involved with vocal learning in zebra finches. J Comp Neurol 358, 260-278.

Kao, M.H., Wright, B.D., and Doupe, A.J. (2008). Neurons in a forebrain nucleus required for vocal plasticity rapidly switch between precise firing and variable bursting depending on social context. J Neurosci 28, 13232-13247.

Keller, G.B., and Hahnloser, R.H. (2009). Neural processing of auditory feedback during vocal practice in a songbird. Nature 457, 187-190.

Kelley, D.B., and Nottebohm, F. (1979). Projections of a telencephalic auditory nucleus-field L- in the canary. J Comp Neurol 183, 455-469.

Kozhevnikov, A.A., and Fee, M.S. (2007). Singing-related activity of identified HVC neurons in the zebra finch. J Neurophysiol 97, 4271-4283.

Lei, H., and Mooney, R. (2010). Manipulation of a central auditory representation shapes learned vocal output. Neuron 65, 122-134.

Leonardo, A. (2004). Experimental test of the birdsong error-correction model. Proc Natl Acad Sci U S A 101, 16935-16940.

Leonardo, A., and Fee, M.S. (2005). Ensemble coding of vocal control in birdsong. J Neurosci 25, 652-661.

Luo, M., Ding, L., and Perkel, D.J. (2001). An avian basal ganglia pathway essential for vocal learning forms a closed topographic loop. J Neurosci 21, 6836-6845.

Mandelblat-Cerf, Y., and Fee, M.S. (2014). An automated procedure for evaluating song imitation. PLoS One 9, e96484.

Mandelblat-Cerf, Y., Las, L., Denisenko, N., and Fee, M.S. (2014). A role for descending auditory cortical projections in songbird vocal learning. Elife 3.

Marler, P. (1997). Three models of song learning: evidence from behavior. J Neurobiol 33, 501-516.

McAlonan, G.M., Robbins, T.W., and Everitt, B.J. (1993). Effects of medial dorsal thalamic and ventral pallidal lesions on the acquisition of a conditioned place preference: further evidence for the involvement of the ventral striatopallidal system in reward-related processes. Neuroscience 52, 605-620.

Mello, C.V., Vates, G.E., Okuhata, S., and Nottebohm, F. (1998). Descending auditory pathways in the adult male zebra finch (Taeniopygia guttata). J Comp Neurol 395, 137-160.

Mogenson, G.J., Jones, D.L., and Yim, C.Y. (1980). From motivation to action: functional interface between the limbic system and the motor system. Progress in neurobiology 14, 69-97.

Moreines, J.L., Owrutsky, Z.L., and Grace, A.A. (2017). Involvement of Infralimbic Prefrontal Cortex but not Lateral Habenula in Dopamine Attenuation After Chronic Mild Stress. Neuropsychopharmacology 42, 904-913.

Murdoch, D., Chen, R., and Goldberg, J.H. (2018). Place preference and vocal learning rely on distinct reinforcers in songbirds. Sci Rep 8, 6766.

Nottebohm, F., Alvarez-Buylla, A., Cynx, J., Kirn, J., Ling, C.Y., Nottebohm, M., Suter, R., Tolles, A., and Williams, H. (1990). Song learning in birds: the relation between perception and production. Philos Trans R Soc Lond B Biol Sci 329, 115-124.

Olveczky, B.P., Andalman, A.S., and Fee, M.S. (2005). Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. PLoS Biol 3, e153.

Ottenheimer, D., Richard, J.M., and Janak, P.H. (2018). Ventral pallidum encodes relative reward value earlier and more robustly than nucleus accumbens. Nat Commun 9, 4350.

Patton, M.H., Bizup, B.T., and Grace, A.A. (2013). The infralimbic cortex bidirectionally modulates mesolimbic dopamine neuron activity via distinct neural pathways. J Neurosci 33, 16865-16873.

Person, A.L., Gale, S.D., Farries, M.A., and Perkel, D.J. (2008). Organization of the songbird basal ganglia, including area X. J Comp Neurol 508, 840-866.

Reiner, A., Perkel, D.J., Mello, C.V., and Jarvis, E.D. (2004). Songbirds and the revised avian brain nomenclature. Ann N Y Acad Sci 1016, 77-108.

Richard, J.M., Ambroggi, F., Janak, P.H., and Fields, H.L. (2016). Ventral Pallidum Neurons Encode Incentive Value and Promote Cue-Elicited Instrumental Actions. Neuron 90, 1165-1173.

Richard, J.M., Stout, N., Acs, D., and Janak, P.H. (2018). Ventral pallidal encoding of reward-seeking behavior depends on the underlying associative structure. eLife 7, e33107.

Samejima, K., and Doya, K. (2007). Multiple representations of belief states and action values in corticobasal ganglia loops. Ann N Y Acad Sci 1104, 213-228.

Schmidt, R.A., Lee, T., Winstein, C., Wulf, G., and Zelaznik, H. (2018). Motor Control and Learning, 6E (Human kinetics).

Schultz, W., and Romo, R. (1987). Responses of nigrostriatal dopamine neurons to high-intensity somatosensory stimulation in the anesthetized monkey. J Neurophysiol 57, 201-217.

Smith, K.S., Tindell, A.J., Aldridge, J.W., and Berridge, K.C. (2009). Ventral pallidum roles in reward and motivation. Behav Brain Res 196, 155-167.

Sober, S.J., Wohlgemuth, M.J., and Brainard, M.S. (2008). Central contributions to acoustic variation in birdsong. J Neurosci 28, 10370-10379.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement learning: an introduction (Cambridge, MA: MIT Press).

Tachibana, Y., and Hikosaka, O. (2012). The primate ventral pallidum encodes expected reward value and regulates motor action. Neuron 76, 826-837.

Takahashi, Y., Schoenbaum, G., and Niv, Y. (2008). Silencing the critics: understanding the effects of cocaine sensitization on dorsolateral and ventral striatum in the context of an actor/critic model. Front Neurosci 2, 86-99.

Takahashi, Y.K., Langdon, A.J., Niv, Y., and Schoenbaum, G. (2016). Temporal Specificity of Reward Prediction Errors Signaled by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. Neuron 91, 182-193.

Takahashi, Y.K., Roesch, M.R., Wilson, R.C., Toreson, K., O'Donnell, P., Niv, Y., and Schoenbaum, G. (2011). Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. Nat Neurosci 14, 1590-1597.

Tchernichovski, O., Nottebohm, F., Ho, C.E., Pesaran, B., and Mitra, P.P. (2000). A procedure for an automated measurement of song similarity. Anim Behav 59, 1167-1176.

Thelen, E. (1995). Motor development. A new synthesis. Am Psychol 50, 79-95.

Tian, J., Huang, R., Cohen, J.Y., Osakada, F., Kobak, D., Machens, C.K., Callaway, E.M., Uchida, N., and Watabe-Uchida, M. (2016). Distributed and Mixed Information in Monosynaptic Inputs to Dopamine Neurons. Neuron 91, 1374-1389.

Tian, J., and Uchida, N. (2015). Habenula Lesions Reveal that Multiple Mechanisms Underlie Dopamine Prediction Errors. Neuron 87, 1304-1316.

Tindell, A.J., Berridge, K.C., and Aldridge, J.W. (2004). Ventral pallidal representation of pavlovian cues and reward: population and rate codes. J Neurosci 24, 1058-1069.

Troyer, T.W., and Doupe, A.J. (2000). An associational model of birdsong sensorimotor learning I. Efference copy and the learning of song syllables. J Neurophysiol 84, 1204-1223.

Tsai, H.C., Zhang, F., Adamantidis, A., Stuber, G.D., Bonci, A., de Lecea, L., and Deisseroth, K. (2009). Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. Science 324, 1080-1084.

Tumer, E.C., and Brainard, M.S. (2007). Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. Nature 450, 1240-1244.

Vallentin, D., and Long, M.A. (2015). Motor Origin of Precise Synaptic Inputs onto Forebrain Neurons Driving a Skilled Behavior. J Neurosci 35, 299-307.

Vates, G.E., Broome, B.M., Mello, C.V., and Nottebohm, F. (1996). Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches. J Comp Neurol 366, 613-642.

Wise, R.A., and Schwartz, H.V. (1981). Pimozide attenuates acquisition of lever-pressing for food in rats. Pharmacol Biochem Behav 15, 655-656.

Woolley, S.C. (2019). Dopaminergic regulation of vocal-motor plasticity and performance. Current opinion in neurobiology 54, 127-133.

Woolley, S.C., Rajan, R., Joshua, M., and Doupe, A.J. (2014). Emergence of context-dependent variability across a basal ganglia network. Neuron 82, 208-223.

Xiao, L., Chattree, G., Oscos, F.G., Cao, M., Wanat, M.J., and Roberts, T.F. (2018). A Basal Ganglia Circuit Sufficient to Guide Birdsong Learning. Neuron 98, 208-221 e205.

Yagishita, S., Hayashi-Takagi, A., Ellis-Davies, G.C., Urakubo, H., Ishii, S., and Kasai, H. (2014). A critical time window for dopamine actions on the structural plasticity of dendritic spines. Science 345, 1616-1620.

Yang, H., de Jong, J.W., Tak, Y., Peck, J., Bateup, H.S., and Lammel, S. (2018). Nucleus Accumbens Subnuclei Regulate Motivated Behavior via Direct Inhibition and Disinhibition of VTA Dopamine Subpopulations. Neuron 97, 434-449 e434.

Yu, A.C., and Margoliash, D. (1996). Temporal hierarchical control of singing in birds. Science 273, 1871-1875.

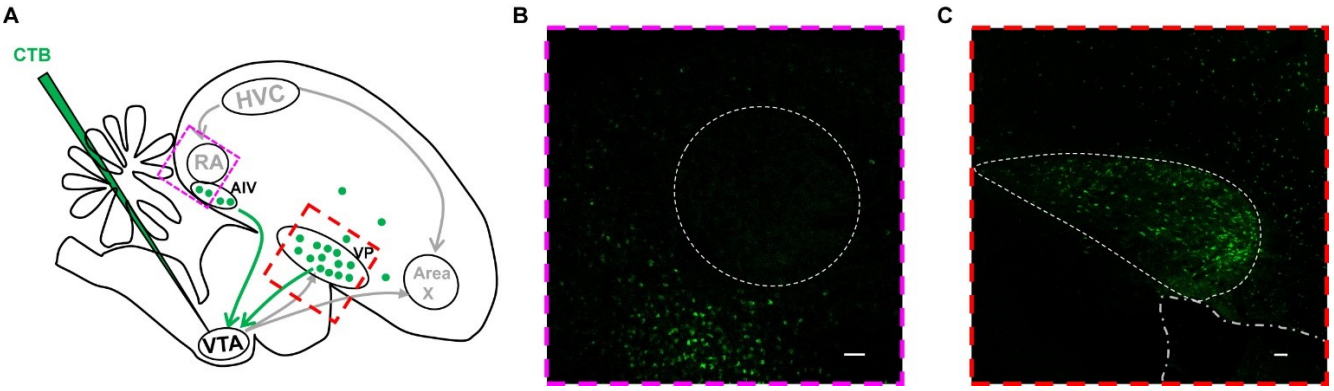

Figure 2.S1. AIV and VP projects to VTA. Related to Figure 2.1A.

(**A**) Schematic of tracing experiment. Injection of CTB into VTA retrogradely labeled cell bodies in AIV and VP, as previously reported (Gale et al., 2008; Mandelblat-Cerf et al., 2014). (**B**) Expanded view of magenta square from (**A**). VTA-projecting neurons are visible in AIV, which surrounds boundaries of RA, denoted by dashed white lines. (**C**) Expanded view of red square from (**A**). VTA-projecting neurons are visible in VP, bounded by dashed white lines, and overlying striatum. Scale bar in (**B**) and (**C**), 50 microns.

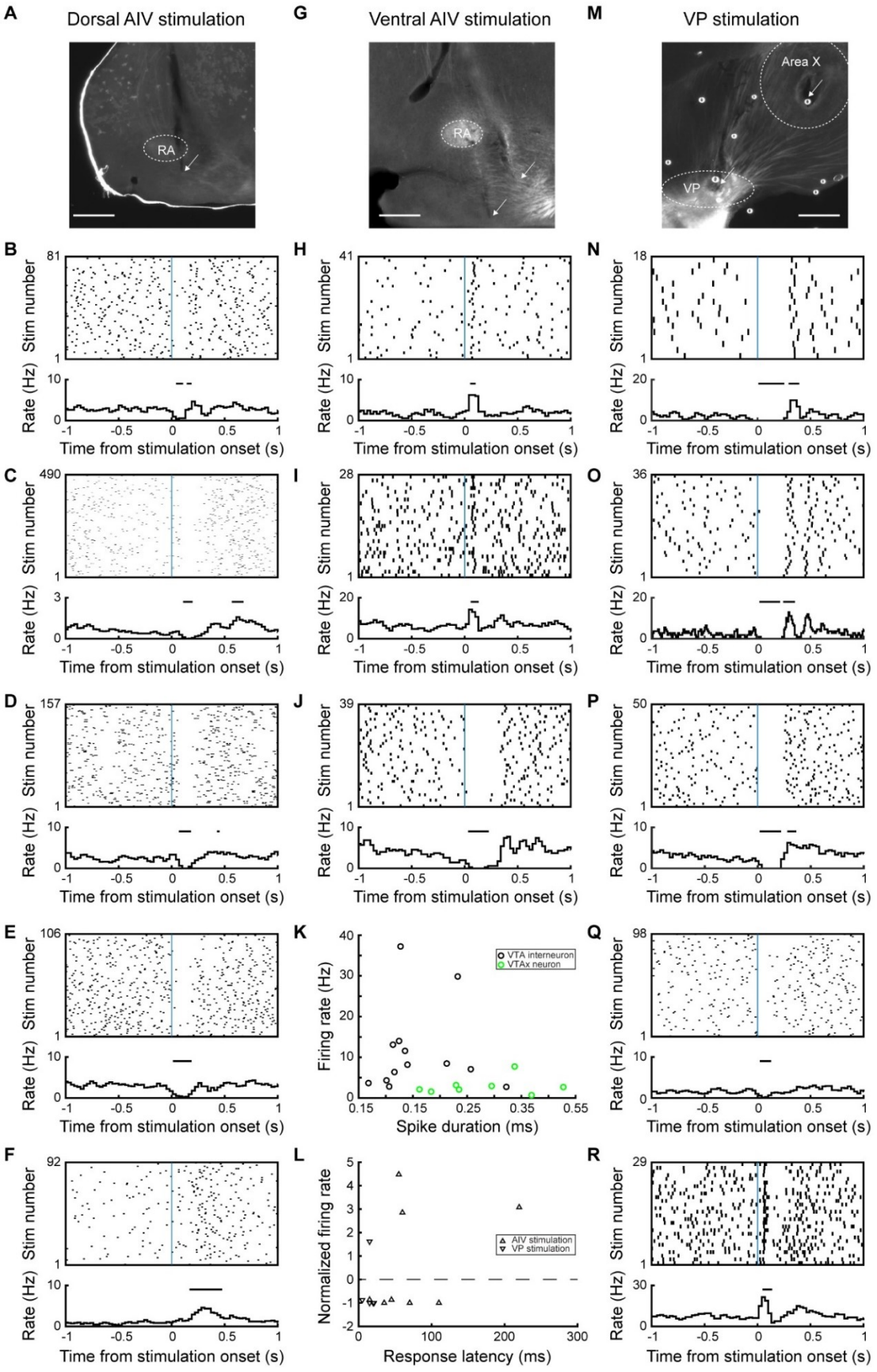
A
Dorsal AIV stimulation
RA
G
Ventral AIV stimulation
RA
M
VP stimulation
Area X
VP
B
H
N
C
I
O
D
J
P
E
K
Q
F
L
R
Stim number
Rate (Hz)
Time from stimulation onset (s)
Firing rate (Hz)
Spike duration (ms)
VTA interneuron
VTAx neuron
Normalized firing rate
Response latency (ms)
AIV stimulation
VP stimulation

Figure 2.S2. Auditory cortical and ventral pallidal stimulation drive diverse changes in VTAx neuron firing. Related to Figure 2.1G.

(**A**-**J**) Stimulating AIV drives diverse responses in VTAx firing. (**A** and **G**) Stimulation electrode location in dorsal (**A**) or ventral (**G**) AIV (arrows). (**B**-**F**) Stimulation onset aligned raster plots and rate histograms for dorsal AIV stimulations. (**B**) is the same neuron as Figure 2.1E. Horizontal bars represent significant suppression or activation (z-test, $p<0.05$, Methods). (**H**-**J**) Same as (**B**) for ventral AIV stimulation. (**K**) Spike duration and baseline firing rate for all recorded VTA neurons (VTAx neurons shown in green). (**L**) Scatter plot showing normalized response to stimulation versus latency to response for all recorded VTAx neurons (Upward and downward triangles show responses to AIV and VP stimulation, respectively) (**M**) Stimulation electrode location in VP and Area X (arrows). (**N**-**R**) Same as (**B**) for VP stimulation. (**N**) is the same neuron as Figure 2.1F. Scale bars in (**A**), (**G**), and (**M**), 500 microns.

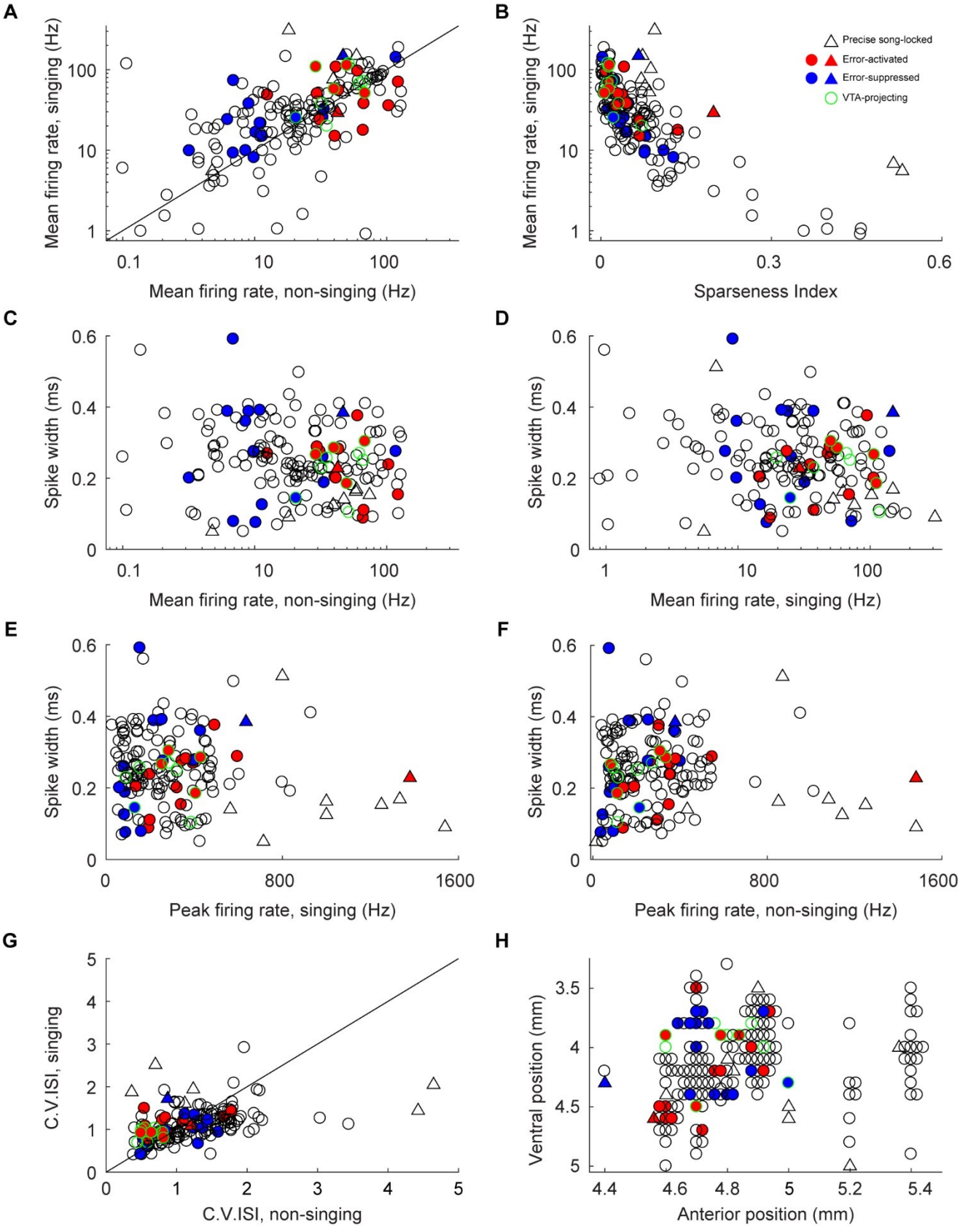
A
Mean firing rate, singing (Hz)
100
10
1
0.1
10
100
Mean firing rate, non-singing (Hz)
B
Mean firing rate, singing (Hz)
100
10
1
0
0.3
0.6
Sparseness Index
Precise song-locked
Error-activated
Error-suppressed
VTA-projecting
C
Spike width (ms)
0.6
0.4
0.2
0
0.1
10
100
Mean firing rate, non-singing (Hz)
D
Spike width (ms)
0.6
0.4
0.2
0
1
10
100
Mean firing rate, singing (Hz)
E
Spike width (ms)
0.6
0.4
0.2
0
0
800
1600
Peak firing rate, singing (Hz)
F
Spike width (ms)
0.6
0.4
0.2
0
0
800
1600
Peak firing rate, non-singing (Hz)
G
C.V.ISI, singing
5
4
3
2
1
0
0
1
2
3
4
5
C.V.ISI, non-singing
H
Ventral position (mm)
3.5
4
4.5
5
4.4
4.6
4.8
5
5.2
5.4
Anterior position (mm)

Figure 2.S3. Firing patterns of VP neurons. Related to Figure 2.3.

(**A**-**H**) VP neurons exhibited diverse firing pattern (**A**-**F**), sparseness during singing (**B**), spike width (**C**-**F**), discharge variability (**G**) and are intermingled in space (**H**). Triangles: precisely time-locked neurons. Solid red fill: error-activated neurons. Solid blue fill: error-suppressed neurons. Green outline: VTA-projecting neurons.

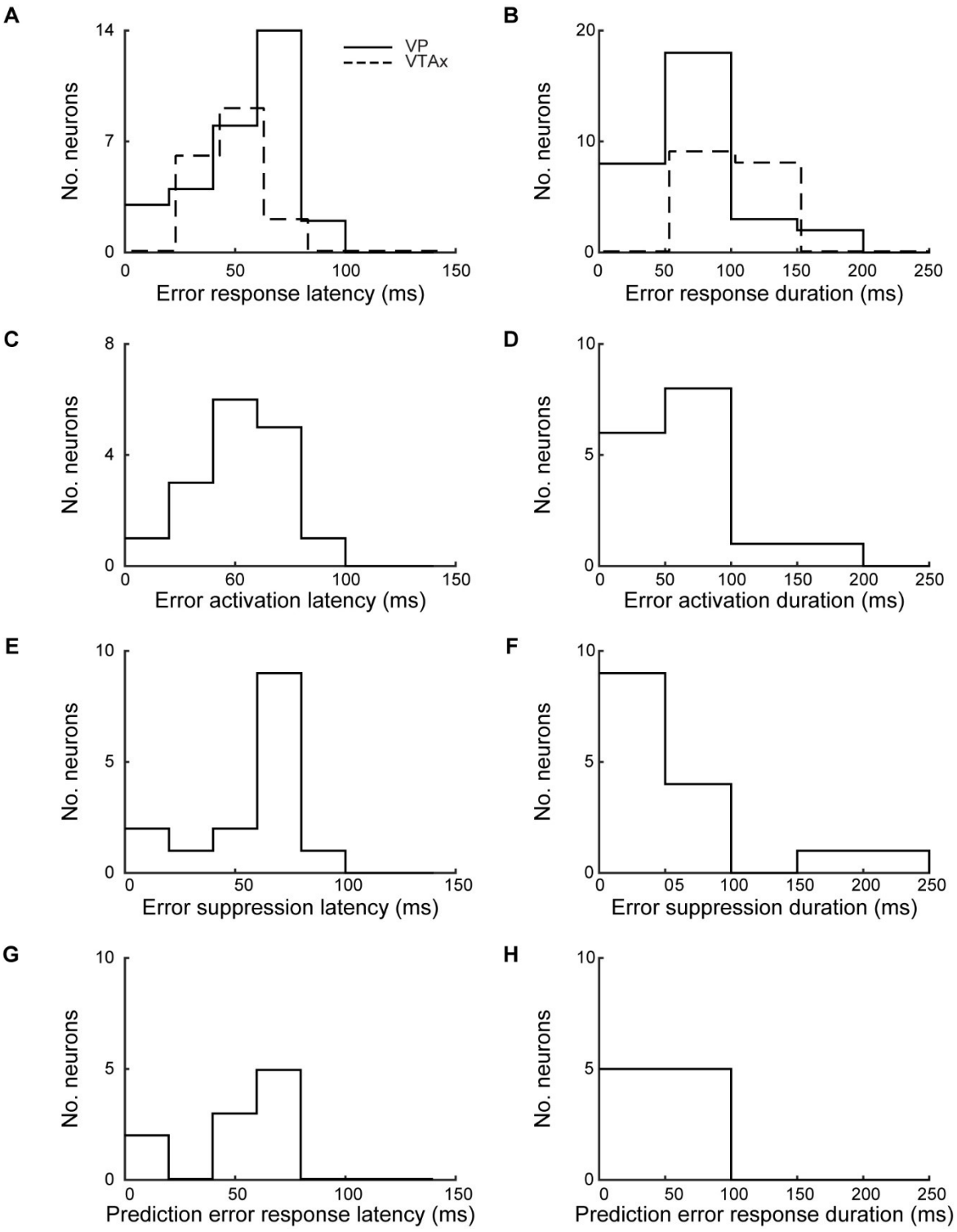

Figure 2.S4. Latency and duration of VP error responses. Related to Figure 2.3E.

(**A**) Distribution of latencies for VP error responsive neurons (solid line) and VTA dopamine neurons (dashed line) from (Gadagkar et al., 2016). (**B**) Duration of VP error response (solid line) and VTA error responses (dashed line). (**C**-**D**) As (**A**-**B**) for error activated VP neurons. (**E**-

F) As (**A-B**) for error suppressed VP neurons. (**G-H**) As (**A-B**) for VP neurons that show prediction error response.

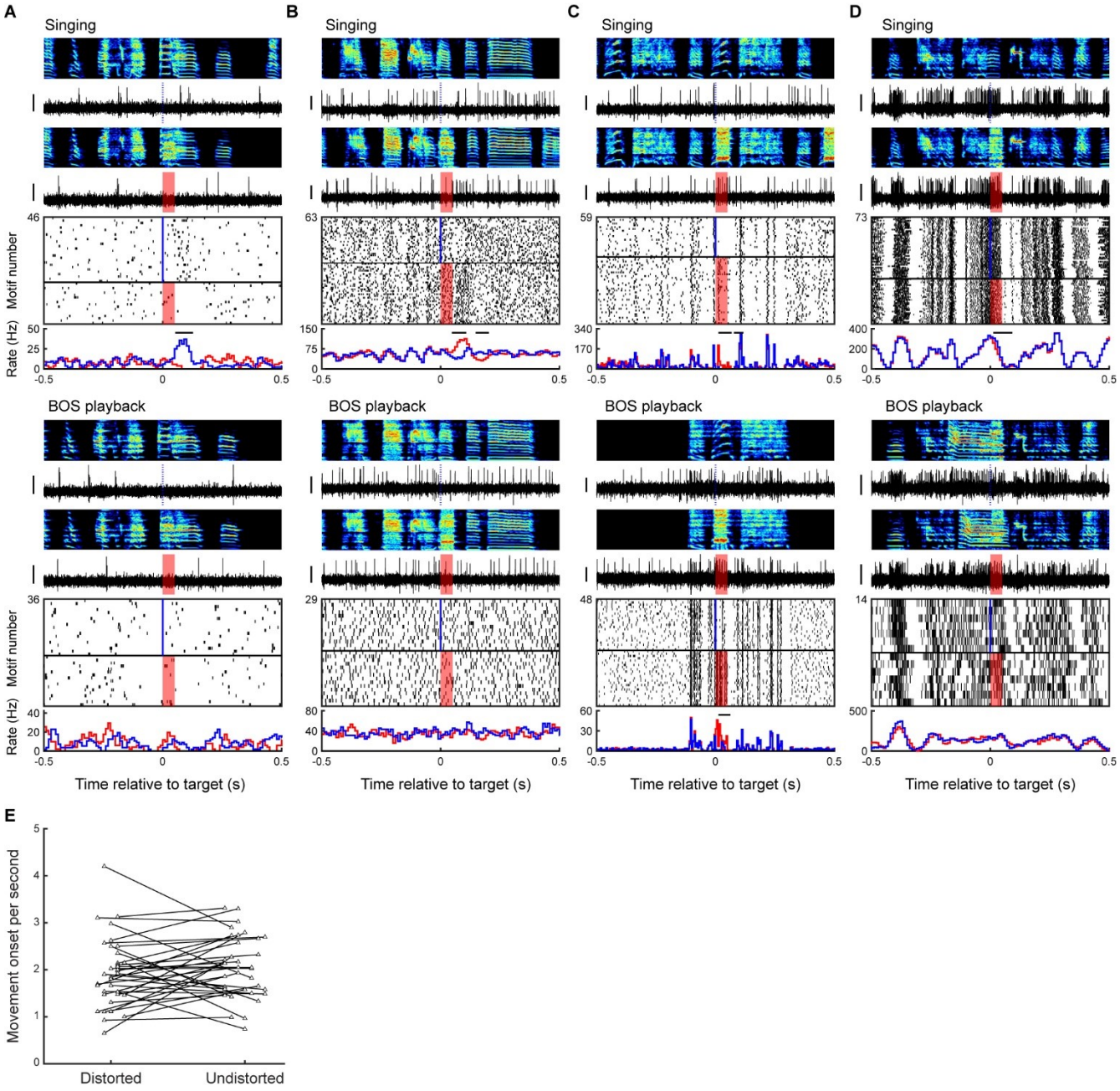

Figure 2.S5. VP neurons can exhibit error signals that occur specifically during singing and was not attributable to animal movement. Related to Figure 2.3.

(A-D) Examples of four neurons recorded during active singing, top row, and during passive playback of bird's own song (BOS), bottom row. Data plotted as in Figure 2.3B for all neurons

in both conditions. Shown are examples of neurons that exhibited error responses only during singing (**A** and **B**) and the only two type 3 song locked neurons recorded in the dataset, which exhibited strong auditory responses to playback of BOS (**C** and **D**). Neuron in (**B**) is the VP_{VTA} neuron from Figure 2.5F. **(I)** Analysis of DAF-related movement responses across all birds. Each line represents data from one bird with average rate of movement onsets in 150 ms following distorted and undistorted syllables. There was no difference in movement onset rate between distorted and undistorted motifs ($p>0.05$ in 35/35 birds, WRS test.)

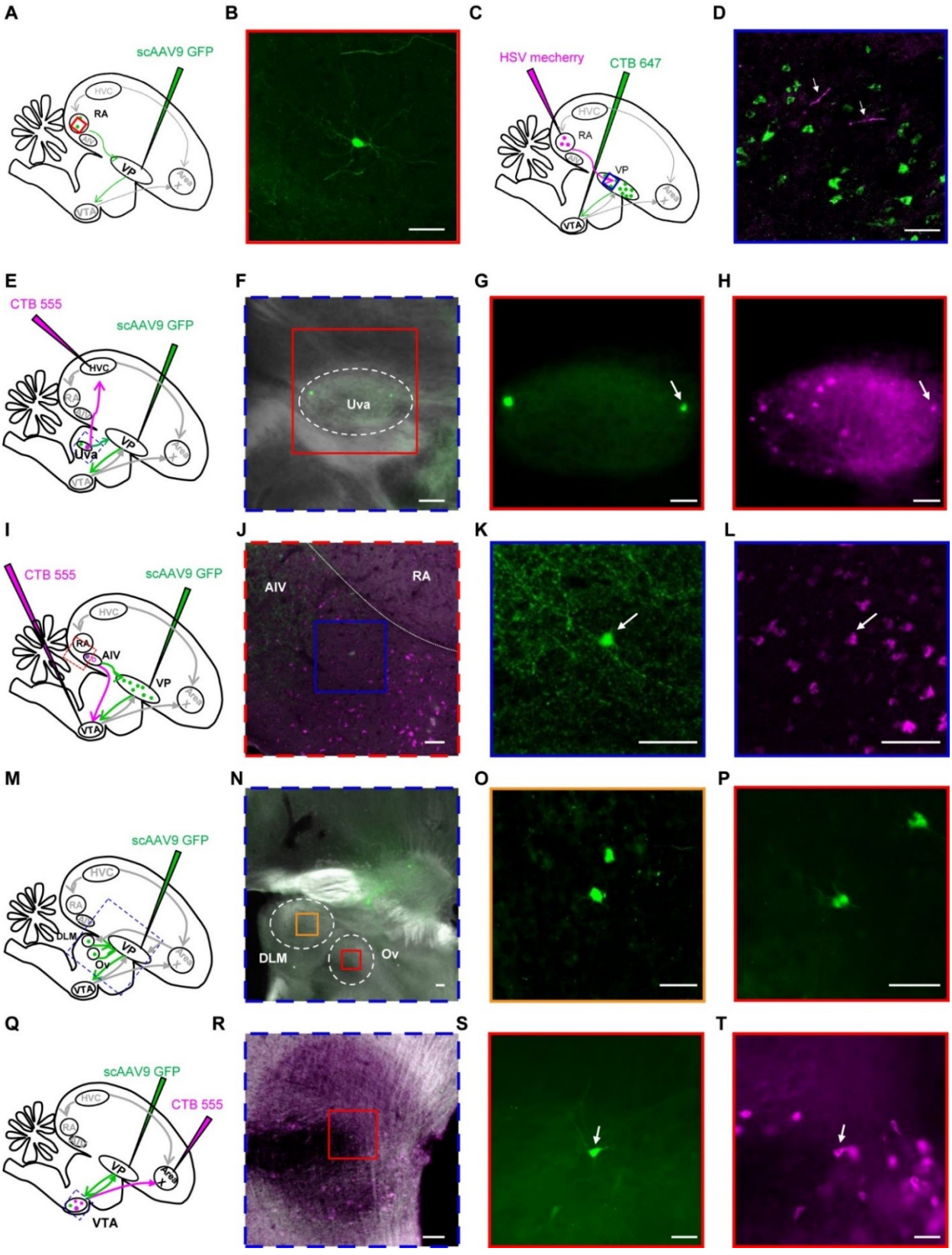

Figure 2.S6. RA, Uva, AIV, DLM, Ov, and VTAx project to VP. Related to Figure 2.6B.

To label VP-projecting neurons, sparse retrograde scAAV9 virus was injected into VP (**A**,**E**,**I**,**M**,**Q**). (**A** to **D**) RA projects to VP. (**B**) Fluorescently labeled RA neuron following VP injection. (**C**) Anterograde virus (HSV mCherry) was injected into RA and CTB 647 was injected into VTA. (**D**) Photomicrograph of part of VP, with VTA-projecting neurons (green) in close proximity to RA axons (purple). (**E** to **H**) HVC-projecting nucleus Uvaeformis (Uva) thalamic neurons project to VP. CTB-555 was injected into HVC (**E**). Retrogradely labeled VP-projecting Uva neuron (arrow, **G**) was co-labeled with HVC-projecting Uva neuron (arrow, **H**). (**I** to **L**) VTA-projecting AIV neurons also project to VP. CTB-555 was injected into VTA (**I**). Retrogradely labeled VP-projecting AIV neuron (arrow, **K**) is co-labeled with VTA-projecting AIV neuron (arrow, **L**). (**M** to **P**) Area X recipient thalamus DLM and primary auditory thalamus Ovoidalis (Ov) project to VP. Retrogradely labeled DLM neurons (**O**) and Ov (**P**). (**Q** to **T**) Area X-projecting VTA dopamine neurons (VTAx) send collaterals to VP. CTB-555 was injected into Area X (**Q**). Retrogradely labeled VP-projecting VTA neuron (arrow, S) was co-labeled with X-projecting VTA neuron (arrow, **T**). Scale bars are 50 microns.

CHAPTER 3

MOVEMENT SIGNALING IN VENTRAL PALLIDUM AND DOPAMINERGIC MIDBRAIN IS GATED BY BEHAVIORAL STATE IN SINGING BIRDS

Abstract

Movement-related neuronal discharge in ventral tegmental area (VTA) and ventral pallidum (VP) is inconsistently observed across studies. One possibility is that some neurons are movement-related and others are not. Another possibility is that the precise behavioral conditions matter - that a single neuron can be movement related under certain behavioral states but not others. We recorded single VTA and VP neurons in birds transitioning between singing and non-singing states, while monitoring body movement with microdrive-mounted accelerometers. Many VP and VTA neurons exhibited body movement-locked activity exclusively when the bird was not singing. During singing, VP and VTA neurons could switch off their tuning to body movement and become instead precisely time-locked to specific song syllables. These changes in neuronal tuning occurred rapidly at state boundaries. Our findings show that movement-related activity in limbic circuits can be gated by behavioral context.

Introduction

Reward related signaling in ventral tegmental (VTA) and ventral pallidal (VP) regions is strongly state-dependent. For example, when an animal is hungry, a food predicting cue can drive dopamine release. But when sated, the DA system may be unresponsive (Ahn and Phillips, 1999; Papageorgiou et al., 2016). Activity of VTA and VP neurons is also strongly driven by movements unrelated to reward (Barter et al., 2015; Brooks, 1986; Engelhard et al., 2019; Jin and Costa, 2010). It remains unknown how these non-reward, movement related signals depend on what an animal is actually doing.

Zebra finches engage in 'bouts' of singing on and off during a typical day. Both during and outside these singing bouts, finches exhibit brief movements such as orienting their head and hopping from perch to perch (Eckmeier et al., 2008; Williams, 2001). Recently, we and others discovered that VP and VTA neurons encode singing-related neural activity, including performance error signals important for song learning (Chen et al., 2019; Gadagkar et al., 2016; Hisey et al., 2018; Kearney et al., 2019; Xiao et al., 2018). These error signals functionally resembled reward prediction error signals observed in the limbic system (Humphries and Prescott, 2010). We also discovered neurons with precisely time-locked firing to specific syllables in VP (Chen et al., 2019).

Here we investigate the movement-related firing properties of VP and VTA neurons and how they depend on whether birds are singing or not. In both VP and VTA, we discovered neurons that change their tuning to movement as birds transition from non-singing to singing states, demonstrating a gating mechanism for movement representation in limbic circuits commonly associated with reward and performance evaluation.

Results

Measuring neural activity and movement as birds transition into and out of singing states

To test if neural activity is correlated with movement timing, we recorded movements with accelerometers attached to head-mounted microdrives (Figure 3.1A). In the recording chamber most movements were head movements associated with orienting and whole body movements during hops. These transient movements occurred during both singing and non-singing periods in the day. Movements were more likely to occur right before onset of syllables (Figure 3.1B, peak movement onset probability 35 ± 2 ms before syllable onset, significant in 42/71 birds, assessed using bootstrap, methods), as previously reported (Gadagkar et al., 2016). At the level of singing bouts, movement probability peaked right before a bout of singing (Figure 3.1C, peak movement onset probability 38 ± 0.2 ms before bout onset, significant in 61/71 birds, bootstrap), and troughed after singing (Figure 3.1D, not significant). During singing, movements were smaller and shorter on average (mean duration 108 ± 2 ms during singing vs. 148 ± 2 ms during non-singing, Figure 3.1E). To control for these differences, we include only those movements that have similar duration and amplitude in subsequent analysis (Figure 3.1E, methods).

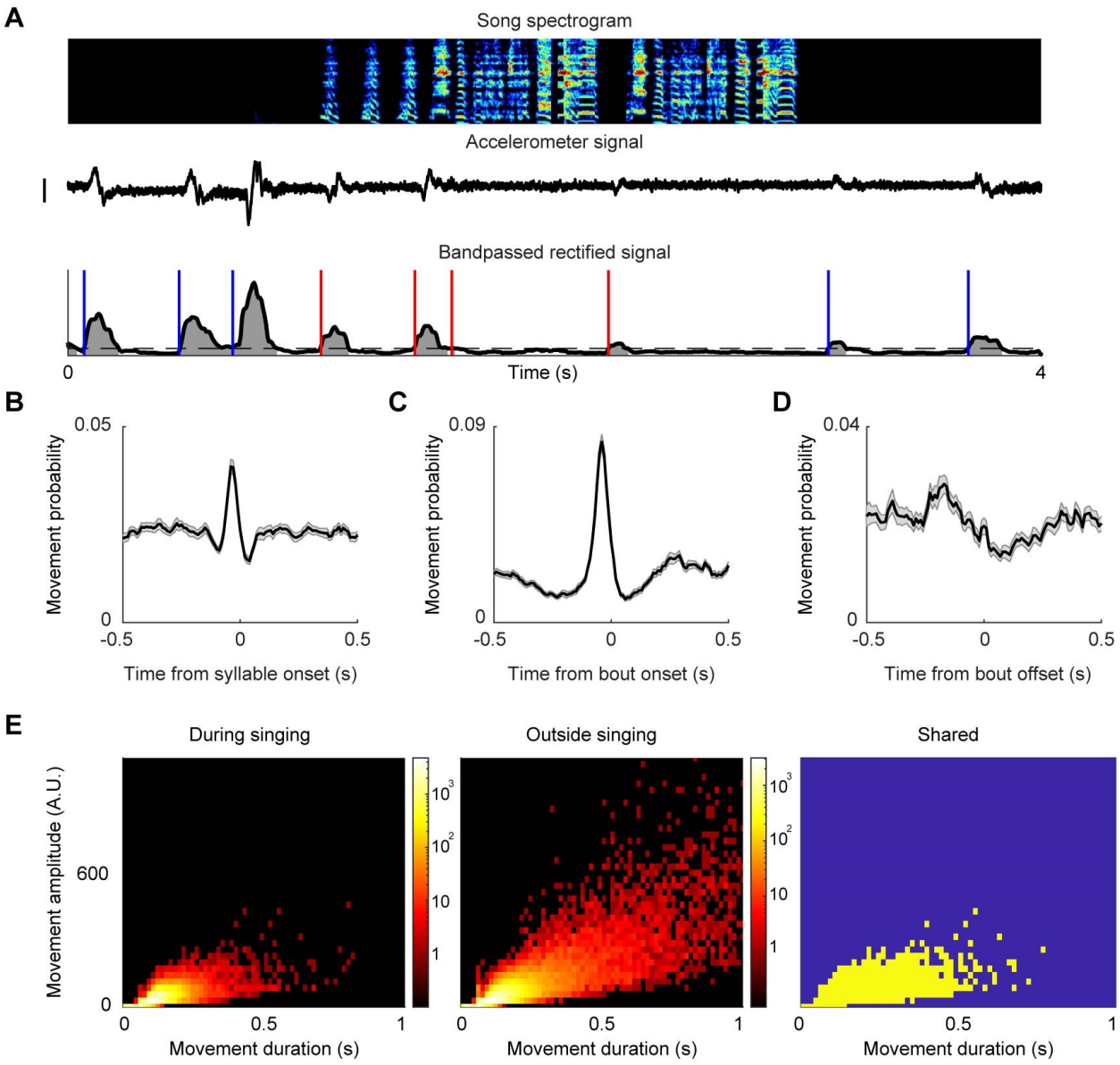

Figure 3.1. Measuring neural activity and movement as birds transition into and out of singing states.

(**A**) (Top to bottom) song spectrogram, accelerometer signal, and bandpassed rectified accelerometer signal. Blue lines indicate onset of movements outside singing, red lines indicate onsets of movements during singing. Scale bar: 0.05mV. (**B**) Average probability of movement onsets around syllable onsets (N = 71 birds). (**C**) Same as B for song bout onsets. (**D**) Same as B for song bout offsets. (**E**) Left and middle: distribution of duration and amplitude of movements during and outside singing for all birds (n=71 birds, 42398 movements in singing, 138422 movements outside singing). Amplitude calculated as area under the curve of bandpassed

rectified accelerometer signal (gray in A). Color axis: number of movements. Right: Same axes as left, with substantially shared bins between movements during singing and outside singing in yellow.

VP and VTA neurons encode movement timing

We recorded VP and VTA neurons as birds transitioned between singing and non-singing states (n=145 VP neurons, n=147 VTA neurons, n=71 birds) (Chen et al., 2019; Gadagkar et al., 2016). Many VP and VTA neurons exhibited activity that was precisely time-locked to movements outside of singing (47/145 neurons in VP, 92/147 neurons in VTA; significance of movement response assessed by comparing rate extrema against randomly time-shuffled data, with threshold p=0.05, methods). Most neurons exhibited brief rate increases after movement onsets (latency to rate increase: 7.9 ± 3.9 ms, duration: 77.3 ± 3.6 ms. n=119/139 movement related neurons, Figure 3.2C), but neurons could also exhibit phasic decreases prior to movements (n=20 neurons).

During our recordings, we controlled perceived error with distorted auditory feedback (DAF) (Andalman and Fee, 2009; Tumer and Brainard, 2007). On randomly interleaved renditions of 'target' syllables, a brief, 50 millisecond snippet of sound was played through speakers surrounding the bird. In previous studies, we found that some VTA and VP neurons discharged differently to distorted and undistorted song renditions (Chen et al., 2019; Gadagkar et al., 2016). A subset of these 'error' responsive neurons also exhibited movement-locked discharge (Figure 3.2C,D, n=4/23 VTA; n=7/28 VP error neurons were movement related, $p<0.05$, bootstrap, methods). We also previously identified VP neurons with precise song-locked discharge, operationally defined as intermotif pairwise correlation coefficient (IMCC) of 0.3 or higher (methods). Interestingly, VTA neurons that were not error reponsive (termed VTAother in

(Gadagkar et al., 2016)) could also exhibit highly song-locked discharge (n=6/147), reported here for the first time. These 'song timing' neurons could also be movement-locked outside of singing (Figure 3.2C,D, n=4/6 in VTA; n=4/6 in VP, $p<0.05$, bootstrap, methods).

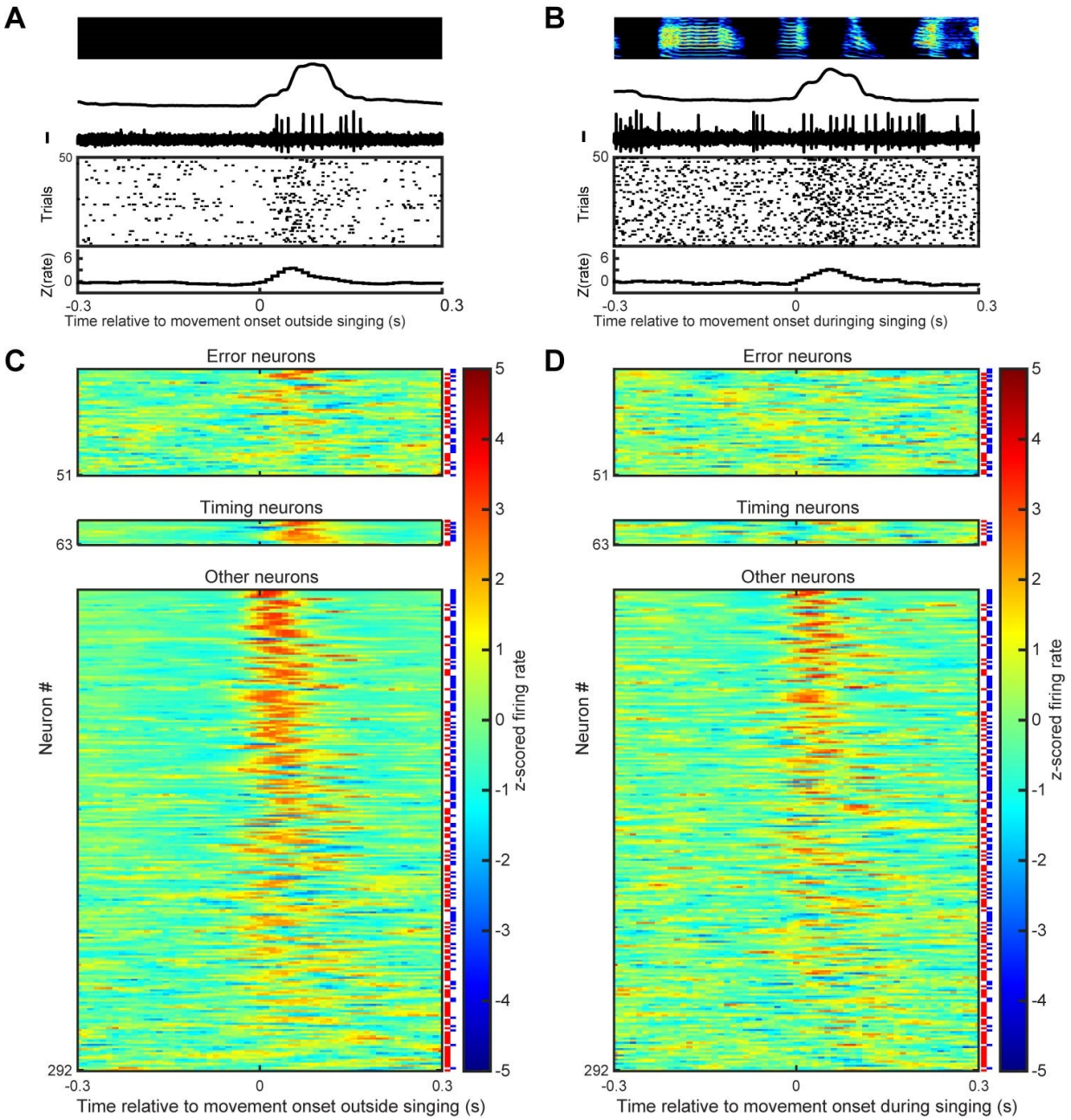

Figure 3.2. VP and VTA neurons can exhibit movement-locked activity.

(**A**) Example VP neuron recorded outside singing. Top to bottom: spectrogram, bandpassed rectified accelerometer signal, neural voltage trace, raster plot of detected spikes, z scored firing rate. All plots are aligned to onsets of movements. Scale bar for neural activity is 0.25mV. (**B**) Same neuron as in A but for movement onsets during singing. (**C**) Z-scored firing rate aligned to movement onsets outside singing for all neurons, separated to error responsive, song timing related, and others. VP and VTA neurons are indicated by red and blue lines to the right of each row. Each group is sorted by maximum absolute response to movement onset. (**D**) Same as C for movements during singing, with the same sorting as in C.

Movement-locked activity can depend on behavioral state

To test if movement responses changed during singing, we compared movement related changes in firing rate between singing and non singing states. In VP and VTA, many neurons were movement-locked outside of singing, but dramatically lost their tuning to similar movements during singing (Figure 3.3, movement response significant outside but not during singing in N=61/139 neurons).

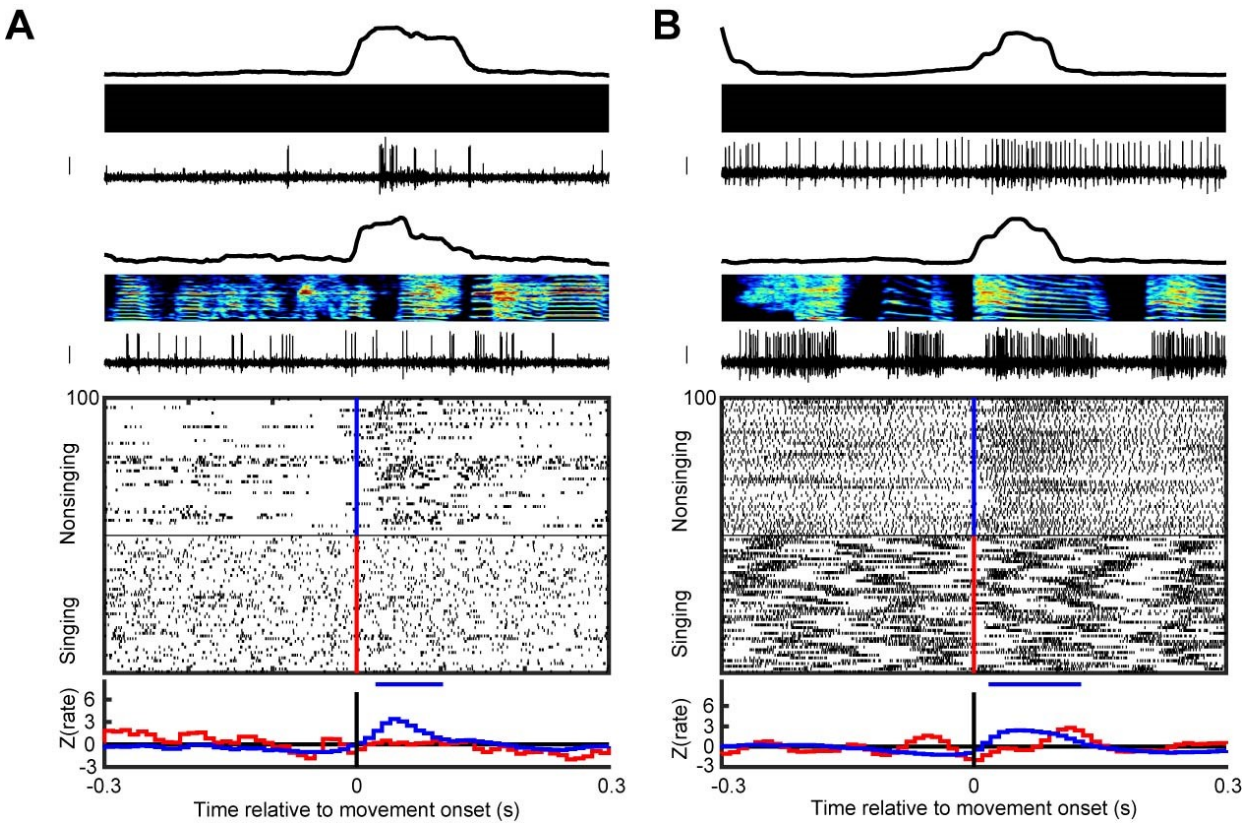

Figure 3.3. Example neurons that exhibit movement selectivity in the non-singing state only.

(**A**) (Top to bottom) bandpassed rectified accelerometer signal, spectrogram, voltage trace from an example VP neuron during non-singing and singing states, raster plot of spiking activity, z-scored firing rate histogram. Blue/red lines in the raster plot indicate movement onset during non-singing/singing. Horizontal bars indicate the duration of significant response to movement (z test, see methods). Scale bar for neural activity is 0.25mV. (**B**) Same as A for an example VTA neuron.

Movement related neurons switch to song-locked firing during singing

In both VP and VTA, movement related neurons could be precisely time locked to song during singing (Figure 3.4A,B, ref Chen). Because song timing and movement timing is correlated during singing (c.f. Figure 3.1B,C in ref (Gadagkar et al., 2016) a purely movement locked neuron could show spurious correlation to song timing. To test this, we compared the

magnitude of movement aligned rate modulation between singing and non-singing states. Movement modulation was quantified by the maximum in absolute value of z-scored rate histogram, and song-locked firing was quantified by intermotif cross correlation (Goldberg and Fee, 2010; Kao et al., 2008; Olveczky et al., 2005). Most neurons with significant time-locked firing to song showed higher movement selectivity outside singing than during singing (n=9/12 timing neurons with IMCC>0.3; n=69/93 song-locked neurons with significant IMCC. Figure 3.4C).

To test if change in movement selectivity could occur immediately at state boundaries, we computed movement-aligned rate histograms for those movements surrounding state transitions (n=53 neurons with at least 10 transitions tested). Whereas the last movements occurring immediately before song bouts were correlated with bursts of firing, the first movements within a song bout were not (Figure 3.4D,F, 493 transitions, 17/53 neurons tested, Methods). Similarly, the movements immediately following bout offsets were reliably associated with a burst (Figure 3.4E,F, 647 transitions, 14/53 neurons tested). Thus the change in tuning to movement can occur extremely rapidly (~0.1 second timescale) at transitions to and from singing state.

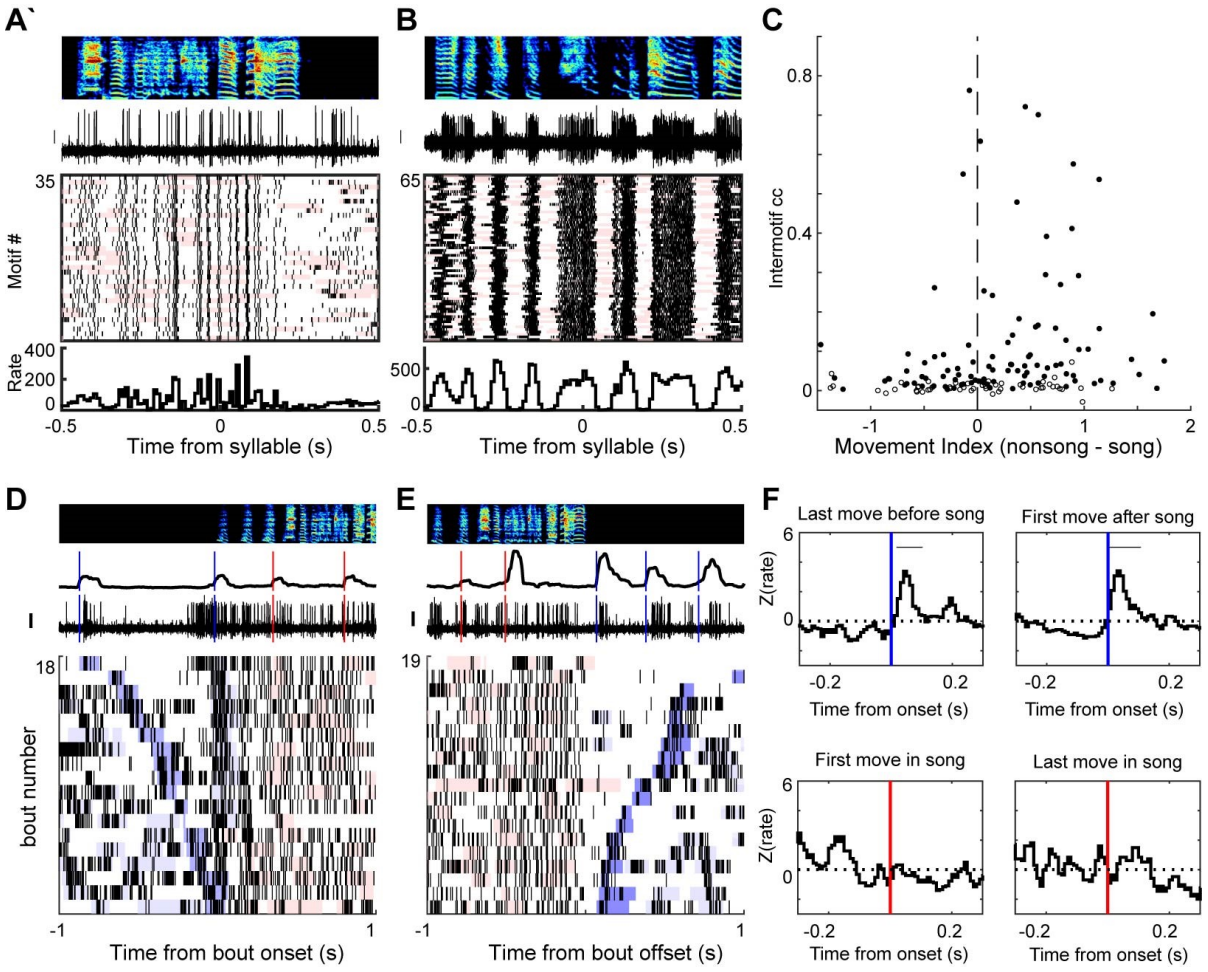

Figure 3.4. Example neurons that switch their tuning between song timing and movement at singing state boundaries.

(**A**) Movement related VP neuron time-locked to song. Same neuron as in Figure 3.3A. Top to bottom: song spectrogram, neural signal, rasters, and rate histogram aligned to song. Pink shades indicate movements during singing. Scale bar for neural activity is 0.25mV. (**B**) Same as A for the VTA neuron in Figure 3.3B. (**C**) Scatterplot of intermotif cross correlation and difference in movement index between non singing and singing states. Each dot is a movement related neuron. Filled circles are neurons with significant song-locked activity ($p<0.05$, methods). (**D**) Example neural activity at transitions from non singing to singing, same neuron as A. Top to bottom: song spectrogram, bandpassed rectified accelerometer signal, neural signal, and rasters, aligned to onset of singing bouts. Blue and red lines indicate movement onsets outside and during singing. Blue shades in raster indicate movements outside song, with darker blue for the last movement before bout onset. Red shades indicate movements during singing. Raster sorted by the timing of

the last movement before song. (**E**) Same as D, but for transitions from singing to non singing. Raster sorted by timing of first movement after song. (**F**) Top: z scored firing rate of neuron in D-E, aligned to movements immediately before (left) and after song (right). Bottom: same as top, for first and last movements during song. Blue/red bars indicate movement onsets outside/during song. Horizontal bars indicate significant response ($p<0.05$, z-test, methods).

Discussion

By recording neural firing during singing and non-singing states in freely behaving birds, we discovered that movement-locked firing in VTA and VP neurons can be gated off during singing. In addition, neurons that were precisely time-locked to movements during non-singing became instead time-locked to song syllables (and not to body movements) during singing. These changes in tuning to distinct behavior modules could occur within tens of milliseconds at state boundaries. While basal ganglia and cerebellar neurons are known to be able to be differentially tuned to internally versus externally (e.g. cue-driven) movements (Strick et al., 1993; Van Donkelaar et al., 1999), to our knowledge such dramatic and rapid change in limbic tuning to movement across behavioral states has never been reported.

While the functional role of these state dependent movement representations is unclear, one possibility is that the context-dependent switch of representations between movement timing and song timing may reflect a common underlying algorithm that evaluates the quality of *any* motor program currently being produced, for example hop/orienting movements during non-singing versus syringeal movements during song. In this scenario, an outcome-weighted timing signal (such as a dopaminergic performance error signal) could be used to compute the predicted quality of ongoing performance, independent of the modality of the motor program.

Still, the switching of representation presents a puzzle for downstream neurons. For example, a neuron receiving input from a song timing related neuron during singing can reliably decode the time of song. However, as soon as singing stops, this same recipient neuron should also switch its decoding algorithm, for the incoming signal has changed. If this is the case, another input indicating the state change will be required. While the possible source of this proposed gating signal remains to be found, we note that previously discovered 'SongOn' and 'SongOff' neurons in VP (c.f. Figure 3.4 in (Chen et al., 2019)) which turn on or off their

activity during singing, exactly like an ‘isSinging’ gate. The potential local connections between these cell types within VP is unknown.

One caveat in our study is that we were unable to fully distinguish between types of movements - e.g. hopping versus neck rotations. Instead, we have measured acceleration at the level of the head, and computed onset and offset timing of movements. While birds appear to move in similar ways during singing and non-singing periods (Yuan and Bottjer, 2020), it’s possible that there are subtle systematic differences between ostensibly similar movements when performed during singing and non-singing. Future work with high speed video will be required to test this possibility. Notwithstanding, the complete cessation of time-locked firing to any type of movement during singing was striking.

Finally, in new analyses for the present paper, we discovered that non-error encoding VTA neurons (termed ‘VTAother’ in (Gadagkar et al., 2016)) could exhibit precise song-locked discharge (Figure 3.2C, 4B). Because error signals in the Area X projecting VTA neurons are temporally precise, timing signals in VTA could play a role in shaping dopaminergic signals important for song learning.

Methods

Subjects, surgery and histology

Subjects were 71 male zebra finches 74-355 days old singing undirected song. 61/71 birds and 269/291 neurons were new analyses of previously published datasets (Chen et al., 2019; Gadagkar et al., 2016). Animal care and experiments were approved by the Cornell Institutional Animal Care and Use Committee. All surgeries were performed with isoflurane anesthetization. Custom made microdrives carrying an accelerometer (Analog Devices AD22301), linear actuator (Faulhaber 0206 series micromotor) and homemade electrode arrays (5 electrodes, 3-5 MOhms, microprobes.com) were implanted into VP and VTA. VP implants (35 birds) were targeted using coordinates (4.4-5.4A, 1.1-1.5L, 3.5V, head angle 20 degrees). VTA implants (36 birds) were targeted using antidromic methods with stimulation in Area X (5.6A, 1.5L, 2.65V, head angle 20 degrees). After each experiment, small electrolytic lesions (30 μA for 60 s) were made with one of the recording electrodes. Brains were then fixed, cut into 100 μm thick sagittal sections for histological confirmation of stimulation electrode tracks and reference lesions.

Singing and non-singing states

We separately analyzed neural activity and movement patterns in singing and non-singing states, and during transitions between these states, as previously described (Goldberg et al., 2010; Goldberg and Fee, 2010). Bouts of singing was defined as consecutive syllables produced with gaps shorter than 300 ms. Non singing states were silent periods at least 300 ms away from syllables. In analysis of movement outside singing, only movements with onsets at least 300 ms away from song were included.

Quantification of movement

An accelerometer (Analog Devices AD22301) was mounted on microdrives to detect gross body movements as described previously (Chen et al., 2019; Gadagkar et al., 2016). Briefly, movement onsets and offsets were determined by threshold crossings of the band-passed, rectified accelerometer signal. We further quantify the amplitude of each movement as the area under the curve of this signal (Figure 3.1A).

Probability of movement onset was estimated for 10ms bins by the fraction of trials in which movement onsets were detected. To assess the significance of peaks in these probability functions, we compared the highest probability peak 1000 surrogate probability functions generated by randomly time-shifting movement onset relative to syllable onsets. Probability peaks exceeding the 95th percentile of surrogate probability maximum were considered significant.

To select similar movements shared between singing and non-singing states, we computed the joint distribution of duration and amplitude for all detected movements from each bird, and restricted subsequent analysis to those movements within 5-95th percentile in both dimensions for both conditions.

To quantify movement locked neural response, we computed z-scored firing rates aligned to movement onsets during singing and non singing states. Movement index was defined as the highest absolute z score within 100 ms before or after movement onsets. To assess the significance of these movement-locked rate modulations, we compared the highest rate peak and lowest nadir in movement onset-aligned rate histogram to 1000 surrogate rate histograms generated by randomly time-shifting spike trains. Rate peaks exceeding the 95th percentile of

surrogate rate maximum and rate nadirs below the 5th percentile of surrogate rate minimum were considered significant.

To calculate the latencies and durations of movement responses, spiking activity within ±300 ms relative to movement onset was binned in a moving window of 10 ms with a step size of 5 ms. Each bin was tested against all the bins in the first 200 ms using a z-test. Response onset (latency) was defined as the first bin for which the next 4 consecutive bins were significantly different from the prior activity (z-test, $P < 0.05$); response offset was defined as the first bin after response onset for which the next 2 consecutive bins did not differ from the prior ($P > 0.05$, z-test). Response duration was the difference between the offset and the onset times

To quantify movement-related responses at song state boundaries, we computed z-scored firing rates aligned to movement onsets using only movements either immediately before or after onsets and offsets of singing bouts. Those state dependent neurons that had at least 10 trials of each transition type were included in this analysis. Significant response to state boundaries was assessed with bootstrap method as above, and the duration of significant responses were quantified using z-test as above (Figure 3.4F).

Analysis of neural activity

Neural signals were band-passed filtered (0.25-15 kHz) in homemade analog circuits and acquired at 40 kHz using custom Matlab software. Spike sorting was performed offline using custom Matlab software (courtesy Dmitriy Aronov). Firing rate histograms were constructed with 10 ms bins and smoothed with a 3-bin moving average.

Song timing related activity

Intermotif pairwise correlation coefficient (IMCC) was used to identify neurons that had highly time-locked firing to song motifs (timing neurons), as previously described (Chen et al 2019). Motif aligned IFR was time warped to the median duration of all motifs, mean-subtracted, and smoothed with a Gaussian kernel of 20 ms SD, resulting in $\boldsymbol{r_i}$ for each motif. IMCC was defined as the mean value of all pairwise CC between $\boldsymbol{r_i}$ as follows:

$$IMCC = \frac{1}{N_{pairs}} \sum_{j>i}^{N_{pairs}} CC_{ij}$$

$$CC_{ij} = \frac{\boldsymbol{r_i} \cdot \boldsymbol{r_j}}{\sqrt{\boldsymbol{r_i}^2 \boldsymbol{r_j}^2}}$$

To assess the significance of IMCC values, we compared the true IMCC value to 1000 surrogate IMCC values generated by randomly time-shifting spike trains. IMCC values were considered significant if greater than the 95th percentile of the surrogate values.

Error-related neurons

VP and VTA neurons were classified as error responsive (error neurons in Figure 3.2) from previous studies (Chen et al., 2019; Gadagkar et al., 2016). Briefly, birds received syllable-targeted distorted auditory feedback (DAF) on randomly interleaved renditions. We compared target aligned activity between distorted and undistorted renditions, and those neurons with significant difference in firing following DAF were labeled as error neurons.

REFERENCES

Ahn, S., and Phillips, A.G. (1999). Dopaminergic correlates of sensory-specific satiety in the medial prefrontal cortex and nucleus accumbens of the rat. Journal of Neuroscience 19, RC29-RC29.

Andalman, A.S., and Fee, M.S. (2009). A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. Proc Natl Acad Sci U S A 106, 12518-12523.

Barter, J.W., Li, S., Lu, D., Bartholomew, R.A., Rossi, M.A., Shoemaker, C.T., Salas-Meza, D., Gaidis, E., and Yin, H.H. (2015). Beyond reward prediction errors: the role of dopamine in movement kinematics. Front Integr Neurosci 9, 39.

Brooks, V.B. (1986). How does the limbic system assist motor learning? A limbic comparator hypothesis. Brain Behav Evol 29, 29-53.

Chen, R., Puzerey, P.A., Roeser, A.C., Riccelli, T.E., Podury, A., Maher, K., Farhang, A.R., and Goldberg, J.H. (2019). Songbird Ventral Pallidum Sends Diverse Performance Error Signals to Dopaminergic Midbrain. Neuron 103, 266-276 e264.

Eckmeier, D., Geurten, B.R., Kress, D., Mertes, M., Kern, R., Egelhaaf, M., and Bischof, H.-J. (2008). Gaze strategy in the free flying zebra finch (Taeniopygia guttata). PLoS One 3.

Engelhard, B., Finkelstein, J., Cox, J., Fleming, W., Jang, H.J., Ornelas, S., Koay, S.A., Thiberge, S.Y., Daw, N.D., and Tank, D.W. (2019). Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. Nature 570, 509-513.

Gadagkar, V., Puzerey, P.A., Chen, R., Baird-Daniel, E., Farhang, A.R., and Goldberg, J.H. (2016). Dopamine neurons encode performance error in singing birds. Science 354, 1278-1282.

Goldberg, J.H., Adler, A., Bergman, H., and Fee, M.S. (2010). Singing-related neural activity distinguishes two putative pallidal cell types in the songbird basal ganglia: comparison to the primate internal and external pallidal segments. J Neurosci 30, 7088-7098.

Goldberg, J.H., and Fee, M.S. (2010). Singing-related neural activity distinguishes four classes of putative striatal neurons in the songbird basal ganglia. Journal of Neurophysiology 103, 2002-2014.

Hisey, E., Kearney, M.G., and Mooney, R. (2018). A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. Nature neuroscience, 1.

Humphries, M.D., and Prescott, T.J. (2010). The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. Prog Neurobiol 90, 385-417.

Jin, X., and Costa, R.M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. Nature 466, 457-462.

Kao, M.H., Wright, B.D., and Doupe, A.J. (2008). Neurons in a forebrain nucleus required for vocal plasticity rapidly switch between precise firing and variable bursting depending on social context. J Neurosci 28, 13232-13247.

Kearney, M.G., Warren, T.L., Hisey, E., Qi, J., and Mooney, R. (2019). Discrete Evaluative and Premotor Circuits Enable Vocal Learning in Songbirds. Neuron.

Olveczky, B.P., Andalman, A.S., and Fee, M.S. (2005). Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. PLoS Biol 3, e153.

Papageorgiou, G.K., Baudonnat, M., Cucca, F., and Walton, M.E. (2016). Mesolimbic dopamine encodes prediction errors in a state-dependent manner. Cell reports 15, 221-228.

Strick, P., Hoover, J., and Mushiake, H. (1993). Role of the cerebellum and basal ganglia in voluntary movement.

Tumer, E.C., and Brainard, M.S. (2007). Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. Nature 450, 1240-1244.

Van Donkelaar, P., Stein, J., Passingham, R., and Miall, R. (1999). Neuronal activity in the primate motor thalamus during visually triggered and internally generated limb movements. Journal of Neurophysiology 82, 934-945.

Williams, H. (2001). Choreography of song, dance and beak movements in the zebra finch (Taeniopygia guttata). Journal of Experimental Biology 204, 3497-3506.

Xiao, L., Chattree, G., Oscos, F.G., Cao, M., Wanat, M.J., and Roberts, T.F. (2018). A Basal Ganglia Circuit Sufficient to Guide Birdsong Learning. Neuron 98, 208-221 e205.

Yuan, R.C., and Bottjer, S.W. (2020). Multi-dimensional tuning in motor cortical neurons during active behavior. bioRxiv.

CHAPTER 4

ACTOR-CRITIC REINFORCEMENT LEARNING IN THE SONGBIRD

Abstract

It feels ‘rewarding’ to ace your opponent on match point. Here we propose common mechanisms underlie reward and performance learning. First, when a singing bird unexpectedly hits the right note, its dopamine (DA) neurons are activated as when a thirsty monkey receives an unexpected juice reward. Second, these DA signals reinforce vocal variations much as they reinforce stimulus-response associations. Third, limbic inputs to DA neurons signal the predicted quality of song syllables much like they signal the predicted reward value of a place or a stimulus during foraging. Finally, songbirds may solve infamous problems in reinforcement learning such as credit assignment and catastrophic forgetting with node perturbation and consolidation of reinforced vocal patterns in motor cortical circuits. Consolidation occurs downstream of a canonical ‘actor-critic’ circuit motif that learns to maximize performance quality in essentially the same way it learns to maximize reward: by computing and learning from prediction errors.

Main Text

Edward Thorndike captured the essence of reinforcement learning in his Law of Effect: 'Responses that produce a satisfying effect in a particular situation become more likely to occur again in that situation, and responses that produce a discomforting effect become less likely to occur again in that situation (Thorndike, 1911).' Learning requires three pieces of information: (1) the response ('action') an animal makes; (2) the situation (or 'state') in which the action is taken; and (3) the evaluation of the outcome (effect).

Studies of hungry or thirsty animals learning for rewards have clarified how these three pieces of information are processed in dopamine-basal ganglia (BG) circuits during reinforcement learning (RL). Ventral tegmental area (VTA) dopamine (DA) neurons signal the outcome in the form of 'reward prediction error' (RPE): the difference between actual and predicted reward (Schultz, 2007). DA neurons exhibit bursts in response to unexpected rewards (surprisingly good outcomes) and pauses when a predicted reward is omitted (disappointingly bad outcomes) (Figure 4.1a). In 'actor-critic' (AC) models, these DA signals control synaptic plasticity throughout the basal ganglia, including a ventral 'critic' with outputs back to the midbrain and a dorsal 'actor' with outputs to the motor system (Joel et al., 2002) (Figure 4.2a). Both subdivisions implement DA-modulated plasticity to weigh cortical (or thalamic) inputs (which encode the situation, or 'state') according to their reward value (Yagishita et al.). DA-modulated plasticity in the critic computes predicted value of a state, i.e. how much reward to expect in a given situation. Predicted state-value signals, manifest as ventral striato-pallidal responses to conditioned stimuli or reward-associated places (Humphries and Prescott, 2010), provide VTA with *prediction* information necessary to compute RPE (Tian et al., 2016). VTA projects back to the critic (to update predicted state-value) and also to the 'actor'. DA-modulated plasticity in the 'actor' weighs each state-action pair according to its predicted quality (or Q

value). Somehow, motor circuits downstream of dorsal striatum convert Q into reward-maximizing action, i.e. the policy (Figure 4.2a)(Ito and Doya, 2011).

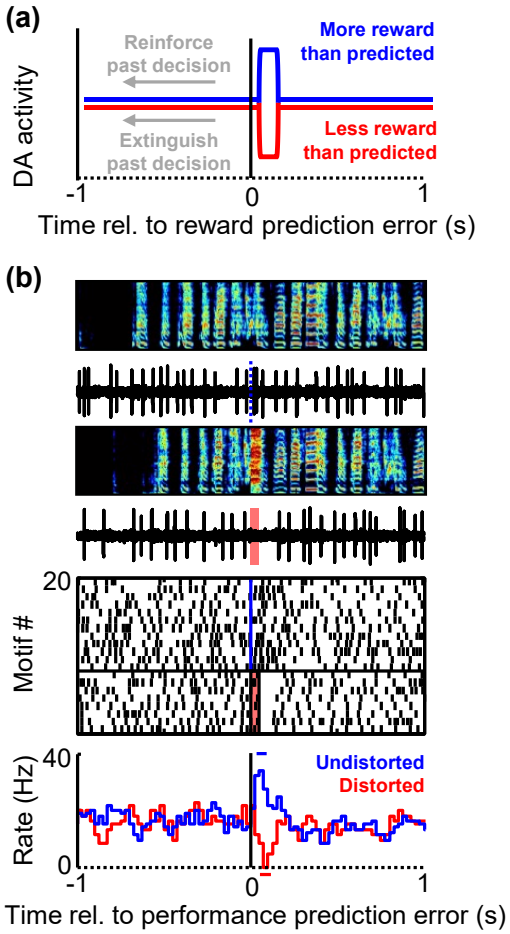

Figure 4.1. Dopaminergic error signals in singing birds.

(a) DA neurons signal better-than and worse-than predicted reward outcomes with phasic activations (blue) and suppressions (red). **(b)** A DA neuron recorded in a singing bird. Perceived error was controlled by playing 50 ms distorted auditory feedback (DAF) during half of the renditions of a 'target' syllable. Spectrograms with time-aligned voltage traces show responses of a DA neuron during undistorted (top) and distorted trials (bottom). Rasters and histograms show suppressions following distortions (red) and activations following the precise song time-step when error was predicted to occur but did not occur (blue). Reproduced from (Gadagkar et al.).

Intrinsically motivated song learning

Like human speech, birdsong is a complex sequence learned by matching ongoing performance to an internal goal. Juvenile zebra finches memorize a tutor song, begin to babble, and gradually learn over weeks to sing the tutor song. Songbirds have a specialized 'song system,' and its output RA (robust nucleus of the arcopallium), is a motor cortex-like nucleus with topographic outputs to brainstem motor neurons (Figure 4.3). For simplicity, RA can be imagined as a piano keyboard, in which the spatial position of a neuron relates to the vocal muscle it will innervate. RA gets inputs from LMAN (lateral magnocellular nucleus of the anterior nidopallium) and HVC (proper name). LMAN exhibits stochastic neural activity, drives vocal babbling, contributes to trial-to-trial variability in adults, and projects topographically to RA (i.e. a 'key' in LMAN has a corresponding key in RA) (Kao and Brainard, 2006; Olveczky et al., 2005). HVC exhibits stereotyped 'chains' of neural activity that drive the correspondingly stereotyped adult song (Lynch et al., 2016; Picardo et al., 2016). HVC axons ignore RA topography (i.e. span the entire keyboard), so that a single HVC axon can, in principle, learn to strike any key (Fee and Goldberg, 2011).

An actor-critic (AC)-inspired framework for song learning

As birds mature from vocal babbling to stereotyped adult song, control of RA firing (and therefore vocal output) gradually transfers from LMAN to HVC (Aronov et al., 2011; Garst-Orozco et al., 2014). Though the practicing bird does not receive external rewards for 'hitting the right note,' we propose that song learning proceeds, at least in part (Hahnloser and Ganguli, 2013) (Box 1), via an RL-like algorithm implemented in an 'actor-critic' circuit motif inside the songbird BG (Figure 4.2b).

Box 1. Are supervised and unsupervised algorithms also implemented in birdsong? *Unsupervised learning:* During babbling the activity of a 'chhh'-producing motor neuron will be reliably correlated with a 'chhh' receptive auditory neuron. And a 'bb' motor neuron will similarly correlate with a 'bb' auditory neuron. Simple Hebbian learning rules could create paired forward and inverse models. In the forward model, the motor system 'tells' the sensory system what is about to happen, so that the sensory consequences of movements can be predicted. In the inverse model, the sensory system can 'call upon' the motor system to produce the desired output (Hahnloser and Ganguli, 2013). Reciprocal connections between HVC and the auditory system could instantiate these internal models (Roberts et al., 2017). In fact, a forward model could be important for extracting prediction error signals upstream AIV and VTA. *Supervised learning*: Supervised error signals encode precisely how an outcome differed from the target, which also specifies the necessary correction (e.g. the pitch was too low, so next time move it up). Learning from supervised error signals requires an inverse model that can implement the correction. So far there is surprisingly little evidence that supervised mechanisms play a role in birdsong, as birds even appear to solve pitch-shifting experiments with DA reinforcement mechanisms (Saravanan et al., 2019).

Songbird DA neurons encode RPE-like song evaluation signals.

AC models construct and learn from DA RPE signals. To test for RPE-like signals during singing, we recorded antidromically-identified VTA neurons that project to Area X (VTAx) while controlling perceived error with distorted auditory feedback (DAF). DAF is a 50-ms snippet of sound with the same amplitude and spectral content as normal zebra finch song that is known to drive DA and Area X-dependent reinforcement of undistorted syllable variants (Andalman and Fee, 2009; Tumer and Brainard, 2007) (Ali et al., 2013; Hoffmann et al., 2016). VTAx neurons, known to be dopaminergic, exhibited pauses after distortions (sounded bad) and bursts after undistorted renditions of target syllables (sounded good). Importantly, burst magnitude depended on past error probability: if one syllable was distorted with high probability, and different one with low probability, DA bursts were larger following the (more surprising) undistorted renditions of the high probability target (Gadagkar et al., 2016). **Thus VTAx DA neurons signal performance prediction error: the difference between how good a syllable**

sounded and how good it was predicted to sound based on recent practice. To compute error, DA neurons need information, at each time-step, about 'just heard' auditory error as well as predicted error (predicted syllable quality).

Auditory cortex sends 'actual' (just heard) error signals to VTA.

A hierarchy of auditory cortical areas converges in a high-order VTA-projecting cortical area called AIV (ventral intermediate arcopallium). VTA-projecting AIV neurons (AIVvta) exhibit bursts in response to DAF during singing (Mandelblat-Cerf et al., 2014). Electrical microstimulation of AIV drives pauses in VTAx neurons (Chen et al., 2019) and optogenetic activation of the AIV-VTA pathway extinguishes syllable variations (just like phasic suppression of the VTA-X pathway does) (Hisey et al., 2018; Kearney et al., 2019; Xiao et al., 2018). **AIV can signal auditory error and drive pauses in DA firing (Figure 4.2b, lower left).** AIV may be functionally analogous to anterior cingulate cortex, which also may send performance error signals to VTA (Kolling et al., 2016).

Ventral pallidum (VP) sends predicted syllable quality signals to VTA.

Songbird VP is a mixed striatopallidal nucleus (Gale et al., 2008) that may function analogous to the critic (Chen et al., 2019; Kearney et al.). VP is necessary for learning and receives inputs from Uva, a thalamic nucleus that sends song time-step information to HVC, and also from VTAx neurons. DA-modulated plasticity of Uva inputs could weigh time-steps according to their past error. For example, consider a song with three time-steps t_1, t_2, t_3. If t_2 is reliably correlated with error, then DA pauses (driven by AIV, as described above) would be coincident with those Uva inputs active at t_2. Then DA-modulated plasticity would re-weigh

these synapses, resulting in a representation in VP of error-weighted timing or, equivalently, predicted syllable quality (Figure 4.2b, upper left). Consistent with this idea, most antidromically-identified **VPvta neurons exhibited pauses in firing immediately *before* the song time-step associated with past error, exactly consistent with a predicted syllable quality signal (Figure 4.2b) (Chen et al., 2019).**

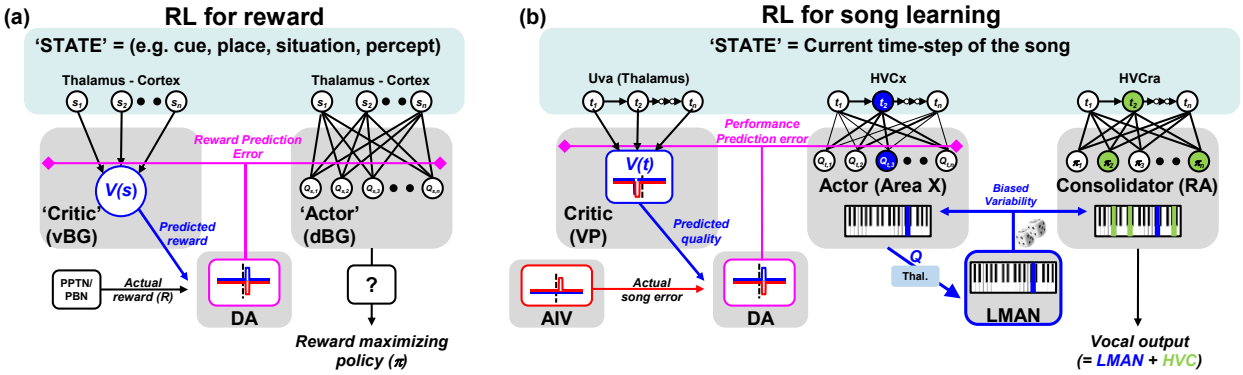

Figure 4.2. Actor-critic RL for both reward and song learning.

(a) The environment provides current state information, S, and current reward, r. The 'Actor' learns the quality of state/action pairs (Q(s,a)) that get converted into the reward maximizing action given the state (i.e. the *policy* $\pi(a|S)$). DA-weighted state representations in the critic compute the predicted state-value, $V(s)$. DA neurons signal RPE by taking the difference between actual, R, and predicted, $V(s)$, reward. **(b)** Lower left: VTA-projecting AIV neurons encode auditory error, e.g. bursts following DAF (Mandelblat-Cerf et al., 2014). Inset schematizes firing rates during distorted (red) and undistorted (blue) renditions; vertical dashed line denotes the time-step 'targeted' with DAF (as in Figure 4.1B). AIV bursts drive pauses in VTAx neurons through local VTA inhibition (not shown) (Chen et al., 2019; Kearney et al., 2019). The DA error signal (pink line) goes to both VP ('Critic,' left) and Area X ('Actor,' right). DA modulated plasticity in VP could weigh time-step (i.e. 'state') information according to past error. With an eiligibility trace (time during which a synapse is eligible for plasticity, Yagishita et al., 2014), this would explain why most VPvta neurons exhibited pauses right before the DAF target time (Chen et al., 2019). This predicted quality signal, similar to predicted state value in classic 'critic' circuits, could help VTAx neurons compute prediction error. DA-modulated plasticity in Area X, schematized as a keyboard due to its topographic organization,

could learn the quality of each state/action pair (*Q(s,a)*). For example Area X could learn that striking blue key at t2 is high quality, and relay this signal through DLM to LMAN, resulting in a premotor 'bias' signal that, together with the HVC-driven motor program, produces the vocal output. LMAN bias, if consistently associated with better-than-predicted song outcomes, is consolidated over days into the motor program (e.g. the Area X and LMAN-driven blue key in RA eventually becomes an HVC-driven green key, lower right) (Andalman and Fee, 2009; Warren et al., 2011).

Put simply, imagine an animal foraging a familiar environment in search of food. It will have a memory of where it got rewards, resulting in a place-dependent reward prediction. Now imagine a bird practicing a song with many syllables. It will similarly have a memory of when in the song it made mistakes, resulting in a syllable-dependent error prediction. Thus we view VP's role in computing the predicted quality of syllables as conceptually similar to its long-established role as a 'critic' that computes the predicted reward value of cues or places.

HVC provides 'state' information in the form of what 'time it is' in the song.

HVC time-steps (Figure 4.3c) are ideal 'state' representations because song policy is, in essence, learning *what* piano key to press *when* (Fee and Goldberg, 2011). Importantly, because the axons of single HVC or VTA neurons span the entirety of Area X, any key can in principle be learned to be struck at any given time-step (Fee and Goldberg, 2011).

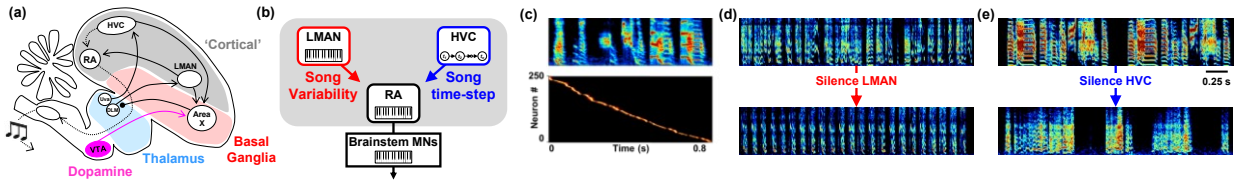

Figure 4.3. Variability- and timing- generating pathways in the song system.
(a-b) The HVC→RA 'motor' pathway exhibits stereotyped neural activity that drives adult song; the LMAN→RA pathway drives vocal variability and babbling. The basal ganglia nucleus Area X receives DA inputs from VTA projects to LMAN via a motor thalamic nucleus called DLM. **(c)** HVC synfire chains track time-step in song. Spectrogram of a song motif plotted above a heatmap of HVC bursting activity. Each row shows the smoothed firing rate of a single burst. **(d)** Spectrogram from a babbling bird before (top) and after (bottom) LMAN inactivation. Note loss of vocal variability, revealing stereotyped, HVC-driven song elements. **(e)** Spectrogram from adult bird before (top) and after (bottom) HVC inactivation. Note elimination of song structure, revealing LMAN-driven vocal variability. Data from (Lynch et al., 2016). AIV and VP also project to VTA (not shown).

Area X as part of the 'actor' that computes Q values.

Area X gets three main inputs: LMAN (which provides a copy of the vocal 'guess' it made through RA), HVC (time-step) and VTAx (DA RPE). In 2011, we proposed a specific learning rule in Area X based on dopamine modulated corticostriatal plasticity: If an HVC and LMAN input to a striatal medium spiny neuron (MSN) are co-active, then an eligibility trace (Etrace) is transiently activated that 'tags' the HVC-MSN synapse (Yagishita et al., 2014) (Box 2). If the LMAN activation was a 'lucky guess,' then a phasic DA burst will occur that coincides with the eligibility trace (Box 2), which would strengthen the tagged HVC-MSN synapse (we proposed the LMAN input implements node perturbation and is not itself plastic, see below). Across Area X, this process would compute Q, the quality of each action in a given state (i.e. the quality of each key strike at each time-step of the song). At each time-step of the song, a new vector of Q values (of length n=number of keys) could then *bias* LMAN to strike a new combination of high quality keys. Consistent with this model, lesions to Area X or its DA inputs block bias (Ali et al., 2013; Hoffmann et al., 2016), and optogenetic activation of DA terminals

in Area X at a specific time-step reinforces immediately preceding vocal variations (Hisey et al., 2018; Xiao et al., 2018). **Thus DA-modulated plasticity in Area X can compute the quality of pressing each key at each time-step (i.e the quality of each state/action pair, or Q(s,a)).**

The idea of Area X as a Q network makes additional predictions. If performance at a given time-step is poor no matter what action is taken, then the predicted quality of all state/action pairs will be low. This occurs naturally during vocal babbling, before the onset of learning, when the predicted quality of all syllables is likely to be low. At this learning stage, all Area X output neurons exhibit phasic activations at syllable onsets, exactly where state representations (HVC activity) are also concentrated (Okubo et al., 2015; Pidoux et al., 2014). As Area X outputs are inhibitory, we hypothesize that this is Area X's way of saying it has not yet learned any policy to promote. These pre-syllable activations rapidly go away over days of singing experience, as birds have an opportunity to learn which state/action pairs produce good outcomes (Goldberg et al., 2010). This hypothesis also predicts that all Area X output neurons recorded in adult birds should exhibit phasic rate increases prior to a song time-step targeted with DAF with 100% probability, an easy experiment.

Consolidation in the HVC-RA pathway.

In AC models used in machine learning, deletion of the actor would have a devastating effect on task performance. Yet song remains intact following lesions to Area X or its downstream thalamo-cortical pathway (Area X-DLM-LMAN). Seminal experiments showed how LMAN-driven 'bias' is consolidated into the HVC-RA pathway (Andalman and Fee, 2009; Warren et al., 2011). When DAF was delivered only at low-pitch renditions of a target syllable, birds learned over hours to move that syllable's pitch up. When LMAN was inactivated at the

end of this day, the pitch of the target syllable immediately returned to the morning's value, showing that Area X rapidly learned to *bias* LMAN to push the song away from error (i.e. to 'educate' LMAN guesses). Yet after days of sustained pitch-up bias, LMAN inactivation no longer caused a pitch shift, meaning that the pitch-up bias had been transferred from LMAN to the HVC-RA synapse (i.e. the LMAN bias was consolidated into the HVC-RA pathway (Figure 4.2b lower right and legend, i.e. the blue key in RA will turn into a green key).

Songbird variations on the classic actor-critic may solve classic problems in RL

A first unique feature of the songbird architecture is that the actor (Area X) sits upstream of a 'variability-generator' (LMAN) which in turn projects to a 'consolidator' (RA). These added thalamocortical layers between the actor and motor output may help solve two important problems in RL: credit assignment and catastrophic forgetting.

Box 2. How might a relatively slow DA reinforcement signal improve a fast behavior? Birds can produce reliable acoustic fluctuations with ~5-10 ms precision, the same duration as an HVC burst, the schematized 'time-step' in our model (Lynch et al., 2016). Yet the DA reinforcement signal is ~50 ms delayed from auditory error and lasts ~100 ms (Gadagkar et al., 2016). How might this relatively slow signal appropriately reinforce past vocalizations? Several lines of evidence suggest that an eligibility trace in the spines of Area X MSNs last around ~0.1 seconds. In carefully implemented distorted auditory feedback experiments, the Brainard group discovered that DAF only reinforces vocal variations in the immediately preceding 0.1 seconds (Tumer and Brainard, 2007). They also discovered that ~0.1 second duration 'chunks' of song are reinforced even when DAF is targeted with millisecond precision to specific syllable trajectories (Charlesworth et al., 2011). Although it may seem optimal to independently evaluate every ~5 ms time-step, we propose that a courser evaluation system may work for birdsong. Acoustic structure of a syllable is largely a function of air pressure and muscle activation in the syrinx, and therefore song production is better understood as a continuous trajectory through syringeal state space rather than transition between discrete states. Because neither air pressure nor muscle configuration can be instantly transformed, the action at each time-step constrains what new configurations are possible in the next. For example, an input to the syrinx that drives a 5 Hz increase in pitch would only produce the desired 500Hz output when the preceding pitch was 495 Hz. We hypothesize that reinforcing a larger chunk of consecutive actions could reduce the dimensionality of search space and improve learning.

Solving credit assignment with LMAN-dependent node perturbation in Area X.

A classic problem in RL is credit assignment: after an error, how does the brain know which of its millions of synapses need to be changed? The error-backpropagation algorithm used in machine learning updates each synaptic weight based on its known unique contribution to behavioral output, but this might not be biologically plausible. An alternative approach is node-perturbation, which measures the change in error caused by local stochastic fluctuations in neural activity (Lillicrap et al., 2020). Node perturbation provided the inspiration for our proposed

learning rule in Area X: Area X MSNs detect which 'guesses' (from LMAN) at which time-steps (from HVC) result in better-than-predicted outcome (from VTA) (Fee and Goldberg). In this model, only the HVC-MSN synapse is plastic, and the LMAN input is there to provide a 'copy' of the perturbation to vocal output caused by LMAN's collateral in RA. A recent connectomics study identified a micro-architecture in Area X ideally suited to implement node perturbation: HVC-MSN synapses were primarily on dendritic spines and exhibited size distributions with structural hallmarks of Hebbian plasticity (Kornfeld et al., 2020). Meanwhile, LMAN-MSN synapses were primarily on dendritic shafts, making them more suited for gating HVC-spine plasticity than for undergoing plasticity themselves. More broadly, this study provides a functional logic for shaft and spine synapses throughout the nervous system.

Solving 'catastrophic forgetting' with consolidation in RA.

Another classic problem of motor sequence reinforcement learning is knowing *when* to allow for plastic changes to a sequence. For example, when learning new songs on the piano, it might be maladaptive to enable plasticity in synapses important for producing songs that you already know. In artificial neural networks, this is known as 'catastrophic forgetting'. Synaptic weight changes that maximize performance of newly learned behaviors can impair previously learned ones. This problem can be solved with 'elastic weight consolidation' - a process that protects synaptic weights that are useful for already-learned behaviors (Kirkpatrick et al., 2017). Consolidation in RA may reduce catastrophic forgetting in several ways. First, after the HVC-RA pathway 'takes over' control of a specific part of the song, Area X synapses are free to learn (or unlearn) new policies without degrading ongoing vocal performance. Area X policies can 'bias' LMAN variability and, only if a bias is stable for days, will it get consolidated into the

HVC-RA pathway (Andalman and Fee, 2009; Warren et al., 2011). Second, neurogenesis of RA-projecting HVC neurons occurs throughout song learning, which could enable weight changes of new HVC-RA connections to occur without altering previously learned ones (Scott and Lois, 2007). Third, plasticity of existing HVC-RA synapses could be gated by uncertainty – such that reliably well executed time-steps of the motor sequence are 'protected.' For example, if the bird repeatedly makes mistakes (or is distorted) at one 'difficult' time-step in the song, the predicted error is high at that specific time-step. Importantly, cholinergic inputs to RA and HVC come from VP (where predicted error signals are known to reside), inhibition in HVC is reduced during new syllables (Vallentin et al., 2016), and cholinergic signaling in RA is required for synaptic plasticity and for song learning (Puzerey et al., 2018; Salgado-Commissariat et al., 2004). We predict that RA projecting VP neurons exhibit bursts of activity immediately before error-prone time-steps (i.e. DAF-targeted) of the song. We predict that acetylcholine 'tells' motor cortex when a time-step with an uncertain outcome is about to occur, enabling synaptic plasticity important for consolidation specifically at this time-step of the sequence. We also predict a specific Ach-modulated heterosynaptic learning rule in RA: If LMAN, HVC and cholinergic inputs to an RA neuron are reliably coactive, then strengthen the connection strength between HVC-RA. This rule would enable an HVC time-step to 'take control' of striking a high quality key specifically at low quality time-steps, and at the same time would 'protect' existing HVC-RA synapses at reliably high quality time-steps. We hypothesize that cholinergic uncertainty signals in mammalian motor cortex could serve a similar function (Ramanathan et al., 2015; Yu and Dayan, 2005).

Nonstandard 'actor-to-critic' projections may implement advantage actor-critic.

Curiously, we found that parts of the proposed 'actor' pathway (Area X, DLM and RA) project to 'the critic' VP (see also (Gale et al., 2008)), revealing projections from actor back to critic not required in standard AC models. Notably, a growing family of 'advantage actor critic' (A2C) algorithms could make use of such projections (Mnih et al., 2016). In contrast to classic AC where the RL signal is the difference between reward received and the predicted state value V(s), in A2C, the RL signal to the actor (for policy update) additionally considers the 'advantage,' i.e. the difference between the predicted quality of the action taken (i.e. the Q value Q(s,a)) and the predicted value of the state (V(s)). VP could inherit Q information from Area X, could compute V(s) as described in Figure 4.2b, and could compute the advantage as the difference between the two. This advantage could be relayed to VTA to influence its reinforcement signal.

This idea predicts that the songbird could compare actual song quality not just to the quality predicted at each time-step (V(s)), but additionally to the predicted quality of the action taken at that time-step (Q(s,a)), analogous to what an A2C network does. A simple experiment to test this possibility would be to record VTAx DA neurons while manipulating the advantage, $A(s_t,a_t)$. This could be done by implementing conditional distorted auditory feedback in which only low-pitch variants of a specific target syllable are distorted (Andalman and Fee, 2009; Tumer and Brainard, 2007). On rare catch trials, low pitch variants would instead be left undistorted. If DA neurons signal the difference between the actual outcome to the predicted outcome given the state (the target time of the song), then the magnitude of DA bursts would be the same for all undistorted target renditions, regardless of which syllable variant was produced. But if DA activity has information about the advantage, then DA bursts following undistorted renditions of the low pitch variants may be larger than bursts following undistorted renditions of

high pitch variants (because low pitch renditions have been associated with histories of more error). Future recordings of VTAx neurons could therefore constrain which variant of AC-like algorithms is realized in the songbird.

Summary

Many open questions remain in songbirds. Foremost, it remains unknown how auditory pathways upstream AIV compare the song to the tutor. This process may occur in reciprocal connections between auditory areas and HVC and likely involves both efference-copy and tutor-memory guided cancellation and evaluation of predicted acoustic outcomes (Hahnloser and Ganguli, 2013; Mackevicius and Fee, 2018; Roberts et al., 2017). Second, our model fails to capture the real complexity of BG circuits (e.g. distinct cell types and pathways) and oversimplifies how DA signals are constructed and used. For example, songbirds and mammals share indirect pathways and striatal interneuron classes whose roles in learning remain unclear (Goldberg et al., 2010; Goldberg and Fee, 2010). And because VTA-projecting neurons in mammals and birds encode an incredible diversity of motor, reward, and error-related signals, it remains unclear how relatively homogenous DA error signals are computed from mixed inputs (Chen et al.; Tian et al.). Finally, our model focuses on learning in adult birds where clear-cut time-step representations already exist in the HVC chain. It remains unclear what neural mechanisms enable HVC chains to develop in the first place (Okubo et al., 2015).

Comparative approaches can distinguish general principles from behavior-, effector-, and species-specific solutions to motor learning problems, and can also generate new hypotheses. For example, we predict that placing the 'actor' upstream of a 'guesser' and a 'consolidator' (as the bird's do with LMAN and RA) could lead to improved machine implementation of sequence learning. We also believe that the utility of the AC framework in song learning, reward-based learning, and machine learning suggests a general principle for computing and learning from prediction errors.

REFERENCES

Ali, F., Otchy, T.M., Pehlevan, C., Fantana, A.L., Burak, Y., and Olveczky, B.P. (2013). The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. Neuron 80, 494-506.

Andalman, A.S., and Fee, M.S. (2009). A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. Proc Natl Acad Sci U S A 106, 12518-12523.

Aronov, D., Veit, L., Goldberg, J.H., and Fee, M.S. (2011). Two distinct modes of forebrain circuit dynamics underlie temporal patterning in the vocalizations of young songbirds. Journal of Neuroscience 31, 16353-16368.

Charlesworth, J.D., Tumer, E.C., Warren, T.L., and Brainard, M.S. (2011). Learning the microstructure of successful behavior. Nat Neurosci 14, 373-380.

Chen, R., Puzerey, P.A., Roeser, A.C., Riccelli, T.E., Podury, A., Maher, K., Farhang, A.R., and Goldberg, J.H. (2019). Songbird Ventral Pallidum Sends Diverse Performance Error Signals to Dopaminergic Midbrain. Neuron 103, 266-276 e264.

Fee, M.S., and Goldberg, J.H. (2011). A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. Neuroscience 198, 152-170.

Gadagkar, V., Puzerey, P.A., Chen, R., Baird-Daniel, E., Farhang, A.R., and Goldberg, J.H. (2016). Dopamine neurons encode performance error in singing birds. Science 354, 1278-1282.

Gale, S.D., Person, A.L., and Perkel, D.J. (2008). A novel basal ganglia pathway forms a loop linking a vocal learning circuit with its dopaminergic input. J Comp Neurol 508, 824-839.

Garst-Orozco, J., Babadi, B., and Ölveczky, B.P. (2014). A neural circuit mechanism for regulating vocal variability during song learning in zebra finches. Elife 3, e03697.

Goldberg, J.H., Adler, A., Bergman, H., and Fee, M.S. (2010). Singing-related neural activity distinguishes two putative pallidal cell types in the songbird basal ganglia: comparison to the primate internal and external pallidal segments. J Neurosci 30, 7088-7098.

Goldberg, J.H., and Fee, M.S. (2010). Singing-related neural activity distinguishes four classes of putative striatal neurons in the songbird basal ganglia. Journal of Neurophysiology 103, 2002-2014.

Hahnloser, R., and Ganguli, S. (2013). Vocal learning with inverse models. In Principles of Neural Coding, S. Panzeri, and P. Quiroga, eds. (Boca Raton, FL: CRC Taylor and Francis), pp. 547-564.

Hisey, E., Kearney, M.G., and Mooney, R. (2018). A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. Nature neuroscience, 1.

Hoffmann, L.A., Saravanan, V., Wood, A.N., He, L., and Sober, S.J. (2016). Dopaminergic Contributions to Vocal Learning. J Neurosci 36, 2176-2189.

Humphries, M.D., and Prescott, T.J. (2010). The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. Prog Neurobiol 90, 385-417.

Ito, M., and Doya, K. (2011). Multiple representations and algorithms for reinforcement learning in the cortico-basal ganglia circuit. Current opinion in neurobiology 21, 368-373.

Joel, D., Niv, Y., and Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives. Neural Netw 15, 535-547.

Kao, M.H., and Brainard, M.S. (2006). Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. J Neurophysiol 96, 1441-1455.

Kearney, M.G., Warren, T.L., Hisey, E., Qi, J., and Mooney, R. (2019). Discrete Evaluative and Premotor Circuits Enable Vocal Learning in Songbirds. Neuron.

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A.A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., et al. (2017). Overcoming catastrophic forgetting in neural networks. Proc Natl Acad Sci U S A 114, 3521-3526.

Kolling, N., Wittmann, M.K., Behrens, T.E., Boorman, E.D., Mars, R.B., and Rushworth, M.F. (2016). Value, search, persistence and model updating in anterior cingulate cortex. Nature neuroscience 19, 1280.

Kornfeld, J., Januszewski, M., Schubert, P., Jain, V., Denk, W., and Fee, M.S. (2020). An anatomical substrate of credit assignment in reinforcement learning. BioRxiv.

Lillicrap, T.P., Santoro, A., Marris, L., Akerman, C.J., and Hinton, G. (2020). Backpropagation and the brain. Nature Reviews Neuroscience, 1-12.

Lynch, G.F., Okubo, T.S., Hanuschkin, A., Hahnloser, R.H., and Fee, M.S. (2016). Rhythmic Continuous-Time Coding in the Songbird Analog of Vocal Motor Cortex. Neuron 90, 877-892.

Mackevicius, E.L., and Fee, M.S. (2018). Building a state space for song learning. Curr Opin Neurobiol 49, 59-68.

Mandelblat-Cerf, Y., Las, L., Denisenko, N., and Fee, M.S. (2014). A role for descending auditory cortical projections in songbird vocal learning. Elife 3.

Mnih, V., Badia, A.P., Mirza, M., Graves, A., Lillicrap, T., Harley, T., Silver, D., and Kavukcuoglu, K. (2016). Asynchronous methods for deep reinforcement learning. In International Conference on Machine Learning, pp. 1928-1937.

Okubo, T.S., Mackevicius, E.L., Payne, H.L., Lynch, G.F., and Fee, M.S. (2015). Growth and splitting of neural sequences in songbird vocal development. Nature 528, 352-357.

Olveczky, B.P., Andalman, A.S., and Fee, M.S. (2005). Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. PLoS Biol 3, e153.

Picardo, M.A., Merel, J., Katlowitz, K.A., Vallentin, D., Okobi, D.E., Benezra, S.E., Clary, R.C., Pnevmatikakis, E.A., Paninski, L., and Long, M.A. (2016). Population-Level Representation of a Temporal Sequence Underlying Song Production in the Zebra Finch. Neuron 90, 866-876.

Pidoux, M., Bollu, T., Riccelli, T., and Goldberg, J.H. (2014). Origins of basal ganglia output signals in singing juvenile birds. J Neurophysiol, jn 00635 02014.

Puzerey, P.A., Maher, K., Prasad, N., and Goldberg, J.H. (2018). Vocal learning in songbirds requires cholinergic signaling in a motor cortex-like nucleus. J Neurophysiol.

Ramanathan, D.S., Conner, J.M., Anilkumar, A.A., and Tuszynski, M.H. (2015). Cholinergic systems are essential for late-stage maturation and refinement of motor cortical circuits. J Neurophysiol 113, 1585-1597.

Roberts, T.F., Hisey, E., Tanaka, M., Kearney, M.G., Chattree, G., Yang, C.F., Shah, N.M., and Mooney, R. (2017). Identification of a motor-to-auditory pathway important for vocal learning. Nat Neurosci 20, 978-986.

Salgado-Commissariat, D., Rosenfield, D.B., and Helekar, S.A. (2004). Nicotine-mediated plasticity in robust nucleus of the archistriatum of the adult zebra finch. Brain Res 1018, 97-105.

Saravanan, V., Hoffmann, L.A., Jacob, A.L., Berman, G.J., and Sober, S.J. (2019). Dopamine depletion affects vocal acoustics and disrupts sensorimotor adaptation in songbirds. eNeuro 6.

Schultz, W. (2007). Behavioral dopamine signals. Trends Neurosci 30, 203-210.

Scott, B.B., and Lois, C. (2007). Developmental origin and identity of song system neurons born during vocal learning in songbirds. Journal of Comparative Neurology 502, 202-214.

Thorndike, E.L. (1911). Animal Intelligence (Darien, CT: Hafner).

Tian, J., Huang, R., Cohen, J.Y., Osakada, F., Kobak, D., Machens, C.K., Callaway, E.M., Uchida, N., and Watabe-Uchida, M. (2016). Distributed and Mixed Information in Monosynaptic Inputs to Dopamine Neurons. Neuron 91, 1374-1389.

Tumer, E.C., and Brainard, M.S. (2007). Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. Nature 450, 1240-1244.

Vallentin, D., Kosche, G., Lipkind, D., and Long, M.A. (2016). Inhibition protects acquired song segments during vocal learning in zebra finches. Science 351, 267-271.

Warren, T.L., Tumer, E.C., Charlesworth, J.D., and Brainard, M.S. (2011). Mechanisms and time course of vocal learning and consolidation in the adult songbird. J Neurophysiol 106, 1806-1821.

Xiao, L., Chattree, G., Oscos, F.G., Cao, M., Wanat, M.J., and Roberts, T.F. (2018). A Basal Ganglia Circuit Sufficient to Guide Birdsong Learning. Neuron 98, 208-221 e205.

Yagishita, S., Hayashi-Takagi, A., Ellis-Davies, G.C., Urakubo, H., Ishii, S., and Kasai, H. (2014). A critical time window for dopamine actions on the structural plasticity of dendritic spines. Science 345, 1616-1620.

Yu, A.J., and Dayan, P. (2005). Uncertainty, neuromodulation, and attention. Neuron 46, 681-692.

CHAPTER *5*

NEURAL ACTIVITY RELATED TO VOCAL AND NONVOCAL SOCIAL BEHAVIOR IN THE PARROT

Abstract

We depend heavily on vocal communication for building and maintaining relationships. Vocal imitation has evolved in only a few groups of animals, including humans, songbirds, parrots, and others. The specialized songbird neural circuitry known as the song system has provided a great deal of insight into both vocal control and vocal learning. Much less is known for the neural mechanism of vocal control and learning in the parrot. Here I present the first neural recordings from a vocal motor cortical area, central nucleus of the anterior arcopallium (AAC), from singing budgerigars (*Melopsittacus undulatus*). Remarkably, AAC neurons responded to both vocalizations and body movements during social interactions.

Introduction

Parrots are known for their outstanding ability for vocal imitation (Cruickshank et al., 1993; Dooling et al., 1987; Hile et al., 2000; Hile and Striedter, 2000) and complex social behavior (Brockway, 1964a, b; Hobson et al., 2014; Ikkatai et al., 2016). Another avian vocal learner group, the songbirds, have evolved specialized anterior and posterior neural pathways for learning and producing songs, respectively. An anterior nucleus LMAN (lateral magnocellular nucleus of the nidopallium) actively generates variability during vocal learning and maintenance. A posterior nucleus HVC (proper name) generates a stereotyped sequence acquired through song development. Both nuclei project to RA (robust nucleus of the arcopallium), which controls song production by its projection to the hypoglossal nucleus in the brain stem (Fee and Goldberg, 2011; Olveczky et al., 2011). Whether an analogous specialization in song production and learning exists in the parrot is not well understood.

The parrot brain contains a set of nuclei that form a song system by virtue of their connection to the vocal apparatus (Paton et al., 1981), neurochemical markers (Ball, 1994), involvement in vocalization as shown by lesion studies (Brauth et al., 1997), and expression profile of immediate early genes during vocalizations (Chakraborty et al., 2015; Jarvis and Mello, 2000). In contrast to songbirds, where song nuclei are surrounded by other structures not involved in song production, parrot brain contains a shell song system, a second set of specialized neural tissue surrounding each of the main song nuclei (Chakraborty et al., 2015). The functional significance of this core-shell organization is unclear.

The central nucleus of the anterior arcopallium (AAC) is the final output of the forebrain song system which projects to the brainstem vocal center hypoglossal nucleus (Paton et al., 1981). Bilateral lesion of AAC abolishes learned calls (Brauth et al., 1997). Here we perform multi-channel neural recording in AAC and surrounding AAC shell in pairs of freely interacting

budgerigars and demonstrate for the first time AAC neurons exhibit precisely time-locked activity to both vocalizations and body movements.

Results

AAC neurons exhibit time-locked activity to contact calls and warble song

Budgerigar vocalizations include brief, stereotyped, individually unique contact calls (Figure 5.1b), and longer lasting warbles with variable duration and acoustic features (Figure 5.1a,c,e). In single and multiunit activity, many AAC neurons were time-locked to vocalizations (n = 156/280 neurons from two birds significantly modulated by contact call onsets). We aligned AAC neural activity to the onset of calls and warbles and found that most neurons responded to calls (Figure 5.1b,d). Some neurons also responded to warble elements (Figure 5.1c). Response to calls could be activations (n = 89 neurons) or suppression (n = 67 neurons). Consistent with a premotor role of AAC, contact call aligned activation were mostly before call onsets (latency to onset of activations: -25 ± 4.1 ms, n = 89 neurons; latency to onset of suppressions, -12 ± 4.6 ms, n = 67 neurons).

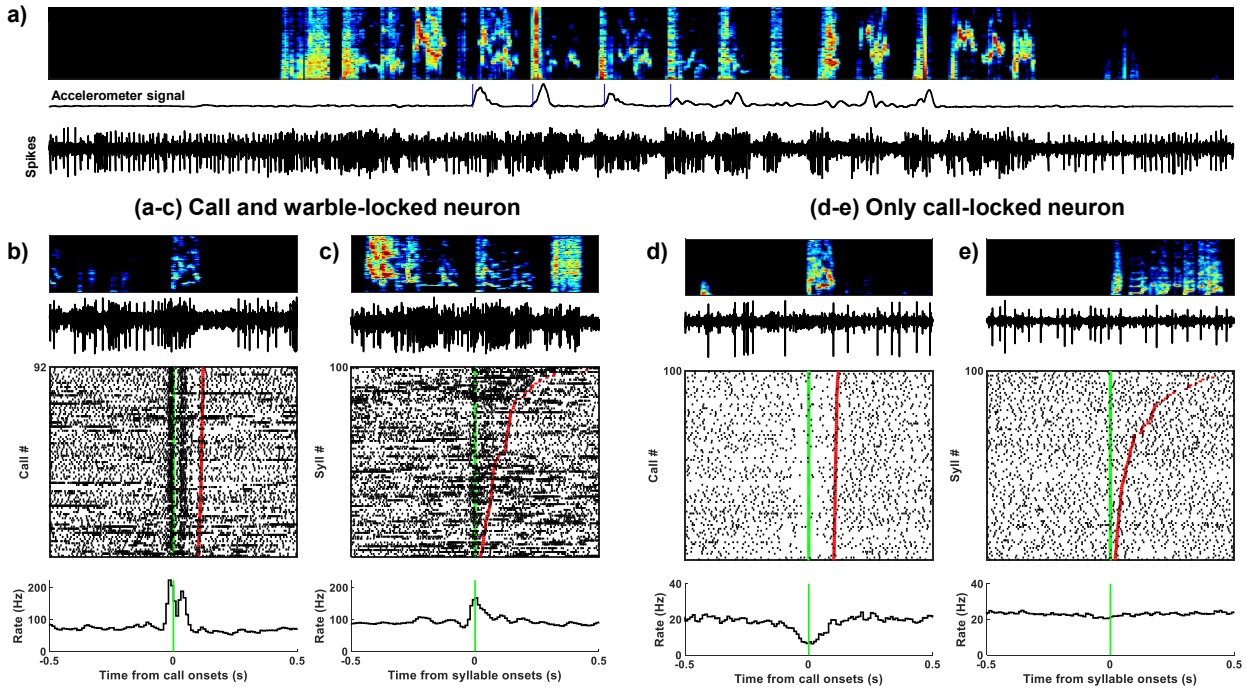

Figure 5.1 Example call- and warble-locked AAC neurons.

(**a**) Spectrogram, processed accelerometer signal, and spiking activity of an AAC neuron during five seconds of a male's warble song. Note the transition from tonic to bursty discharge during singing. (**b**) Top to bottom: spectrogram, spikes, raster, and rate histogram of the neuron from panel a, aligned to the onset of vocal calls. (**c**) Data plotted as in b for the same neuron aligned to warble syllable onsets. Rasters are sorted by duration, and red line indicates syllable offset. (**d-e**) Data from a different AAC neuron plotted as in (**b-c**). Note this neuron is suppressed during calls, and unmodulated during warble.

AAC neurons exhibit movement modulated activity

Neurons in the songbird vocal motor cortex RA are specialized for vocalizations and do not respond to body movement (Leonardo and Fee, 2005; Sober et al., 2008; Yu and Margoliash, 1996). We wondered if the same principle would apply to parrot AAC. To test this, we aligned the activity of call responsive AAC neurons to onsets of movements detected by the accelerometer. Surprisingly, we found a subset of AAC neurons responded to both contact calls and movements (N = 48/156 call responsive neurons from 2 birds also had significant response to the bird's own movement onsets, Figure 5.2).

We considered whether AAC neurons could respond to the vocal and non-vocal behavior of a conspecific as animals interact. Paired recording allowed us to align neural activity from one bird to calls produced by another bird. The contact calls also share acoustic similarity. Some AAC neurons were modulated by the other bird's calls (n = 30/156 call responsive neurons). None of the recorded AAC neurons had significant modulation by the other bird's movements.

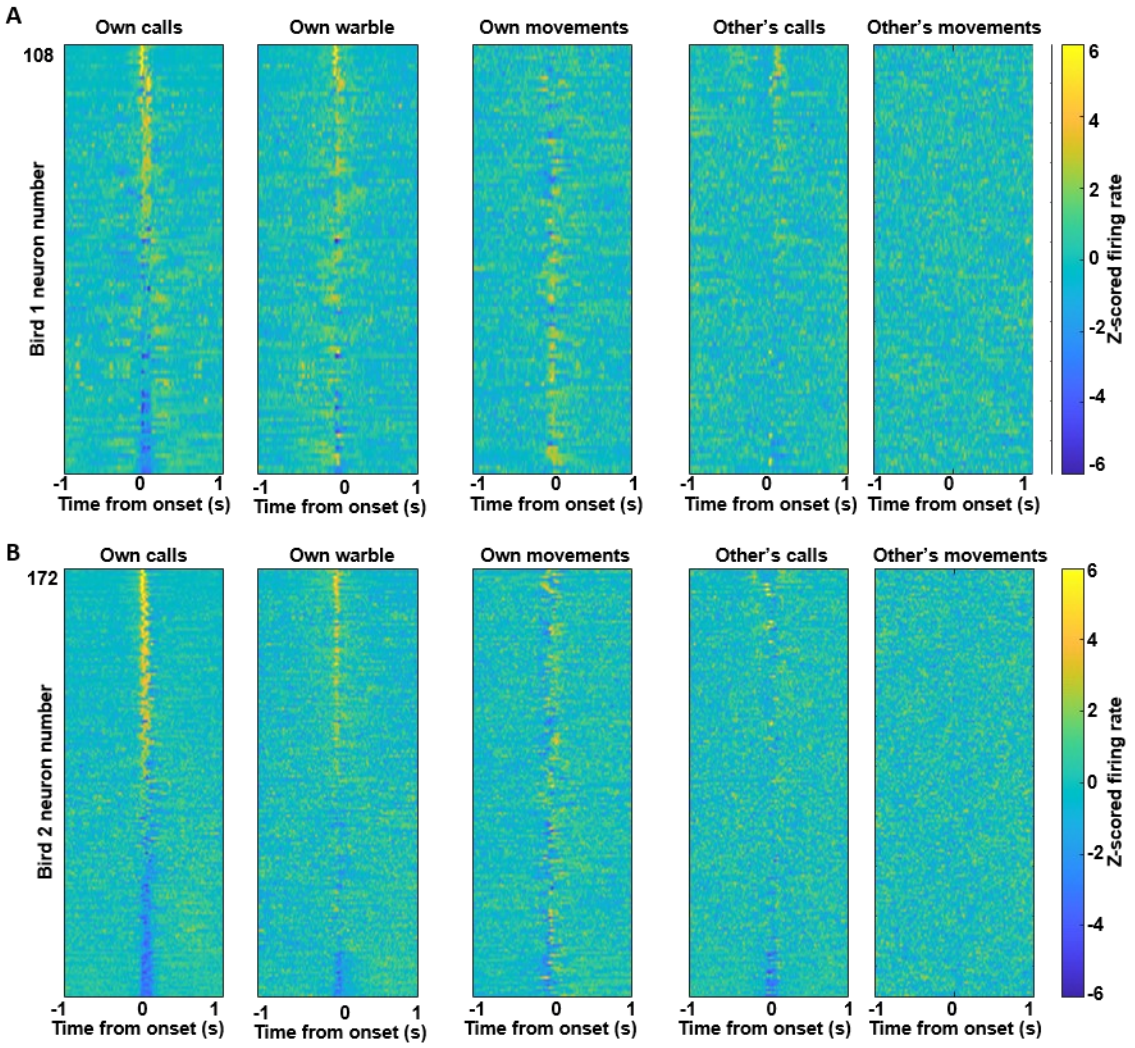

Figure 5.2 AAC neural responses to bird's own calls, warble song, movements, and calls and movements generated by a partner bird.

a). From left to right: each row plots the z-scored firing rate aligned to the onset of the bird's own calls, warble song elements, movements, and the other bird's calls and movements. All panels are sorted by the maximum value within 100 ms from call onsets (first panel).

b). Same as a) for neural activity recorded from the partner bird of a). Sorting is based on the bird's own calls.

Discussion

Using in vivo electrophysiology and accelerometry, we provide the first neural recordings from behaving parrots. Initial analysis supports a role of AAC in vocalizations. Approximately half of all AAC neurons recorded were highly time-locked to contact calls. This is in line with the anatomical organization of the parrot song system, and supports the idea that AAC controls vocalizations by its projection to the brainstem.

One caveat to interpreting whether these signals directly control AAC's downstream targets is that we do not have access to the projection target of recorded neurons. Future experiments can resolve this issue by either antidromic stimulation of downstream targets, or an opto-tagging approach.

Intriguingly, we also found movement related activity in the same neurons that are time-locked to syllables in AAC. This finding suggests the budgerigar song system, which evolved independently from the songbird song system, may operate on distinct principles. In the songbird song system, the output nucleus RA fires precisely timed bursts during song production but otherwise maintains a tonic firing rate outside of singing. Parrots may have evolved a hybrid system where AAC controls not only song but also some body movements by its broad projections to the brainstem motor centers (Brauth et al., 1997).

The proportion of movement responsive neurons reported here is likely to be an underestimate of true movement/vocal hybrid neurons. If there are neurons selective for specific kinds of movements such as social gestures, or gestures using distinct muscle groups, our broad analysis could result in false negatives. Future work is required to carefully classify the type of movement.

We hypothesize the parrot AAC has a function in producing communicative nonvocal gestures. The human Broca's area is thought to support hand and orofacial gestures (Nishitani et

al., 2005). The fact that parrots have a rich repertoire of social interactions, combined with our finding of representation of both vocalization and head movement, suggests the neural control of vocalization and non-vocal gestures may overlap in parrots in a similar way to Broca's area. Further analysis of movements with careful categorization based on video or accelerometry is required to elucidate this potential role of AAC in social gestures.

REFERENCES

Ball, G.F. (1994). Neurochemical specializations associated with vocal learning and production in songbirds and budgerigars. Brain, behavior and evolution 44, 234-246.

Brauth, S.E., Heaton, J.T., Shea, S.D., Durand, S.E., and Hall, W.S. (1997). Functional Anatomy of Forebrain Vocal Control Pathways in the Budgerigar (Melopsittacus undulatus). Annals of the New York Academy of Sciences 807, 368-385.

Brockway, B.F. (1964a). Ethological Studies of the Budgerigar (Melopsittacus Undulatus): Non-Reproductive Behavior. Behaviour 22, 193-222.

Brockway, B.F. (1964b). Ethological Studies of the Budgerigar: Reproductive Behavior. Behaviour 23, 294-323.

Chakraborty, M., Walloe, S., Nedergaard, S., Fridel, E.E., Dabelsteen, T., Pakkenberg, B., Bertelsen, M.F., Dorrestein, G.M., Brauth, S.E., Durand, S.E., and Jarvis, E.D. (2015). Core and Shell Song Systems Unique to the Parrot Brain. PLoS One 10, e0118496.

Cruickshank, A.J., Gautier, J.-P., and Chappuis, C. (1993). Vocal mimicry in wild African Grey Parrots Psittacus erithacus. Ibis 135, 293-299.

Dooling, R.J., Gephart, B.F., Price, P.H., McHale, C., and Brauth, S.E. (1987). Effects of deafening on the contact call of the budgerigar, Melopsittacus undulatus. Anim Behav 35, 1264-1266.

Fee, M.S., and Goldberg, J.H. (2011). A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. Neuroscience 198, 152-170.

Hile, A.G., Plummer, T.K., and Striedter, G.F. (2000). Male vocal imitation produces call con vergence during pair bonding in budgerigars, Melopsittacus undulatus. Anim Behav 59, 1209-1218.

Hile, A.G., and Striedter, G.F. (2000). Call convergence within groups of female budgerigars (Melopsittacus undulatus). Ethology 106, 1105-1114.

Hobson, E.A., Avery, M.L., and Wright, T.F. (2014). The socioecology of Monk Parakeets: Insights into parrot social complexity. Auk 131, 756-775.

Ikkatai, Y., Watanabe, S., and Izawa, E.I. (2016). Reconciliation and third-party affiliation in pair-bond budgerigars (Melopsittacus undulatus). Behaviour 153, 1173-1193.

Jarvis, E.D., and Mello, C.V. (2000). Molecular mapping of brain areas involved in parrot vocal communication. J Comp Neurol 419, 1-31.

Leonardo, A., and Fee, M.S. (2005). Ensemble coding of vocal control in birdsong. J Neurosci 25, 652-661.

Nishitani, N., Schurmann, M., Amunts, K., and Hari, R. (2005). Broca's region: From action to language. Physiology 20, 60-69.

Olveczky, B.P., Otchy, T.M., Goldberg, J.H., Aronov, D., and Fee, M.S. (2011). Changes in the neural control of a complex motor sequence during learning. Journal of Neurophysiology 106, 386-397.

Paton, J.A., Manogue, K.R., and Nottebohm, F. (1981). Bilateral Organization of the Vocal Control Pathway in the Budgerigar, Melopsittacus-Undulatus. J Neurosci 1, 1279-1288.

Sober, S.J., Wohlgemuth, M.J., and Brainard, M.S. (2008). Central contributions to acoustic variation in birdsong. J Neurosci 28, 10370-10379.

Yu, A.C., and Margoliash, D. (1996). Temporal hierarchical control of singing in birds. Science 273, 1871-1875.