Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

James N. McGregor Date

Shared mechanisms of auditory and non-auditory vocal learning in the songbird brain

By

James N. McGregor
Doctor of Philosophy

Neuroscience

Samuel Sober, Ph.D.
Advisor

Gordon Berman, Ph.D.
Committee Member

Michael Brainard, Ph.D.
Committee Member

Brian Dias, Ph.D.
Committee Member

Robert Liu, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Shared mechanisms of auditory and non-auditory vocal learning in the songbird brain

By

James N. McGregor
B.S., Brandeis University, MA, 2014

Advisor: Samuel Sober, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Neuroscience
2021

Abstract

Shared mechanisms of auditory and non-auditory vocal learning in the songbird brain
By James N. McGregor

The brain uses sensory feedback to guide changes in motor output. This process, known as sensorimotor learning, underlies the ability to learn complex skills necessary for animal survival, such as speech. A variety of sensory sources (auditory and non-auditory) of information are crucial for guiding vocal learning in both humans and songbirds. Also, a specialized neural pathway that underlies vocalizations has evolved in both species. However, the neural mechanisms that process non-auditory sensory information to guide vocal learning are unknown. Here, we study whether the specialized vocal neural circuit in songbirds processes exclusively auditory information to guide adaptive changes in vocal motor output. We do so by assessing the necessity of specific songbird brain regions within this vocal learning pathway for auditory and non-auditory vocal learning. We found that songbirds are capable of adapting elements of their song in response to non-auditory sensory signals. We also found that a cortical-basal ganglia circuit and its dopaminergic input are required for non-auditory vocal learning. Thus, the specialized neural circuitry for vocal learning in songbirds does not process exclusively auditory feedback. Instead, it processes sensory information from a variety of different sources to drive adaptive changes in vocal motor output. Due to the numerous analogies between human and songbird vocal neural pathways, we believe that this work improves our knowledge of how neural circuits underlie sensorimotor learning across species.

Shared mechanisms of auditory and non-auditory vocal learning in the songbird brain

By

James N. McGregor
B.S., Brandeis University, MA, 2014

Advisor: Samuel Sober, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Neuroscience
2021

ACKNOWLEDGEMENTS

Science is extremely rewarding yet also highly challenging. I would not have been able to tackle these challenges and be where I am today if it were not for the care and support from a number of people. I would like to express my deepest gratitude to these people.

First, I would like to thank my Ph.D. advisor, Dr. Samuel Sober. There are countless ways in which you were an outstanding advisor: your ability to design experiments, your interpersonal skills, your friendliness, your training in how to make a delicious cocktail, etc. There are two areas of mentorship I'd like to highlight in particular. First is your level of consistent support both personally and professionally. Second is your tremendous enthusiasm for science. This enthusiasm was infectious and kept me highly motivated to continue to do research even when challenges arose. I feel like I have developed a great deal as both a scientist and as a person being a member of the Sober lab, and for this I am extremely grateful.

I'd also like to thank the members of my thesis committee, Dr. Gordon Berman, Dr. Michael Brainard, Dr. Brian Dias, and Dr. Robert Liu. You all have also been highly supportive over the years. You've provided helpful feedback on my dissertation work that significantly improved the quality of our experimental design and interpretation of results. You've also pushed me to think deeply about the implications of my work across fields, which helped me develop my critical thinking abilities and grow as a scientist. You've all also been very generous with your time and very willing to offer both academic and career advice.

I'd like to thank the members of the Sober lab: Kyle Srivastava, David Nicholson, Lukas Hoffmann, Alynda Wood, Varun Saravanan, Andrea Pack, Amanda Jacob, Ben Kuebrich, Bryce Chung, Rachel Conn, Sena Agezo, MacKenzie Wyatt, Emily Berthiaume, Conor Kelly, David Hercules, Laura Waters-Goggins, Connor Gallimore, Sean O'Connell, Abi Grassler, Sherry Ye, Raveena Chhibber, Kailash Nagapudi, Samir

Chowdhury, Matt Williams, Leila Pascual, Kyle Thomas, Joyce Li, and Shreya Rana. You have all provided insightful feedback, helpful career advice, and important training. In addition, it has been a pleasure just to show up and work with all of you. We've always had great, fun conversations (let's go Celtics!), the lab has always felt like a collaborative environment, and we've developed friendships that I deeply cherish. I hope to stay in touch with all of you long into the future and look forward to swinging by the Sober lab posters at the next SfN.

I'd like to specifically mention that Amanda Jacob was particularly helpful in developing the histological techniques used throughout this dissertation work and Abi Grassler was a close collaborator for multiple years who contributed significantly to numerous experiments described in this dissertation. I'd like to also thank Michael Brainard and Paul Jaffe, who collaborated with us on this dissertation work and offered data in addition to excellent advice and feedback.

On a personal note, I certainly would not be where I am today if it weren't for my family. I'd like to thank my parents, Bob and Joyce McGregor, and my brother, Mikey McGregor, for always being a source of both love and comedy. I look up to all of you and I feel like you've taught me how to take my work very seriously but to not forget to find the humour in the little things outside of work. As our favorite jazz instructor would say, "you can't be too serious all of the time, otherwise you just end up becoming... serious." I appreciate our text conversations, visits, annual Cape trips, and weekly Zoom meetings where we play Codenames (which have helped to keep me sane this past year during the pandemic). I look forward to forcing you all to call me "doctor".

I'd like to thank my friends, many of whom were kind enough to watch and participate in my oral defense presentation. I feel very fortunate to have such hilarious, awesome friends from back home in Northborough/Southborough, from college, and from graduate school. I'm really glad that I've gotten to know all of you and I cherish

the memories of the time we've spent together. Finally, I'd like to thank my wonderful and supportive girlfriend, April Ratliff, for exploring Atlanta with me, making goofy jokes with me, reinvigorating my love of 90s alternative music, and teaching me how to throw a scoober.

Contents

List of Figures

Chapter 1

Introduction

1.1 Sensorimotor Learning

The brain receives sensory feedback and uses this information to modify future behaviors. This process, known as sensorimotor learning, is the basis of how animals acquire many complex skills (Krakauer et al., 2017; Doya, 2000; Krakauer and Mazzoni, 2011; Shadmehr et al., 2010). For example, sensorimotor learning is how a basketball player like Steph Curry learns to activate his muscles in a coordinated fashion to aim and shoot a basketball towards the net. If he misses the shot, his brain can then use the sensory feedback it receives to modify future motor commands to adjust his aim and hopefully make the next shot. This sensorimotor learning process is also how human babies acquire speech. Speech is a skilled behavior that requires complex, coordinated activation of muscles that control the mouth, tongue, and larynx, and it is necessary for adequate quality of life (Doupe and Kuhl, 1999; Marler, 1970; Goldstein et al., 2003). Human babies initially produce babbling noises. Their brains receive sensory information from the adults they hear, as well as the babbling sounds they produce, and they gradually learn to change their motor commands until they can produce words and sentences (Fitch, 2000). This endows humans with the remarkable ability

to communicate incredibly complicated, nuanced ideas.

The neural mechanisms underlying sensorimotor learning have been the focus of scientific research for decades (Kahl, 1878; Stratton, 1897; Von Helmholtz, 1925). Specialized cortical-basal ganglia neural circuits that have evolved in humans and songbirds have been implicated in implementing vocal learning (Hikosaka et al., 2002; Graybiel, 1998; Doupe and Kuhl, 1999). Although information from multiple sensory modalities informs behavior, the vast majority of studies on sensorimotor learning have focused on one sensory modality at a time (i.e., the visual feedback Steph Curry receives from shooting a basketball, or the auditory feedback a person receives during speech). It is unknown whether the neural pathways for sensorimotor learning are segregated, such that specialized neural circuits that drive specific behaviors process only sensory information most relevant to that behavior, or if these pathways instead integrate sensory information from a variety of sources. We will study these questions by analyzing specialized neural circuits for vocal learning in songbirds. The primary scientific question addressed in this dissertation is: *Do vocal learning circuits, which have evolved in humans and songbirds to underlie the production of primarily acoustic behaviors, solely process auditory feedback to guide motor learning, or do they have the ability to integrate sensory information from other sources?*

In this introduction, I will first detail the behavioral studies on sensorimotor learning (Section 1.1.1) and then describe literature on how the brain implements sensorimotor learning (Section 1.1.2 and Section 1.1.3), all with a primary focus on mammalian systems. I will then shift the focus to songbirds (Section 1.2), which are a model system for studying the neural circuit underpinnings of sensorimotor learning and are the focus of all of the research performed during my time in graduate school. I will provide an overview on how songbirds use sensory feedback to modify behavior (Section 1.2.1 and 1.2.2), then I will describe how songbird brain circuitry underlies this sensorimotor learning process (Section 1.3).

1.1.1 Sensorimotor learning in mammalian systems

Due to its importance in the acquisition of skilled behaviors, sensorimotor learning has been the focus of neuroscience research for a long period of time. Over a century ago, Hermann Von Helmholtz and others observed that people were capable of adapting their motor output (e.g. pointing to a target) to compensate for distortions of their visual feedback produced by wearing prism goggles (Kahl, 1878; Stratton, 1897; Von Helmholtz, 1925). In early studies, subjects were asked to point to a target. If the prism goggles shifted a subject's visual field to the right, subjects initially pointed too far to the right immediately after putting on the goggles. They then gradually adjusted to the goggles and could successfully point to the target, indicating that the brain relies on sensory feedback to update future motor commands. Upon removal of the goggles, subjects pointed too far in the opposite direction of the visual shift (in this case, to the left). They then adapted to these errors and were able to once again successfully point to the target. The observation that a subject's motor output immediately following the removal of a sensory perturbation is different from their motor output pre-perturbation is known as the after-effect (Von Helmholtz, 1925; Gibson, 1933; Fernandez-Ruiz and Diaz, 1999). After-effects suggest that the learned motor adaptation induced by sensory perturbation remains stored in the brain and biases motor output in the compensatory direction even after the removal of the prism goggles (Redding et al., 2005; Fernandez-Ruiz and Diaz, 1999). Over time, the study of motor adaptation to prism goggles has extended to more complex actions that require coordinated movement, such as aiming and throwing a ball at a visual target (Martin et al., 1996). Other sets of experiments showed that subjects compensated their arm reaching movements in response to perturbed visual feedback provided through a computer monitor (Krakauer et al., 2000; Krakauer, 2009). This work demonstrated that the brain can compensate for visual perturbations other that simply prism goggle-induced shifts in visual fields. This compensatory motor adjust-

ment in response to changes in sensory feedback is a form of sensorimotor learning called sensorimotor adaptation. Together, these early studies showed that brains process visual information about errors produced by a wide variety of complex behaviors and alter motor output to adjust to those errors.

Sensorimotor adaptation also occurs in response to perturbations of sensory modalities other than vision. For example, human subjects were asked to make arm movements towards a target point while in darkness (to diminish visual feedback) (Lackner and Dizio, 1994). The environment was then rotated in order to produce Coriolis forces that deterred the subjects' reaches in one direction. The subjects learned to compensate for this inertial force against their arms to successfully reach towards the target, even without visual feedback. This demonstrates that proprioceptive feedback from motor commands to the arm is sufficient to drive adaptive compensation of movement in response to non-visual sensory perturbations.

Non-visual sensorimotor adaptation was probed in finer detail through the use of end-point forces applied directly to the hand to perturb reaches. In these experiments, force fields (experimentally-imposed mechanical forces upon the subjects arms) were applied as the subjects were performing arm reaches to a target (Thoroughman and Shadmehr, 2000; Shadmehr and Mussa-Ivaldi, 1994). Specifically, human subjects were instructed to perform similar reaching movements by holding and moving a robot manipulandum from a starting location to a particular end target by moving their arm (Thoroughman and Shadmehr, 2000; Shadmehr and Mussa-Ivaldi, 1994). A mechanical force field was applied to the manipulandum as the subjects were performing the reaches in order to impose errors on the outcomes of the reaches. Immediately after the application of the force field, subjects arm movements were altered from the typical trajectories they would follow when freely moving the manipulandum. Gradually, the subjects learned to compensate for the force field and once again produce similar movement trajectories to the trajectories their arms followed prior to applica-

tion of the force field. Similar to other sensorimotor learning paradigms, this training also results in a negative after-effect when the force field is suddenly removed after subjects had learned to compensate for it. The negative after-effect produced in this particular experimental paradigm shows that this effect has been consistently observed across many different experiments (Diedrichsen et al., 2010; Thoroughman and Shadmehr, 2000; Shadmehr and Mussa-Ivaldi, 1994). Further, these studies clearly established that the brain must process information from multiple sensory modalities to adaptively change motor output.

In order to study how feedback from multiple sensory modalities interact to guide adaptations in motor output, another set of studies had subjects perform arm reaching movements while in a virtual reality environment, where visual feedback was experimentally altered (Sober and Sabes, 2003, 2005). Subjects therefore received dissociated visual and proprioceptive feedback about their current arm position. Computational modeling of the arm movements in response to the shifted visual feedback revealed that the brain relies more heavily on one sensory modality than the other at different points during the planning of the reaching movement, demonstrating that the brain performs complex computations to adjust motor output in response to sensory signals from multiple sensory modalities.

These studies of sensorimotor adaptation have been performed across multiple species, even into non-mammalian systems, allowing researchers to assess principles of sensorimotor learning that may be relevant across species. For example, rodents were trained to grasp a joystick with its paws and manipulate the joystick to move it onto a virtual target, similar to the previous studies where human subjects grasped and pulled a robot manipulandum (Mathis et al., 2017). These mice learned to compensate their reaching movements after a lateral force field was imposed on the joystick. The mice also displayed negative after-effects similar to those described by Von Helmholtz decades earlier: when the force field was removed after the mice

learned to compensate for it, the mice then displayed significantly altered movement trajectories in the same direction as the force field. Also, even some non-mammalian species, such as songbirds, are capable of adapting their motor output in response to experimentally-imposed sensory errors (Sober and Brainard, 2009; Saravanan et al., 2019). Adult songbirds adaptively changed their vocal output in response to perturbed auditory feedback, and also showed an after-effect immediately after the auditory perturbation was stopped. The similarities between sensorimotor adaptation in humans, primates, rodents, and songbirds, suggest that sensorimotor adaptation serves a very important role in shaping behavior across species.

In addition to perturbations that induce sensory error signals, explicit reinforcing (rewarding or aversive) cues can also shape behavior. A reinforcement cue is an external signal that explicitly indicates either a good or bad outcome from a behavior, which then increases or decreases, respectively, the likelihood an animal will perform that same behavior in the future (Schultz, 1998; Thorndike, 1970; Wolpert et al., 2011; Schultz, 2017). This is subtly different than the sensory error signals used to induce motor learning that were described previously (Sober and Brainard, 2009; Shadmehr and Mussa-Ivaldi, 1994; Krakauer and Mazzoni, 2011; Von Helmholtz, 1925; Wolpert et al., 2011), which inform the subject not only of the existence of an error, but also of the direction of the error (i.e., an arm reach missed a target to the left, or an element of song is a higher pitch than usual). Reinforcement cues, such as a juice reward or an electric shock, only inform the subject of the positive or negative outcome of a motor output. These cues can drive an animal to modify its behavior to maximize rewarded outcomes and minimize aversive outcomes (Schultz, 1998; Thorndike, 1970). This process can also shape motor learning, as reinforcement cues can drive approach behaviors, increase the speed of a movement, drive an animal to spend time near environmental structures associated with rewards, or shape the learning of complex motor skills (Schultz, 1998; Bindra, 1968; Izawa and Shadmehr, 2011; Huang et al.,

2011).

1.1.2 Neural circuits for sensorimotor learning in mammals

Increasing evidence has linked cortical-basal ganglia circuits in particular with sensorimotor learning. Cortical-basal ganglia circuits are brain pathways found across species that involve connections between cortical regions and subcortical regions within the basal ganglia. These circuits have been implicated in a variety of aspects of learning and behavior. Here, I will describe the neural underpinnings of sensorimotor learning in more detail, with a particular focus on cortical-basal ganglia circuits.

Clinical studies of cortical-basal ganglia circuits and motor learning

Neurological disorders often provide insight into how specific brain regions and brain circuits, when functioning improperly, produce deficits in motor learning. For example, Huntington's Disease (HD) results in widespread neurological damage, but there are especially severe abnormalities in cortical-basal ganglia neural circuits (Vonsattel et al., 1985; Dayalu and Albin, 2015). The symptoms most commonly associated with HD surround motor performance: people with HD often have difficulty performing coordinated movements, they have difficulty walking and speaking, and they can have hyperkinesia (Vonsattel et al., 1985; Dayalu and Albin, 2015). However, motor learning is also impaired by HD. For example, people with HD have difficulty in performing sensorimotor adjustments during arm reaching tasks (Smith and Shadmehr, 2005).

Parkinson's disease (PD) also provides insights into how cortical-basal ganglia neural circuits underlie motor control and sensorimotor learning. Again, PD is most commonly associated with symptoms related to motor control, such as tremors, rigidity, bradykinesia (slowness of movements), postural instability, and difficulty with coordination (Crossman, 1987; Lotharius and Brundin, 2002; Barbeau, 1962; Graybiel,

2000; Jankovic, 2008). However, PD causes deficits not only in motor control, but also in motor learning. For example, people with PD demonstrate a significantly impaired ability to adapt to visuomotor transformations (experimentally altered visual feedback) when making reaching movements (Paquet et al., 2008). The same individuals showed improved sensorimotor adaptation abilities in these experiments after being treated with L-DOPA. Also, people with PD display impaired abilities to successfully adapt when a force field is applied while performing arm reaches (Mongeon et al., 2013). People with PD also demonstrate impaired sensorimotor adaptation during experiments in which prism goggles alter visual feedback (Stern et al., 1988). Of particular interest, another symptom of PD disease is impaired speech production (Ho et al., 1998; Lieberman et al., 1992; Benke et al., 2000; Canter, 1963; Scott and Caird, 1983). This is such a robust symptom that speech signal processing algorithms are capable of predicting PD symptom severity and discriminate PD subjects from healthy controls, which could potentially serve as a diagnostic tool (Tsanas et al., 2012). People with PD often have a number of speech impairments, including reduced vocal amplitude and reduced respiratory control, that require speech therapy to improve (Duffy, 2013; Liotti and Ramig, 2003; Mollaei et al., 2013; Abur et al., 2018; Diamond et al., 1987). In one set of experiments, people with PD produced vocalizations while they received perturbed auditory feedback (the pitch of their vocalizations was shifted) (Liu et al., 2012). People with PD displayed no vocal compensation, as well as increased vocal variability, whereas people who did not have PD were able to adaptively shift the pitch of their vocalizations to compensate for the perceived auditory errors. These researchers suggested that the speech symptoms caused by PD may be due to impaired sensorimotor learning capabilities more generally, thereby suggesting a link between sensorimotor learning and speech (which will be discussed in detail in Section 1.1.3).

One of the defining characteristics of PD is the loss of dopaminergic cells within

the Substantia Nigra pars compacta (SNc), which provides dopaminergic input to the basal ganglia (Lotharius and Brundin, 2002; Barbeau, 1962, 1969; Bernheimer et al., 1973; Albin et al., 1989; Jankovic, 2008). While no known medical intervention is capable of fully stopping disease progression, the most successful treatments to date at ameliorating the movement deficits caused by PD include the administration of L-DOPA (L-3,4-dihydroxyphenylalanine) and deep brain stimulation (Barbeau, 1969; Jankovic, 2008; Cotzias et al., 1969; Jenner, 2008; Jankovic and Stacy, 2007; Carlsson and Carlsson, 1990). L-DOPA is a naturally occurring compound that is a precursor to dopamine. Dopamine cannot cross the blood brain barrier and therefore cannot be provided as a treatment for people suffering from disorders that result in a depletion of dopamine neurons (Hardebo and Owman, 1980). However, L-DOPA can cross the blood brain barrier, so when it is provided as medical treatment, it can enter the brain, get converted into dopamine, and increase dopamine concentrations in the brain (Cotzias et al., 1969; Jenner, 2008). The result of L-DOPA treatment is often a significant initial improvement in the motor symptoms associated with PD, yet these changes are not long-lasting (Barbeau, 1969; Jankovic, 2008; Cotzias et al., 1969; Jenner, 2008; Carlsson and Carlsson, 1990). Deep brain stimulation (DBS) is another one of the most effective treatments for Parkinson's disease. DBS entails the surgical implantation of stimulating electrodes in specific brain regions, such as thalamus, globus pallidus internal (GPi), subthalamic nucleus (STN) (all of which are a part of thalamocortical-basal ganglia circuitry) (Deuschl et al., 2006; Bronstein et al., 2011; Rodriguez-Oroz et al., 2005; Jankovic, 2008; Graybiel, 2000). Treatment consists of persistent electrical stimulation through the implanted electrodes to alter neural activity in these brain regions (Deuschl et al., 2006; Bronstein et al., 2011; Rodriguez-Oroz et al., 2005). This treatment results in improvement in motor symptoms associated with PD (Deuschl et al., 2006; Bronstein et al., 2011; Rodriguez-Oroz et al., 2005). Because damage to cortical-basal ganglia circuits produces movement

symptoms in PD, and multiple therapeutic interventions that target cortical-basal ganglia circuits help to ameliorate movement symptoms, cortical-basal ganglia neural circuits have been linked to motor control and motor learning.

Studies of cortical-basal ganglia circuits and motor learning in animal models

A variety of experimental approaches in animal model systems have also revealed neural circuit mechanisms for motor learning. Anatomical studies have shown that the basal ganglia send a large amount of projections to the motor and prefrontal areas of cortex (Graybiel, 2000). Also, a number of "loops" are formed: different regions of cortex project to subcortical structures, like the basal ganglia and thalamus (Hikosaka et al., 1999, 2002; Graybiel, 2000; Hoover and Strick, 1999; Parent and Hazrati, 1995). These subcortical structures then project back to numerous different regions of the cortex. Therefore, neuroscientists have speculated that the function of these cortical-basal ganglia loops could be to support a variety of different elements of behavior, such as motor planning, motor control, movement sequencing, and learning (Graybiel, 2000; Doya, 2000). For example, some researchers have hypothesized that sensory information in the cortex is integrated with reward value and/or match to the internal sensorimotor template in the basal ganglia, then fed back to the cortex to influence motor output, thereby driving motor learning (Hikosaka et al., 2002).

Lesions or blockade of activity within specific brain regions have demonstrated the necessity of cortical-basal ganglia circuitry for motor learning. For example, chemical blockade of the preSMA resulted in impaired abilities to learn new button press sequences (Nakamura et al., 1999). Also, chemical injection of 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) into the nigrostriatal neural pathways caused significant dopamine depletion in the basal ganglia of monkeys (Matsumoto et al., 1999). This manipulation of basal ganglia circuitry led to impaired ability to

learn a sequential button-press task, as well as impaired ability to perform skillful arm reaching maneuvers. Additionally, reversible blockade of the striatum causes deficits in sequence learning (Miyachi et al., 1997). One study assessed the necessity of motor cortex in motor learning by training rats to perform a learned, skilled behavior (lever presses with a particular timing) (Kawai et al., 2015). Lesions of motor cortex did not impair the ability of the rats to perform this skilled behavior if they had already learned to perform the behavior successfully. Lesions of motor cortex performed prior to learning the behavioral task did impair the ability to learn to perform the task successfully.

Electrophysiological studies in behaving animals have revealed that the activity of neurons within cortical-basal ganglia circuits correlates with aspects of motor learning. For example, neural recordings of neurons in the supplementary motor area (a motor cortical region) during behavior have shown that the activity levels of these neurons significantly increase when animals are performing a learned, skilled movement (Rosenbaum et al., 1983; Willingham, 1998; Shima and Tanji, 2000). To investigate this phenomenon further, monkey subjects were trained (through trial and error) to perform a series of button presses in the correct order (Hikosaka et al., 1995). Neurons in the pre-supplemental motor area (preSMA) were strongly activated when the animals were learning new button press sequences, but not when the animals were performing an already learned sequences of presses. Changes in neural activity patterns occur in mammalian sensorimotor cortex during the learning of a skilled behavior (Monfils and Teskey, 2004; Petersen et al., 1998). Also, chronic recordings from individual neurons within the striatum were performed in rats as they learned to move through a maze enclosure (Jog et al., 1999). The activity of populations of striatal neurons changed dramatically during the initial learning of this task, then remained largely stable during weeks of subsequent performance of the already learned task. Recordings of a subpopulation of striatal interneurons demonstrated that this class of

neurons develop phasic, coordinated responses following motor learning, which may influence the activity of projection neurons to guide behavior (Graybiel, 2000). Also, recordings of both motor cortical and striatal neurons during motor skill learning revealed that there was significant recruitment of task-responsive neurons over the course of skill learning (Costa et al., 2004).

Motor learning is linked with cellular and molecular changes in the brain associated with plasticity and learning. Analysis of synapses in M1 have shown that structural, synaptic changes occur in M1 during motor learning (Klintsova and Greenough, 1999). For instance, motor learning induces new dendritic spine formation as well as dendritic spine growth in a variety of subpopulations of motor cortical neurons (Xu et al., 2009; Yang et al., 2009). Also, the number of synapses per neuron increases in rat motor cortex in response to motor skill learning (Kleim et al., 1996). Similar synaptic modifications indicative of long-term potentiation (a well-studied neural plasticity mechanism strongly associated with learning (Malenka and Nicoll, 1999; Whitlock et al., 2006)) occur in primary motor cortex following motor learning (Kida and Mitsushima, 2018; Avanzino et al., 2015; Rioult-Pedotti et al., 2000). Also, chemical injections into motor cortex that reduce the rate of protein synthesis in neurons produce deficits in motor skill learning, indicating that intact protein synthesis in motor cortex is required for the acquisition of skilled movements (Luft et al., 2004)

These studies (along with numerous other studies not mentioned here) have used a variety of experimental approaches across multiple species to show that cortical-basal ganglia circuits are important for motor learning. The details of how these cortical-basal ganglia loops underlie specific aspects of behavior, such as sequence learning, motor control, sensorimotor adaptation, or learning in response to different training paradigms is still the subject of active research.

Dopaminergic neural circuits and motor learning

Dopaminergic neural pathways play a particularly important role in shaping animal behavior. This has been most heavily studied in the context of operant conditioning and reinforcement learning: when animals change their behavior in response to rewarding or aversive feedback in order to maximize good outcomes and minimize bad outcomes in the future. These processes are relatively well-understood cases of how sensory experiences can shape behavior. To measure how neural activity patterns correlate with aspects of reinforcement learning, electrophysiological recordings of VTA dopamine neurons were performed to study neural activity patterns while a monkey underwent an operant conditioning paradigm (Schultz, 1998; Schultz et al., 1992). Monkeys were presented with a sensory cue (a tone), then an unexpected reward (juice), all while VTA dopamine neurons were being recorded. The firing rates of VTA dopamine neurons phasically increased immediately following this unexpected juice reward (Fig. 1.1, top). After performing this operant training repeatedly, the monkeys learned to expect juice reward following the predictive cue. At this time point, the firing rates of VTA dopamine neurons phasically increased in response to the sensory cue and did not change during the presentation of the (now expected) juice reward (Fig. 1.1, middle). After training, when the monkeys had learned to associate the predictive sensory cue with reward, the researchers began to omit the juice reward after presenting the monkey with the cue. The firing rates of VTA dopamine neurons were phasically suppressed in response to this omission of expected reward at the exact point in time when the animals learned to expect the reward (Fig. 1.1, bottom).

Careful analysis and computational approaches suggested that these response patterns of VTA dopamine neurons exhibited qualities consistent with the encoding of reward prediction error (RPE). RPE is the difference between the reward an animal expects to receive and the reward it actually receives (Schultz et al., 1997; Schultz and

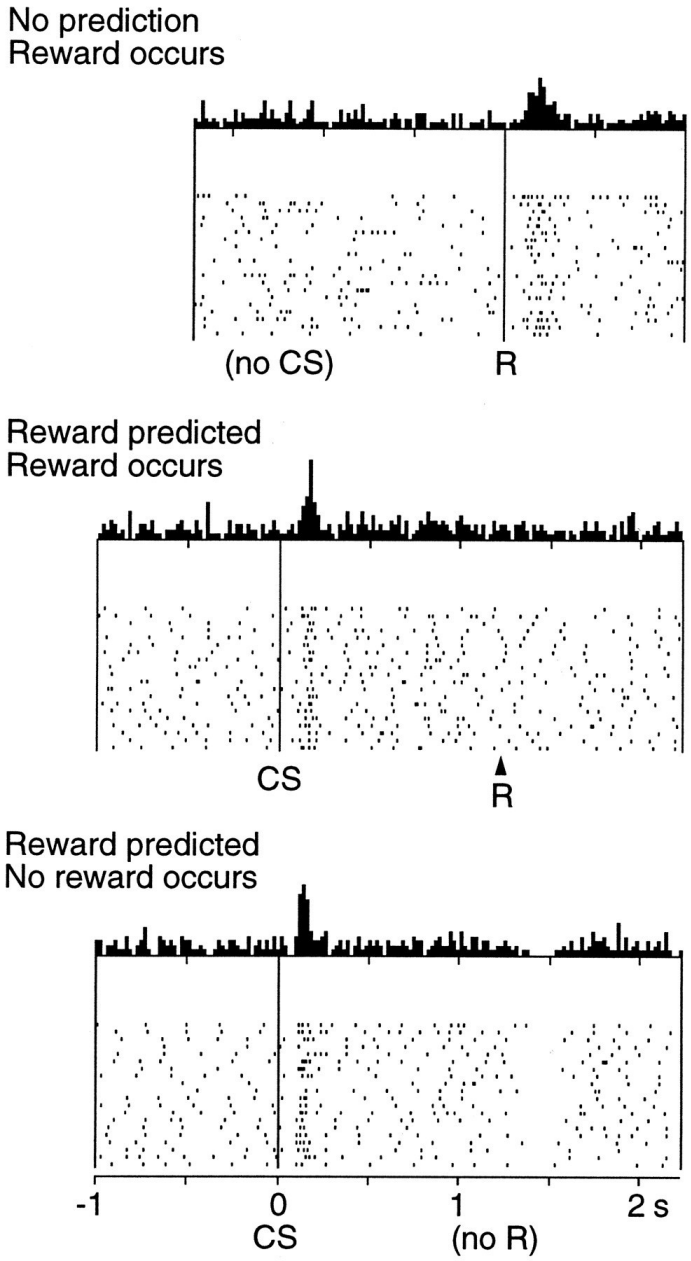


Figure 1.1: Electrophysiological recordings of VTA dopamine neurons. Top: a monkey is not yet trained to associated a conditioned stimulus (CS) with a reward (R). Example dopamine neuron firing rate increases immediately following receiving an unexpected reward (positive prediction error). Middle: After training, a monkey has learned to associate the CS with reward. Example dopamine neuron firing rate now increases following the CS but does not change in response to receiving an expected reward (neutral prediction error). Bottom: After training, a monkey associates a CS with an R, but this time the R is omitted. The example dopamine neuron is phasically suppressed at the time point when the reward was expected but not received (negative prediction error). From Schultz, W. (1998). Predictive reward signal of dopamine neurons. Journal of Neurophysiology, 80(1), 1-27. Reprinted with permission from American Journal of Physiology.

Dickinson, 2000; Schultz, 2017). An RPE-encoding signal should display increased firing in response to unexpected reward, where the difference between expected reward and received reward is large. An RPE-encoding signal should also display no change in firing in response to an expected reward, because the expected reward and actual reward are equal, so the difference would be zero. Finally, this signal should display a decrease in firing in response to the omission of an expected reward, when the actual reward is less than the expected reward. The actual response patterns of VTA dopamine neurons were entirely consistent with these parameters (Schultz, 1998; Schultz et al., 1997).

Future studies that altered the probability and magnitude of rewards further supported the hypothesis that VTA dopamine neurons encode RPE. For example, researchers hypothesized that an RPE-encoding signal should be modulated based on the magnitude of the reward, since larger rewards would produce larger differences between actual and expected reward. Accordingly, the size of the phasic increase in firing rates of VTA dopamine neurons correlates significantly with the magnitude of the reward provided (the mL of juice given) (Fiorillo et al., 2003). Also, when the probability of providing a reward was experimentally altered, the magnitude of the VTA dopamine response scaled in relation the "surprisingness" of the reward (Tobler et al., 2005). That is, when the animal was trained in a contingency with a lower probability of receiving reward (the animal had a low expectation of reward), the phasic increase in firing of their VTA dopamine neurons was very large (presumably because the difference between received reward and expected reward was quite large). On the other hand, when the probability of receiving reward was very high (the animal had a higher expectation of reward), the phasic increase in VTA dopamine neuron firing was smaller.

More recently, optogenetic approaches have allowed for the testing of the causal role VTA dopamine firing plays in reward learning. Optogenetics involves viral and

genetic techniques that cause neurons to produce special, light-sensitive ion channel proteins (Boyden et al., 2005; Deisseroth, 2011; Yizhar et al., 2011). These ion channels are expressed in the cell membrane and, in response to light, they allow ions to pass through, which produces an action potential. This gives experimenters the ability to control the activity of genetically-defined subpopulations of neurons in behaving animals (Zhang et al., 2007). Using optogenetics to control the neural activity of VTA dopamine neurons revealed that experimentally stimulating VTA dopamine neurons in an RPE-like pattern was sufficient to drive reward learning (Steinberg et al., 2013; Kravitz et al., 2010; Tsai et al., 2009; Lammel et al., 2012).

While research on the role dopamine plays in learning has focused most heavily on operant conditioning, as described above, dopamine does play an important role in other forms of motor learning as well. For example, lesions of all of VTA completely abolishes motor skill learning, which requires sensory feedback to guide adaptations in motor output but does not involve operant conditioning (Molina-Luna et al., 2009). More precise lesions of VTA dopamine neurons that send projections to primary motor cortex (M1) in rats significantly impairs the acquisition of skilled forelimb reaching movements but does not impact the performance of this motor skill once it if the lesion is perform after the skill was learned (Hosp et al., 2011). Administration of L-DOPA into rat M1 after these VTA lesions recovers motor learning abilities. Also, injections of chemical antagonists of D1-receptors and injections antagonists of D2-receptors each block motor skill learning, indicating that motor skill learning is dependent on dopaminergic signaling through both of these systems. The tonic firing patterns of dopaminergic projections to motor cortex have been hypothesized to be consistent with the support of LTP-like learning mechanisms that support the acquisition of motor skills (Hosp and Luft, 2013).

Additional neural pathways for sensorimotor learning

Neural pathways outside of cortical-basal ganglia circuits also play an important role in sensorimotor learning. For example, the role of somatosensory cortex in sensorimotor adaptation was tested using optogenetic inhibition of somatosensory cortex in mice as they were learning to adapt to a mechanical force field (Mathis et al., 2017). Inhibiting neural activity in somatosensory cortex significantly impaired the ability of the mice to adapt to the force field perturbation, suggesting that somatosensory cortex plays a causal role in the expression of this sensorimotor learning process. Somatosensory cortex receives feedback from the limbs via thalamocortical neural pathways, implicating these brain structures in this form of sensorimotor learning (Bostan and Strick, 2018; Caligiore et al., 2017). Also, people with atrophy or damage to parts of the cerebellum are impaired in their ability to adapt their throws while wearing prism goggles (Wolpert et al., 2011; Martin et al., 1996; Baizer et al., 1999; Diedrichsen et al., 2005; Morton and Bastian, 2006; Tseng et al., 2007). Cerebellum and the neural pathways between cerebellum and cortex have been implicated in sensorimotor learning in other studies, as well (Machado et al., 2015; Albergaria et al., 2018; Becker and Person, 2019; Tseng et al., 2007). The role the cerebellum plays in driving sensorimotor learning is still an exciting area of active research across multiple species and learning paradigms (Nicholson et al., 2018; Machado et al., 2015; Albergaria et al., 2018; Becker and Person, 2019). While the these neural circuits are not the primary focus of this dissertation, these studies nonetheless demonstrate the general principle that neural pathways between cortical and subcortical structures are crucial for processing sensory feedback to adjust motor commands (Martin et al., 1996; Krakauer and Mazzoni, 2011; Shadmehr et al., 2010; Wolpert et al., 2011; Becker and Person, 2019).

1.1.3 Human speech: behavior and neural circuits

One of the most important skilled behaviors that humans produce is speech. Specialized neural circuits have evolved in humans (and in songbirds) that underlie vocal learning (Doupe and Kuhl, 1999; Jarvis, 2019; Hickok and Poeppel, 2007). Here, I will first describe behavioral elements of human vocal learning, then I will discuss the brain circuitry that underlies this phenomenon.

The role of auditory feedback in human speech

Human babies produce babbling sounds, which are often syllables repeated or strung together. Gradually, human babies develop these babbling sounds into vocal sounds that resemble those of adults. After this process of speech acquisition, humans can produce a repertoire of patterned, meaningful vocal sounds throughout the entire duration of their adult lives.

A large body of literature has shown that auditory feedback (of one's own vocalizations as well as the vocalizations of others) is crucially important for human vocal learning (Doupe and Kuhl, 1999). For example, children acquire the same language and dialects as those produced by their parents. Human babies who are born deaf never acquire the ability to produce spoken language (Petitto, 1993). Deaf infants produce babbling noises with abnormal structure and a limited range of sounds compared to infants with intact hearing (Oller and Eilers, 1988; Stoel-Gammon and Otomo, 1986). Children in the process of learning speech who become deaf suffer from severe impairments in speech production (Cowie and Douglas-Cowie, 2011). Even children who become deaf later in childhood will produce impaired speech (Plant and Hammarberg, 1983).

Auditory feedback is not only important for speech acquisition, but also for the maintenance of proper speech function throughout adulthood. Humans who become deaf in adulthood (after fully acquiring speech) display alterations in the structure

of their speech (Cowie and Douglas-Cowie, 2011). However, these effects are far less severe than when deafness occurs early in life, and the later in life an adult becomes deaf, the less the severity of the deterioration in speech production that occurs. Also, studies performed in human adults where experimenters manipulated the subjects' auditory feedback showed that delayed auditory feedback of an adult's own vocalizations causes disruptions to ongoing vocalizations (vocal slowing, pauses, repetitions) (Howell and Archer, 1984; Lee, 1950). Even more subtle manipulations of auditory feedback, such as altering the feedback of individual syllables, result in impaired speech production in adult humans (Houde and Jordan, 1998).

The role of non-auditory sensory feedback for human speech

Non-auditory feedback is also important for human vocal learning, yet it has received far less research attention compared to auditory feedback. Social interaction (which involves auditory and non-auditory sensory feedback) plays an important role in speech acquisition. During social interaction between adults and babies, adults will modify their typical speech patterns as well as their other behaviors, which appear to help support the speech acquisition process of the baby (Locke and Snow, 1997; Kuhl, 2007). Studies were performed where the responsiveness of parents to their infants' vocalizations was experimentally controlled (Goldstein et al., 2003). Specifically, one group of parents would smile, touch, and interact with their infants like normal when the infants vocalized. In another group, parents performed the same reactions, but experimenters triggered the timing of these reactions so they did not align with infant vocalization. The infants who received social experiences immediately contingent on their vocalizations produced significantly more vocalizations, and their vocalizations were more similar to adult speech. Importantly, the social interactions provided by the parents in this study were non-auditory. In another study, some infants were exposed to auditory feedback of a different language from adults only through a tele-

vision, while other infants were able to interact acoustically and socially (in-person) with native speakers of the new language (Kuhl et al., 2003). The infants who received in person, social feedback produced more accurate, adult-like vocalizations.

Another classic case of non-auditory feedback influencing human vocalizations is the McGurk effect. Humans watch an audio-visual illusion, where their perceived auditory experience from hearing another speaker is of one particular syllable but their visual experience from watching a video of the speaker talking is of a different syllable (McGurk and MacDonald, 1976). People will report perceiving incorrect syllables, often a mixture of the two (auditory and visual). Although this dissertation will focus more on the learning and production of speech and less on the perception of speech, the McGurk effect nonetheless demonstrates that human brains process information from multiple sensory modalities and use this information for perception relevant for speech.

Somatosensory feedback plays a key role in guiding human vocalizations as well. The brain receives proprioceptive feedback from vocal muscles and mechanoreceptors in the skin and vocal tissue. The fact that speech does not complete degrade in deaf adults indicates that non-auditory feedback can be used to maintain vocal output (Tremblay et al., 2003). Careful experimental studies of the role of somatosensory feedback for vocal learning have been performed. For example, a robot was used to create a somatosensory perturbation by producing a mechanical force against the jaw of human subjects while they were speaking (Tremblay et al., 2003; Nasir and Ostry, 2008). This perturbation is unique in that it interferes with somatosensory feedback from vocalizations without affecting speech or any other source of auditory feedback. Subjects gradually learned to adapt to this perturbation and correct for the mechanical force against their jaw. This shows that the brain processes somatosensory feedback related to the positions of body parts that produce speech, and it adaptively changes motor output, even without auditory feedback from vocalizations. People

who were deaf (and therefore clearly received no auditory feedback) were also capable of learning to adapt to this perturbation, suggesting that somatosensory feedback alone is capable of driving sensorimotor adaptation. To determine how feedback from multiple sensory modalities interact to shape behavior, experiments were performed that placed somatosensory feedback and auditory feedback at odds with one another (Lametti et al., 2012). Researchers performed the same robotic manipulation of the jaw position during speech in human subjects. Simultaneously, these human subjects were provided altered auditory feedback, because the frequency of particular elements of speech were experimentally altered and played it back through headphones the subjects were wearing. These experiments revealed that auditory feedback was not dominant in guiding behavior. Instead, some subjects favored somatosensory feedback and compensated further in the adaptive direction of the somatosensory perturbation than in the adaptive direction of the auditory perturbation.

Neural circuits for human speech

Both human and songbird brains have evolved specialized neural pathways for the learning and control of vocal output (Jarvis, 2019) (Fig. 1.2). Both species have neural mechanisms for the fine control of the larynx (or syrinx in songbirds) as well as vocal articulators or effectors. Vocal effectors are the parts of the body involved in producing vocalization (mouth, jaw, tongue in humans, beak in songbirds) (Bouchard et al., 2013; Riede and Goller, 2010).

In humans, early work led by Paul Broca and Karl Wernicke sought to determine the brain regions necessary for speech. This work found that neurological damage to specific parts of the brain led to specific speech impairments (Konopka and Roberts, 2016; Dronkers et al., 2007). Damage to a portion of the inferior frontal cortex (now known as Broca's area) led to deficits in the ability to produce speech (but no deficits in speech perception) (Dronkers et al., 2007; Amunts et al., 1999). Damage

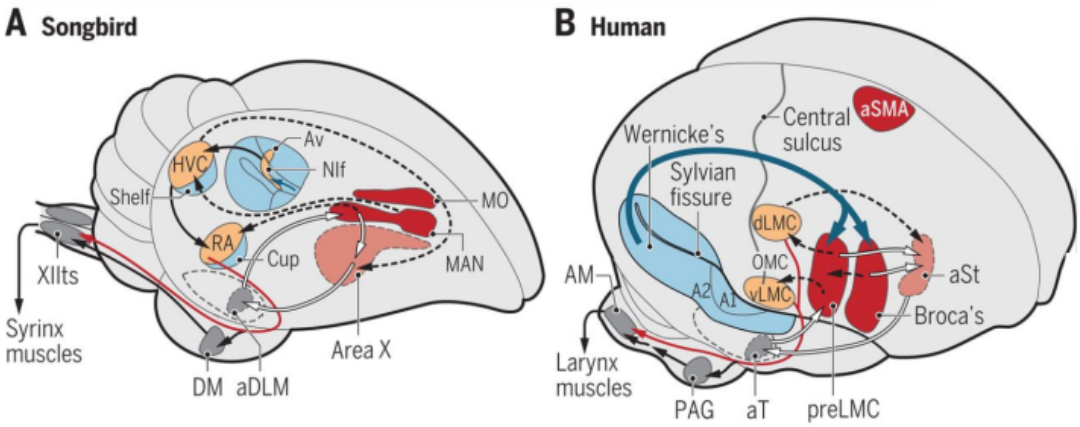


Figure 1.2: Neural circuits for vocal learning across species. A. Songbird vocal learning circuit. B. Human vocal learning circuit. Analogous brain regions between species are in the same color. Orange regions are the posterior vocal motor pathway. Red regions are the anterior vocal pathway. Dashed arrows show connections between these two pathways. Red arrows show the specialized direct projections from motor cortex to brainstem vocal motor neurons only present in vocal learners. Blue regions are auditory regions. Blue arrows are pathways for auditory input to enter the specialized vocal learning pathway. A1, primary auditory cortex; A2, secondary auditory cortex; aDLM, anterior dorsolateral medial nucleus of the thalamus; Ai, intermediate arcopallium; Am, nucleus ambiguus; aSMA, anterior supplementary motor area; aSt, anterior striatum speech area; aT, anterior thalamus speech area; Av, avalanche; CMM, caudal medial mesopallium; CSt, caudal striatum; DM, dorsal medial midbrain nucleus; HVC, a letter-based name; L2, Field L2; dLMC, dorsal laryngeal motor cortex; vLMC, ventral laryngeal motor cortex; preLMC, premotor laryngeal motor cortex; OMC, oral motor cortex; MAN, magnocellular nucleus of the nidopallium; MO, mesopallium oval nucleus; NCM, nidopallium, caudal medial part; NIf, nidopallium interfacial nucleus; NLC, nidopallium, lateral caudal; PAG, periaqueductal gray; RA, robust nucleus of the arcopallium; XIIts, 12th vocal motor nucleus, tracheosyringeal part. From Jarvis, E. D. (2019). Evolution of vocal learning and spoken language. Science, 366(6461), 50-54. Reprinted with permission from AAAS.

to a portion of the superior temporal gyrus caused deficits in speech perception (but not vocal motor production) (Blank et al., 2002). Since then, further studies have advanced our understanding of the complex neural pathways for vocal learning in humans.

Areas throughout human cortex, thalamus, and basal ganglia are thought to be important for human speech learning. Regions of auditory cortex are important for processing auditory feedback to guide speech learning and maintenance. There are specific regions of human auditory cortex that, when measured with fMRI, respond

distinctly to subjects hearing speech (Norman-Haignere et al., 2015). These auditory cortical regions, as well as other cortical regions (including somatosensory cortex) send projections to specialized, downstream speech circuits (Petkov and Jarvis, 2012; Jarvis, 2004). For example, connections exist between auditory cortical regions important for speech, such as secondary auditory cortex (A2) and Wernicke's area, and motor cortical regions associated with speech, such as Broca's area (Friederici et al., 2017). These connections have been hypothesized to play a crucial role in vocal production, speech perception, and vocal learning in humans.

Human motor cortex is also important for the motor control of vocal effectors. Motor cortical regions in humans, such as the dorsal and ventral laryngeal motor cortices (dLMC and vLMC) and the oral-facial motor cortex (OMC), are important for speech production (Jarvis, 2019). Also, premotor cortical regions, such as the LMC and Broca's area, are important for vocal learning (Jarvis, 2007, 2004; Dichter et al., 2018). Neuroscientists have hypothesized that LMC in human brains projects to brainstem vocal motor neurons, which then project to vocal musculature, allowing for the fine motor control necessary to produce complex, coordinated movements necessary for vocalization (Jarvis, 2019). This pathway is highly analogous to the songbird motor pathway (which I explain in detail in section 1.3.1), where regions of songbird pallium project to a brainstem nucleus that then projects to muscles involved in song production (Jarvis, 2007; Mooney, 2009a).

The basal ganglia also play a crucial role in speech production in humans, which is particularly relevant for comparative studies across species, as the songbird vocal learning neural circuit is a cortical-basal ganglia loop. Neurological diseases that affect the basal ganglia in particular (HD and PD) consistently produce deficits in speech production and motor control of vocal effectors (Konopka and Roberts, 2016). Interestingly, when a human mutant transgene associated with HD is expressed in songbirds, these songbirds developed dramatic deficits in vocal production, including

impaired vocal learning, stuttering, and degradation of patterned song structure (Liu et al., 2015). Also, genetic studies have revealed that the FOXP2 transcription factor gene is particularly important for human speech. Mutations of this gene produce impaired speech development and vocal abilities in humans, yet often leave other motor functions unimpaired (Fisher and Scharff, 2009). For example, FOXP2 mutations produce impaired ability to control vocal effectors in a coordinated fashion to produce proper vocalizations. FOXP2 is expressed throughout cortex, basal ganglia, and thalamus, providing further evidence that these neural pathways are important for human vocal learning (Fisher and Scharff, 2009; Ferland et al., 2003; Lai et al., 2003). FOXP2 is expressed at particularly high levels in human basal ganglia during early stages of human development - when speech acquisition is underway (Teramitsu et al., 2004). FOXP2 is also expressed in songbird brains, particularly in regions within the cortical-basal ganglia song learning pathway, and the expression levels of FOXP2 in songbirds that change their song seasonally correlates with the time of year associated with high levels of vocal plasticity (Haesler et al., 2004; Scharff and Nottebohm, 1991). Also, genetic knockdowns of FOXP2 in developing songbirds impairs vocal learning abilities (Haesler et al., 2007).

Thus far, I have described the literature that shows that both auditory and non-auditory feedback strongly shape vocal learning in humans, and that unique neural circuits have evolved in human brains to underlie vocal learning. However, how the brain processes auditory and non-auditory sensory feedback to drive vocal motor learning is not well understood and performing scientific studies on this topic is highly challenging. First of all, the ability to perform invasive techniques in human brains is highly restricted (Jarvis, 2007). For instance, neuroscientists are not able to easily lesion brain regions and observe the resulting impact on speech production or implant invasive recording devices to analyze neural activity in particular neural circuits during vocal learning. Second, there is a limited number of animal model

systems that neuroscientists can use to study the neural mechanisms of vocal learning (Doupe and Kuhl, 1999). Although many animal species are able to produce vocalizations, very few species rely upon a sensorimotor learning process to acquire this capability (where they use feedback from their own vocalizations and information from the vocalizations of others in their species to guide learning). In the next section, I will shift my focus onto one of the most successful animal models for studying the neural mechanisms of vocal learning: the songbird (Doupe and Kuhl, 1999; Jarvis, 2004; Brainard and Doupe, 2002). Though birdsong is clearly not as complex as human language, there are clear analogies between the two species in the motor control of vocal musculature, the production of patterned vocal structures, and use of sensorimotor learning to acquire vocalizations (Marler, 1970; Grieser and Kuhl, 1989; Doupe and Kuhl, 1999; Jarvis, 2004; Brainard and Doupe, 2002, 2013; Jarvis, 2019). These similarities lead us to believe that studies of songbird vocal learning will be of relevance for our understanding of human brains.

1.2 Songbirds as a model system for studying sensorimotor learning

We address how the brain processes different sensory modalities to drive motor learning by using a unique model system ideally suited for this question, the Bengalese finch (a type of songbird). Bengalese finches are well-suited for this purpose for two primary reasons. First, songbirds rely on both auditory and non-auditory sensory feedback to learn an ethological, skilled behavior (song) (Brainard and Doupe, 2002, 2000a; West and King, 1988; Mann et al., 1991; Konishi, 1965; Immelmann, 1969). In many songbird species, including Bengalese finches, song is a learned, complex behavior that juveniles acquire during their development (Fig. 1.3, A) (Brainard and Doupe, 2002; Mooney, 2009a,b; Brainard and Doupe, 2000a). During this sensori-

motor learning process, juvenile songbirds listen to the song produced by their adult tutor (typically their father) (Fig. 1.3, B), and gradually learn to copy this tutor song (Fig. 1.3, C). By the time a male zebra finch or Bengalese finch reaches adulthood, at approximately 90-100 days post hatch (d.p.h.), they typically produce a well-copied version of the tutor song that they learned, and they produce this crystallized song stably throughout the course of their adult lives (Fig. 1.3, D). This crystallized, adult finch song consists of individual syllables, which are 50-100 ms long, unique, acoustic elements of the song with stable spectral structure that are separated by silent intervals, demonstrating the complexity of this behavior (Fig. 1.3, B-E) (Leonardo and Konishi, 1999; Zann, 1996). Auditory feedback is required for this juvenile song learning process (Fig. 1.3, E). Adult songbirds naturally produce many renditions, or bouts, of their song each day, which allows for experimenters to study the relationship between rendition-to-rendition behavioral variability and learning (Brainard and Doupe, 2002). Also, birdsong is a courtship behavior produced only by male songbirds, so all experiments throughout this dissertation use male songbirds.

The second reason we use songbirds to address the central question of this dissertation is that songbirds are a uniquely effective model system for studying how neural circuits process sensory feedback to drive motor learning (Brainard and Doupe, 2000a). Very few animal species outside of humans and songbirds undergo sensorimotor learning to acquire complex vocalizations. Songbirds are a tractable model system with well-characterized neural circuitry that underlies vocal learning, allowing for the scientific investigation of the neural mechanisms underlying vocal motor learning.

Throughout this section, I will discuss in detail how songbirds use sensory feedback to guide motor output. Similar to scientific research on humans, there is extensive literature on how auditory feedback drives vocal learning in songbirds, which I will describe in Section 1.2.1. However, non-auditory sensory feedback is also crucial for guiding behavior across both humans and songbirds. Thus, I will describe the less ex-

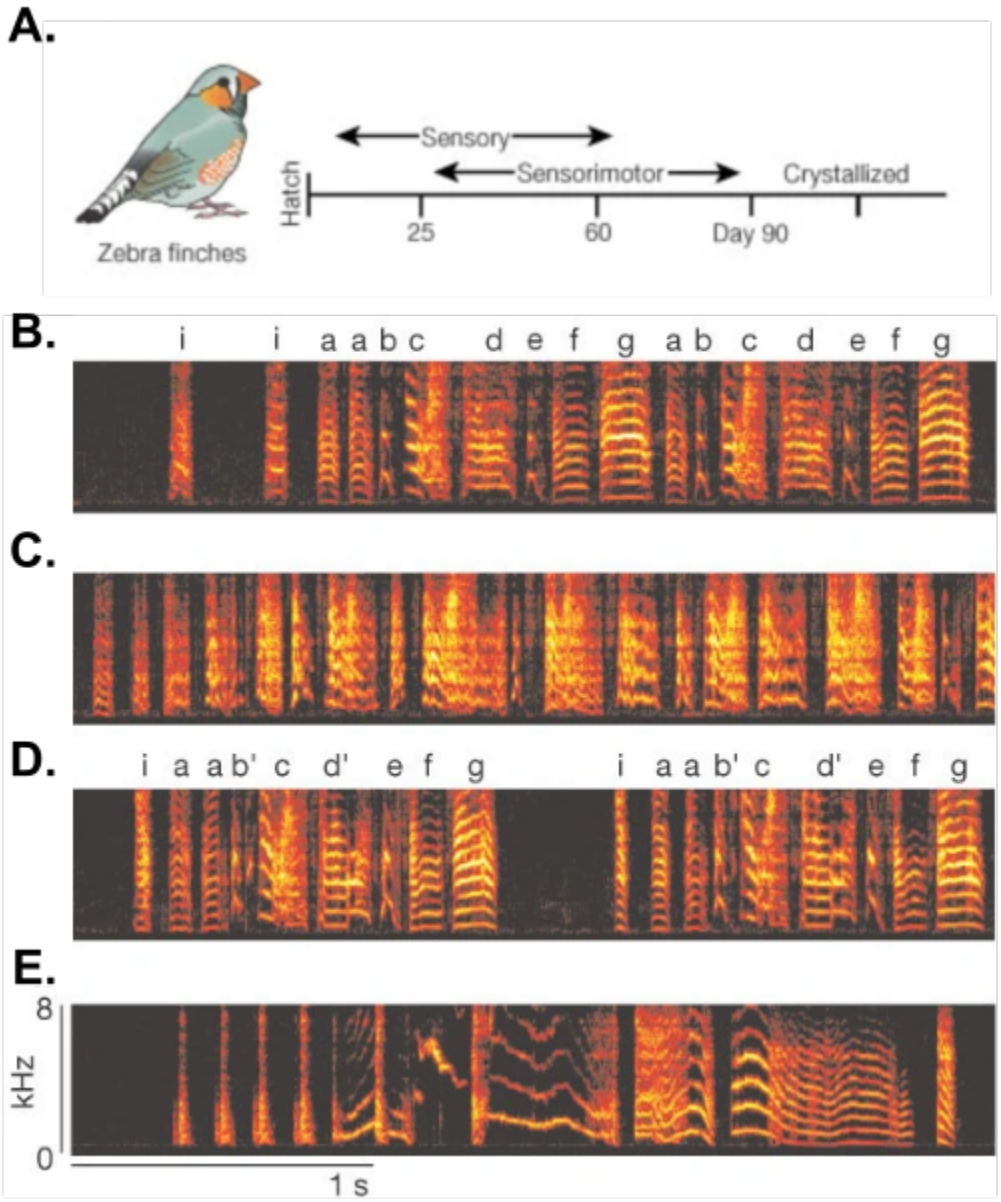


Figure 1.3: A) Schematic of the Zebra finch song learning process. During the sensorimotor phase (approximately 30-90 days post hatch), the juvenile Zebra finch learns to copy the adult tutor song they are exposed to. After 90 days post hatch, the Zebra finch produces a crystallized adult song, which is remarkably stable over their adult lives. B) An example of a song spectrogram of a rendition of a song produced by an adult male Zebra finch tutor. C) An example of a song produced by a juvenile Zebra finch, tutored by the adult bird in (A), early in the sensorimotor learning process. D) Song produced by the same juvenile Zebra finch in (C), late in the sensorimotor learning process (close to adulthood). E) Song of a different zebra finch that was raised in acoustic isolation. From Brainard, Michael S., and Allison J. Doupe. "What songbirds teach us about learning." Nature 417.6886 (2002): 351-358. Reprinted with permission from Springer Nature.

tensive, yet important, literature on how non-auditory sources of sensory information shapes songbird vocal behavior in Section 1.2.2.

1.2.1 Auditory feedback plays a particularly important role in shaping songbird vocal output

Auditory feedback is necessary for the juvenile song acquisition process

Early pioneering work in the songbird field demonstrated that auditory feedback is necessary for song learning in juvenile birds. In these studies, the cochlea was surgically removed in juvenile songbirds before the onset of song or subsong in order to deafen these birds and assess the role auditory feedback plays in the typical song development process (Konishi, 1964). The song that the deaf birds developed had numerous alterations compared to control birds: the deaf birds produced songs that tended to have fewer syllables, a wider range of frequencies, larger inter-syllable durations, and more variation in inter-syllable durations, indicating that these birds did not develop the typical temporal patterning of song seen in control birds. Also, the frequency structure of the song syllables was abnormal and the rendition-to-rendition consistency of the structure of the song syllables was impaired. In another similar study, White-crowned Sparrows were deafened both before they started singing and after they started singing but before they fully learned to copy their tutor songs (Konishi, 1965). When these songbirds were deafened before starting singing, they were completely unable to copy their tutor song. When the birds were deafened after they started singing but before the full song learning process was completed, they still learned to partially copy the tutor song, but the songs these birds produced as adults lacked typical patterning and structure seen in songbirds who were not deafened. These early studies demonstrated the necessity of auditory experience for the juvenile song acquisition process.

Impaired auditory feedback produces deficits in maintenance of song production in adult songbirds

Auditory feedback is not only crucial for juvenile song acquisition but also for song maintenance in adult songbirds. Surgical removal of the cochlea was performed in songbirds, however instead of deafening juvenile birds, adult songbirds were deafened after the song crystallization process had completed (Nordeen and Nordeen, 1992). This deafening led to gradual changes in song performance, whereas control (undeafened) birds demonstrated highly stable song output over long time periods. Specifically, the temporal patterns of the songs of deafened adult songbirds became abnormal, the average intersyllable interval was significantly increased after deafening, the occurrences of stuttering increased significantly, and the deafened birds stopped singing a large number of the unique song syllables found within their song prior to deafening. Both the harmonic structure of individual song syllables and the sequence structure of the overall song were perturbed by deafening. In another study, auditory feedback was perturbed during song performance in adult zebra finches (Leonardo and Konishi, 1999). Rather than deafening the adult birds, the researchers instead developed a computer program to monitor ongoing song performance in real-time and delivered feedback signals (via a speaker) as the songbirds were performing song. The feedback signals that were delivered were of naturally occurring sounds, but were played back so as to interfere with the typical auditory experience of the adult birds, which would be to hear the unperturbed auditory feedback of their own song production. The songs of the songbirds who received perturbed auditory feedback gradually deteriorated - the stability of the song sequence decreased significantly, the spectral organization of song changed, individual song syllables were sometimes deleted or distorted, and there was a significant increase in the occurrences of stuttering. After the removal of the artificial, interfering auditory playback, all of the temporal and spectral changes in song, including stuttering and syllable sequence instability,

gradually reversed until the adult birds once again accurately produced their typical, crystallized song observed prior to experimentation. Similar experiments were performed, but the auditory feedback interference was targeted to individual song syllables rather than to the entire song. This produced gradual alterations in typical harmonic syllable structure in perturbed birds compared to control birds. Overall, this paper demonstrated that adult zebra finches retain the ability to modify song in response to auditory feedback even after crystallization, and that this vocal plasticity affects both the temporal pattern of song and the spectral structure of individual syllables. Thus, the remarkable stability of adult song performance over time is not because the song motor program in the brain remains impervious to change after crystallization, but rather because the songbird brain constantly monitors auditory feedback and uses this feedback to maintain adequate vocal performance over the course of the songbirds' lives. Other work demonstrated similar findings in other species of songbirds by showing that adult Bengalese finches also require normal auditory feedback in order to maintain crystallized song structure throughout adulthood (Okanoya and Yamaguchi, 1997; Woolley and Rubel, 1997).

Pitch-contingent auditory cues are sufficient to drive vocal learning in adult songbirds

Adult songbirds are also capable of learning in response to pitch-contingent, auditory feedback cues. (Tumer and Brainard, 2007; Hoffmann et al., 2016; Warren et al., 2011; Andalman and Fee, 2009; Ali et al., 2013). Ongoing song performance, specifically the natural, trial-to-trial variations in the fundamental frequency of individual song syllables, was monitored (Fig. 1.4, A) (Tumer and Brainard, 2007). Short, loud bursts of white noise were provided through a speaker in real time only in response to a subset of these pitch variations (Fig. 1.4, B). These white noise bursts are thought by some to be a reinforcement cue: they are externally provided, presumably (mildly) aversive,

and they do not clearly rely on sensory error signals. This pitch-contingent white noise feedback produced rapid and adaptive changes in the pitch of the targeted syllable that decreased the frequency of triggering the white noise feedback. These adaptive changes in song output were precisely locked to the pitch of the targeted syllable, and not to other features of song, showing that the adult songbirds were capable of associating particular variations in their vocal output with auditory feedback signals from the white noise bursts. These results demonstrated not only that auditory feedback cues are sufficient to drive vocal learning in adult songbirds, but also that the trial-to-trial variability in skilled motor behaviors can be used to support continuous learning and maintenance of optimal behavioral performance.

Pitch-contingent white noise feedback, when provided extremely consistently at a particular time point within the production of an individual song syllable, can drive extremely precise, adaptive modifications of syllable structure (Charlesworth et al., 2011). Specifically, songbirds most strongly modify a portion of the individual song syllable being targeted by white noise at the exact time-point within the syllable when the white noise bursts most often occur. These results show that the brain uses auditory feedback to drive precise, millisecond-timescale modifications of behavior.

Also, these white noise feedback cues are sufficient to drive adaptive changes in syllable sequence in adult songbirds (Warren et al., 2012). In this study, rather than providing pitch-contingent white noise feedback, the researchers provided sequence-contingent bursts of white noise by only targeting the occurrence of a particular sequence of song syllables. Adult songbirds learned to modify their song sequence to reduce the frequency of the white noise feedback, showing that auditory feedback cues can change elements of song structure other than simply the fundamental frequency of individual syllables.

Numerous studies have since replicated these findings and used this methodology to investigate the neural basis of vocal learning in response to white noise feedback

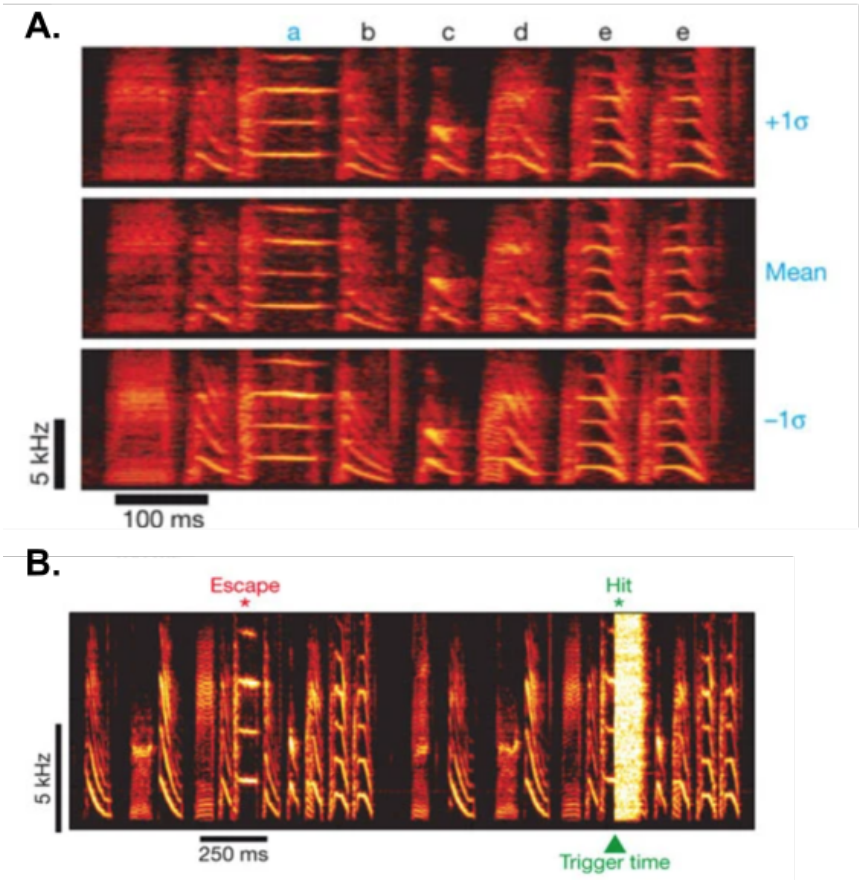


Figure 1.4: A) Song spectrograms from 3 example renditions of song produced by an adult Bengalese finch. The song of this bird consisted of unique song syllables, labelled "a, b, c, d, e, e". The three example songs are shown for which the fundamental frequency of syllable "a" ranged 2 standard deviations of the baseline distribution of fundamental frequencies for this particular syllable. B) Renditions of syllable "a" where the fundamental frequency was above a specified threshold triggered white noise bursts ('hit') that were played through speakers in the bird's cage. From Tumer, E., Brainard, M. Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. Nature 450, 1240–1244 (2007). https://doi.org/10.1038/nature06390. Reprinted with permission from Springer Nature.

(Hoffmann et al., 2016; Warren et al., 2011; Andalman and Fee, 2009; Ali et al., 2013; Gadagkar et al., 2016; Charlesworth et al., 2012), as will be described later in this introduction.

Auditory error signals are sufficient to drive vocal learning in adult songbirds

As I discussed in Section 1.1.1, subtle differences exist between motor learning in response to explicit, external, reinforcement cues and in response to sensory error signals (deviations in sensory feedback from internal goals). After establishing that adult vocal learning can be shaped by auditory reinforcing cues (white noise bursts) (Tumer and Brainard, 2007), researchers then assessed whether adult songbirds, like mammals, can undergo sensorimotor error correction by modifying motor output in a compensatory manner in response to auditory errors signals different than white noise bursts (Sober and Brainard, 2009; Saravanan et al., 2019). Custom-built, lightweight headphones were created and surgically placed over the ears of individual, adult Bengalese finches. The songbird's songs were then recorded, fed through sound-processing hardware to general artificial shifts in pitch of the song syllables, then played back through speakers in the headphones in real-time so the auditory feedback the songbird received during ongoing song performance was experimentally manipulated. The songbirds significantly shifted the pitch of the experimentally targeted song syllables in a compensatory manner (in the opposite direction of the imposed, experimentally-produced pitch shifts) in order to correct for the experience of auditory errors. These findings show that songbirds learn to alter their vocal motor output when auditory errors (a mismatch between experienced auditory feedback from ongoing song production and an internal, memorized song template) occur. Further, these studies suggest that adult songbirds actively maintain the stability of their crystallized song over the course of their lives by constantly engaging in this process of sensory error

correction.

1.2.2 Non-auditory sensory information can shape songbird behavior

Non-auditory sensory information guides vocal learning in juvenile songbirds

Although much attention has been paid to the role of auditory feedback in guiding vocal learning, non-auditory feedback also plays an important role in shaping vocal development in both juvenile songbirds and human infants (King et al., 2005; Lipkind et al., 2013). As described in Section 1.1.3, human infants initially produce babbling noises and gradually learn speech from their parents through a process of vocal imitation that relies on auditory, visual, and somatosensory sensory feedback (Kuhl and Meltzoff, 1996; Meltzoff and Moore, 1983; Tremblay et al., 2003). Similarly, juvenile songbirds initially produce babbling sounds and gradually learn to imitate the song of their adult tutor through a sensorimotor learning process that depends on information from multiple sensory modalities (Price, 1979; Doupe and Kuhl, 1999; West and King, 1988). Here, I will discuss the experimental evidence showing that non-auditory sensory feedback profoundly modulates songbird behavior.

Juvenile songbirds use visual, auditory, and somatosensory sensory information to learn song from their tutors during the juvenile song acquisition process (West and King, 1988; Mann et al., 1991). As I described earlier, juvenile birds use auditory feedback to shape their vocalizations to match the song of their tutor (Fig. 1.3, E) (Konishi, 1965; Immelmann, 1969; Doupe and Kuhl, 1999; Brainard and Doupe, 2002). In addition, juvenile songbirds use visual feedback to guide the song acquisition process (West and King, 1988; Mann et al., 1991). For instance, female visual displays affect male song learning in juvenile cowbirds (West and King, 1988). In this study,

male juvenile cowbirds were housed with non-singing female cowbirds. The females produced a visual display (a wing stroke), that was elicited by specific vocalizations by the male cowbirds. When the male birds observed these female visual displays, they responded by moving towards the female, altering their posture, observing the females' wings, and modifying their vocal output. The female visual displays not only caused the male birds to sing more frequently, but they also influenced the male birds to produce the specific vocal patterns that elicited the wing strokes, suggesting that visual feedback is sufficient to reinforce modifications to specific elements of song.

Further work demonstrated that non-auditory sensory feedback not only reinforces elements of song, but also enhances the song acquisition process in juvenile songbirds. Juvenile zebra finches were provided non-vocal, visual signals from adult females immediately contingent on the production of vocalizations early on in juvenile song development (Carouso-Peck and Goldstein, 2019). These visual signals were provided by playing a video of an adult female bird performing a fluff-up, which is a behavior performed by adult female zebra finches (consisting of erecting their feathers and performing high-frequency, side-to-side movements of the upper body) that is, like the wing-strokes of cowbirds described previously, commonly performed in response to attractive, "good" performances of song. Juvenile songbirds that received this song-contingent visual information learned song significantly more accurately and more quickly, produced songs upon reaching adulthood that more closely matched the song of their tutor, and produced adult songs that were significantly lower in entropy, than control birds, which received identical visual signals from females that were not immediately contingent on the production of vocalizations, indicating that juvenile songbirds are capable of linking social reinforcement with vocal output to help guide and optimize the song learning process during development. Also, juvenile zebra finches that are deprived of information from multiple sensory modalities and are instead provided solely auditory feedback learn to copy tutor songs less accurately

than juvenile songbirds who receive feedback from multiple modalities (Chen et al., 2016). In this study, juvenile zebra finches were deprived of natural social interaction with their tutor (which is, by nature, an experience that involves not only acoustic signals but also information from other sensory modalities (Bottjer and Johnson, 1997)). These juvenile songbirds were then allowed to socially interact with their adult tutor. Deprived songbirds who experienced even as little as one day of normal, multimodal interaction with a tutor showed significantly increased amounts of song learning (measured by the similarity of their song compared to the tutor song) compared to juvenile zebra finches that are deprived of social interaction entirely and not provided the one day of social interaction. Importantly, the group of juvenile finches that were allowed to both visually and acoustically interact with their tutors learned significantly more than the juvenile finches that were only allowed to acoustically interact with their tutors. Interestingly, this study also showed that adult male songbirds alter their song when singing to juveniles, a phenomenon hypothesized to help facilitate the juvenile song acquisition process and further demonstrating the importance of complex, multimodal, social interactions in shaping both juvenile and adult behavior in order to facilitate juvenile song learning.

Sensory feedback from multiple modalities can shape adult songbird vocal behavior

Feedback from auditory and non-auditory sources not only drives juvenile song learning, but also influences adult vocal output. Social interactions between adult songbirds, which, again, involve multiple sensory modalities (Bottjer and Johnson, 1997), influence adult vocal behavior. For example, male songbirds change their rate of song production based on the quantity of and familiarity with nearby birds (Adar et al., 2008; Bischof et al., 1981; Matheson et al., 2016). Male songbirds also consistently alter particular aspects of their song structure when singing in social settings

compared to singing alone. Directed song is when male songbirds sing aimed at a female bird (Morris, 1954; Jarvis et al., 1998), it often occurs in conjunction with the performance of a courtship dance (Hessler and Doupe, 1999). Directed song involves significant changes in vocal behavior (Kao et al., 2005; Kao and Brainard, 2006; Ölveczky et al., 2005). On the other hand, undirected song occurs when a male songbird sings while alone or not oriented towards another particular bird. (Morris, 1954; Dunn and Zann, 1996; Hessler and Doupe, 1999; Woolley and Doupe, 2008). The variability in the fundamental frequency of individual song syllables in adult zebra finches is significantly greater during undirected song than during directed song (Kao et al., 2005; Kao and Brainard, 2006; Ölveczky et al., 2005). Also, the number of introductory elements leading into song is increased during directed song, the number of repetitions of the stereotyped motif per bout is increased during directed song, the stereotypy of syllable sequence is increased during directed song, and the tempo of song is faster during during directed song compared to undirected song (Sossinka and Böhner, 1980). Ultimately, female birds demonstrate responses to male songbird directed singing both at the level of behavior and at the level of gene expression in specific brain regions (Woolley and Doupe, 2008; Mooney, 2009b). These responses are all in line with the interpretation that directed song is more attractive to the female birds, suggesting that the changes in song output made by the male songbird serve an important, ethological purpose.

In addition to making changes to vocal output in the presence of a female bird, male songbirds will also modulate their song output when in the presence of other male birds as compared to singing alone (Jesse and Riebel, 2012). Specifically, male songbirds will increase the rate of song production when within audible and visible distance of other male conspecifics as compared to when alone, demonstrating that the sensory signals from being in social environments, even in non-courtship settings, can significantly alter vocal output.

Similar to humans (as described in Section 1.1.3) (Tremblay et al., 2003), songbirds rely on somatosensory feedback in particular for shaping vocal motor output. The production of birdsong requires the coordinated, patterned activation of respiratory, vocal organ, and vocal tract muscles - all of which provide mechanoreceptive feedback to the brain (Suthers et al., 2002). In one study, respiratory pressure was perturbed by briefly and unpredictably injecting air into the cranial thoracic air sac during ongoing song performance in both deafened and non-deafened adult northern cardinals (Cardinalis cardinalis). These air injection perturbations resulted in a compensatory reduction in the electrical activity of the abdominal expiratory muscles in both hearing and deafened adult northern cardinals, showing that somatosensory feedback to expiratory muscles is processed by the brain and elicits (in real time) compensatory adjustments in vocal motor output. In fact, even when certain parameters of the air injection resulted in increased duration of one of the song syllables, the following inter-syllable interval decreased, presumably to compensate for the shortened syllable and maintain a stable song tempo. The fact that compensatory motor adjustments were made in both deafened and hearing birds strongly supports the interpretation that these modulations to motor output were based entirely on non-auditory feedback. By monitoring somatosensory feedback from muscles important for vocalization and using this feedback to compensate for unexpected perturbations, the brain could potentially adapt to changes in posture, physical activity, or other external changes that could potentially interfere with the adequate production of vocal output, or even to help maintain stable song performance when auditory feedback is interfered with or eliminated. Thus, the combination of auditory and non-auditory (somatosensory) feedback may help promote the maintenance of stable song production in adult songbirds.

Non-auditory sensory information can shape adult songbird non-vocal behavior

Social interactions and non-auditory sensory feedback also play a key function in shaping adult songbird behavior outside of vocalizations. In many species of songbirds, song is a highly important courtship behavior that often involves the production of complex acoustic outputs in conjunction with non-acoustic behaviors, such as dancing, bobbing, and making patterned changes to body posture (Ota et al., 2015). When male songbirds perform directed song (singing directed towards a female bird), they often make alterations to their orientation, such as turning to face the female bird, posture, and position (Zann, 1996; Williams, 2001). In one study, Blue-capped Cordon-bleus (Uraeginthus cyanocephalus) were recorded with high-speed videography to assess non-acoustic courtship behaviors in both male and female birds (Ota et al., 2015). Both males and females displayed an array of courtship behaviors, including bobbing and rapid step-dancing. Interestingly, both male and female cordon-bleus increased the intensity of their dance behaviors when their potential mate was nearby on the same perch as them, suggesting that in addition to the auditory signals (song and the sound of tapping and dancing) and visual signals (the visual displays involved in moving and dancing behaviors), tactile signals (the vibrations of the perch as the potential mate performs its dance and tapping behavior) are also processed and exerts a measurable effect on behavior. Thus, information from multiple sensory modalities is important for an ethological, complex courtship behavior that involves communication and signaling between male and female birds.

Adult male birds also use a variety of sources of sensory information to perform another important, ethological behavior: the selection between potential female mates (Galoch and Bischof, 2006, 2007). Male and female zebra finches were trained to choose between two images presented to them by hopping to different perches inside their cage. The zebra finches much preferred images of conspecifics of the opposite sex

compared with images of empty cages, indicating that visual information is sufficient to drive songbird behavior. Next, similar experiments were performed, but in this case, auditory playback of sounds of the conspecific bird were provided on some trials in addition to the visual images. In the zebra finches making the selection, the time spent in the testing compartment, frequency of courtship song, amount of beak wiping, and the frequency of making calls were all significantly increased by the presence of both audio and visual information being presented together compared with unimodal feedback, demonstrating the significance of information from multiple sensory modalities in shaping ethologically-relevant, social, non-vocal behavior.

1.3 Songbird neural circuitry for sensorimotor learning

Songbirds are a model system uniquely suited for studying the neural basis of sensorimotor learning largely due to the fact that songbirds have well-characterized thalamocortical-basal ganglia neural circuitry (the AFP) that is highly specialized for driving vocal learning but not overall vocal performance (Brainard and Doupe, 2002; Mooney, 2009a). This has allowed for the investigation of the neural circuit underpinnings of motor learning. Prior research has studied how songbird AFP processes sensory feedback to drive changes in motor output, with a particular focus on auditory-driven vocal learning. Here, I will describe the songbird neural pathways that underlie vocal learning in Section 1.3.1. I will then explain how this specialized neural circuitry underlies auditory-guided vocal learning in Section 1.3.2. Finally, I will discuss the literature on the songbird neural pathways for processing non-auditory sensory information in Section 1.3.3.

1.3.1 Songbird neuroanatomy and the song system

Before discussing how songbird neural circuitry underlies vocal learning, I will first provide a general overview of the structure of the songbird brain here. The surface portion of songbird brain is referred to as "pallium", which includes subdivisions: hyperpallium, mesopallium, nidopallium, and arcopallium (Jarvis et al., 2005; Reiner et al., 2004). The large-scale structure of songbird pallium has important differences from mammalian cortex (e.g., mammalian cortex is multilayered, songbird pallium is not). However, many similarities in cell types, microcircuit structure, and function exist between the two systems (Brainard and Doupe, 2013; Fee and Goldberg, 2011; Mooney, 2009a). For example, like mammalian cortex, songbird pallium receives sensory input from the thalamus, processes sensory information, and sends descending projections to motor neurons of the brainstem and spinal cord for motor control and motor learning (Zeier and Karten, 1971; Karten and Shimizu, 1989; Vates et al., 1996; Wild, 1997a; Jarvis et al., 2005; Vicario, 1991; Wild, 1993; Nottebohm et al., 1976). Also, the expression patterns of brain-derived neurotrophic factor (BDNF), glutamate receptor mGluR2, and other glutamate receptors are similar throughout songbird pallium and mammalian cortex (Wada et al., 2004; Li and Jarvis, 2001), further strengthening the idea that songbird and mammalian brain systems are similar.

Of particular relevance for this dissertation, numerous similarities exist between the neural pathways for speech in humans and the neural pathways for vocalizations in songbirds. Songbird neural circuitry that supports vocal production and vocal learning involves a number of brain nuclei (brain regions with particularly dense neuronal cell populations in one defined area) throughout songbird pallium, basal ganglia, and thalamus (Fee and Goldberg, 2011; Jarvis et al., 2005; Reiner et al., 2004). For example, auditory pallial regions in songbirds, like auditory cortex in mammals, are important for processing auditory feedback and relaying that information to other brain regions (Mello et al., 1992). Electrophysiological recordings of neurons in Field

L, an auditory pallial brain nucleus, show that these neurons respond to the white noise bursts described in Section 1.2.1, suggesting Field L plays a role in processing that particular source of auditory feedback to drive vocal learning (Keller and Hahnloser, 2009).

HVC (used as a proper name) and the robust nucleus of the arcopallium (RA) are other pallial brain regions that, while not the primary focus of this dissertation, are nonetheless highly important for song production. HVC sends neural projections to RA, which then send projections to a brainstem nucleus nucleus hypoglossus pars tracheosyringealis (nXIIts), which then sends direct neural projections to vocal musculature involved in producing song (Wild, 1993; Gahr, 2000; Wild, 1997b). This pathway is known as the vocal motor pathway, and it is particularly important for the control of vocal musculature and song performance. Lesions of either HVC or RA in adult songbirds will severely impair or abolish the ability to perform song. Electrophysiological recordings of HVC neurons have shown that individual HVC neurons consistently fire at distinct points in song (Hahnloser et al., 2002; Long and Fee, 2008). This firing pattern has been described as a sparse code, where each individual HVC neuron encodes distinct time points within song, and that the population of a large number of HVC neurons encodes the entire time-course of a song motif (Hahnloser et al., 2002; Long and Fee, 2008; Picardo et al., 2016). Researchers have hypothesized that the firing pattern of HVC neurons that project to downstream song learning circuits could be useful for driving adaptive changes to particular elements within song, such as the structure of individual syllables (Fee and Goldberg, 2011). RA is thought to be analagous to mammalian primary motor cortex because of its functional involvement in the vocal motor control of song and its anatomical projections to brainstem nucleus nXIIts, which then sends projections directly to vocal muscles (Sober et al., 2008; Jarvis et al., 2005; Wild, 1997b). Also, individual neurons within RA display firing patterns that correlate significantly with song parameters,

such as syllable pitch, spectral entropy, and amplitude, suggesting they play a role in governing song structure (Sober et al., 2008).

Songbirds also have a specialized neural pathway for vocal learning that exerts its influence on vocal output through its connections to the vocal motor pathway (via projections to RA) (Brainard and Doupe, 2002; Mooney, 2009a). This song learning pathway, called the Anterior Forebrain Pathway (AFP), is a thalamocortical-basal ganglia loop that underlies only song learning but not song production. The lateral magnocellular nucleus of the anterior nidopallium (LMAN) is a pallial brain region and is the output nucleus of the AFP: LMAN sends projections to RA that are crucial for song learning (Charlesworth et al., 2012). Area X is a songbird basal ganglia nucleus that is a part of the AFP and has numerous analogies to mammalian basal ganglia. For example, both songbird Area X and mammalian basal ganglia receive dense dopaminergic innervation from the ventral tegmental area (VTA) and the substantia nigra pars compacta (SNc), two midbrain regions that contain neurons that produce dopamine and send projections to multiple other brain regions (Gale and Perkel, 2010b; Gadagkar et al., 2016; Hoffmann et al., 2016; Gale and Perkel, 2005; Person et al., 2008; Jarvis et al., 2005). Area X and mammalian basal ganglia also both contain GABAergic medium spiny neurons (that express dopamine receptors and send projections to other brain regions), contain interneurons, and form connections between cortex and thalamus important for motor learning (Gale and Perkel, 2010a; Sasaki et al., 2006; Person et al., 2008; Carrillo and Doupe, 2004; Gale and Perkel, 2005; Goldberg et al., 2010; Goldberg and Fee, 2010; Luo and Perkel, 1999b,a; Medina and Reiner, 1995; Doya, 2000; Graybiel, 1998, 2000). Area X neurons and mammalian basal ganglia neurons display similar electrophysiological properties as well (Bottjer, 1993; Bottjer and Alexander, 1995; Casto and Ball, 1994; Luo and Perkel, 1999a,b; Perkel et al., 2002; Medina and Reiner, 1995). However, differences do exist between songbird Area X and mammalian basal ganglia. For instance, Area X projection

neurons are sent directly to the dorsolateral nucleus of the anteriorthalamus (DLM), which is the thalamic song system nuclei in songbirds (Jarvis et al., 2005), whereas mammalian striatum sends projections to a separate brain structure, the pallidum, which then projects to thalamus (Carrillo and Doupe, 2004; Farries and Perkel, 2002; Medina and Reiner, 1995). Also, mammalian basal ganglia has distinct neuronal populations that express only D1 receptors (a subclass of dopamine receptors) or only D2 receptors, which constitute separate neural pathways through the basal ganglia, whereas songbird Area X neurons typically express both D1 and D2 receptors (Gerfen et al., 1990; Gale and Perkel, 2005, 2010b; Kubikova et al., 2010; Xiao et al., 2020; Casto and Ball, 1994). Despite these differences, the numerous similarities in gene expression, cell type, neuroanatomy, and function, between songbird and mammalian systems indicates that studies of songbird cortical-basal ganglia circuitry are likely to help advance our understanding of how these neural circuits support behavior in mammalian species as well (Brainard and Doupe, 2013; Mooney, 2009a; Jarvis et al., 2005; Brainard and Doupe, 2002; Jarvis, 2019). The AFP and its role in vocal learning are at the center of the work performed throughout this dissertation.

Specialized neural pathways for vocal learning are rare amongst animal species, yet they are found in humans and songbirds (two of vocal learning species). How specific elements within these circuits evolved across different species is the subject of much speculation and active research (Jarvis, 2019). What is clear is that specialized vocal pathways have evolved in both humans and songbirds, that these pathways are crucial for supporting skilled behaviors like speech, and that it is extremely challenging to study these pathways due to technical limitations involved in studying human brains. There is no animal model that has evolved neural pathways for vocal learning and have neural circuitry as well-understood and experimentally tractable as the songbird.

1.3.2 Songbird neural circuitry for processing auditory feedback to drive vocal learning

The role of songbird AFP in juvenile song acquisition

Songbird AFP plays an important role in the juvenile song acquisition process. Early work in this area focused on understanding the role that LMAN, the output nucleus of the AFP, played in song learning in juvenile songbirds. At the time, LMAN was associated with song learning due to its expression of hormones related to vocalizations in males (Arnold et al., 1976; Arnold and Saltiel, 1979). To investigate the role LMAN plays in song learning, bilateral electrolytic lesions of LMAN were made in juvenile songbirds (Bottjer et al., 1984). These lesions resulted in significant impairments in song learning capabilities: the songs produced by lesioned birds typically consisted of a small number of highly abnormal notes, were produced at low amplitudes, lacked the typical frequency modulations seen in control birds, and lacked typical bout duration and phrasing throughout adulthood. Further, these LMAN lesions in juvenile birds produced more dramatic impairments the earlier in the song acquisition process the lesions were performed. On the other hand, songbirds who did not undergo LMAN lesions until adulthood did not display any disruptions to the performance normal, crystallized song. Together, these results indicated that LMAN plays a crucial role in song learning, but not in song production.

The role that Area X, the basal ganglia nucleus of the AFP, plays in juvenile song learning was tested (Sohrabji et al., 1990). At the time, Area X was understood to be a prominent, very large, sexually dimorphic nucleus (Nottebohm and Arnold, 1976) that had anatomical connections to other vocal control brain regions (Bottjer et al., 1989; Nottebohm and Arnold, 1976; Nottebohm et al., 1982; Okuhata and Saito, 1987), and therefore was hypothesized to be important for vocal learning. To directly test this hypothesis, bilateral, electrolytic lesions of Area X were performed in juvenile

zebra finches undergoing the song acquisition process (Sohrabji et al., 1990). This method of performing Area X lesions indiscriminately and highly effectively killed cells within Area X. These lesions significantly impaired song learning in juvenile zebra finches, yet they did not produce a pronounced effect on normal song production when performed in adult zebra finches, demonstrating that, like LMAN, Area X is required for juvenile song learning but not overall song production.

Later work demonstrated that pharmacological inactivation of LMAN in juvenile zebra finches significantly and reversibly decreased the acoustic and sequence variability of juvenile song (Ölveczky et al., 2005). Also, electrophysiological recordings of individual LMAN neurons performed in singing juvenile zebra finches displayed high variability of spiking patterns across song renditions. These results established a role for LMAN specifically in generating the behavioral variability necessary for the reinforcement learning process, which is necessary to evaluate different motor outputs and select for the most optimal behaviors over time. The role of LMAN in adult vocal variability will be discussed later in this section.

The role of songbird AFP in processing white noise cues to drive adult vocal plasticity

In addition to its role in the juvenile song acquisition process, the AFP also plays a crucial role in adult songbird vocal learning (Brainard and Doupe, 2002, 2013). As I will describe below, LMAN, the output nucleus of the AFP, is vital for both processing auditory feedback to drive changes in vocal output and generating the behavioral variability necessary for the implementation of vocal learning. Area X, the basal ganglia nucleus of the AFP, also plays a necessary role in driving songbird vocal plasticity.

LMAN is important for the vocal plasticity observed after deafening in adult songbirds (Brainard and Doupe, 2000b, 2002). Typically, surgical, bilateral removal of the

cochlea (which deafens the birds) in adult zebra finches causes a gradual degradation in song quality and, in some cases, causes the songs being produced to barely resemble the patterned, crystallized song structure typically produced by adult zebra finches (Nordeen and Nordeen, 1992). However, if LMAN is electrolytically lesioned just before the removal of the cochlea, the degradation of crystallized song is significantly decreased, such that even after 100 days the song structure and sequence remain largely intact (Brainard and Doupe, 2000b). These results suggest that disrupting auditory feedback in adult songbirds (via deafening) leads to the generation of an instructive signal that drives changes in song, and that process is dependent on the AFP. This study was one of the earliest studies to provide direct evidence for LMAN and the AFP playing a necessary role in evaluating auditory feedback to drive vocal learning in adult songbirds.

LMAN is also crucial for the expression of white noise learning - the experimental paradigm where white noise feedback cues are provided during ongoing song performance in a pitch-contingent manner to differentially reinforce the pitch of a targeted syllable. As explained earlier, this experimental approach produces significant, adaptive shifts in the pitch of the target syllable, indicating that adult songbirds were capable of processing auditory signals to drive vocal learning (Tumer and Brainard, 2007; Hoffmann et al., 2016). Electrolytic as well as chemical lesions of LMAN in adult zebra finches significantly impair the ability of these songbirds to express vocal learning in response to white noise feedback (Ali et al., 2013). Similarly, bilateral electrolytic lesions performed in Area X in adult zebra finches significantly impair the ability of these birds to express white noise-driven vocal learning. Together, these results indicate that the AFP is not only necessary for juvenile song acquisition, but also for vocal learning in adult songbirds.

Pharmacological inactivation of LMAN neurons during white noise learning paradigms has revealed more specific ways by which the AFP helps drive adult vocal learning

(Warren et al., 2011; Andalman and Fee, 2009). In these experiments, microdialysis probes were surgically inserted above LMAN in adult songbirds, then online, pitch-contingent, white noise feedback was provided to target syllables. Through the microdialysis probes, pharmacological agents, such as TTX or muscimol, were diffused into LMAN to inhibit neural activity in the region. When this pharmacological inactivation of LMAN was performed after one day of the pitch-contingent white noise feedback, the pitch of the target syllable reverted back to the baseline pitch prior to initiating white noise-driven learning, suggesting that neural activity in LMAN plays a necessary role in biasing motor output towards adaptive outcomes early in the learning process. Interestingly, when these same pharmacological inactivations were performed after several days of white noise training, the magnitude by which the pitch of the target syllable reverted back towards baseline was significantly less compared to when the LMAN inactivations were performed after only one day of white noise training. These results suggest that a portion of the learning that occurs during prolonged exposure to the white noise training paradigm is consolidated in a neural pathway outside of LMAN.

Another line of study that has revealed the mechanisms by which LMAN contributes to the learning process has focused on understanding the role of LMAN in generating variability in motor output that is necessary for implementing reinforcement learning. For example, when bilateral, electrolytic lesions of LMAN in adult Bengalese finches were performed, the variability of the fundamental frequency of individual song syllables had significantly decreased postlesion compared to prelesion (Hampton et al., 2009). In congruence with this finding, bilateral, electrical microstimulation of LMAN during ongoing song performance in adult zebra finches produces rapid, systematic changes in syllable structure, whereas microstimulation of control brain regions outside of LMAN do not produce the same changes to ongoing song performance (Kao et al., 2005). The systematic changes in song structure observed in

these experiments occurred with very low latency relative to the time of microstimulation and were not related to overall song structure or syllable sequence. Instead, song-triggered microstimulation of LMAN produced significant shifts in the pitch of the targeted syllable, indicating that neural activity within LMAN causally drives changes in ongoing motor commands, specifically by affecting syllable pitch but not overall song production. Further, electrophysiological recordings of individual LMAN neurons were performed during two different experimental conditions: when a male songbird sings along (undirected song) and when the male songbird sings to a female (directed song). The trial-to-trial variability in the fundamental frequency of individual song syllables was significantly lower during directed song compared to undirected song. Correspondingly, the trial-to-trial variability in neural activity patterns of individual LMAN neurons was significantly lower during directed song compared to undirected song, and the magnitude of the changes in LMAN neural firing variability and the changes in fundamental frequency variability were significantly correlated with each other. Prior to LMAN lesions in adult male zebra finches, the trial-to-trial variability in the fundamental frequency of song syllables was significantly greater during directed song than during undirected song, and, after LMAN lesions, this context-dependent difference was eliminated, suggesting that LMAN is necessary for producing naturally occurring changes in song variability. The white noise experiments described earlier, where pitch-contingent white noise bursts were provided to a subset of the naturally distribution of syllable pitches and significant changes in the fundamental frequency of targeted syllable were observed, indicate the importance of naturally occurring variability of the moment-to-moment structure of song syllables in motor learning (Tumer and Brainard, 2007). LMAN produces variability in song output, which can then be evaluated via auditory feedback, and then good outcomes (i.e., escapes from white noise) can be reinforced and bad outcomes (i.e., white noise bursts are triggered) can be learned to be avoided, thereby modulating future vocal

output.

Area X, the basal ganglia nucleus of songbird AFP, also plays a crucial role in songbird vocal learning. As described briefly above, bilateral electrolytic lesions of Area X in adult zebra finches significantly impair white noise-driven learning (Ali et al., 2013). Area X receives a very large anatomical projection from dopaminergic neurons in VTA, which in turn receive projections from auditory brain regions analogous to auditory cortex in mammals (Person et al., 2008). These auditory cortical brain regions were hypothesized to encode the comparison between ongoing song performance and the memorized tutor song template and then send this information to VTA, which then could convert this information into an error signal useful for reinforcement learning and implement plasticity in song circuits (Fee and Goldberg, 2011). In order to assess whether dopaminergic input to Area X plays a role in song learning, bilaterally injections of 6-OHDA (a neurotoxin that selectively kills catecholiminergic neurons) were made into Area X, and the effect of these dopamine lesions on song performance was measured (Fig. 1.5 A) (Hoffmann et al., 2016). They found that these reductions in dopaminergic input to Area X in adult Bengalese finches did not have an effect on number of songs produced, song sequence, overall song structure, or trial-to-trial variability in the fundamental frequency of song syllables postlesion compared to prelesion. Prelesion, these birds were capable of expressing significant amounts of vocal learning in response to pitch-contingent white noise feedback. Postlesion, this white noise-driven vocal learning was significantly impaired, indicating that intact dopaminergic input to Area X is required for expressing white noise-driven vocal learning in adult songbirds (Fig. 1.5 B).

Electrophysiological recordings of individual VTA dopamine neurons that project to Area X have revealed the patterns of neural activity that encode information relevant for vocal learning (Gadagkar et al., 2016). In these experiments, adult male zebra finches underwent white noise training to learn to associate a white noise burst

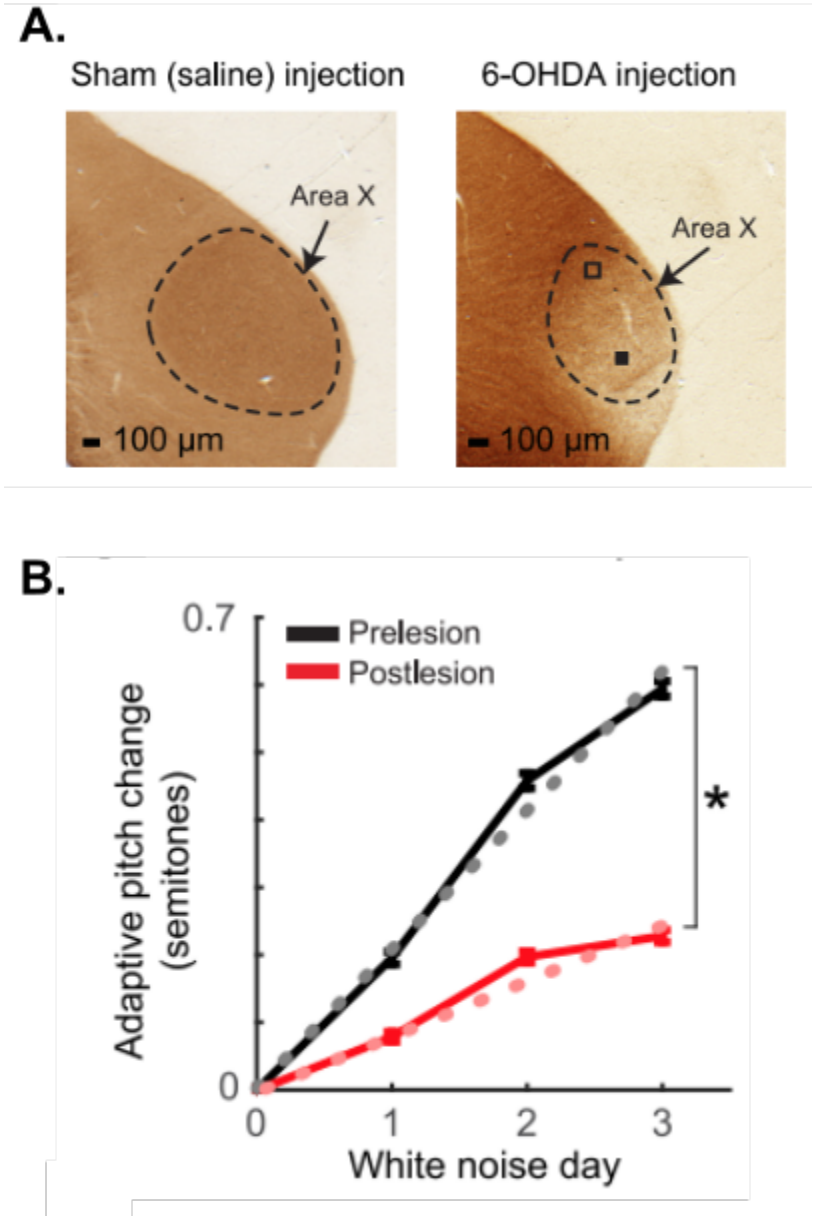


Figure 1.5: A) Comparison of TH stain, which marks dopaminergic fibers, in Bengalese finch brains that received sham injections (left) and 6-OHDA injections(right). 6-OHDA-injected brains shows a reduction in the optical density of the TH stain B) Adult Bengalese finches underwent 3 days of white noise training in which pitch-contingent white noise feedback was provided in real time during the production of certain renditions of a targeted syllable. Black and red traces represent the pitch (in semitones) of the targeted syllable before and after 6-OHDA injections, respectively. Pitch changes in the adaptive direction (upwards in experiments when shifting the pitch up results in less frequent white noise feedback, downwards in experiments where shifting the pitch down results in less frequent white noise feedback) are plotted as positive values. From Hoffmann, L. A., Saravanan, V., Wood, A. N., He, L., and Sober, S. J. (2016). Dopaminergic contributions to vocal learning. Journal of Neuroscience, 36(7), 2176-2189. Reprinted with permission from Journal of Neuroscience.

with positive and negative outcomes from song performance. Then, antidromically-identified VTA-to-Area X projection neurons were recorded in vivo. When the birds sang, white noise feedback was provided on random renditions of a targeted song syllable (not in a pitch-contingent manner) and neural spiking patterns of the VTA dopamine neurons were collected (Fig. 1.6 A, B). The firing rates of these VTA dopamine neurons significantly increased on trials during which the white noise feedback was not provided, and the firing rates significantly decreased on trials during which the white noise feedback was provided, suggesting that these neurons encoded a performance error signal - the difference between expected outcome of the ongoing song performance and the sensory feedback from the actual outcome of the song performance (Fig. 1.6 C). Further, when the probability of providing white noise feedback was altered, it revealed that the phasic increase in the firing rates of VTA dopamine neurons was significantly greater following more surprising renditions of the higher probability target compared to the lower probability, less surprising target syllable. These results revealed that songbird VTA dopamine neurons that project to Area X encode a performance error signal thought to play a crucial role in implementing learning (Fee and Goldberg, 2011), thereby providing further evidence that dopaminergic input to Area X is important for vocal learning and suggesting a potential computational mechanism for driving vocal learning.

The role of the AFP in sensory error correction in adult songbirds

The role of dopaminergic input to Area X in driving songbird sensory error correction has also been assessed (Saravanan et al., 2019). Briefly, researchers applied miniature headphones over adult Bengalese finches' ears, as described earlier (Sober and Brainard, 2009). Then, in real-time during ongoing song performance, they played back recordings of the bird's song with artificially imposed errors to manipulate the auditory feedback the songbirds received. They observed gradual adaptations to vo-

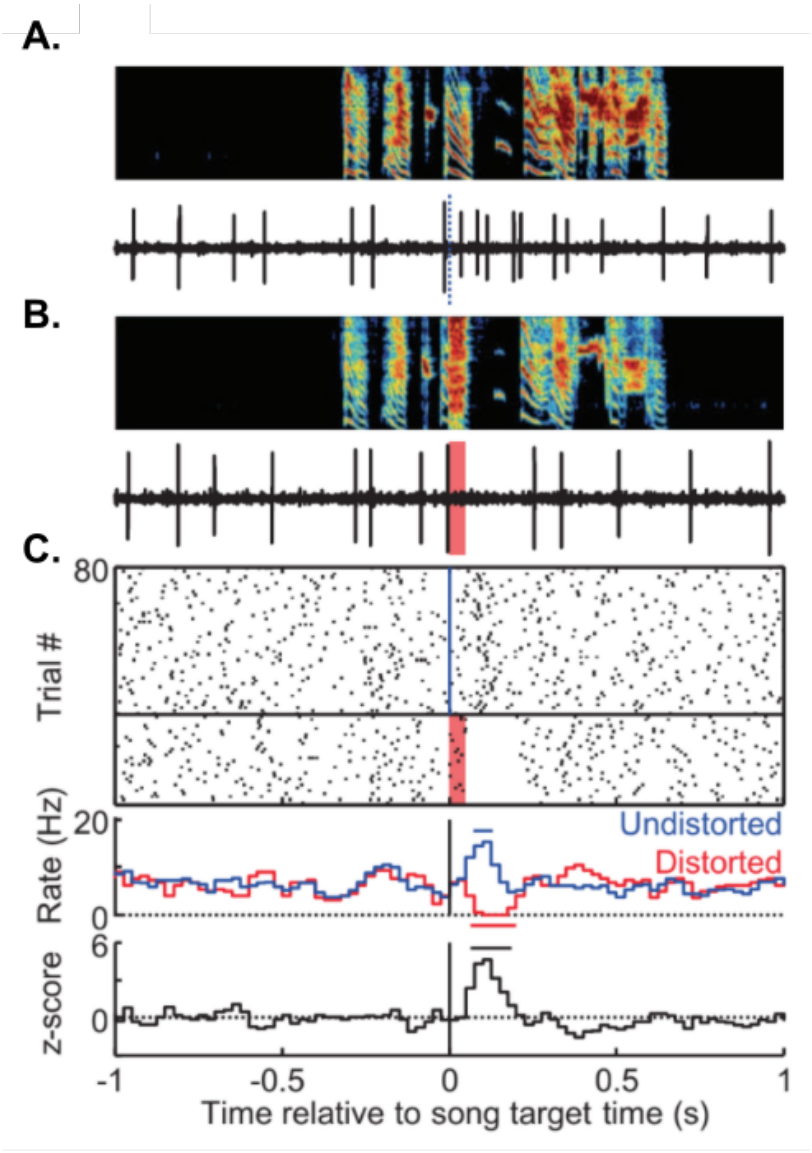


Figure 1.6: A) Spectrogram of an example song rendition sung by an adult Zebra finch (top) and the voltage trace of an individual VTA neuron (bottom) aligned to song. On this rendition of song, no distorted auditory feedback was provided. The blue lines show the point in song when distorted auditory feedback sometimes occurs on other song renditions. B) Song rendition sung by the same bird (top), and the voltage trace of the example VTA neuron, during a rendition where distorted auditory feedback was provided (red shading). C) A raster plot of VTA neuron spiking activity during undistorted and distorted trials (top), the corresponding rate histograms (middle), and the z-scored difference between undistorted and distorted rate histograms (bottom). All plots are aligned to the onset of the target syllable. Horizontal bars in histograms indicate significant deviations from baseline ($P < 0.05$, z test). From 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(6317), 1278-1282. Reprinted with permission from AAAS.

cal output to correct for the imposed auditory errors. Next, 6OHDA injections were performed bilaterally in Area X to lesion dopaminergic input to Area X (Hoffmann et al., 2016), then the songbirds underwent the headphones training paradigm again to assess their ability to perform sensory error correction. The songbirds learned to adaptively modify the fundamental frequency of their song syllables in response to auditory errors significantly less following the lesions of dopaminergic input to Area X compared to prelesion. This suggests that intact dopaminergic input to Area X is not only required for expressing vocal learning in response to an external cue, like a white noise burst, but it is also required for expressing vocal learning during what is clearly a sensorimotor adaptation paradigm, where the songbirds learn to compensate for differences between auditory feedback of ongoing song performance and their internal, memorized song template.

1.3.3 Songbird neural circuitry for processing non-auditory sensory feedback

Even though non-auditory sensory feedback plays an important role in shaping both human and songbird behavior (discussed in sections and 1.1.3 and 1.2.2, respectively), the neural circuit mechanisms for processing non-auditory feedback to drive learning are poorly understood compared to the neural mechanisms underlying auditory-driven vocal learning. The AFP is one of the best-studied sensorimotor learning neural pathways, yet the role the AFP plays in processing non-auditory sensory feedback to guide vocal learning is unknown. Here, I will discuss the current understanding in the field about how songbird brains process non-auditory feedback.

Anatomical studies have provided clues for how sensory information from various sources travels through the brain and might enter the song learning circuit. These anatomical studies used anterograde and retrograde tracing techniques in zebra finches to visualize the neural projections to and from specific brain regions (Paterson

and Bottjer, 2017). They found that LMAN, which carries song-related information, sends anatomical projections to the dorsal caudolateral nidopallium and the ventral arcopallium, two brain regions that also receive inputs from other brain regions surrounding the nidopallium that process information from multiple sensory modalities, suggesting a potential neural circuit pathway for integrating song-related information with multimodal sensory information to guide behavior. Also, the visual tectofugal system, which carries information originating from the retina, sends dense projections that terminate in the nidopallium in zebra finches (the portion of the brain that contains LMAN) (Krützfeldt and Wild, 2004; Wild and Gaede, 2016). Also, projections from this tectofugal pathway are sent to nucleus uvaeformis (Uva) of the posterior thalamus in zebra finches, which plays a role in the neural control of song and sends projections to nuclei upstream of the AFP (Wild and Gaede, 2016). Areas of zebra finch forebrain receive sensory information through a hypoglossal projection, somatosensory information about the beak and tongue through a trigeminal projection, and auditory sensory information through the intermediate nucleus of the lateral lemniscus (Wild and Farabaugh, 1996; Wild and Williams, 1999). Finally, perhaps the most convincing anatomical evidence that songbird AFP may receive input from multiple sensory modalities is that the anterior nidopallium in other bird species, such as pigeon and chick, receives anatomical projections from auditory, visual, and somatosensory brain regions analagous to cortex (Wild, 1994; Ahumada-Galleguillos et al., 2015; Dubbeldam et al., 1997; Güntürkün and Kröner, 1999; Helduser et al., 2013; Kröner and Güntürkün, 1999; Shimizu et al., 1995).

Recent work has interrogated how patterns of neural activity within specific songbird brain regions might encode information relevant for processing non-auditory sensory information. For example, researchers performed electrophysiological recordings of individual neurons within dorsal intermediate arcopallium (AId), a region of the songbird brain that has been linked to juvenile song learning, that is analogous to

motor cortex in mammals, and that receives anatomical projections from other brain regions that process auditory, visual, and somatosensory information (Yuan and Bottjer, 2020). In order to study how individual neurons encode non-vocal songbird behavioral information, these neural recordings were made in freely behaving juvenile zebra finches and were analyzed not only during song production, but also during non-acoustic, naturalistic behaviors, such as hopping, pecking, preening, fluff-ups, beak interactions, scratching, and stretching. The activity of individual AId neurons was significantly modulated by singing, and some individual neurons displayed heterogeneous response patterns during different non-vocal movements. Also, some neurons were modulated not by specific movements, but rather by overall behavioral states, such as active behavior or quiescence. These results demonstrate that neural responses in certain songbird brain regions not only encode vocal behaviors but also encode non-vocal behavior, which provides clues as to how the brain processes non-auditory information to guide learning.

Social interactions between an adult tutor and a juvenile songbird involve auditory and non-auditory feedback, and they modulate songbird behavior and influence learning (Chen et al., 2016). For example, when juvenile songbirds are allowed to interact socially with a live tutor, the song learning process is enhanced. However, how social interactions (and multimodal sensory information) between tutors and juvenile songbirds impacts neural activity and drives attention and learning remains unclear. One study demonstrated that EGR-1 expression, which is a transcription factor that has been used as a cellular marker for neural activity levels (O'Donovan et al., 1999; Aston-Jones and Cohen, 2005), was significantly modulated in catecholaminergic neurons within the locus coeruleus (LC) and ventral tegmental area (VTA) of juvenile songbird brains when exposed to social interactions with a live, singing tutor, compared to untutored juveniles (Chen et al., 2016). Importantly, social interactions with a non-singing tutor did not significantly change the EGR-1

expression in catecholaminergic neurons in the LC and VTA of juvenile songbirds, indicating that multimodal sensory feedback is necessary for the observed changes in neural activity within this neural circuit. The authors suggested that catecholaminergic neurons within LC and VTA integrate multimodal sensory information related to social interactions with tutors to help enhance the learning process.

Also, EGR-1 expression in several brain regions varies between adult Bengalese finches who were exposed to different social contexts (Matheson et al., 2016). EGR-1 expression in interface nucleus of the nidopallium (NIf), HVC, the robust nucleus of the arcopallium (RA), Area X, and LMAN was significantly greater in adult male Bengalese finches performing either undirected or directed song compared to male Bengalese finches who were not producing any song, indicating the song performance elevates neural activity levels throughout these brain regions. EGR-1 expression was higher in HVC, RA, Area X, and LMAN in adult male Bengalese finches who were producing undirected song compared to finches that were producing female directed song, showing that neural activity levels in these song control brain regions is modulated by social context.

In order to study neural mechanisms underlying the influence of social interactions on behavior, studies have analyzed the expression levels of another immediate-early gene, ZENK, in various adult male zebra finch brain regions during female directed and undirected song performance (Jarvis et al., 1998). ZENK is a transcription factor that is activated by neural activity, modulates the expression of other genes, and is activated in songbirds by the performance of song (Chaudhuri, 1997; Stripling et al., 1997; Jarvis and Nottebohm, 1997). The expression levels of ZENK expression in Area X, LMAN, and RA in male songbird brains were significantly higher during undirected song than during female directed song performance, and there was a correlation between the amount of undirected song produced and the level of Area X ZENK expression (Jarvis et al., 1998). Importantly, ZENK expression levels in Area

X, LMAN, and RA of deafened adult male zebra finches performing female directed song were equivalent to the ZENK expression levels in those song nuclei in hearing-intact zebra finches performing female directed song, indicating that non-auditory feedback is responsible for modulating the expression of ZENK in the song nuclei of adult zebra finches.

The noradrenergic system, which responds to complex, multimodal stimuli, is required for this social modulation of ZENK expression patterns in song nuclei in adult male zebra finch brain (Castelino and Ball, 2005). Neurons within the locus coeruleus (LC), which is a brain nuclei responsible for the production of noradrenaline (Bouret and Sara, 2005), show an increase in activity levels in response to multimodal sensory stimuli across a variety of species (Rasmussen et al., 1986; Rasmussen and Jacobs, 1986; Jacobs, 1986; Foote et al., 1980; Aston-Jones et al., 1991, 1994). Even functional neuroimaging studies in humans have demonstrated that activity levels in the noradrenergic system are modulated by multimodal stimuli (Coull et al., 1999). In songbirds, the injection of N-(2-chloroethyl)-N-ethyl-2-bromobenzylamine hydrochloride (DSP-4), a noradrenergic neurotoxin known to effectively deplete norepinephrine in forebrain brain regions of male zebra finches (Barclay et al., 1996), results in a significant increase in the level of ZENK expression in Area X of male zebra finches performing female directed song compared to the level of ZENK expression when the noradrenergic system is left intact (Castelino and Ball, 2005).

Similarly, EGR-1 expression (a proxy for neural activity) in VTA-SNc neurons in adult songbirds is significantly greater during directed singing than during undirected singing, linking the songbird dopaminergic system with the social modulation of gene expression (Hara et al., 2007). Also, in vivo microdialysis in awake, behaving, adult songbirds has demonstrated not only that singing in general is correlated with significantly greater levels of dopamine in Area X compared to non-singing, but also that dopamine levels in Area X are significantly higher during directed singing com-

pared to undirected singing (Sasaki et al., 2006). In vivo blockade of the dopamine transporter eliminates the social context-dependent difference in dopamine levels in Area X. Also, in vivo microdialysis in awake, behaving, adult songbirds has demonstrated not only that singing in general is correlated with significantly greater levels of dopamine in Area X compared to non-singing, but also that dopamine levels in Area X are significantly higher during directed singing compared to undirected singing. In vivo blockade of a dopamine reuptake transporter (DAT), which governs the kinetics of dopamine levels (Gale and Perkel, 2005) eliminates the social context-dependent difference in dopamine levels in Area X.

The patterns of neural activity that may encode social information have also been assessed (Hessler and Doupe, 1999). Neural spiking activity in both LMAN and Area X in male songbirds is significantly greater during undirected singing compared to during female directed singing. Also, rendition-to-rendition variability in spiking activity of LMAN and Area X neurons is significantly greater across many renditions of undirected song compared to many renditions of directed song. This effect of social context on LMAN and Area X neuron activity patterns was not dependent on auditory feedback, as the same effects were observed in songbirds before and after deafening. These results show that social context does not only alter gene expression levels, but also directly modulates neural spiking activity in song system brain nuclei.

1.4 Summary

In this introduction, I first described how sensorimotor learning underlies the acquisition of skilled behaviors in mammals, and how mammalian neural circuits underlie this process. In particular, I focused on human vocal learning, a form of sensorimotor learning that is particularly important for our survival and quality of life. I then discussed songbirds, which are arguably the best animal model for studying vocal motor

learning, and how they use auditory and non-auditory feedback to guide behavior. Finally, I discussed the songbird neural mechanisms important for auditory and non-auditory learning. It is clear that both auditory and non-auditory information are crucial for guiding vocal learning in humans and in songbirds. However, we know very little about how non-auditory feedback is processed by vocal circuits to drive vocal learning. I will now address an open question in the field: *Do the specialized vocal learning circuits that have uniquely evolved in humans and songbirds solely process auditory feedback, do they have the ability to process information from other sensory modalities?* I address this question using songbirds as a model system to study the neural circuit underpinnings of vocal learning, with hopes that these findings may help to inform our understanding of how neural circuits support vocal learning more generally in other systems, like in humans.

Chapter 2

Shared mechanisms of auditory and non-auditory vocal learning in the songbird brain

2.1 Abstract

Songbirds and humans have specialized neural circuitry that underlies vocal learning. Auditory feedback is crucial for shaping vocal output, and numerous prior studies have revealed neural mechanisms for auditory-guided vocal learning. Non-auditory information also profoundly shapes vocal output, yet the neural circuit mechanisms that underlie non-auditory vocal learning are unclear. Here, we assessed whether and how non-auditory feedback guides vocal learning in songbirds. After establishing that non-auditory feedback drives adaptive changes in vocal motor output, we assessed the role of a specialized vocal learning neural circuit in this novel form of non-auditory learning. We found that this specialized song learning circuit and its dopaminergic inputs are necessary for non-auditory vocal learning, demonstrating that this neural pathway is not specialized exclusively for processing auditory feedback to reshape

vocal behavior. The ability of the songbird vocal learning pathway to process both auditory and non-auditory information to guide changes to vocal output may be a general principle for the neural systems that have evolved to support vocal learning across species.

2.2 Introduction

A fundamental goal of neuroscience research is to understand how the brain uses sensory feedback to drive adaptive changes to motor output. This process of sensorimotor learning underlies the acquisition of skilled behaviors and is supported by specific neural pathways (Graybiel, 1998; Hikosaka et al., 2002). One such skilled behavior acquired through sensorimotor learning is human speech. Humans are one of the few species that uses sensory feedback from their own vocalizations and sensory feedback from the vocalizations of others to guide vocal motor learning (Doupe and Kuhl, 1999; Brainard and Doupe, 2002). Both auditory and non-auditory sources of sensory feedback are crucial for proper speech acquisition: deafened infants display impaired speech learning (Cowie and Douglas-Cowie, 2011; Oller and Eilers, 1988; Stoel-Gammon and Otomo, 1986), and hearing-intact infants who are provided with non-auditory feedback from their parents contingent on the production of vocalizations show enhanced speech acquisition (Goldstein et al., 2003). Human adults use non-auditory (somatosensory) feedback from vocal effectors to maintain proper vocal output throughout adulthood (Tremblay et al., 2003; Nasir and Ostry, 2008). Humans have evolved specialized neural pathways that underlie vocal learning, but it is unknown these whether these neural circuits process non-auditory sensory information to guide vocal motor learning (Jarvis, 2019).

We address these questions by using a unique model system ideally suited for the study of vocal learning, the Bengalese finch (a type of songbird). Like humans, song-

birds use auditory and non-auditory feedback to learn and maintain an ethological, skilled, vocal behavior (West and King, 1988; Mann et al., 1991; Chen et al., 2016). For example, deafened juvenile songbirds fail to properly acquire song during development, and the songs of deafened adult songbirds degrade over time, demonstrating the importance of auditory feedback for song learning and maintenance (Konishi, 1964, 1965; Nordeen and Nordeen, 1992; Brainard and Doupe, 2000b). Also, juvenile songbirds use visual and somatosensory sensory information to learn song from their adult tutors (Carouso-Peck and Goldstein, 2019; Bottjer and Johnson, 1997). Juvenile songbirds that are deprived of visual feedback and only provided acoustic feedback from their tutors learn to copy tutor songs less accurately than songbirds who receive both auditory and visual feedback (Chen et al., 2016). Adult songbirds also use non-auditory somatosensory feedback (from respiratory muscles associated with vocalization) to compensate for perturbations during song performance (Suthers et al., 2002).

Songbirds are one of the only animal models that, like humans, have evolved a specialized vocal neural pathway, allowing for the precise interrogation of the neural circuit mechanisms for vocal motor learning (Doupe and Kuhl, 1999; Jarvis, 2019). However, prior research on this topic has almost exclusively focused on one particular form of sensory feedback (auditory). These studies have revealed that songbird brains have a cortical-basal ganglia circuit, the Anterior Forebrain Pathway (AFP), that underlies auditory-guided vocal learning but not vocal production (Fig. 2.1 A) (Nordeen and Nordeen, 1993; Brainard and Doupe, 2013, 2000b; Mooney, 2009a). For example, lesions of LMAN (the output nucleus of the AFP) impair the juvenile song acquisition process (Bottjer et al., 1984) and prevent adult vocal plasticity in response to auditory feedback cues (Ali et al., 2013). Also lesions of dopaminergic input into the Area X (the basal ganglia nucleus of the AFP), also impair adult songbird vocal learning in response to the pitch-contingent delivery of auditory cues

(white noise bursts) (Hoffmann et al., 2016). Although recent work has demonstrated that songbird AFP receives anatomical projections from brain regions that process non-auditory sensory information (Paterson and Bottjer, 2017), it remains unknown whether auditory feedback is given a privileged role by this song learning neural circuit, or if non-auditory information is also processed by this circuit to drive vocal learning.

First, we tested whether adult songbirds can adaptively modify specific elements of their song structure in response to non-auditory feedback (Fig. 2.1 B, top). We developed a novel training paradigm in adult Bengalese finches. Our approach uses non-auditory cues (cutaneous electric shocks), which we deliver during ongoing song performance, to differentially reinforce specific elements of song (the pitch of an individual syllable). We used these non-auditory cues because they are fast enough to reinforce (in real time) specific elements of ongoing song performance and they allow for a very similar experimental approach as another well-established paradigm that uses auditory cues to drive changes in syllable pitch in adult songbirds (Tumer and Brainard, 2007; Andalman and Fee, 2009; Warren et al., 2011; Hoffmann et al., 2016). These prior studies that used auditory cues to drive adult songbird vocal learning have revealed important insights into how the songbird AFP processes auditory feedback to drive adaptive changes in vocal motor output (Warren et al., 2011; Andalman and Fee, 2009; Charlesworth et al., 2012; Hoffmann et al., 2016; Gadagkar et al., 2016). By using a similar experimental approach but with a non-auditory cue instead of an auditory one, we can effectively compare whether and how the songbird AFP processes these two different sources of sensory inputs.

We next tested the neural circuit mechanisms underlying non-auditory vocal learning (Fig. 2.1 B, middle). The AFP, the songbird vocal learning pathway, might be highly specialized for processing solely auditory feedback to drive vocal learning, especially considering the auditory nature of song output and the high level of importance

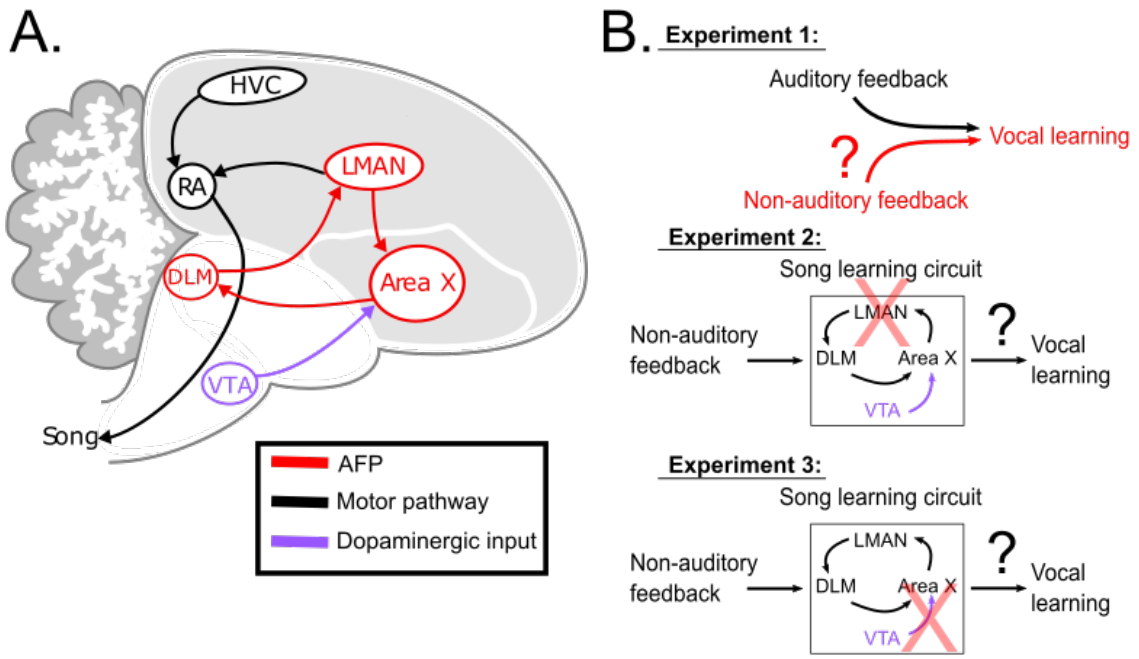


Figure 2.1: A. Songbird brain circuitry. Brain nuclei of the motor pathway – the neural circuit for vocal production – are black. Brain nuclei of the Anterior Forebrain Pathway (AFP) – the neural circuit for vocal learning – are red. VTA (purple) provides dopaminergic input into Area X, the basal ganglia nucleus of the AFP. B. The three primary hypotheses tested in this paper. In the first set of experiments, we tested whether non-auditory input plays a role in vocal output (Experiment 1). In the second set of experiments, we assessed the necessity of LMAN for non-auditory vocal learning (Experiment 2). In the third set of experiments, we tested the necessity of dopaminergic projections to Area X for non-auditory vocal learning (Experiment 3).

auditory feedback in particular for songbird behavior (Kuebrich and Sober, 2015). Prior studies have proposed that auditory and non-auditory feedback are processed by segregated neural pathways (Murdoch et al., 2018), and that the function of the AFP is to compensate for auditory performance errors by comparing auditory feedback from ongoing song performance with an internal goal for how song should sound (Fee and Goldberg, 2011). We tested the hypothesis that songbird AFP processes non-auditory feedback for vocal learning, as opposed to being specialized for the sole purpose of processing auditory feedback. We did so by determining the necessity of LMAN (the output nucleus of the AFP) for non-auditory vocal learning.

Finally, we assessed the role of dopaminergic neural circuitry in non-auditory vocal learning (Fig. 2.1 B, bottom). Recent work has established the contributions of dopamine to auditory-guided vocal learning. For example, lesions of dopaminergic input to Area X impair vocal learning in response to auditory cues in adult songbirds (Hoffmann et al., 2016). Moreover, dopamine neurons in songbird VTA respond to auditory stimuli associated with song learning (Gadagkar et al., 2016). We therefore assessed the role of dopaminergic neural circuitry for non-auditory vocal learning. We tested the hypothesis that dopaminergic neural circuitry underlies the processing of non-auditory feedback for vocal learning. We did so by performing selective lesions of dopaminergic input to Area X in adult songbirds to determine the necessity of this circuitry for non-auditory vocal learning.

2.3 Materials and Methods

All subjects were adult (>100 days old) male Bengalese finches (Lonchura striata var. domestica). All procedures were approved by Emory University's Institutional Animal Care and Use Committee. All singing was undirected (in the absence of a female bird) throughout all experiments.

2.3.1 Delivery of non-auditory sensory feedback

To deliver non-auditory feedback signals to freely-behaving songbirds during ongoing song performance, we first performed a surgery prior to any experimentation. Stainless steel wires were uninsulated at the tip (2-4 mm) and implanted subcutaneously on the bird's scalp. In 7 out of all 28 birds used across all experiments performed, wires were implanted intramuscularly in the birds' necks instead of on their scalps. The wires were soldered onto a custom-made circuit board that, during surgery, was placed on the bird's skull using dental cement. The circuit was connected to an electric stimulator (A-M Systems Isolated Pulse Stimulator), which produced pitch-contingent electrical currents through the wires implanted on the bird. We set the duration of these electric shocks to 50 ms, which was a long enough duration to overlap with a large portion of the targeted syllable, yet a short enough duration to avoid interfering with following song syllables. We typically set the magnitude of electric current used for producing the shocks to 100-350 µA, which is behaviorally salient (the first few instances of electric shock interrupt song), yet subtle enough as to not produce any body movements or signs of distress. Acute effects of electrical shock on song structure, such as pitch, amplitude, entropy, or syllable sequence, were assessed to ensure these non-auditory cues produced no immediate, systematic, acoustic effects. This ensures that any observed gradual changes to song structure in response to electric shock are due to non-auditory learning.

2.3.2 Vocal learning paradigm and song analysis

Experimental testing of vocal learning was performed by driving adaptive changes in the fundamental frequency (pitch) of song syllables. To do so, we delivered pitch-contingent, non-auditory feedback (cutaneous electric shock) to freely-behaving songbirds in real time during song performance. We followed the same experimental protocols as experiments using white noise feedback to drive vocal learning (Tumer and

Brainard, 2007; Hoffmann et al., 2016), except we used cutaneous electric shocks instead of white noise bursts. After surgically implanting the fine-wire electrodes, we recorded song continuously for three days without providing any experimental feedback (electric shocks or white noise bursts). We refer to this period as "baseline" (Fig. 2.2 A).

On the last (third) day of baseline, we measured the pitch of every rendition of the target syllable sung between 10 a.m. and 12 p.m. We set a fixed pitch threshold based on the distribution of these pitches, such that we would provide sensory feedback only when the pitch of a rendition of the target syllable was above the 20th percentile of the baseline distribution ("hit"), and all renditions outside of this range did not trigger any feedback ("escape"). In this case, an adaptive vocal change would therefore be to change the pitch of the target syllable down, thereby decreasing the frequency of triggering electric shocks. In other experiments, we triggered feedback on all renditions below the 80th percentile of the baseline pitch distribution. In this case, an adaptive vocal change would be to change the pitch of the target syllable up. For each experiment, we randomly selected which of these two contingencies we employed so we could assess bidirectional adaptations in vocal motor output. In a subset of experiments, we used the 90th percentile and 10th percentile pitch values to set the pitch threshold. Importantly, we also randomly withheld triggering feedback on 10 % of syllable renditions, regardless of syllable pitch or the experimental pitch-contingency. This allows us to compare syllable renditions that did or did not result in electric shocks to assess any acute effects of this form of feedback on syllable structure.

At 10 a.m. on the fourth day of continuous song recording, we began providing these pitch-contingent electric shocks in real time, targeted to specific song syllables sung within a specified range of pitches. We refer to this time period as "electric shock training" (Fig. 2.2 A). We used custom LabVIEW software to continuously

record song, monitor song for specific elements indicative of the performance of the target syllable, perform online, rapid pitch calculation, and trigger feedback in real time. The computers running this software were connected to an electric stimulator. When the electric stimulator received input from the LabVIEW software, it would then trigger a 50 ms burst of electric current through the implanted wire electrodes. During electric shock training, we continuously recorded song and provided pitch-contingent electric shocks at the set fixed pitch threshold for three days. During these three days, every time the bird sang within the "hit" range, an electric shock was immediately triggered.

After three days of electric shock training, we stopped providing electric shocks but continued recording unperturbed song for six additional days. We refer to this period as "washout" (Fig. 2.2 A). During washout, we consistently observed spontaneous pitch restoration back to baseline across all experiments, which is in congruence with results from numerous white noise learning experiments (Tumer and Brainard, 2007; Warren et al., 2011; Andalman and Fee, 2009; Hoffmann et al., 2016). This allows for multiple experiments to be performed from similar baseline conditions in the same individual songbird.

In 14 out of all 28 birds used throughout this study, we performed both white noise training and electric shock training in the same individual birds (Fig. 2.2 A). After the end of electric shock training and six days of washout (when the pitch of the target syllable had restored to baseline levels), we performed the exact same experimental protocol, but we used white noise feedback instead of electric shocks. We could then compare learning in response to two different sources of sensory feedback in the same individual subject. We also sometimes reversed the order of experimentation by performing white noise experiments first and electric shock experiments second. The order of experimentation was randomly decided for each songbird before beginning any white noise or electric shock training.

For all LMAN lesion (Fig. 2.3 A) and 6-OHDA lesion experiments(Fig. 2.4 A), we performed an electric shock training experiment prelesion. After six days of washout, we then performed surgery to lesion the neural circuit of interest. We then performed another electric shock experiment in the same individual bird using the exact same protocol we used prelesion. For all of these lesion electric shock experiments, we used the aforementioned electric shock training paradigm, but with one slight alteration: We extended the number of days of electric shock training and introduced a new methodology for setting the pitch threshold on these extended days of training. We still set a fixed pitch threshold based on analysis of the pitch distribution from the final day of baseline and performed three days of electric shock training using this fixed pitch threshold. We refer to this portion of the lesion experiments as "fixed" because the pitch threshold for determining electric shock feedback remained the same for all 3 days. Rather than stopping electric shock training at this point, we instead continued providing pitch-contingent electric shocks for an additional 1-5 days. In the morning (at 10 a.m.) on each of these extended days of electric shock training, we changed the pitch threshold to the 20th or 80th percentile (consistent with the initial contingency) of the pitch distribution of all renditions of the target syllable sung between 8 A.M. to 9:30 A.M. on that same day. As the bird changed the pitch of the target syllable in the adaptive direction, the new pitch thresholds continued to be set further and further in the adaptive direction to drive greater amounts of learning. We refer to these additional days as "staircase". After 1-5 days of staircase training, we stopped providing electric shocks and began the washout portion of the experiment. We used this experimental approach for both prelesion and postlesion experiments in our LMAN, 6-OHDA, and Sham data sets. Importantly, although the number of days of staircase varied between individual birds, for each individual bird we matched the same number of prelesion days of staircase and postlesion days of staircase to ensure that ,in both experimental conditions, the bird had an equivalent

amount of time and opportunity to learn.

Custom-written MATLAB software (The MathWorks) was used for song analysis. On each day of every experiment, we quantified important song features, such as the pitch, amplitude, and spectral entropy, of all renditions of the targeted syllable produced between 10 A.M. and 12 P.M. We did so to account for potential circadian effects on song production. To ensure a level of consistency in number of target syllable renditions measured on each day of an experiment, and to have a minimum number of syllable renditions necessary to get an accurate measure of average syllable pitch, we checked that at least 30 renditions of the target syllable were sung within the 10 A.M. to 12 P.M. window. If there were less than 30 renditions of the target syllable, we extended the time window for song analysis by 1 hour in both directions (9 A.M. to 1 P.M.) and then reassessed to see if there were at least 30 syllable renditions. If not, we continued this process of extending the time window by 1 hour until 30 song renditions were in that day's data set. Daily targeting sensitivity (hit rate) and precision (1- false-positive rate) were measured in all experiments to ensure accurate targeting of the specific target syllable (and not accidentally targeting different song syllables). During the pitch-contingent feedback portion of the experiment, a subset (10%) of randomly selected target syllables did not trigger feedback, regardless of syllable pitch. These "catch trials" allowed for the quantification and comparison of syllable features, such as pitch, amplitude, and entropy between trials when feedback was provided and trials when feedback was not provided. Pitch changes were quantified in units of semitones as follows: s = 12*log2(h/b), where s is the pitch change (in semitones) of the syllable, h is the average pitch (in Hertz) of the syllable, and b is the average baseline pitch (in Hertz) of the syllable.

2.3.3 Analysis of Variability in Syllable Pitch

We compared pitch variability pre- and postlesion using methods described in prior literature (Kao et al., 2005; Kao and Brainard, 2006; Hampton et al., 2009). We analyzed all song renditions (within the 10 A.M. - 12 P.M. time window) performed on the final day of baseline prelesion and on the final day of baseline postlesion. We did so in our LMAN lesion experimental group as well as our 6-OHDA lesion experimental group. To measure the variability in pitch of the song syllables, we calculated the Coefficient of Variation (C.V.) for the pitch of each syllable using the following formula: C.V. = (Standard Deviation/Mean) * 100.

2.3.4 LMAN Lesions

Birds were anesthetized under ketamine and midazolam and were mounted in a stereotax. The beak angle was set to 20° relative to the surface level of the surgery table. For stereotactic targeting of specific brain regions (in this case, LMAN), anterior-posterior (AP) and medial-lateral (ML) coordinates were found relative Y0, a visible anatomical landmark located at the posterior boundary of the central venous sinus in songbirds. Dorsal-ventral (DV) coordinates were measured relative to the surface of the brain. Bilateral craniotomies were made at the approximate AP coordinates 4.9 mm to 5.7mm and ML coordinates 1.5 mm to 2.5 mm. A lesioning electrode was then inserted 1.9 mm to 2.1 mm below the brain surface. These stereotactic coordinates targeted locations within LMAN. We then passed 100 µA of current for 60-90 seconds at 5-6 locations in LMAN in both hemispheres in order to electrolytically lesion the areas. This methodology was based on prior work involving LMAN lesions and LMAN inactivations (Ali et al., 2013; Warren et al., 2011; Andalman and Fee, 2009; Charlesworth et al., 2012; Kao et al., 2005; Kao and Brainard, 2006; Hampton et al., 2009). In sham operated birds, we instead performed small lesions in brain areas dorsal to LMAN. Again, this was consistent with methodology from prior studies

(Ali et al., 2013; Kao et al., 2005; Kao and Brainard, 2006).

Birds recovered within two hours of surgery and began singing normally (at least 30 renditions of target syllable within 2 hours) typically 3 to 8 days after surgery. Lesions were confirmed histologically using cresyl violet staining after completion of behavioral experimentation. In tissue from sham operated birds, we identified Area X and LMAN based on regions of denser staining as well as well-characterized anatomical landmarks (Karten et al., 2013).

2.3.5 6-OHDA Lesions

Birds were anesthetized using ketamine and midazolam and were mounted in a stereotax, where the beak angle was set to 20° relative to the surface level of the surgery table. Isofluorine was used in later hours of the surgery to maintain an anesthetized state. Bilateral craniotomies were made above Area X from the approximate AP coordinates 4.5 mm to 6.5mm and ML coordinates 0.75 mm to 2.3 mm.

In each hemisphere, we inserted a glass pipette containing a a 6-OHDA solution (see below) and made 12 pressure-injections in a 3 mm x 4 mm grid between AP coordinates 5.1 mm and 6.3 mm, ML coordinates 0.9 mm and 2.2 mm and the DV coordinate 3.18 mm. Additional bilateral 6-OHDA injections were made at the AP coordinate 4.8 mm, ML coordinate +/- 0.8 mm, and DV coordinate 2.6 mm from the brain surface to lesion the most medial portion of Area X. Each injection consisted of 13.8 nL of 6-OHDA solution, injected at a rate of 23 nL/s at each site. The pipette was kept in place for 30 seconds after each injection and was then slowly removed.

Again, birds recovered within two hours of surgery and began singing normally (at least 30 renditions of target syllable within 2 hours) typically 3 to 8 days after surgery. 6-OHDA solution was prepared using 11.76 mg 6-OHDA-HBr and 2 mg ascorbic acid in 1 mL of 0.9% normal saline solution. The solution was light-protected after preparation to prevent oxidation.

2.3.6 Histology

Between 14 and 54 days after surgery, birds were injected with a lethal dose of ketamine and midazolam and were perfused. Tissue was post-fixed in 4% paraformaldehyde at room temperature for 4-16 hours and then moved to a solution of 30% sucrose for at least one day at 4°C for cryoprotection. Then, brain tissue was sliced in 40 µm sections. A chromogenic tyrosine hydroxylase (TH) stain was used to quantify the depletion of catecholaminergic fiber innervations in tissue collected from 6-OHDA lesioned birds, and Nissl and fluorescent NeuN staining was used to assess the density of cell bodies in tissue from LMAN lesioned and sham operated birds. For one bird in the 6-OHDA lesioned group, a Nissl stain was performed on alternate tissue sections to ensure no cell death occurred as a result of the lesion.

For TH immunohistochemistry, tissue was incubated overnight in a primary anti-TH antibody solution. The tissue was next incubated in biotinylated horse anti-mouse secondary antibody solution for 1 hour. Then, the tissue was submerged in a diaminobenzidine (DAB) solution (2 DAB tablets, Amresco E733 containing 5 mg DAB per tablet, 20 mL Barnstead H2O, 3 µmL H2O2) for less than 5 minutes for vizualization. The DAB solution was prepared 1h prior to use. Tissue was washed, mounted and coverslipped using Permount mounting medium.

Birds were perfused between 14 and 54 days after surgery using 10% heparin in 0.9% normal saline solution (0.9% NaCl and deionized water) followed by 4% paraformaldehyde. Tissue was post-fixed in 4% paraformaldehyde at room temperature for 4-16 hours and then moved to a solution of 30% sucrose for at least one day at 4 degrees C for cryoprotection. The tissue was sliced in 40 µm sections using a freezing sliding microtome and a chromogenic TH stain was performed. For one bird a Nissl stain was performed on alternate tissue sections to ensure no cell death occurred as a result of the lesion.

2.3.7 Tyrosine Hydroxylase Stain

Between each incubation tissue was washed with 0.1 M phosphate buffer (PBS) (23 g dibasic sodium phosphate, 5.25 g monobasic sodium phosphate, and 1 L deionized H2O) 3 times for 10 min each. Tissue was first washed and then incubated in 0.3% H_2O_2 for 30 min and then 1% $NaBH_4$ for 20 min, followed by overnight incubation in a primary anti-tyrosine hydroxylase antibody solution. The tissue was next incubated in biotinylated horse anti-mouse secondary antibody solution for 1 h, then incubated in avidin-biotin-complex (ABC) solution for 1 h that had been prepared 30 min prior to use. The tissue was then submerged in a diaminobenzidine (DAB) solution for less than 5 minutes. Tissue was then washed, mounted and coverslipped using Permount mounting medium. These TH stains mark neurons expressing TH, which are catecholaminergic.

2.3.8 Nissl Stain

Tissue was washed in 0.1 M PBS three times for 10 minutes and was then mounted. The slides were incubated in citrisolv twice for 5 min each, then delipidized in the following ethanol concentrations for two minutes each: 100%, 100%, 95%, 95%, and 70%. The tissue was briefly (less than 15 s) rinsed in deionized water, then was incubated in cresyl violet (665 µmL glacial acetic acid, 1 g cresyl violet acetate, and 200 mL deionized water) for 30 min. The tissue was rinsed in deionized water, then briefly (less than 15 seconds) submerged in the following ethanol concentrations for 2 min each: 70%, 95%, 95%, 100%, and 100%. The tissue was then incubated in citrisolv twice for 5 min. The tissue was coverslipped using Permount mounting medium. These nissl stains mark neuronal cell bodies.

2.3.9 NeuN Antibody Stain

Between each incubation tissue was washed with 0.1 M PBS 3 times for 10 min each. Tissue was incubated in primary antibody solution (4 mL EMD Millipore guinea pig anti-NeuN Alexa Fluor 488 antibody, 6 mL Triton X-100, 20 mL normal donkey serum (NDS) and 1.95 mL 0.1 M PBS) overnight. The tissue was then washed and incubated in a secondary antibody solution (10 mL Jackson Labs donkey anti-guinea pig (DAG), 6 mL Triton X-100, and 1.975 mL 0.1 PBS) overnight. Tissue was then washed, mounted and coverslipped with Flurogel mounting medium. Slides were sealed with lacquer. Images were taken under a widefield microscope (BioTek Lionheart FX, Sony ICX285 CCD camera, Gen5 acquisition software, 1.25x magnification, 16-bit grayscale).

2.3.10 Lesion Analysis

Analysis of lesions was based on previously published methodology (Hoffmann et al., 2016; Saravanan et al., 2019). Images of stained tissue sections were obtained using a slide scanner and were converted into 8-bit grayscale images in ImageJ. In our group of control (unlesioned) birds, where no lesions or invasive brain operations were performed, Area X stains darker than surrounding striatum in TH-DAB-stained tissue due to a higher density of catecholaminergic inputs in Area X (Hoffmann et al., 2016). The baseline level of stain darkness can vary from bird to bird. Therefore, rather than directly comparing the stain density of lesioned and unlesioned tissue, the ratio of the stain density of Area X to that of the surrounding striatum (OD ratio) was calculated to determine whether the concentration of catecholaminergic fibers was decreased. Prior work demonstrated that the vast majority of catecholaminergic input to Area X is dopaminergic (Hoffmann et al., 2016).

For each section of tissue containing Area X, a customized ImageJ macro was used to select regions of interest (ROIs) within Area X and within a portion of striatum

outside Area X by manually outlining Area X and selecting a circular 0.5 mm-diameter region of striatum anterior to Area X. Pixel count and optical density (OD) of each ROI were measured, and the density of TH-positive fibers was calculated using the ratio of the OD of Area X to the OD of non-X-striatum.

The cumulative distribution of OD ratios for control birds was used to construct a 95% confidence interval and determine the threshold for lesioned tissue. 6-OHDA-lesioned tissue in which the OD ratio fell below the 5th percentile of control tissue had a significantly reduced TH-positive fiber density.

2.3.11 Statistical Testing

All error bars presented in the main text represent SEM. When assessing whether a significant amount of vocal learning occurred in one experiment, we used one-sample t-tests to compare the mean pitch on the final day of training vs zero. To assess whether a significant difference in amount of learning occurred within an individual bird pre- vs postlesion, we used paired t-tests. To assess for significance between distributions of target syllable pitches on various days of the experiment (Baseline, shock, washout), we used a 2-sample KS test.

Each experimental group had at least five birds, and for each bird, the target syllable was typically repeated well over 30 times a day. Therefore, the structure of our data is hierarchical, so error accumulates at different levels (birds and syllable iterations). Simply grouping all the data together ignores the non-independence between samples and underestimates the error. To address this issue, we employed the hierarchical bootstrap method to measure SEM and calculate p-values for adaptive pitch change data sets (Saravanan et al., 2020; Efron and Tibshirani, 1994). For each experimental day we calculated normalized pitch values (in semitones) for each day of a learning experiment (normalized to the mean pitch on the final baseline day during that particular experiment). We then generated a population of 10000 bootstrapped

means. To generate each individual subsample, we resampled across each level of hierarchy in our data (first resampled among the birds, then for each selected bird, we resampled among syllable iterations). The standard deviation of this population of bootstrapped means provides an accurate estimate of the uncertainty of the original data (Saravanan et al., 2019). Thus, the SEM values (which are used for error bars) we report when employing the hierarchical bootstrap method are equal to this standard deviation.

To calculate p-values and determine significance for comparing my data to zero using the hierarchical bootstrap method, we calculated p_{boot}: the proportion of boostrapped means greater than zero compared to the total number of bootstrapped means. Using an acceptable type-1 error rate of .05, any value of this p_{boot} ratio greater than .975 indicates the mean was significantly greater than zero and any value less than .025 indicates the mean was significantly less than zero. p_{boot} values between .025 and .975 indicate no significant difference between the data set and zero. Because we measure adaptive pitch changes in semitones, which are a normalized measure of pitch change where baseline is set to zero, this method of calculating p_{boot} was employed in all instances where it was necessary to assess whether there was a significant change in pitch at the end of training compared to baseline (zero).

We also sometimes sought to determine significance for the comparison of two means rather than what was previously described (where we assess significance between one mean compared to baseline (zero). We used a similar hierarchical bootstrap statistical methodology and calculated p_{boot}. The key difference is that, rather than measuring the proportion of resampled means greater than or less than zero, we instead calculate a joint probability distribution for the means of the two resampled data sets. We measured the percentage of this joint probability distribution that was above one side of the unity line. This percentage is the p_{boot} values we report in these instances. If the proportion of this joint probability distribution that falls above the

unity line is greater than .975, it indicated a significantly greater mean of data set 1 over data set 2. If the percentage of the joint probability distribution that was above the unity line was less than .025, it indicated a significantly lower mean of data set 1 compared to data set 2. p_{boot} values between .025 and .975 indicate no significant difference between the two data sets. This method was employed in all instances where it was necessary to assess whether the learning magnitudes (adaptive pitch changes by the end of training) were significantly different pre- vs postlesion (or pre- vs postsham) or across experimental conditions (e.g., postsham vs postlesion or post LMAN lesion vs post 6-OHDA lesion).

In both forms of p_{boot} calculation, the lowest statistical limit for p_{boot} is $p_{boot} < 10^{-4}$, due to resampling 10^4 times to create bootstrapped means. The highest possible limit for p_{boot} is $p_{boot} > (1.000 - 10^{-4})$, for the same reason.

2.4 Results

2.4.1 Non-auditory feedback is sufficient to drive adult vocal learning

We tested whether non-auditory feedback drives vocal learning by providing pitch-contingent electric shocks through a set of wire electrodes on the scalps of adult songbirds. Before initiating electric shock training, we continuously recorded song without providing any feedback for three days (baseline) (Fig. 2.2 A). Every day, songbirds naturally produce many renditions of song, which consists of repeated patterns of unique vocal gestures, called syllables (Fig. 2.2 B, top). For one "target" syllable in each experimental subject, we quantified rendition-to-rendition variability in the fundamental frequency of each occurrence of this syllable on the final baseline day (Fig. 2.2 B, top). To differentially reinforce the pitch of a target syllable, we determined a range of pitches within this baseline distribution (either all pitches above

the 20th percentile or all pitches below the 80th percentile), and then triggered the delivery of cutaneous electric shocks in real time when the pitch of the target syllable fell within this range (Fig. 2.2 B, bottom). We performed this pitch-contingent electric shock training continuously for three days.

For example, in one experiment (shown in Fig. 2.2 A-D), electric shocks were triggered every rendition of the target syllable that had a pitch above 2.13 kHz (the 20th percentile of the baseline distribution) for three days. In this example experiment, the songbird gradually changed the pitch of the targeted syllable downwards (the adaptive direction), such that electric shocks were triggered less frequently (Fig. 2.2 C). In other experiments where the adaptive direction of pitch change is upwards, we triggered electric shock whenever the target syllable pitch was below the 80th percentile of this distribution. In the example experiment, at the start of the first day of electric shock training, 80% of syllable renditions resulted in electric shock and 20% of syllable renditions resulted in escapes. On the third (final) day of electric shock training, escapes occurred on over 60% of target syllable renditions and the entire distribution of pitches had changed significantly in the adaptive direction, indicating that a significant amount of vocal learning occurred in this example experiment (Fig. 2.2 D; 2-sample KS test to assess the difference between baseline and end of electric shock training, p = 1.1776e-12). We then stopped triggering electric shock feedback and continued to record unperturbed song for six additional days (washout). After six days of washout, there was no significant difference between the distribution of target syllable pitches at the end of washout compared to baseline (Fig. 2.2 D; 2-sample KS test, p = 0.606). For analysis of washout across all experiments, see Supplemental Fig. 2.5 in Section 2.6.

In order to assess whether non-auditory feedback is sufficient to drive vocal learning across multiple songbirds, we first measured the adaptive pitch change (in semitones) for each individual experiment. Semitones are a normalized measurement of a

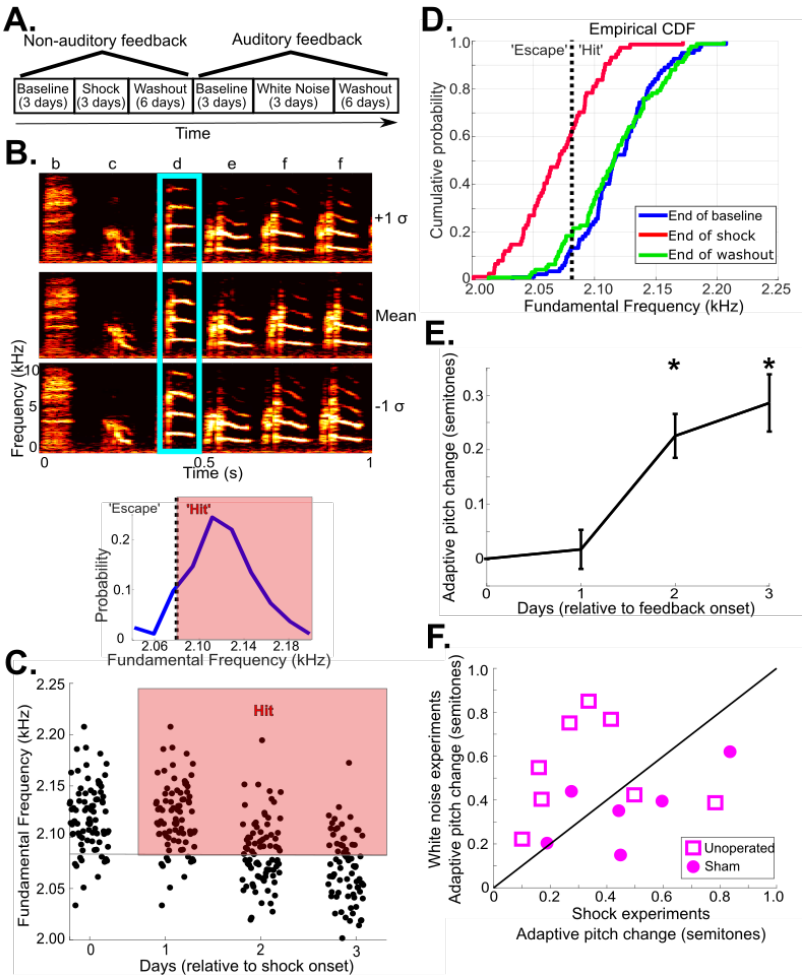


Figure 2.2: Non-auditory feedback drives vocal learning. A. Timeline of auditory and non-auditory vocal learning experiments in the same bird. The order of experiments varied between birds. B. (Top) Spectrograms showing example renditions of birdsong on one day. Syllables are labeled by an experimenter (letters b through f) and a target syllable was chosen ("d"). (Bottom) Pitch threshold was set based on the baseline pitch distribution. Electric shocks were provided in real time during renditions of the target syllable above a chosen pitch threshold ("hit"). C. Birds underwent three days of electric shock training with a fixed pitch threshold. Each dot represents the pitch of one rendition of the target syllable (from same experiment shown in B). Every rendition in the 'hit' range triggered an immediate electric shock. D. CDFplot showing the probability a value of pitch from a distribution falls at or below the value on the x-axis. The different pitch distributions are from different portions of the same example experiment in B and C. End of electric shock training was significantly greater than baseline (2-sample KS test, p = 1.178e-12). End of washout distribution was not significantly different from baseline (2-sample KS test, p = 0.606) E. Adaptive pitch change (in semitones) of the target syllables during electric shock training (n = 13 experiments). The mean change during training was significantly greater than baseline (p_{boot} >(1.000 - 10^{-4}) on training days 2 and 3, where p_{boot} >.975 indicates the mean is significantly greater than baseline). F. Learning magnitudes (adaptive pitch change by end of training) during white noise training compared to during electric shock training (n = 14 experiments). Each dot represents the learning magnitudes from an individual bird. No significant difference in learning magnitudes during electric shock training vs during white noise training (paired t-test, p = 0.313).

change in pitch in the adaptive direction (i.e., the direction that results in less frequent triggering of electric shocks). We next we employed a hierarchical bootstrap approach to measure SEM and assess significance (see Methods) (Saravanan et al., 2019, 2020; Crowley, 1992; Efron and Tibshirani, 1994). We do so because this method more accurately quantifies the error in hierarchical data (e.g., many renditions of a target syllable collected across multiple birds). Briefly, to determine significance, we calculated the the proportion of bootstrapped means (p_{boot}) that were >0.975. This indicates that the mean pitch of this data set is significantly greater than baseline (zero). If p_{boot} <0.025, then the mean pitch of this data set is significantly less than baseline (zero). We performed this statistical test on our hierarchical data collected across 13 electric shock experiments, 12 birds, and at least 30 renditions of the target syllable per day. We found that the mean pitch (in semitones) of the target syllables showed a significant, adaptive change from baseline on days two and three of electric shock training (Fig. 2.2 E; p_{boot} >(1.000 - 10^{-4}), limit due to resampling 10^4 times, on electric shock training days 2 and 3. Here, p_{boot} >.975 indicates a mean significantly greater than baseline). This demonstrates that non-auditory feedback is sufficient to driving vocal learning in adult songbirds.

To compare vocal learning in response to different sources of sensory feedback (auditory and non-auditory), we performed multiple learning experiments - one electric shock and one white noise - in 8 out of the 12 individual birds from this data set (Fig. 2.2 A). We randomized the order of white noise training and electric shock training for the birds who underwent both training paradigms. We also included 6 sham operated birds from a later set of experiments in this particular analysis. We did so because the sham operated birds had intact song systems and underwent both electric shock and white noise training.

Consistent with prior studies (Tumer and Brainard, 2007; Hoffmann et al., 2016; Ali et al., 2013), by the end of white noise training, the adaptive pitch change (in

semitones) across all white noise experiments performed in unoperated birds (birds who had wire electrodes surgically implanted but received no invasive brain procedures like sham lesions) was significantly greater than baseline (zero) (Supplemental Fig. 2.6 A; $p_{boot} > (1.000 - 10^{-4})$ on all three white noise training days, where $p_{boot} > .975$ indicates the mean is significantly greater than baseline). In the separate experimental group of birds that underwent sham lesions, we also observed significant adaptive pitch changes in response to white noise bursts, as expected (Supplemental Fig. 2.6 B, $p_{boot} > (1.000 - 10^{-4})$ on all three white noise training days). Interestingly, we observed individual variability in learning magnitudes (adaptive pitch change at the end of training) during electric shock and white noise experiments (Fig. 2.2 F). However, we found no systematic differences between learning magnitude during electric shock training and the learning magnitude during white noise training (Fig. 2.2 F; paired t-test, $p = 0.313$). The result that learning magnitude during white noise training is roughly comparable to learning magnitude during electric shock training suggests that similar neural circuitry may underlie these two forms of vocal learning.

2.4.2 LMAN is required for non-auditory vocal learning

We next investigated the neural circuitry that processes non-auditory feedback to drive vocal learning. To assess whether the AFP is required for non-auditory vocal learning, we measured the effect of lesions of LMAN, the output nucleus of the AFP, on learning magnitude in response to non-auditory feedback. We performed electric shock training experiments in the same individual birds before and after bilateral, electrolytic LMAN lesions or sham lesions (Fig. 2.3 A, $n = 5$ birds). To perform electric shock training in this group of experiments, we used the same protocol described previously, except we extended the period of electric shock training by 1-5 days. During this extended training period (called "staircase"), we set a new pitch threshold each morning to drive even greater amounts of learning (see Methods). We observed

that this methodology drove significant amounts of learning in adult songbirds with intact song systems prelesion (Fig. 2.3 B).

Behavioral measures indicated that LMAN was effectively lesioned in the birds in our LMAN lesion data set. LMAN lesions in adult songbirds produce a significant decrease in the trial-to-trial variability in song syllable pitch (Kao et al., 2005; Kao and Brainard, 2006; Hampton et al., 2009). To assess lesion-induced changes in the variability of syllable pitch between conditions (LMAN lesion and sham), we calculated the CV of syllable pitch pre- and postlesion (see Methods). We found that LMAN lesions induced a significant decrease in syllable CV (Fig. 2.3 C; paired t-test, $p = 0.002$). Sham lesions did not induce a significant change in syllable CV (Fig. 2.3 C; paired t-test, $p = 0.911$). The lesioned-induced change in syllable CV (post/pre) was significantly greater for LMAN lesioned birds than for sham (Supplemental Fig. 2.7 A; 2 sample KS test, $p = 0.006$). We also performed postlesion white noise training across conditions (LMAN lesion and sham) (Supplemental Fig. 2.6 B). After LMAN lesions, songbirds did not significantly change the pitch of the target syllable from baseline (zero) (p_{boot} >0.223 on all of the final four days of training). Postsham, songbirds did significantly change the pitch of the target syllable in the adaptive direction (p_{boot} <10<-4 on each of the final four days of training). This indicates that our LMAN lesions induced significant deficits in auditory vocal learning, consistent with previous work that demonstrated that electrolytic LMAN lesions inhibit auditory vocal learning (Ali et al., 2013).

Histological analysis (see Methods) further confirmed the lack of LMAN in stained brain tissue from birds in our LMAN lesion data set. Our histology methodology followed previous literature involving LMAN lesions (Kao et al., 2005; Ali et al., 2013; Hoffmann et al., 2016). Briefly, we performed Nissl stains to stain for neuronal cell bodies in brain slices after experiments were complete (Supplemental Fig. 2.8, A). We then calculated the optical density ratio of the region containing LMAN compared

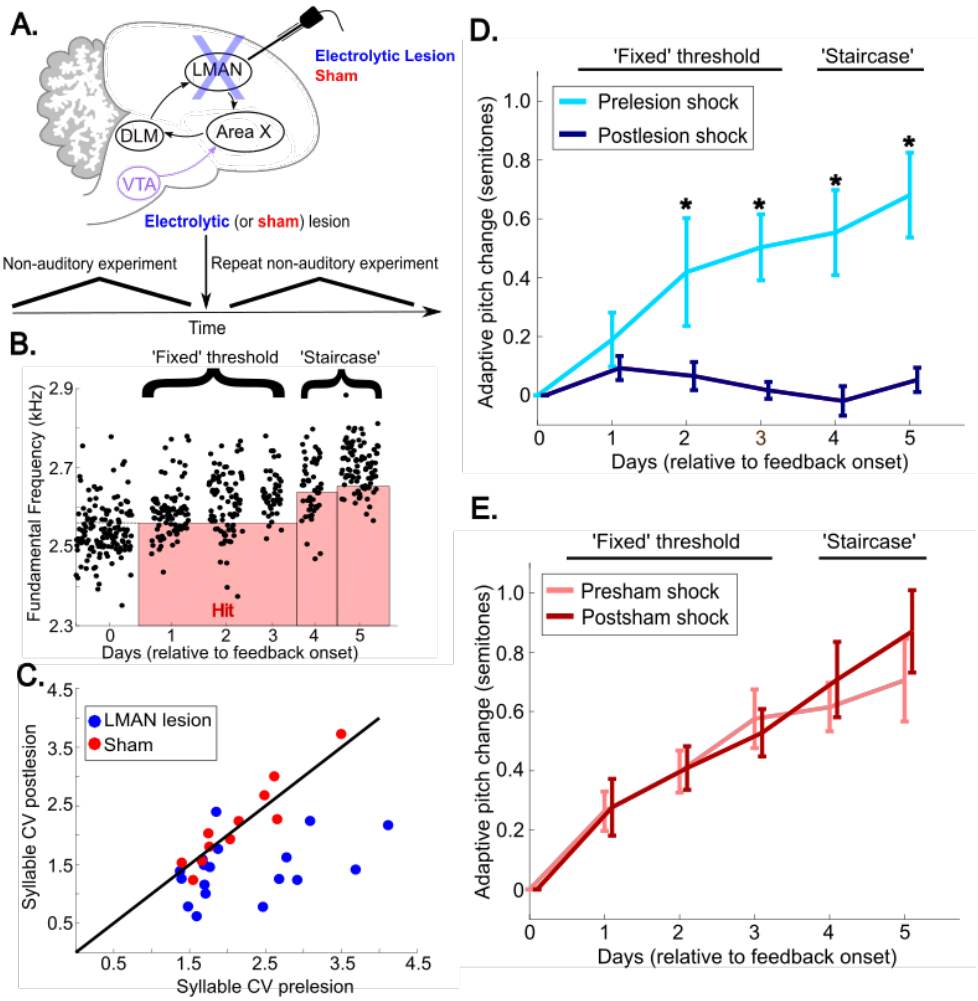


Figure 2.3: LMAN is required for non-auditory vocal learning. A. Electrolytic lesions of LMAN were performed. Pre- and postlesion experimental timeline in individual birds. B. Prelesion experiment. Electric shock training consisted of three days of using a fixed pitch threshold (determined from the baseline pitch distribution) then additional days of staircase where the pitch threshold was changed (based on the pitch distribution from that same morning) once per day. Each dot represents the pitch of a rendition of the target syllable. Any pitch that fell in the 'hit' range resulted in an electric shock. C. CV of song syllables pre- vs postlesion and pre- vs postsham. LMAN lesions induced a significant reduction in syllable CV (paired t-test, $p = 0.0015$). Sham lesions did not induce a significant change in syllable CV (paired t-test, $p = 0.9106$) D. Adaptive pitch change (in semitones) during electric shock training ($n = 5$ LMAN lesioned birds). Prelesion learning magnitude was significantly greater than baseline ($p_{boot} > (1.000 - 10^{-4})$ on each of the final four days of training, where $p_{boot} > .975$ indicates the mean is significantly greater than baseline). Postlesion learning magnitude did not significantly differ from baseline ($0.297 < p_{boot} < 0.660$ on each of the final four days of training, where $0.025 < p_{boot} < 0.975$ indicates no significant difference). Prelesion learning magnitude was significantly greater than postlesion learning magnitude ($p_{boot} < 10^{-4}$ on each of the final three days of training). E. Adaptive pitch change (in semitones) during electric shock training ($n = 6$ sham operated birds). Learning magnitudes were significantly greater than baseline both pre- and postsham ($p_{boot} > (1.000 - 10^{-4})$ on each of the final four days of training, where $p_{boot} > .975$ indicates the mean is significantly greater than baseline). Learning magnitudes pre- vs postsham did not significantly differ ($0.143 < p_{boot} < 0.582$ on each of the final four days of training).

to background (a pallial region outside of LMAN) (Supplemental Fig. 2.8, B). The distribution of OD ratios from LMAN lesioned tissue was significantly less than the OD ratios from sham lesioned tissue (Supplemental Fig. 2.8 C; 2 sample KS test, $p < 0.001$). This suggests that the density of neuronal cell bodies within LMAN was reduced following electrolytic lesions compared to following sham. Exactly like a prior study, we also qualitatively assessed each slice of brain tissue to measure the percentage of intact LMAN remaining in the tissue (Ali et al., 2013). We found that all of our LMAN lesioned birds had 80-100% of LMAN lesioned in both hemispheres. Together, these behavioral and histological results suggest that our lesion technique effectively damages LMAN neurons, and our sham technique effectively spares neurons in the region.

LMAN lesions significantly impaired non-auditory vocal learning. Prelesion, songbirds adaptively changed the pitch of the target syllable away from baseline in response to non-auditory feedback ($p_{boot} > (1.000 - 10^{-4})$ on all of the final four days of training, where $p_{boot} > 0.975$ indicates the data are significantly greater than baseline)(Fig. 2.3 D). Postlesion, this non-auditory vocal learning was abolished in those same birds ($0.297 < p_{boot} < 0.660$ on all of the final four days of training, where $0.025 < p_{boot} < 0.975$ indicates no significant difference, n = 5 birds) (Fig. 2.3 D). Learning magnitude prelesion was significantly greater compared to learning magnitude postlesion ($p_{boot} < 10^{-4}$ on each of the final three days of training). Both pre- and postsham, songbirds displayed significant amounts of learning during electric shock training (Fig. 2.3 E, $p_{boot} > (1.000 - 10^{-4})$ on all of the final four days of training for both presham and postsham datasets, where $p_{boot} > 0.975$ indicates the data are significantly greater than baseline, n = 6 birds). Also, the learning magnitudes during electric shock training did not significantly differ pre- vs postsham ($0.143 < p_{boot} < 0.582$ on the final four days of training).

We also directly compared the lesion-induced change in learning magnitudes be-

tween conditions (LMAN lesion vs sham) (see Supplemental figures, Section 2.6). First, we calculated learning magnitude at the end of the fixed threshold training period across conditions. The lesion-induced change in learning magnitude (post – pre) for LMAN lesioned birds was significantly greater than sham (Supplemental Fig. 2.7 B; 2 sample KS test, $p = 0.0361$). Next, we calculated learning magnitude at the end of the extended "staircase" portion of electric shock training across conditions. The lesion-induced change in learning magnitude (post – pre) for LMAN lesioned birds calculated at this time point was also significantly greater than sham (Supplemental Fig. 2.7 C; 2 sample KS test, $p = 0.0038$). These results indicate that LMAN is required for non-auditory vocal learning in adult songbirds, suggesting that non-auditory sensory feedback is processed by the AFP.

2.4.3 Dopaminergic input to Area X is required for non-auditory vocal learning

We next assessed dopaminergic contributions to non-auditory vocal learning. Learning magnitude during electric shock training was assessed before and after bilaterally lesioning dopaminergic projections in Area X, the basal ganglia nucleus of the AFP, in individual songbirds (Fig. 2.4 A, $n = 5$ birds). Selective lesions of dopaminergic projections in Area X were performed via bilateral 6-OHDA injections in Area X (see Methods)(Fig. 2.4 B, C). This approach has previously been shown to selectively lesion dopaminergic inputs to Area X without damaging non-dopaminergic cells (Hoffmann et al., 2016; Saravanan et al., 2019).

In order to confirm the effectiveness of our 6-OHDA injections at lesioning dopaminergic input to Area X, we quantified the extent of the reduction of catecholaminergic fiber innervation within Area X after completing the behavioral experimentation in each bird (Hoffmann et al., 2016; Saravanan et al., 2019). To visualize dopaminergic innervation, we labeled tissue with a common biomarker for catecholaminergic cells

(Fig. 2.4 B). To determine whether the concentration of dopaminergic fibers in Area X had decreased, we measured the optical density ratio (OD): the ratio of the stain density of Area X to the stain density of the surrounding striatum. The OD of Area X in 6-OHDA lesioned brain tissue was, on average, 44.7% lower than the OD of Area X in brain tissue from control birds. The distribution of all OD ratios from all of our 6-OHDA lesioned tissue was significantly lower than that of our control (unlesioned) birds (Fig. 2.4 C; 2 sample KS test, p = 5.75e-9). Also, OD ratios from individual 6-OHDA lesioned brains decreased compared to control (Supplemental Fig. 2.9, A). These results are similar to previous reports that used 6-OHDA injections to lesion dopaminergic input to Area X (Hoffmann et al., 2016; Saravanan et al., 2019), and they indicate that our 6-OHDA injections successfully lesioned dopaminergic input to Area X.

Depletion of dopaminergic input to Area X significantly impaired non-auditory vocal learning. Prelesion, songbirds adaptively changed the pitch of the target syllable during electric shock training (p_{boot} >0.990 on all of the final four days of training, where p >0.975 indicates the data are significantly greater than baseline) (Fig. 2.4 E). Postlesion, these same songbirds were not able to adaptively change the pitch of the target syllable during electric shock training (0.067 <p_{boot} <0.019 on days 1-4 of training and p_{boot} <10^{-4} on the final day of training, where p_{boot} <0.025 indicates the mean of the dataset is significantly less than baseline, n = 5 birds). Learning magnitude prelesion was significantly greater compared to learning magnitude postlesion (p_{boot} <10^{-4} on each of the final three days of training). These results suggest that dopaminergic input to Area X is required for non-auditory vocal learning.

We again measured the variability of syllable pitch pre- and postlesion by calculating syllable CV. The pitch variability of the targeted syllable was not significantly affected by dopaminergic lesions in Area X (Fig. 2.4 D; paired t-test, p = 0.6192). This finding is consistent with prior work using similar 6-OHDA injections to lesion

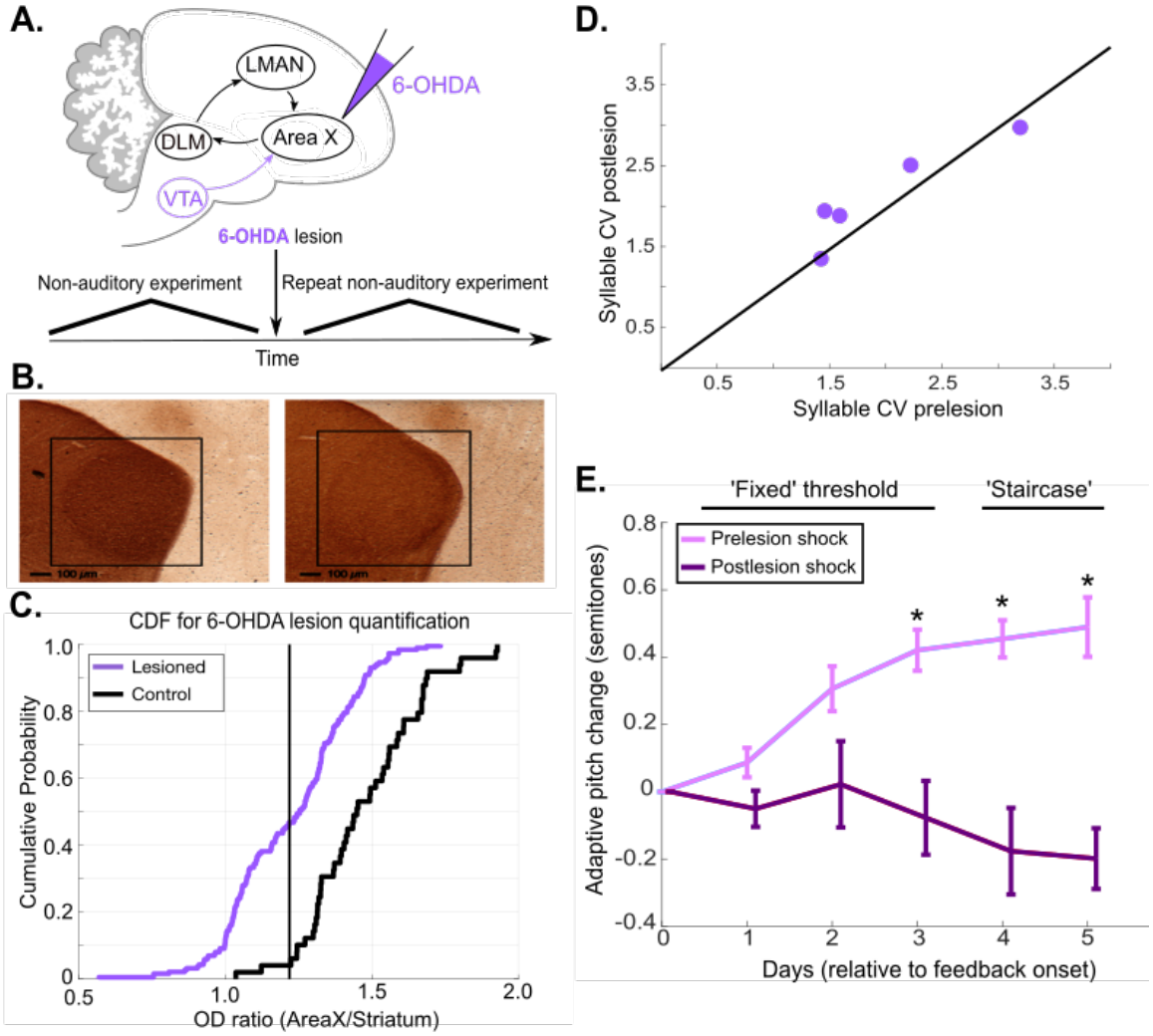


Figure 2.4: Dopaminergic input to Area X is required for non-auditory vocal learning. A. 6-OHDA injections into Area X were performed. Pre- and postlesion experimental timeline in individual birds. B. Example TH stains from control (left) and 6-OHDA lesioned (right) brain sections. The optical density (OD) ratio was calculated by comparing the TH stain in Area X (black) to that of outside striatum. C. Comparison of OD ratios in 6-OHDA lesioned and unlesioned (Control) birds. Roughly 50 % of lesioned sections had an OD ratio below the 95th percentile in control tissue (vertical line). D. CV of song syllables did not significantly differ pre- vs postlesion (paired t-test, $p = .6192$) E. Adaptive pitch change (in semitones) during electric shock training ($n = 5$ 6-OHDA lesioned birds). Prelesion learning magnitude was significantly greater than baseline ($p_{boot} > 0.990$ on each of the final four days of training, where $p_{boot} > .975$ indicates the mean is significantly greater than baseline). Postlesion learning magnitude did not significantly differ from baseline except for on the final day, when the mean had changed in the anti-adaptive direction ($p_{boot} > 0.067$ on training days 1-4, $p_{boot} < 10^{-4}$ on training day 5). Prelesion learning magnitude was significantly greater than postlesion learning magnitude ($p_{boot} < 10^{-4}$ on each of the final three days of training).

dopaminergic to Area X (Hoffmann et al., 2016). Prior work has strongly linked dopamine in songbird AFP with the generation of variability in syllable pitch in adult songbirds (Leblois et al., 2010; Murugan et al., 2013; Simonyan et al., 2012; Sasaki et al., 2006). It is likely that our dopamine lesion methodology, which spares about 50% of the dopaminergic input to Area X (Hoffmann et al., 2016), is sufficient to impair vocal learning but insufficient to impair dopamine-mediated generation of syllable variability. The result that our dopamine lesions impair vocal learning but not vocal variability suggests that learning deficits observed following lesions of AFP circuits are not simply due to decreased pitch variability.

Lesion size was quantified by determining the proportion of 6-OHDA lesioned tissue that had an OD ratio of Area X to non-X striatum that was less than the fifth percentile of OD ratios in control tissue. There was not a significant correlation between lesion size and the lesion-induced change in learning magnitude (post-pre) (Supplemental Fig. 2.10, $R2 = 0.019$, $p = 0.137$).

2.5 Discussion

Our results demonstrate that non-auditory feedback drives vocal learning in adult songbirds, and that the AFP and its dopaminergic inputs are required for non-auditory vocal learning. We first demonstrated that adult songbirds learn to adaptively change the pitch of their song syllables in response to cutaneous electric shocks (Fig. 2.2). We next demonstrated that LMAN, the output nucleus of the AFP, is necessary for the expression of this non-auditory vocal learning (Fig. 2.3). Finally, we showed that dopaminergic input to Area X, the basal ganglia nucleus of the AFP, is necessary for non-auditory vocal learning (Fig. 2.4). These results show that songbird AFP is not specialized just for processing auditory feedback for vocal learning, as has previously been hypothesized (Murdoch et al., 2018). Instead, our results demon-

strate that the AFP processes auditory feedback as well as non-auditory feedback to drive vocal learning. The fact that our electric shock stimulus is different than the direct proprioceptive feedback from vocal muscles or vocal effectors, yet the AFP still underlies vocal learning in response to electric shocks, suggests that the AFP can integrate sensory information from a wide variety of sources of sensory feedback, even those not directly produced by vocalizations.

The result that the AFP processes sensory feedback from electric shocks on the scalp suggests that neural pathways exist for cutaneous somatosensory information to enter the song system. The neuroanatomical pathways for auditory feedback to enter the AFP are well-characterized. For example, recent work has demonstrated that songbird ventral pallidum (VP) receives input from auditory cortical areas, encodes auditory feedback information, and projects to VTA (Chen et al., 2019). This represents a likely pathway by which sensory information from white noise bursts could influence neural activity in VTA, which could then drive changes in the AFP that promote song learning. Comparatively less is known about pathways in the songbird brain that might carry sensory information from the electric shocks. Our results showing that dopaminergic input to Area X (which originates in VTA) is necessary for non-auditory vocal learning suggests that pathways for non-auditory information ultimately project to VTA, where this information can be encoded and transmitted to the AFP to drive learning. For further discussion regarding how cutaneous information from electric shocks may enter parts of the song system other than VTA, see Section 3.1.

Non-auditory feedback plays a profound role in shaping songbird vocal learning in ethological contexts. For example, juvenile songbirds rely on auditory and non-auditory feedback when acquiring songs (Chen et al., 2016; Carouso-Peck and Goldstein, 2019; West and King, 1988). Juvenile songbirds that receive both auditory and visual feedback from live tutors display enhanced song learning than juvenile

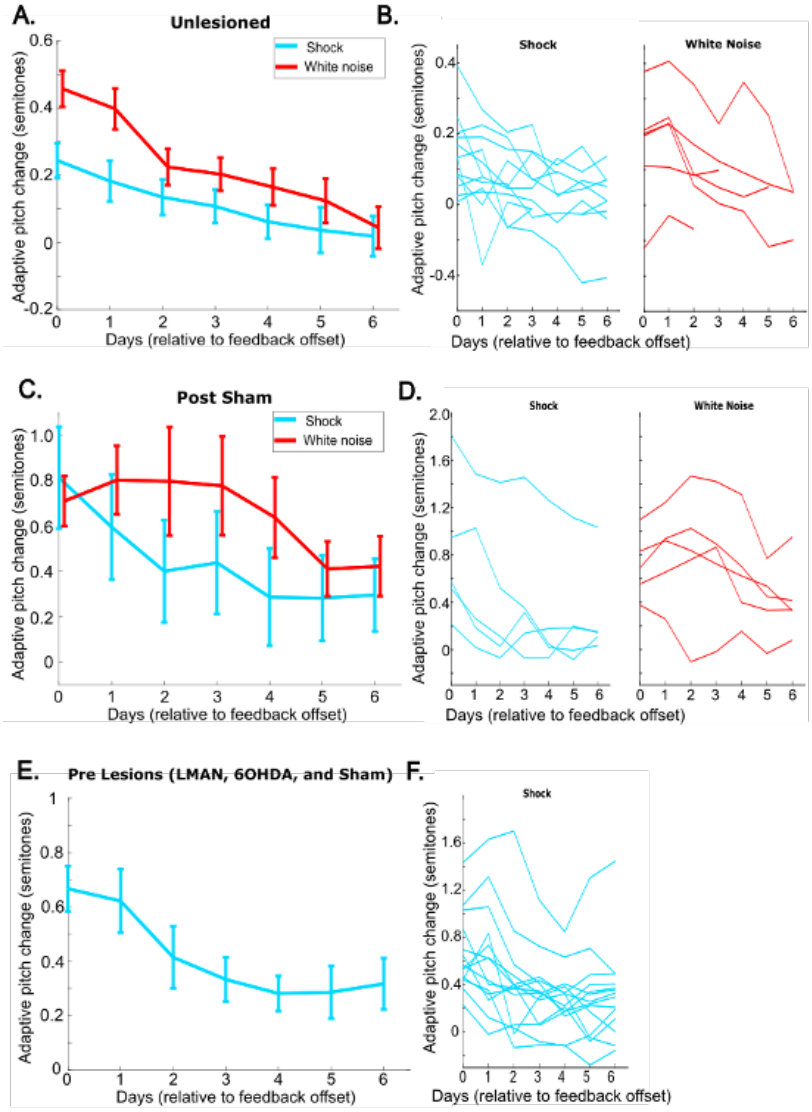
songbirds who only receive auditory feedback from their tutors (Chen et al., 2016). Further, visual displays from adult songbirds positively reinforce the production of specific song structures in juvenile songbirds (West and King, 1988). Our results show that the AFP processes non-auditory feedback to drive vocal learning. This suggests that although electric shocks are not a natural source of sensory feedback during singing, the AFP could play a role in processing other sources of external, non-auditory sensory information (such as the non-auditory signals from social interactions) to guide vocal learning. For further discussion on the role of non-auditory feedback in shaping songbird behaviors beyond juvenile song acquisition, see Section 3.2.

It has been hypothesized that the function of the songbird AFP is to encode auditory performance error: the evaluation of the match between the auditory feedback the songbirds receive and their internal goal for what their song should sound like (based on their stored memory of the tutor song template) (Fee and Goldberg, 2011; Sober and Brainard, 2009; Gadagkar et al., 2016; Saravanan et al., 2019). Some have speculated that the white noise bursts are interpreted by the bird as an auditory performance error: an adult songbird expects to hear the auditory feedback from a well-performed song syllable, but instead hears a loud burst of white noise, which it interprets as a very poorly-performed song syllable (Gadagkar et al., 2016; Chen et al., 2019). Some evidence supports this hypothesis. For example, pitch-contingent white noise bursts provided during song performance drive adaptive vocal changes (Tumer and Brainard, 2007; Hoffmann et al., 2016), but when white noise bursts are provided in nonvocal contexts, such as when a songbird stands on a particular perch (not during song performance), they can positively reinforce place preference (Murdoch et al., 2018). This suggests that white noise is not a generally aversive reinforcement cue. Other studies have suggested that white noise bursts are an aversive reinforcement cue, since white noise bursts are loud and jarring, they sound very dif-

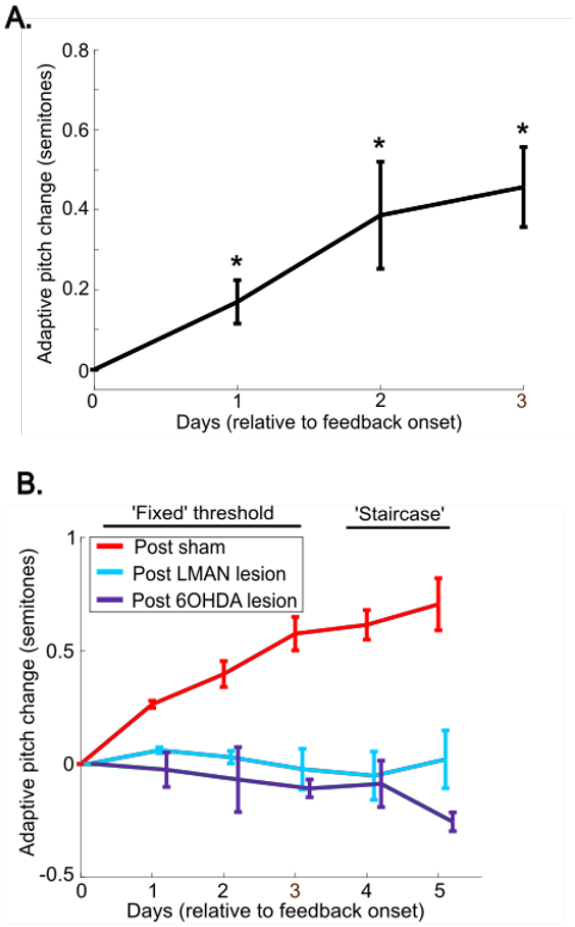
ferent than birdsong, and songbirds will adaptively change their vocalizations to avoid triggering white noise bursts as frequently (Tumer and Brainard, 2007; Charlesworth et al., 2012; Warren et al., 2011; Hoffmann et al., 2016). Although our results do not conclusively prove whether or not white noise bursts drive learning because birds believe they are producing a performance error or because birds find white noise aversive, the cutaneous electric shocks we employed to drive vocal learning are clearly an explicit, external reinforcement cue. That the AFP underlies electric shock-driven learning suggests that the AFP does not solely encode auditory performance error. Instead, the AFP may encode more general information about whether vocal performance resulted in a "good" or "bad" outcome, and it may use this information to drive changes to future motor output.

The numerous analogies between the specialized vocal learning neural circuits that have evolved in songbirds and in humans suggest that our findings may be of relevance for understanding the neural underlying of human speech (Doupe and Kuhl, 1999; Brainard and Doupe, 2002, 2013; Jarvis, 2019). Human speech depends on both auditory and non-auditory sensory information to guide learning, yet very little is known about the neural mechanisms for non-auditory vocal learning (Goldstein et al., 2003; Locke and Snow, 1997; Kuhl, 2007). Our findings show that specialized vocal learning circuitry in songbirds processes non-auditory information to drive vocal learning. We suggest that the analogous vocal circuitry in humans may also underlie non-auditory vocal learning. This neural circuitry in humans may underlie the processing of multimodal sensory signals during social interactions that modulate speech learning (Goldstein et al., 2003; Locke and Snow, 1997; Kuhl, 2007), or the non-auditory, somatosensory feedback from vocal effectors during speech production (Tremblay et al., 2003).

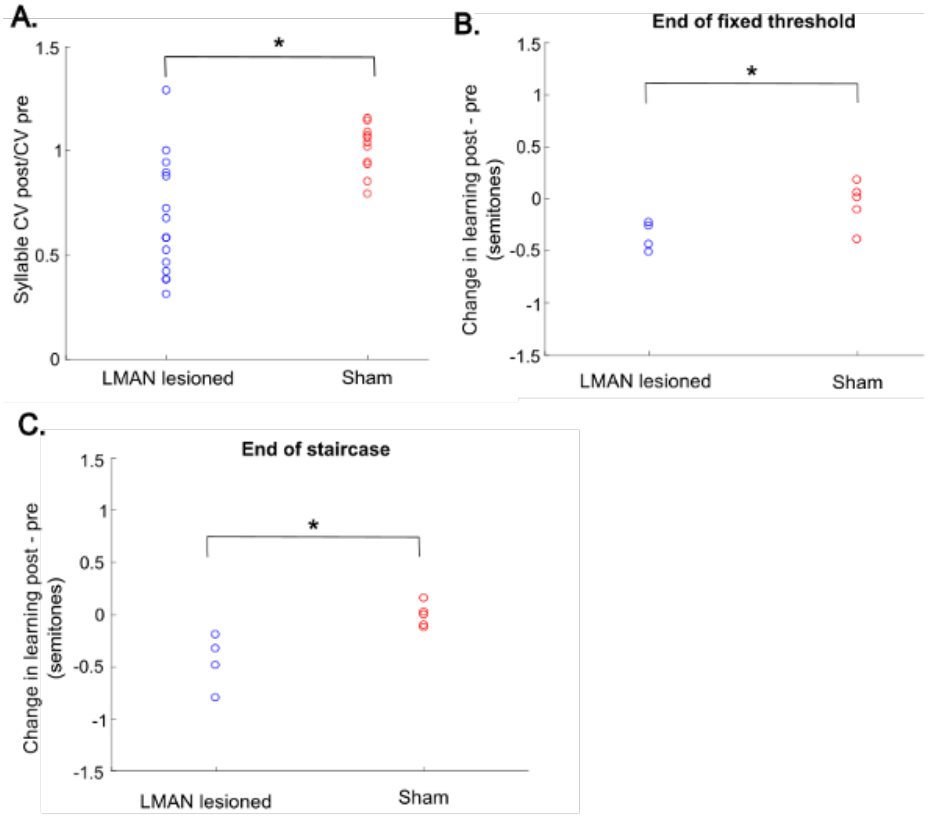
2.6 Supplemental Figures



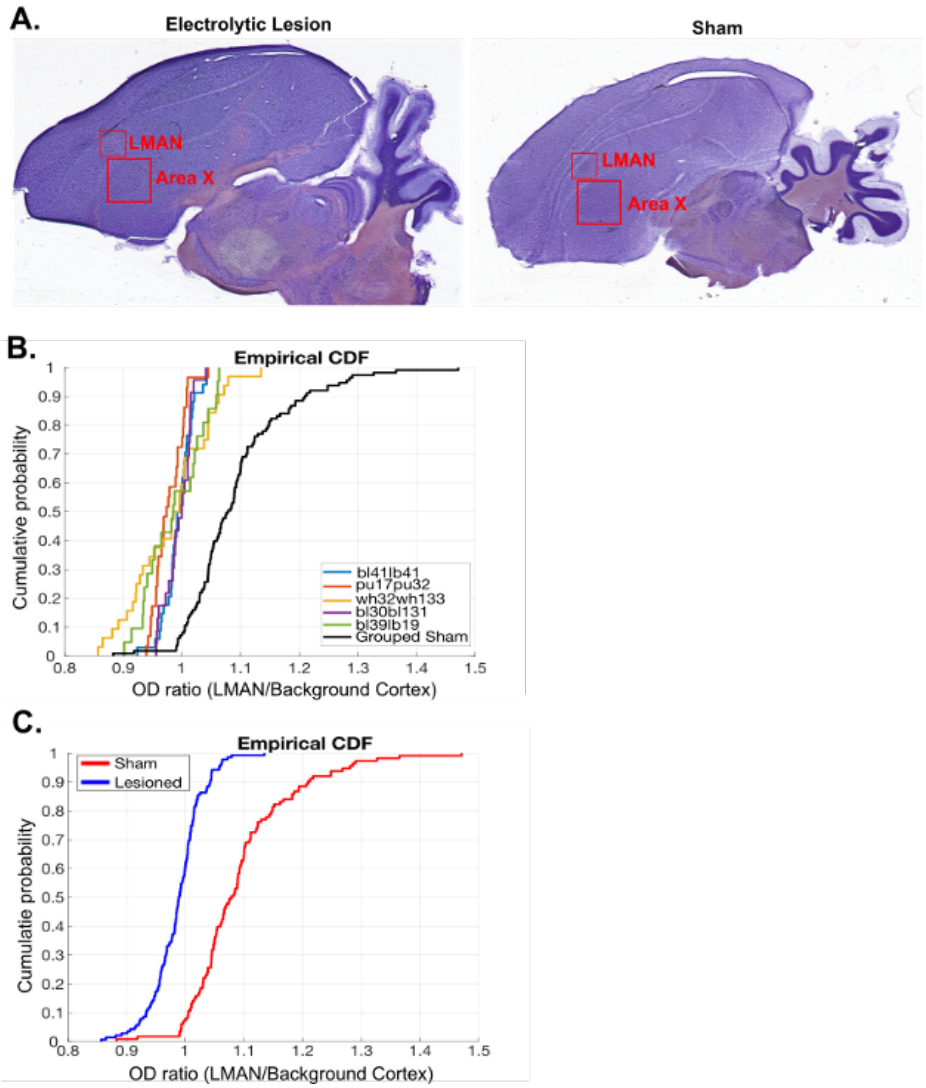
Supplementary Figure 2.5: Rates of washout across different experimental conditions. A. Adaptive pitch change (measured relative to baseline) during washout from the group of birds who received no invasive brain operations (n=13 experiments). Adaptive pitch change did not significantly differ between white noise and electric shock training experiments on any of the days of washout ($0.487 < p_{boot} < 0.541$ on each day of washout, where $0.025 < p_{boot} < 0.975$ indicates no significant difference between means). B. The same washout data from A, except each trace is the data from an individual experiment. C. Adaptive pitch change (measured relative to baseline) during washout from the sham lesioned data set (n=6 experiments). Adaptive pitch change did not significantly differ between white noise and electric shock training experiments on any of the days of washout ($0.370 < p_{boot} < 0.900$) D. The same washout data from C, except each trace is the data from an individual experiment. E. Adaptive pitch change (measured relative to baseline) during washout from all prelesion experiments in birds who received invasive brain operations (presham, pre LMAN lesion, and pre 6-OHDA lesion), n = 16. F. The same washout data from E, except each trace is the data from an individual experiment.



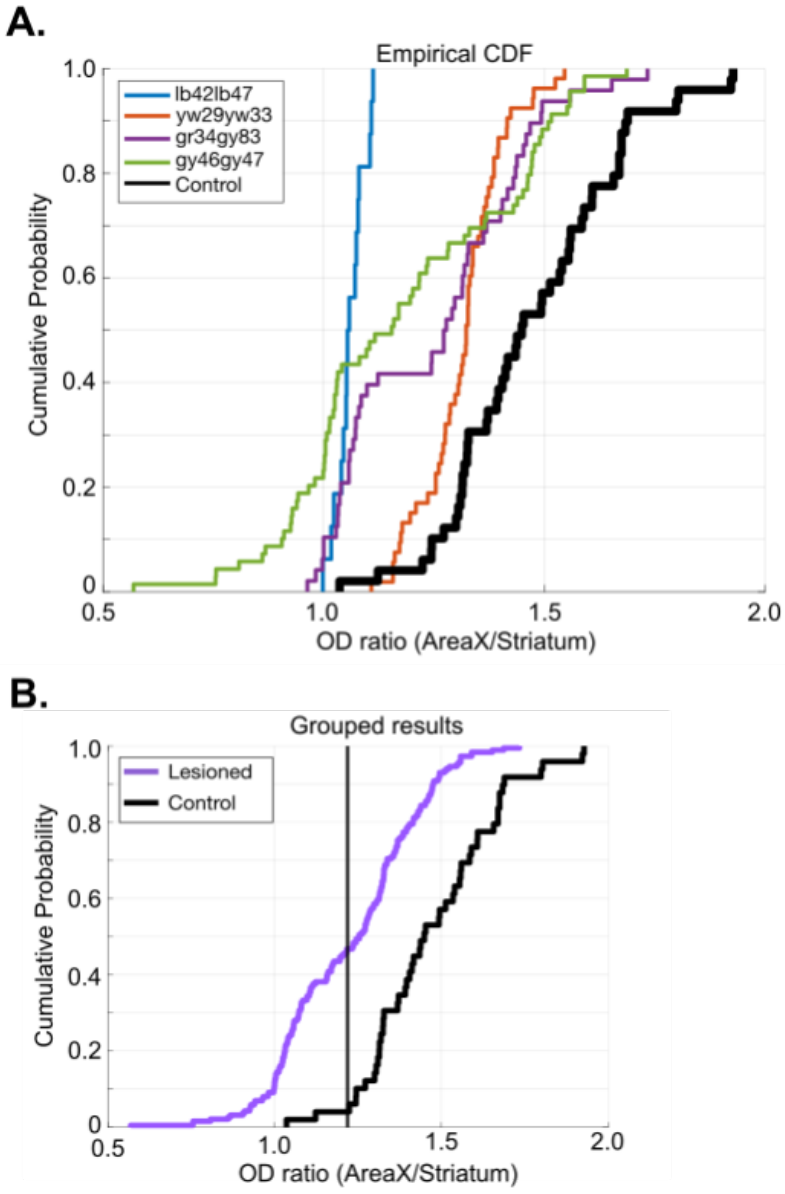
Supplementary Figure 2.6: A. Adaptive change in target syllable pitch (in semitones) during three days of white noise training in unlesioned birds. Error bars represent SEM (see Methods for calculation) for n = 8 birds. p_{boot} >(1.000 - 10^{-4}) on each of the three days of training, where p_{boot}>0.975 indicates the mean is significantly greater than baseline. B. Learning magnitudes (adaptive change in target syllable pitch in semitones) during five days of white noise training in operated birds (Sham, LMAN lesioned, 6-OHDA lesioned). Only postsham learning magnitude was significantly greater than baseline (p_{boot} >(1.000 - 10^{-4}) on each of the final four days of training, where p_{boot} >.975 indicates the mean is significantly greater than baseline). Post LMAN lesion learning magnitudes were significantly less than postsham (p_{boot} <10^{-4} on each of the final four days of training).



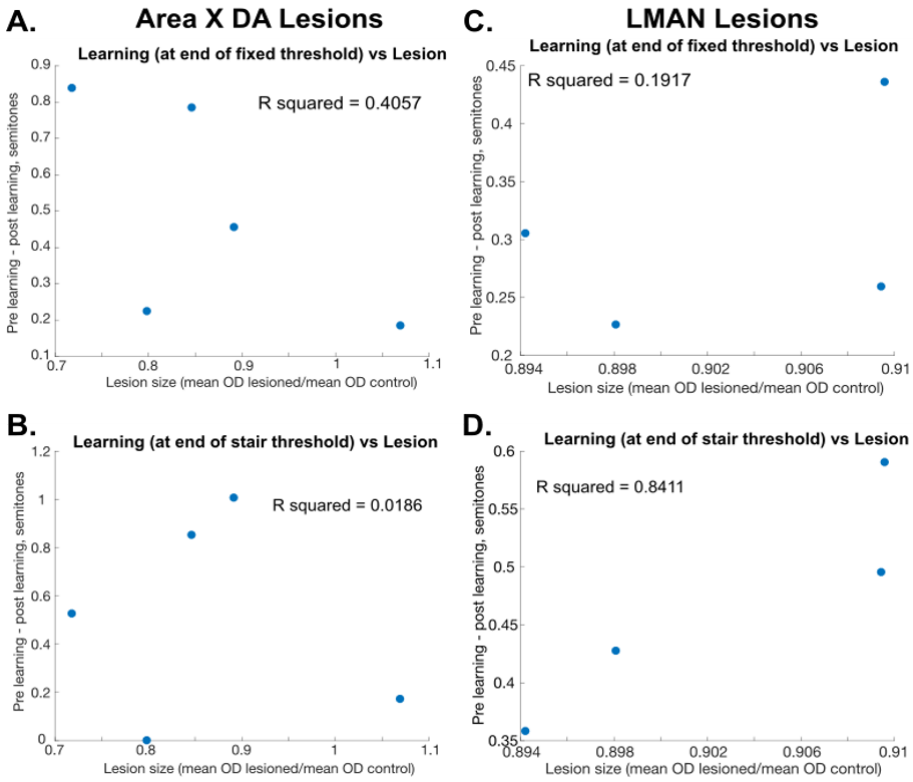
Supplementary Figure 2.7: A. Syllable CV postlesion / syllable CV prelesion between conditions (LMAN lesioned and sham lesioned birds). Each data point represents the CV postlesion / CV prelesion of one individual song syllable. LMAN lesions induced a significant reduction in syllable CV compared to sham (2 sample KS test, p = 0.006) B. Lesion-induced change in learning magnitude (measured at the end of three days of electric shock training) between conditions (LMAN lesioned and sham). The lesion-induced change in learning magnitude (post – pre) for LMAN lesioned birds was significantly greater than sham (2 sample KS test: p = 0.036) C. Lesion-induced change in learning magnitude (measured at the end the extended staircase portion of electric shock training) between conditions (LMAN lesioned and sham). The lesion-induced change in learning magnitude (post – pre) for LMAN lesioned birds was significantly greater than sham (2 sample KS test, p = 0.004)



Supplementary Figure 2.8: LMAN lesion histological analysis. A. Example images of Nissl-stained brain tissue. LMAN lesioned tissue on the left and sham tissue on the right. Red boxes highlight the locations of Area X and LMAN. B. CDF plot of optical density (OD) ratios (OD of LMAN / OD of non-LMAN-pallium) in lesioned and sham operated birds. Each line shows the OD ratios from each individual LMAN lesioned bird, and the black line shows the OD ratios from the grouped sham data set. C. CDF plot of OD ratios in lesioned and sham operated birds. Blue line shows the OD ratios from the grouped LMAN lesion data set and the red line shows the OD ratios from the grouped sham data set (2 sample KS test, $p < 0.001$).



Supplementary Figure 2.9: 6-OHDA lesion histological analysis. A. CDF plot of optical density (OD) ratios (OD of Area X / OD of non-X-striatum) in lesioned and unlesioned (control) birds. Each line shows the OD ratios from each individual 6-OHDA lesioned bird, and the black line shows the OD ratios from the grouped control data set. B. CDF plot of OD ratios in lesioned and control birds. Purple line shows the OD ratios from the grouped 6-OHDA lesioned dataset, and the black line shows the OD ratios from the grouped control dataset (2 sample KS test, $p < 0.001$).



Supplementary Figure 2.10: Comparison of lesion magnitude and learning deficit. A. The difference between the magnitude of learning prelesion and the magnitude of learning postlesion (in both cases, the magnitude of learning is measured at the end of the three days of training with the fixed pitch threshold), compared to the size of the Area X dopamine lesion in 6-OHDA injected birds, measured by the ratio of the mean OD of the lesioned bird to the mean OD of control birds. Each dot represents the results from each individual bird. B. The difference between the magnitude of learning prelesion and the magnitude of learning postlesion (in both cases, the magnitude of learning is measured at the end of not only the three days of training with the fixed pitch threshold but also the two days of staircase threshold), compared to the size of the Area X dopamine lesion in 6-OHDA injected birds, measured by the ratio of the mean OD of the lesioned bird to the mean OD of control birds. Each dot represents the results from each individual bird. C. The difference between the magnitude of learning prelesion and the magnitude of learning postlesion (in both cases, the magnitude of learning is measured at the end of the three days of training with the fixed pitch threshold), compared to the size of the LMAN lesion in electrolytically lesioned birds, measured by the ratio of the mean OD of the lesioned bird to the mean OD of control birds. Each dot represents the results from each individual bird. D. The difference between the magnitude of learning prelesion and the magnitude of learning postlesion (in both cases, the magnitude of learning is measured at the end of not only the three days of training with the fixed pitch threshold but also the two days of staircase threshold), compared to the size of the LMAN lesion in electrolytically lesioned birds, measured by the ratio of the mean OD of the lesioned bird to the mean OD of control birds. Each dot represents the results from each individual bird.

Chapter 3

Extended Discussion and Future Directions

The central scientific question at the heart of this dissertation is: *Do specialized neural circuits that have evolved to support vocal learning only process auditory feedback, or do they process information from a variety of sensory sources?* In Chapter 2, I described a series of experiments aimed at addressing these questions. First, we found that non-auditory feedback (cutaneous electric shocks) can drive changes in vocal output in adult songbirds (Fig. 2.2). Next, we found that intact LMAN is necessary for the expression of this non-auditory form of vocal learning (Fig. 2.3). Finally, we found that dopaminergic input to Area X is required for this non-auditory vocal learning (Fig. 2.4). These results suggest that specialized songbird neural circuitry for vocal learning does not solely process auditory feedback directly relevant for their vocal behavior. Instead, this neural circuit can process information from non-auditory sources of sensory feedback to drive vocal learning the adaptation of vocal motor output (Fig. 3.1 A). Here, I will expand upon the discussion of the implications of these results that I reviewed in Section 2.5, and I will propose future experiments to further test the role of dopaminergic neural dynamics for vocal motor

learning. In my expanded discussion, I will describe the anatomical pathways by which non-auditory information may enter vocal learning circuits (Section 3.1) and how songbird neural circuits could process non-auditory information for ethological behaviors (Section 3.2). Finally, I will propose experiments that primarily focus on testing how dopaminergic neural dynamics may underlie non-auditory vocal learning (Section 3.3).

3.1 Pathways for auditory and non-auditory information to reach the AFP

The result that songbird AFP processes both auditory and non-auditory sensory feedback to drive song learning indicates that neuroanatomical pathways exist for both auditory and non-auditory feedback to enter the vocal learning circuit. In Section 2.5, I described the neural pathways for auditory feedback to enter VTA. Although we demonstrated that this neural circuit is critical for vocal learning (see Section 2.4.3), there are other important neural pathways by which auditory information could enter the song system to influence learning. Songbirds have cochlea that receive auditory information and pass this information along through deep brain structures, through auditory thalamus (Ov), to auditory pallial brain regions. These pallial brain regions, such as Field L and AIV, are analogous to mammalian auditory cortical areas (Nagel and Doupe, 2006; Mello and Clayton, 1994; Mello et al., 1992; Keller and Hahnloser, 2009; Bauer et al., 2008; Fortune and Margoliash, 1992). For example, gene expression patterns are altered in Field L in response to playbacks of song (Mello et al., 1992; Mello and Clayton, 1994) and electrophysiological recordings of individual Field L neurons reveal that they phasically burst in response to the songbird hearing a white noise burst (Keller and Hahnloser, 2009). Field L (and other auditory pallial brain regions) project to HVC, which then projects to Area X, the basal ganglia portion

of the AFP (Jarvis et al., 2005; Fee and Goldberg, 2011). This constitutes a likely neural pathway for auditory signals to enter the AFP, where neurons can then encode this auditory information that can drive vocal learning (Gadagkar et al., 2016; Fee and Goldberg, 2011).

How sensory information from a non-auditory, cutaneous electric shock enters the song system is less clear. In Section 2.5, I explained how the results described throughout Chapter 2 suggest that this sensory information travels through neural pathways that ultimately connect to VTA, which then can modulate the AFP to drive adaptive vocal changes. However, other pathways may allow for cutaneous electric shocks to enter the song system and influence vocal learning. For example, one study paired a foot shock stimulus with the general performance of song in adult songbirds (Jarvis et al., 1995). Following the association between song and foot shock, IEG expression was assessed in different songbird brain regions. There was a significant increase in IEG expression (a proxy for changes in neural activity) in the caudomedial auditory telencephalon (NCM), which projects to HVC, which then projects to Area X. Due to the similarities between cutaneous skin shock and foot shock, our results suggest that this particular neural pathway might carry sensory information related to aversive electric shocks into the song sytem, where this information can then drive learning. In addition, in some bird species the somatosensory system for processing cutaneous tactile information has been mapped out (Wild, 1995; Delius and Bennetto, 1972; Erulkar, 1955; Wild, 1997a; Medina and Reiner, 2000). This pathway involves a somatosensory recipient portion of the dorsal thalamus (DIVA), which then projects through the lateral forebrain to the telencephalon, as well as a primary somatosensory area that exists within a pallial brain structure formerly known as the Wulst (Wild, 1997a; Medina and Reiner, 2000). This pathway sends projections to Uva, a thalamic songbird brain region that receives anatomical projections from brain regions that process multiple sensory modalities, that contains neurons that respond

to both visual and somatosensory stimuli, and that sends projections to HVC (Wild, 1994), which then projects to Area X. Our electric shock cues are likely encoded by the somatosensory system through cutaneous sensory receptors. Thus, our results implicate the aforementioned neural pathway for carrying somatosensory information into the song system in non-auditory vocal learning.

Another neural circuit that our results suggest may be involved in vocal learning includes the songbird brain nucleus called the nucleus taeniae of the amygdala, which is analogous to mammalian amygdala (Jarvis et al., 2005; Ikebuchi et al., 2013). Songbird amygdala receives auditory information regarding vocalizations (Fujii et al., 2016, 2015; Ikebuchi et al., 2013). Songbird amygdala is also involved the recognition and processing of social information, which includes visual and auditory feedback (Klatt and Goodson, 2013; Mayer et al., 2019). In mammals, the amygdala plays a very well-studied role in fear conditioning, which entails the modification of behaviors in response to aversive stimuli (such as electric shocks) (Phillips and LeDoux, 1992; LeDoux et al., 1990; Rogan et al., 1997). We do not believe we are inducing fear conditioning during the experiments described throughout Chapter 2 of this dissertation, because we carefully calibrated the magnitude of electric shock in order to avoid producing off-target behavioral effects, such as willingness to sing. However, given the similarities between mammalian foot shock and the aversive electric shock stimulus used in our experiments, songbird amygdala may also play an important role in this particular form of non-auditory vocal learning.

3.2 How non-auditory information may guide ethological behaviors

Social interactions involve auditory as well as non-auditory sensory experiences and can strongly shape vocal learning in both humans and songbirds. In Section 2.5, I

described how non-auditory signals during social interaction can guide vocal learning in humans and songbirds. Further, I explained how the results described throughout Chapter 2 suggest that specialized songbird neural circuitry for vocal learning may process the non-auditory signals involved in social modulation of vocal learning during development. Here, I suggest that this neural circuitry may also process the non-auditory signals involved in other ethologically-relevant behaviors other than juvenile song acquisition.

Social interactions and, by extension, feedback from multiple sensory modalities (Bottjer and Johnson, 1997), are highly important in shaping vocal output in adult songbirds. Male songbirds change numerous aspects of their song, such as the variability of pitch of song syllables and the stereotypy of syllable sequence, when in the presence of a female bird (Adar et al., 2008; Bischof et al., 1981; Matheson et al., 2016). These song changes are different than the juvenile song learning process, yet they still rely on auditory and non-auditory sensory information to drive alterations in song production during the performance of an ethological behavior (courtship song). Prior studies have associated changes in neural activity in the AFP with the social modulation of song (Kao et al., 2005; Kao and Brainard, 2006). Our results suggest that the AFP might specifically be involved in processing the non-auditory signals during female-directed song to drive changes in vocal motor output.

Song is a courtship behavior crucial for the mating process, and prior research has shown that songbird courtship depends on multimodal sensory interactions (Huang and Hessler, 2008; Ota et al., 2018, 2015). For instance, songbird courtship not only entails the acoustic performance of song, but also is accompanied by visual displays, such as dancing, hopping, and postural realignment. Little is known about whether and how songbird brains encode the multimodal sensory feedback involved in courtship behavior to guide changes to future motor output. Our results that the AFP is involved in processing non-auditory sensory information to drive adaptive

changes to song (an important part of courtship behavior) suggest the AFP might play a role in processing information related to non-vocal elements of courtship.

3.3 Future directions

Our results indicating that songbird dopaminergic neural circuitry is necessary for processing both auditory and non-auditory feedback to drive vocal learning raises the important question of how the dynamics of dopaminergic neural activity encode information relevant for this sensorimotor process. Here, I will outline specific future experiments that could address this broader question. I propose to record electrophysiological signals from individual dopamine neurons (Section 3.3.1), to record dynamics of dopamine release in downstream brain regions (Section 3.3.2), and to assess the correlation between dopaminergic neural dynamics and individual variability in learning magnitude (Section 3.3.3). Finally, I will speculate on potential future experiments outside of the realm of studying dopaminergic neural dynamics (Section 3.3.4).

3.3.1 VTA neuron encoding of auditory and non-auditory feedback

Our results indicating that songbird dopaminergic neural circuitry is necessary for processing both auditory and non-auditory sensory cues raises the question of whether and how individual dopamine neurons encode these different forms of sensory information. There is a rich literature on how individual dopaminergic neurons encode reinforcement signals and promote learning. As explained in detail in Section 1.1.2, the firing rate of dopamine neurons in the ventral tegmental area (VTA) of mammals phasically increases or decreases when an animal receives a better-than-expected or worse-than-expected outcome, respectively (Schultz, 1998; Schultz and Dickinson, 2000; Schultz, 2017; Schultz et al., 1997). In songbirds, midbrain dopamine neu-

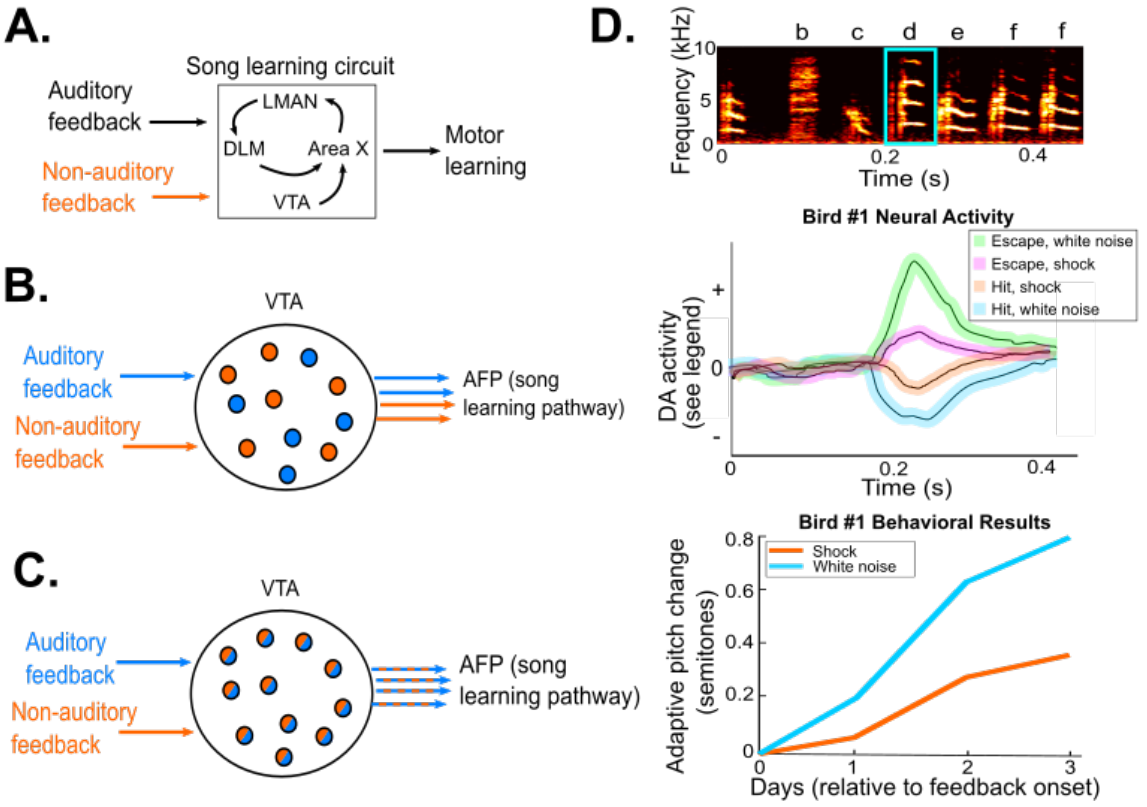


Figure 3.1: Conclusions and future experiments. A. Our experiments demonstrated that non-auditory input can drive vocal learning and that this non-auditory form of vocal learning is dependent on AFP circuitry. B. One potential hypothesis for how VTA dopamine neurons convey auditory and non-auditory reinforcing signals. Under this hypothesis, separate populations of VTA dopamine neurons encode auditory and non-auditory cues. C. Under this second hypothesis, the same populations of VTA dopamine neurons encode both auditory and non-auditory cues. D. Hypotheses for how dopaminergic input to the AFP may encode a learning signal and how this may result in intra-individual variability in the magnitude of song learning. Top figure shows a song spectrogram. The syllable being targeted during separate shock and white noise experiments is outlined. Middle figure shows an example of hypothesized data that would result from doing recordings of dopaminergic neural activity during ongoing song performance during these learning experiments. Dopaminergic neural activity could be measured by fiber photometry or electrophysiology of dopamine cell bodies in VTA, or by measuring extracellular dopamine concentration with optical sensors, such as dLight. In this hypothesized example, this bird has larger upwards deflections in DA activity when escaping white noise and larger downwards deflections when hit by white noise as compared to shock. Bottom figure shows example data from the hypothesized example described previously demonstrating the behavioral results of the song learning experiments (both shock and white noise). In this case, the bird learns to a greater magnitude during white noise experiments as compared to shock experiments. One could also envision an experimental result where neural activity is more sensitive to shock and the bird's behavioral experiments demonstrate greater learning during shock experiments compared to white noise.

rons respond in a similar fashion – they are phasically suppressed when a songbird hears distorted auditory feedback (a worse-than-expected result) and are phasically activated when the songbird does not hear expected distorted auditory feedback (a better-than-expected result) (Gadagkar et al., 2016; Chen et al., 2019). However, the neurons recorded in this paper were only studied in the context of auditory feedback, and they only represent a fraction of the total number of neurons found in songbird VTA. This leaves open the question of whether the same neurons that encode auditory performance error also encode non-auditory information crucial for vocal learning, or if entirely separate subpopulations of VTA dopamine neurons encode auditory and non-auditory sensory feedback signals.

I propose that future experiments should investigate this open question by performing electrophysiological recordings of individual songbird VTA dopamine neurons while engaging in a white noise (auditory) feedback paradigm. Then, in the same individual songbirds, one could also perform the electric shock feedback paradigm while continuing to record from the same individual VTA dopamine neurons. I hypothesize two likely outcomes may arise from these experiments. One outcome would be that some individual dopamine neurons respond (their activity is phasically suppressed or activated during hits or escapes, respectively) during only the white noise training paradigm, and other dopamine neurons respond during only the electric shock training paradigm. This result would suggest that separate subpopulations of VTA dopamine neurons encode separate forms of sensory feedback (Fig. 3.1 B). An alternative result would be that the same individual dopamine neurons respond to both the white noise and electric shock training paradigms, which would suggest that the same subpopulations of VTA dopamine neurons encode different sources of sensory feedback (Fig. 3.1 C). Recent work has shown that mammalian dopamine neurons within the SNc and VTA receive anatomical projections from a wide variety of other brain regions, including motor, somatosensory, medial prefrontal, cingulate, and orbitofrontal cortices, as

well as lateral hypothalamus, preoptic areas, the diagonal band of Broca, and amygdala (Watabe-Uchida et al., 2012). This suggests that VTA dopamine neurons are well positioned anatomically to integrate information from different sensory sources. Also, recordings of neural activity of individual VTA dopamine neurons were made during the performance of behavioral tasks that involved reward, decision making, movement, and navigation (Engelhard et al., 2019). These recordings revealed that some individual dopamine neurons encode multiple task parameters in addition to encoding reward, and others preferentially encode specific non-reward task parameters, suggesting that VTA dopamine neurons display specialized responses to non-reward task parameters. Therefore, both hypotheses of mechanisms of VTA encoding of auditory and non-auditory feedback to drive songbird vocal learning seem plausible and warrant further investigation.

3.3.2 Dopamine release in Area X as a mechanism for electric shock and white noise learning

In addition to electrophysiological recordings of individual VTA dopamine neurons during electric shock and white noise training, one could measure dopaminergic input to Area X during these learning paradigms to assess how dynamics of dopamine release in Area X correlate with vocal learning. Prior work has demonstrated that Area X overall is required for vocal learning in response to white noise feedback (Ali et al., 2013). More targeted experiments demonstrated that dopaminergic input to Area X is required for vocal learning in response to white noise feedback (Hoffmann et al., 2016). Our results demonstrate that dopaminergic input to Area X also plays a necessary role in vocal learning in response to non-auditory, electric shock feedback. Therefore, measuring how patterns of dopaminergic release into Area X change during vocal learning could point to a potential neural mechanism underlying the expression of learning during these experiments. To record dopaminergic input to Area X during

vocal learning experiments, one could perform fiber photometry. Fiber photometry is another relatively new recording technique, which entails the expression of GCaMP in neurons in specific brain regions (as explained earlier) (Gunaydin et al., 2014). Fiber optic cables are then implanted in the brain and can record the fluorescent signal of GCaMP, but instead of recording fluorescent signal in the cell body, one can place the fiber optic cable in downstream brain regions and record fluorescence from the axon terminals of projection neurons in this region. This allows for the recording of genetically and topographically defined subsets of neurons. One could perform this technique to record the neural activity of dopaminergic axon terminals in Area X during behavioral experiments. I hypothesize that the activity of dopaminergic axon terminals in Area X would show similar patterns of activity as previously described – phasic activation on trials when white noise or electric shock feedback is withheld and phasic suppression on trials when white noise or electric shock is provided.

The dynamics of VTA dopamine neuron firing and dopamine release in basal ganglia may separately encode different aspects of behavior relevant for motor learning. Exciting recent work has tested this hypothesis by performing challenging recording techniques in behaving animals to assess how the dynamics of dopamine neural activity and dopamine release in downstream regions correlate with the learning and performance of a complex behavior (Mohebi et al., 2019). Specifically, these researchers simultaneously recorded both electrophysiological signals from dopamine neuron cell bodies in VTA and dopamine concentrations in downstream brain regions during a behavioral training task known as an operant bandit task (Mohebi et al., 2019). This behavioral paradigm involves motivation to engage in the task, decision-making, and reward (Hamid et al., 2016). The measurement of dopamine concentrations in brain regions downstream of VTA during behavior was performed using cutting-edge, genetically-encoded optical sensors (dLight1) that release increased fluorescent signal when dopamine receptor proteins are bound by dopamine molecules (Patriarchi et al.,

2018). This allows for high-speed, high-resolution recording of dopamine dynamics in neural networks in behaving animals. This study revealed that dopamine neuron firing, when measured via electrophysiology of cell bodies in VTA, displayed activity patterns consistent with the encoding of reward prediction error (Mohebi et al., 2019). Surprisingly, downstream dopamine release, measured either through voltammetry or with the high-speed dopamine indicators (dLight1), encoded value-related information, like reward expectation, and not reward prediction error. These results show that the dynamics of dopamine cell spiking and dopamine release from axon terminals differ, and that they encode different aspects of behavior that can guide motor learning. Further, these results suggest that dopamine release in brain regions downstream of VTA may be sculpted by local neural circuits to serve different behavioral functions. Therefore, future studies in songbirds should record dopamine cell body spiking in VTA (Gadagkar et al., 2016) as well as dopamine release in Area X (Patriarchi et al., 2018) to see how dopamine cell spiking and dopamine release correlate with different aspects of vocal learning. I hypothesize that dopamine release in songbird brain regions downstream from VTA may encode information relevant for song learning other than performance error, which is encoded by VTA cell firing (Gadagkar et al., 2016). Further, I hypothesize that different brain regions will shape the dynamics of dopamine release to encode different aspects of song learning. For example, dopamine release in HVC (a songbird brain region important for controlling the timing of song) may encode information more related to timing, such as syllable duration, inter-syllable gap duration, or song sequences (Long and Fee, 2008; Hahnloser et al., 2002).

3.3.3 Neural correlates of individual variability in learning

Individual variability in learning has been observed across species, but it is unclear how specific patterns of neural activity within thalamocortical-basal ganglia circuits

contribute to individual variability in motor output. In our studies, we observed both inter-individual variability in the expression of vocal learning (differences in the magnitude of learning between birds undergoing the same learning paradigm) and intra-individual variability in learning (differences in the magnitude of learning within the same bird undergoing different learning paradigms) (Fig. 2.2 F, see Section 2.4.1). By performing the behavioral learning paradigms described here (electric shock-driven learning and white noise-driven learning) while recording neural activity in the brain regions necessary for the expression of learning during these experiments, one could assess the patterns of neural activity that correlate with individual variability in learning. In these proposed experiments, one could record dopaminergic neural activity in a variety of ways: fiber photometry to measure bulk activity of dopaminergic neuron axons in Area X, electrophysiological recordings of Area X-projecting VTA dopamine neurons, or dLight recordings of dopamine concentration in Area X (Fig. 3.1 D). Dopaminergic neurons that project to Area X convey performance error signals – they are phasically suppressed by auditory reinforcing cues and are phasically activated during the performance of a song syllable when a predicted auditory reinforcement cue does not occur (Gadagkar et al., 2016). I hypothesize similar signals would be observed when similar experiments are performed using auditory feedback and non-auditory feedback. I hypothesize that an individual bird's dopaminergic neural circuits may be more sensitively tuned to responding to auditory cues (white noise) than to non-auditory cues (electric shock), and, in this hypothetical experiment, I would see a greater magnitude of phasic suppression of dopaminergic activity in response to white noise than in response to electric shock and a greater magnitude of phasic activation in response to escapes from white noise than in response to escapes from electric shock (Fig. 3.1 D, middle). I further hypothesize that in this case, where dopamine neurons respond stronger to white noise than they do to shock, I might observe a greater magnitude of learning in response to white noise than in

response to electric shock (Fig. 3.1 D, middle). In the subset of birds that learn more in response to electric shock than they do in response to white noise (Fig. 3.1 D, bottom), I might observe that dopaminergic activity responds more strongly to electric shock than to white noise. A significant correlation between the magnitude of dopaminergic activity and the magnitude of learning would suggest a potential neural mechanism underlying intra-individual variability in learning.

3.3.4 Studying other neural systems involved in sensorimotor learning

So far, I have proposed a detail series of experiments to dissect how different aspects of dopaminergic neural signaling underlie vocal motor learning. However, dopaminergic neural dynamics are only one of many possible aspects of this dissertation that warrant further investigation in the future. Therefore, I will briefly describe two examples of additional lines of future research that could assess how vocal learning neural systems shape behavior.

Our studies have revealed that dopaminergic input to songbird basal ganglia is important for auditory and non-auditory vocal learning, yet it remains unclear how other neurochemical pathways may also guide these processes. The cholinergic system, for example, is a key component of basal ganglia signaling (Graybiel, 1998) plays an important role in motor learning in mammals (Conner et al., 2003, 2005, 2010). The songbird cholinergic system shapes songbird vocal behavior. For instance, experimental activation of this system in behaving adult songbirds produced an number of changes to song production, including an increase in syllable pitch, amplitude, duration, and stereotypy (Jaffe and Brainard, 2020). However, the role the songbird cholinergic system plays in vocal learning is not fully understood. Area X contains cholinergic interneurons, which may sculpt the neural activity of Area X projection neurons during behavior and thereby shape the neural signaling involved

in vocal learning (Woolley and Kao, 2015; Farries and Perkel, 2002; Goldberg et al., 2010; Goldberg and Fee, 2010). Future studies should assess the contributions of the cholinergic system for vocal learning by first selectively manipulating cholinergic interneurons in Area X during auditory and non-auditory vocal learning paradigms. One could inject chemical antagonists of cholinergic signaling within Area X during auditory and non-auditory vocal learning experiments to assess the necessity of the cholinergic in underlying these two forms of vocal learning. I hypothesize that blocking cholinergic signaling will impair both auditory and non-auditory vocal learning. If this manipulation only impaired one form of vocal learning (auditory or non-auditory), it would suggest that different neurochemical pathways process different streams of sensory information to guide learning. One could also inject chemical agonists to see if enhancing cholinergic signaling in Area X also enhances the learning process. I hypothesize that injection of cholinergic agonists in Area X will drive larger magnitudes of adaptive pitch changes in songbirds undergoing auditory and non-auditory guided vocal learning experiments. Finally, future experiments should perform electrophysiological recordings of cholinergic interneurons within Area X during auditory and non-auditory vocal learning (Goldberg et al., 2010; Goldberg and Fee, 2010) to assess how patterns of neural activity in this system encode aspects of behavior relevant for motor learning.

Another important area of future research could study the neural circuit mechanisms for non-vocal, multimodal, ethological behaviors. Courtship is a songbird behavior that involves information from a variety of sensory sources. The production of courtship behaviors involves producing acoustic behaviors (birdsong) as well as non-acoustic behaviors (dances, head bobs, taps) (Zann, 1996; Williams, 2001). Also, feedback from female birds as males perform courtship behaviors involves multiply sensory modalities (Galoch and Bischof, 2006, 2007; Ota et al., 2015). The neural mechanisms underlying this ethological, multimodal behavior are unclear. Future

experiments should therefore elucidate the neural circuit underpinnings of courtship behavior. For example, one could first lesion brain regions hypothesized to underlie the acoustic elements of courtship behavior, like Area X (Leblois et al., 2010; Murugan et al., 2013; Simonyan et al., 2012; Sasaki et al., 2006), and observe any potential changes in non-acoustic aspects of courtship behavior, such as dancing, head bobbing, or posture. I hypothesize that brain areas previously thought to underlie auditory information relevant for courtship will also be required for non-auditory elements of the behavior. Further, one could perform electrophysiological recordings in these brain regions during courtship to assess how patterns of neural activity correlate with dancing, tapping, and other non-vocal courtship behaviors.

More broadly, future experiments performed in humans could help assess the relevance of the work described in this dissertation across species. Our results, described in Chapter 2, demonstrate that non-auditory sources of information are sufficient to drive vocal learning and that a vocal neural pathway is important for processing this non-auditory feedback to guide vocal learning in songbirds. We suggest, due to the similarities between songbird song learning and human speech acquisition, that non-auditory information may be important for the human speech learning process. Further, due to the numerous analogies between songbird and human vocal brain circuits, we suggest that human vocal pathways may process non-auditory sources of information to guide speech acquisition. This may play an important role in encoding social interactions involved in speech learning. Future experiments could test, through fMRI or other brain recording methods, whether the human vocal pathway is important for processing social interactions involved in speech learning.

Bibliography

Abur, D., R. A. Lester-Smith, A. Daliri, A. A. Lupiani, F. H. Guenther, and C. E. Stepp (2018). Sensorimotor adaptation of voice fundamental frequency in parkinson's disease. *PLoS One 13*(1), e0191839.

Adar, E., A. Lotem, and A. Barnea (2008). The effect of social environment on singing behavior in the zebra finch (taeniopygia guttata) and its implication for neuronal recruitment. *Behavioural brain research 187*(1), 178–184.

Ahumada-Galleguillos, P., M. Fernández, G. J. Marin, J. C. Letelier, and J. Mpodozis (2015). Anatomical organization of the visual dorsal ventricular ridge in the chick (gallus gallus): layers and columns in the avian pallium. *Journal of Comparative Neurology 523*(17), 2618–2636.

Albergaria, C., N. T. Silva, D. L. Pritchett, and M. R. Carey (2018). Locomotor activity modulates associative learning in mouse cerebellum. *Nature neuroscience 21*(5), 725–735.

Albin, R. L., A. B. Young, and J. B. Penney (1989). The functional anatomy of basal ganglia disorders. *Trends in neurosciences 12*(10), 366–375.

Ali, F., T. M. Otchy, C. Pehlevan, A. L. Fantana, Y. Burak, and B. P. Ölveczky (2013). The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. *Neuron 80*(2), 494–506.

Amunts, K., A. Schleicher, U. Bürgel, H. Mohlberg, H. B. Uylings, and K. Zilles (1999). Broca's region revisited: cytoarchitecture and intersubject variability. *Journal of Comparative Neurology 412*(2), 319–341.

Andalman, A. S. and M. S. Fee (2009). A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proceedings of the National Academy of Sciences 106*(30), 12518–12523.

Arnold, A. P., F. Nottebohm, and D. W. Pfaff (1976). Hormone concentrating cells in vocal control and other areas of the brain of the zebra finch (poephila guttata). *Journal of Comparative Neurology 165*(4), 487–511.

Arnold, A. P. and A. Saltiel (1979). Sexual difference in pattern of hormone accumulation in the brain of a songbird. *Science 205*(4407), 702–705.

Aston-Jones, G., C. Chiang, and T. Alexinsky (1991). Discharge of noradrenergic locus coeruleus neurons in behaving rats and monkeys suggests a role in vigilance. *Progress in brain research 88*, 501–520.

Aston-Jones, G. and J. D. Cohen (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci. 28*, 403–450.

Aston-Jones, G., J. Rajkowski, P. Kubiak, and T. Alexinsky (1994). Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. *Journal of Neuroscience 14*(7), 4467–4480.

Avanzino, L., N. Gueugneau, A. Bisio, P. Ruggeri, C. Papaxanthis, and M. Bove (2015). Motor cortical plasticity induced by motor learning through mental practice. *Frontiers in behavioral neuroscience 9*, 105.

Baizer, J. S., I. Kralj-Hans, and M. Glickstein (1999). Cerebellar lesions and prism adaptation in macaque monkeys. *Journal of neurophysiology 81*(4), 1960–1965.

Barbeau, A. (1962). The pathogenesis of parkinson's disease: a new hypothesis. *Canadian Medical Association Journal 87*(15), 802.

Barbeau, A. (1969). L-dopa therapy in parkinson's disease: a critical review of nine years' experience. *Canadian Medical Association Journal 101*(13), 59.

Barclay, S. R., C. F. Harding, and S. A. Waterman (1996). Central dsp-4 treatment decreases norepinephrine levels and courtship behavior in male zebra finches. *Pharmacology Biochemistry and Behavior 53*(1), 213–220.

Bauer, E. E., M. J. Coleman, T. F. Roberts, A. Roy, J. F. Prather, and R. Mooney (2008). A synaptic basis for auditory–vocal integration in the songbird. *Journal of Neuroscience 28*(6), 1509–1522.

Becker, M. I. and A. L. Person (2019). Cerebellar control of reach kinematics for endpoint precision. *Neuron 103*(2), 335–348.

Benke, T., C. Hohenstein, W. Poewe, and B. Butterworth (2000). Repetitive speech phenomena in parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry 69*(3), 319–324.

Bernheimer, H., W. Birkmayer, O. Hornykiewicz, K. Jellinger, and F. Seitelberger (1973). Brain dopamine and the syndromes of parkinson and huntington clinical, morphological and neurochemical correlations. *Journal of the neurological sciences 20*(4), 415–455.

Bindra, D. (1968). Neuropsychological interpretation of the effects of drive and incentive-motivation on general activity and instrumental behavior. *Psychological review 75*(1), 1.

Bischof, H.-J., J. Böhner, and R. Sossinka (1981). Influence of external stimuli on the quality of the song of the zebra finch (taeniopygia guttata castanotis gould). *Zeitschrift für Tierpsychologie 57*(3-4), 261–267.

Blank, S. C., S. K. Scott, K. Murphy, E. Warburton, and R. J. Wise (2002). Speech production: Wernicke, broca and beyond. *Brain 125*(8), 1829–1838.

Bostan, A. C. and P. L. Strick (2018). The basal ganglia and the cerebellum: nodes in an integrated network. *Nature Reviews Neuroscience 19*(6), 338–350.

Bottjer, S. W. (1993). The distribution of tyrosine hydroxylase immunoreactivity in the brains of male and female zebra finches. *Journal of neurobiology 24*(1), 51–69.

Bottjer, S. W. and G. Alexander (1995). Localization of met-enkephalin and vasoactive intestinal polypeptide in the brains of male zebra finches (part 1 of 2). *Brain, behavior and evolution 45*(3), 153–165.

Bottjer, S. W., K. A. Halsema, S. A. Brown, and E. A. Miesner (1989). Axonal connections of a forebrain nucleus involved with vocal learning in zebra finches. *Journal of Comparative Neurology 279*(2), 312–326.

Bottjer, S. W. and F. Johnson (1997). Circuits, hormones, and learning: vocal behavior in songbirds. *Journal of neurobiology 33*(5), 602–618.

Bottjer, S. W., E. A. Miesner, and A. P. Arnold (1984). Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science 224*(4651), 901–903.

Bouchard, K. E., N. Mesgarani, K. Johnson, and E. F. Chang (2013). Functional organization of human sensorimotor cortex for speech articulation. *Nature 495*(7441), 327–332.

Bouret, S. and S. J. Sara (2005). Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. *Trends in neurosciences 28*(11), 574–582.

Boyden, E. S., F. Zhang, E. Bamberg, G. Nagel, and K. Deisseroth (2005). Millisecond-timescale, genetically targeted optical control of neural activity. *Nature neuroscience 8*(9), 1263–1268.

Brainard, M. S. and A. J. Doupe (2000a). Auditory feedback in learning and maintenance of vocal behaviour. *Nature Reviews Neuroscience 1*(1), 31–40.

Brainard, M. S. and A. J. Doupe (2000b). Interruption of a basal ganglia–forebrain circuit prevents plasticity of learned vocalizations. *Nature 404*(6779), 762–766.

Brainard, M. S. and A. J. Doupe (2002). What songbirds teach us about learning. *Nature 417*(6886), 351–358.

Brainard, M. S. and A. J. Doupe (2013). Translating birdsong: songbirds as a model for basic and applied medical research. *Annual review of neuroscience 36*, 489–517.

Bronstein, J. M., M. Tagliati, R. L. Alterman, A. M. Lozano, J. Volkmann, A. Stefani, F. B. Horak, M. S. Okun, K. D. Foote, P. Krack, et al. (2011). Deep brain stimulation for parkinson disease: an expert consensus and review of key issues. *Archives of neurology 68*(2), 165–165.

Caligiore, D., G. Pezzulo, G. Baldassarre, A. C. Bostan, P. L. Strick, K. Doya, R. C. Helmich, M. Dirkx, J. Houk, H. Jörntell, et al. (2017). Consensus paper: towards a systems-level view of cerebellar function: the interplay between cerebellum, basal ganglia, and cortex. *The Cerebellum 16*(1), 203–229.

Canter, G. J. (1963). Speech characteristics of patients with parkinson's disease: I. intensity, pitch, and duration. *Journal of speech and hearing disorders 28*(3), 221–229.

Carlsson, M. and A. Carlsson (1990). Interactions between glutamatergic and monoaminergic systems within the basal ganglia-implications for schizophrenia and parkinson's disease. *Trends in neurosciences 13*(7), 272–276.

Carouso-Peck, S. and M. H. Goldstein (2019). Female social feedback reveals non-imitative mechanisms of vocal learning in zebra finches. *Current Biology 29*(4), 631–636.

Carrillo, G. D. and A. J. Doupe (2004). Is the songbird area x striatal, pallidal, or both? an anatomical study. *Journal of Comparative Neurology 473*(3), 415–437.

Castelino, C. B. and G. F. Ball (2005). A role for norepinephrine in the regulation of context-dependent zenk expression in male zebra finches (taeniopygia guttata). *European Journal of Neuroscience 21*(7), 1962–1972.

Casto, J. M. and G. F. Ball (1994). Characterization and localization of d1 dopamine receptors in the sexually dimorphic vocal control nucleus, area x, and the basal ganglia of european starlings. *Journal of neurobiology 25*(7), 767–780.

Charlesworth, J., E. Tumer, T. Warren, and M. Brainard (2011). Learning the microstructure of successful behavior. *Nature neuroscience 14*(3), 373–380.

Charlesworth, J., T. Warren, and M. Brainard (2012). Covert skill learning in a cortical-basal ganglia circuit. *Nature*.

Chaudhuri, A. (1997). Neural activity mapping with inducible transcription factors. *Neuroreport 8*(13), iii–vii.

Chen, R., P. A. Puzerey, A. C. Roeser, T. E. Riccelli, A. Podury, K. Maher, A. R. Farhang, and J. H. Goldberg (2019). Songbird ventral pallidum sends diverse performance error signals to dopaminergic midbrain. *Neuron 103*(2), 266–276.

Chen, Y., L. E. Matheson, and J. T. Sakata (2016). Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proceedings of the National Academy of Sciences 113*(24), 6641–6646.

Conner, J., M. Kulczycki, and M. Tuszynski (2010). Unique contributions of distinct cholinergic projections to motor cortical plasticity and learning. *Cerebral Cortex 20*(11), 2739–2748.

Conner, J. M., A. A. Chiba, and M. H. Tuszynski (2005). The basal forebrain cholinergic system is essential for cortical plasticity and functional recovery following brain injury. *Neuron 46*(2), 173–179.

Conner, J. M., A. Culberson, C. Packowski, A. A. Chiba, and M. H. Tuszynski (2003). Lesions of the basal forebrain cholinergic system impair task acquisition and abolish cortical plasticity associated with motor skill learning. *Neuron 38*(5), 819–829.

Costa, R. M., D. Cohen, and M. A. Nicolelis (2004). Differential corticostriatal plasticity during fast and slow motor skill learning in mice. *Current Biology 14*(13), 1124–1134.

Cotzias, G. C., P. S. Papavasiliou, and R. Gellene (1969). Modification of parkinsonism—chronic treatment with l-dopa. *New England Journal of Medicine 280*(7), 337–345.

Coull, J., C. Büchel, K. Friston, and C. Frith (1999). Noradrenergically mediated plasticity in a human attentional neuronal network. *Neuroimage 10*(6), 705–715.

Cowie, R. and E. Douglas-Cowie (2011). *Postlingually acquired deafness: speech deterioration and the wider consequences*, Volume 62. Walter de Gruyter.

Crossman, A. (1987). Primate models of dyskinesia: the experimental approach to the

study of basal ganglia-related involuntary movement disorders. *Neuroscience 21*(1), 1–40.

Crowley, P. H. (1992). Resampling methods for computation-intensive data analysis in ecology and evolution. *Annual Review of Ecology and Systematics 23*(1), 405–447.

Dayalu, P. and R. L. Albin (2015). Huntington disease: pathogenesis and treatment. *Neurologic clinics 33*(1), 101–114.

Deisseroth, K. (2011). Optogenetics. *Nature methods 8*(1), 26–29.

Delius, J. D. and K. Bennetto (1972). Cutaneous sensory projections to the avian forebrain. *Brain Research 37*(2), 205–221.

Deuschl, G., C. Schade-Brittinger, P. Krack, J. Volkmann, H. Schäfer, K. Bötzel, C. Daniels, A. Deutschländer, U. Dillmann, W. Eisner, et al. (2006). A randomized trial of deep-brain stimulation for parkinson's disease. *New England Journal of Medicine 355*(9), 896–908.

Diamond, S., J. Schneider, and C. Markham (1987). Oral sensorimotor defects in patients with parkinson's disease. *Advances in neurology 45*, 335–338.

Dichter, B. K., J. D. Breshears, M. K. Leonard, and E. F. Chang (2018). The control of vocal pitch in human laryngeal motor cortex. *Cell 174*(1), 21–31.

Diedrichsen, J., T. Verstynen, S. L. Lehman, and R. B. Ivry (2005). Cerebellar involvement in anticipating the consequences of self-produced actions during bimanual movements. *Journal of neurophysiology 93*(2), 801–812.

Diedrichsen, J., O. White, D. Newman, and N. Lally (2010). Use-dependent and error-based learning of motor behaviors. *Journal of Neuroscience 30*(15), 5159–5166.

Doupe, A. J. and P. K. Kuhl (1999). Birdsong and human speech: common themes and mechanisms. *Annual review of neuroscience 22*(1), 567–631.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current opinion in neurobiology 10*(6), 732–739.

Dronkers, N. F., O. Plaisant, M. T. Iba-Zizen, and E. A. Cabanis (2007). Paul broca's historic cases: high resolution mr imaging of the brains of leborgne and lelong. *Brain 130*(5), 1432–1441.

Dubbeldam, J., A. den Boer-Visser, and R. Bout (1997). Organization and efferent connections of the archistriatum of the mallard, anas platyrhynchos l.: an anterograde and retrograde tracing study. *Journal of Comparative Neurology 388*(4), 632–657.

Duffy, J. R. (2013). *Motor Speech disorders-E-Book: Substrates, differential diagnosis, and management*. Elsevier Health Sciences.

Dunn, A. and R. Zann (1996). Undirected song encourages the breeding female zebra finch to remain in the nest. *Ethology 102*(4), 540–548.

Efron, B. and R. J. Tibshirani (1994). *An introduction to the bootstrap*. CRC press.

Engelhard, B., J. Finkelstein, J. Cox, W. Fleming, H. J. Jang, S. Ornelas, S. A. Koay, S. Y. Thiberge, N. D. Daw, D. W. Tank, et al. (2019). Specialized coding of sensory, motor and cognitive variables in vta dopamine neurons. *Nature 570*(7762), 509–513.

Erulkar, S. D. (1955). Tactile and auditory areas in the brain of the pigeon. an experimental study by means of evoked potentials. *Journal of Comparative Neurology 103*(3), 421–457.

Farries, M. A. and D. J. Perkel (2002). A telencephalic nucleus essential for song learning contains neurons with physiological characteristics of both striatum and globus pallidus. *Journal of Neuroscience 22*(9), 3776–3787.

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

Ferland, R. J., T. J. Cherry, P. O. Preware, E. E. Morrisey, and C. A. Walsh (2003). Characterization of foxp2 and foxp1 mrna and protein in the developing and mature brain. *Journal of comparative Neurology 460*(2), 266–279.

Fernandez-Ruiz, J. and R. Diaz (1999). Prism adaptation and aftereffect: specifying the properties of a procedural memory system. *Learning and Memory 6*(1), 47–53.

Fiorillo, C. D., P. N. Tobler, and W. Schultz (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science 299*(5614), 1898–1902.

Fisher, S. E. and C. Scharff (2009). Foxp2 as a molecular window into speech and language. *Trends in Genetics 25*(4), 166–177.

Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in cognitive sciences 4*(7), 258–267.

Foote, S., G. Aston-Jones, and F. Bloom (1980). Impulse activity of locus coeruleus neurons in awake rats and monkeys is a function of sensory stimulation and arousal. *Proceedings of the National Academy of Sciences 77*(5), 3033–3037.

Fortune, E. S. and D. Margoliash (1992). Cytoarchitectonic organization and morphology of cells of the field l complex in male zebra finches (taenopygia guttata). *Journal of Comparative Neurology 325*(3), 388–404.

Friederici, A. D., N. Chomsky, R. C. Berwick, A. Moro, and J. J. Bolhuis (2017). Language, mind and brain. *Nature Human Behaviour 1*(10), 713–722.

Fujii, T., M. Ikebuchi, and K. Okanoya (2015). Neural responses to songbird vocalizations in the nucleus taeniae of the amygdala in bengalese finches. *The Journal of the Acoustical Society of America 138*(3), 1792–1792.

Fujii, T. G., M. Ikebuchi, and K. Okanoya (2016). Auditory responses to vocal sounds in the songbird nucleus taeniae of the amygdala and the adjacent arcopallium. *Brain, behavior and evolution 87*(4), 275–289.

Gadagkar, V., P. A. Puzerey, R. Chen, E. Baird-Daniel, A. R. Farhang, and J. H. Goldberg (2016). Dopamine neurons encode performance error in singing birds. *Science 354*(6317), 1278–1282.

Gahr, M. (2000). Neural song control system of hummingbirds: comparison to swifts, vocal learning (songbirds) and nonlearning (suboscines) passerines, and vocal learning (budgerigars) and nonlearning (dove, owl, gull, quail, chicken) nonpasserines. *Journal of Comparative Neurology 426*(2), 182–196.

Gale, S. D. and D. J. Perkel (2005). Properties of dopamine release and uptake in the songbird basal ganglia. *Journal of neurophysiology 93*(4), 1871–1879.

Gale, S. D. and D. J. Perkel (2010a). Anatomy of a songbird basal ganglia circuit essential for vocal learning and plasticity. *Journal of chemical neuroanatomy 39*(2), 124–131.

Gale, S. D. and D. J. Perkel (2010b). A basal ganglia pathway drives selective auditory responses in songbird dopaminergic neurons via disinhibition. *Journal of Neuroscience 30*(3), 1027–1037.

Galoch, Z. and H.-J. Bischof (2006). Zebra finches actively choose between live images of conspecifics. *Ornithological Science 5*(1), 57–64.

Galoch, Z. and H.-J. Bischof (2007). Behavioural responses to video playbacks by zebra finch males. *Behavioural Processes 74*(1), 21–26.

Gerfen, C. R., T. M. Engber, L. C. Mahan, Z. Susel, T. N. Chase, F. J. Monsma, and D. R. Sibley (1990). D1 and d2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science 250*(4986), 1429–1432.

Gibson, J. J. (1933). Adaptation, after-effect and contrast in the perception of curved lines. *Journal of experimental psychology 16*(1), 1.

Goldberg, J. H., A. Adler, H. Bergman, and M. S. Fee (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. *Journal of Neuroscience 30*(20), 7088–7098.

Goldberg, J. H. and M. S. Fee (2010). Singing-related neural activity distinguishes four classes of putative striatal neurons in the songbird basal ganglia. *Journal of neurophysiology 103*(4), 2002–2014.

Goldstein, M. H., A. P. King, and M. J. West (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences 100*(13), 8030–8035.

Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. *Neurobiology of learning and memory 70*(1-2), 119–136.

Graybiel, A. M. (2000). The basal ganglia. *Current biology 10*(14), R509–R511.

Grieser, D. and P. K. Kuhl (1989). Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology 25*(4), 577.

Gunaydin, L. A., L. Grosenick, J. C. Finkelstein, I. V. Kauvar, L. E. Fenno, A. Adhikari, S. Lammel, J. J. Mirzabekov, R. D. Airan, K. A. Zalocusky, et al. (2014).

Natural neural projection dynamics underlying social behavior. *Cell 157*(7), 1535–1551.

Güntürkün, O. and S. Kröner (1999). A polysensory pathway to the forebrain of the pigeon: the ascending projections of the nucleus dorsolateralis posterior thalami (dlp). *European journal of morphology 37*(0), 185–189.

Haesler, S., C. Rochefort, B. Georgi, P. Licznerski, P. Osten, and C. Scharff (2007). Incomplete and inaccurate vocal imitation after knockdown of foxp2 in songbird basal ganglia nucleus area x. *PLoS Biol 5*(12), e321.

Haesler, S., K. Wada, A. Nshdejan, E. E. Morrisey, T. Lints, E. D. Jarvis, and C. Scharff (2004). Foxp2 expression in avian vocal learners and non-learners. *Journal of Neuroscience 24*(13), 3164–3175.

Hahnloser, R. H., A. A. Kozhevnikov, and M. S. Fee (2002). An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature 419*(6902), 65–70.

Hamid, A. A., J. R. Pettibone, O. S. Mabrouk, V. L. Hetrick, R. Schmidt, C. M. Vander Weele, R. T. Kennedy, B. J. Aragona, and J. D. Berke (2016). Mesolimbic dopamine signals the value of work. *Nature neuroscience 19*(1), 117–126.

Hampton, C. M., J. T. Sakata, and M. S. Brainard (2009). An avian basal ganglia-forebrain circuit contributes differentially to syllable versus sequence variability of adult bengalese finch song. *Journal of neurophysiology 101*(6), 3235–3245.

Hara, E., L. Kubikova, N. A. Hessler, and E. D. Jarvis (2007). Role of the midbrain dopaminergic system in modulation of vocal brain activation by social context. *European Journal of Neuroscience 25*(11), 3406–3416.

Hardebo, J. E. and C. Owman (1980). Barrier mechanisms for neurotransmitter monoamines and their precursors at the blood-brain interface. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society 8*(1), 1–11.

Helduser, S., S. Cheng, and O. Güntürkün (2013). Identification of two forebrain structures that mediate execution of memorized sequences in the pigeon. *Journal of neurophysiology 109*(4), 958–968.

Hessler, N. A. and A. J. Doupe (1999). Social context modulates singing-related neural activity in the songbird forebrain. *Nature neuroscience 2*(3), 209–211.

Hickok, G. and D. Poeppel (2007). The cortical organization of speech processing. *Nature reviews neuroscience 8*(5), 393–402.

Hikosaka, O., H. Nakahara, M. K. Rand, K. Sakai, X. Lu, K. Nakamura, S. Miyachi, and K. Doya (1999). Parallel neural networks for learning sequential procedures. *Trends in neurosciences 22*(10), 464–471.

Hikosaka, O., K. Nakamura, K. Sakai, and H. Nakahara (2002). Central mechanisms of motor skill learning. *Current opinion in neurobiology 12*(2), 217–222.

Hikosaka, O., M. K. Rand, S. Miyachi, and K. Miyashita (1995). Learning of sequential movements in the monkey: process of learning and retention of memory. *Journal of neurophysiology 74*(4), 1652–1661.

Ho, A. K., R. Iansek, C. Marigliani, J. L. Bradshaw, and S. Gates (1998). Speech impairment in a large sample of patients with parkinson's disease. *Behavioural neurology 11*(3), 131–137.

Hoffmann, L. A., V. Saravanan, A. N. Wood, L. He, and S. J. Sober (2016). Dopaminergic contributions to vocal learning. *Journal of Neuroscience 36*(7), 2176–2189.

Hoover, J. E. and P. L. Strick (1999). The organization of cerebellar and basal ganglia outputs to primary motor cortex as revealed by retrograde transneuronal transport of herpes simplex virus type 1. *Journal of Neuroscience 19*(4), 1446–1463.

Hosp, J. A. and A. R. Luft (2013). Dopaminergic meso-cortical projections to m1: role in motor learning and motor cortex plasticity. *Frontiers in neurology 4*, 145.

Hosp, J. A., A. Pekanovic, M. S. Rioult-Pedotti, and A. R. Luft (2011). Dopaminergic projections from midbrain to primary motor cortex mediate motor skill learning. *Journal of Neuroscience 31*(7), 2481–2487.

Houde, J. F. and M. I. Jordan (1998). Sensorimotor adaptation in speech production. *Science 279*(5354), 1213–1216.

Howell, P. and A. Archer (1984). Susceptibility to the effects of delayed auditory feedback. *Perception & Psychophysics 36*(3), 296–302.

Huang, V. S., A. Haith, P. Mazzoni, and J. W. Krakauer (2011). Rethinking motor learning and savings in adaptation paradigms: model-free memory for successful actions combines with internal models. *Neuron 70*(4), 787–801.

Huang, Y.-C. and N. A. Hessler (2008). Social modulation during songbird courtship potentiates midbrain dopaminergic neurons. *PloS one 3*(10), e3281.

Ikebuchi, M., S. Nanbu, K. Okanoya, R. Suzuki, and H.-J. Bischof (2013). Very early development of nucleus taeniae of the amygdala. *Brain, Behavior and Evolution 81*(1), 12–26.

Immelmann, K. (1969). Song development in the zebra finch and other estrildid finches. *Bird vocalizations*, 61–77.

Izawa, J. and R. Shadmehr (2011). Learning from sensory and reward prediction errors during motor adaptation. *PLoS Comput Biol 7*(3), e1002012.

Jacobs, B. L. (1986). Single unit activity of locus coeruleus neurons in behaving animals. *Progress in neurobiology 27*(2), 183–194.

Jaffe, P. I. and M. S. Brainard (2020). Acetylcholine acts on songbird premotor circuitry to invigorate vocal output. *Elife 9*, e53288.

Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. *Journal of neurology, neurosurgery & psychiatry 79*(4), 368–376.

Jankovic, J. and M. Stacy (2007). Medical management of levodopa-associated motor complications in patients with parkinson's disease. *CNS drugs 21*(8), 677–692.

Jarvis, E. D. (2004). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences 1016*, 749.

Jarvis, E. D. (2007). Neural systems for vocal learning in birds and humans: a synopsis. *Journal of Ornithology 148*(1), 35–44.

Jarvis, E. D. (2019). Evolution of vocal learning and spoken language. *Science 366*(6461), 50–54.

Jarvis, E. D., O. Güntürkün, L. Bruce, A. Csillag, H. Karten, W. Kuenzel, L. Medina, G. Paxinos, D. J. Perkel, T. Shimizu, et al. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience 6*(2), 151–159.

Jarvis, E. D., C. V. Mello, and F. Nottebohm (1995). Associative learning and stimulus novelty influence the song-induced expression of an immediate early gene in the canary forebrain. *Learning & memory 2*(2), 62–80.

Jarvis, E. D. and F. Nottebohm (1997). Motor-driven gene expression. *Proceedings of the National Academy of Sciences 94*(8), 4097–4102.

Jarvis, E. D., C. Scharff, M. R. Grossman, J. A. Ramos, and F. Nottebohm (1998). For whom the bird sings: context-dependent gene expression. *Neuron 21*(4), 775–788.

Jenner, P. (2008). Molecular mechanisms of l-dopa-induced dyskinesia. *Nature Reviews Neuroscience 9*(9), 665–677.

Jesse, F. and K. Riebel (2012). Social facilitation of male song by male and female conspecifics in the zebra finch, taeniopygia guttata. *Behavioural processes 91*(3), 262–266.

Jog, M. S., Y. Kubota, C. I. Connolly, V. Hillegaart, and A. M. Graybiel (1999). Building neural representations of habits. *Science 286*(5445), 1745–1749.

Kahl, R. (1878). Selected writings of hermann helmholtz.

Kao, M. H. and M. S. Brainard (2006). Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *Journal of neurophysiology 96*(3), 1441–1455.

Kao, M. H., A. J. Doupe, and M. S. Brainard (2005). Contributions of an avian basal ganglia–forebrain circuit to real-time modulation of song. *Nature 433*(7026), 638–643.

Karten, H. J., A. Brzozowska-Prechtl, P. V. Lovell, D. D. Tang, C. V. Mello, H. Wang, and P. P. Mitra (2013). Digital atlas of the zebra finch (taeniopygia guttata) brain: A high-resolution photo atlas. *Journal of Comparative Neurology 521*(16), 3702–3715.

Karten, H. J. and T. Shimizu (1989). The origins of neocortex: connections and lamination as distinct events in evolution. *Journal of Cognitive Neuroscience 1*(4), 291–301.

Kawai, R., T. Markman, R. Poddar, R. Ko, A. L. Fantana, A. K. Dhawale, A. R. Kampff, and B. P. Ölveczky (2015). Motor cortex is required for learning but not for executing a motor skill. *Neuron 86*(3), 800–812.

Keller, G. B. and R. H. Hahnloser (2009). Neural processing of auditory feedback during vocal practice in a songbird. *Nature 457*(7226), 187–190.

Kida, H. and D. Mitsushima (2018). Mechanisms of motor learning mediated by synaptic plasticity in rat primary motor cortex. *Neuroscience research 128*, 14–18.

King, A. P., M. J. West, and M. H. Goldstein (2005). Non-vocal shaping of avian song development: Parallels to human speech development. *Ethology 111*(1), 101–117.

Klatt, J. D. and J. L. Goodson (2013). Oxytocin-like receptors mediate pair bonding in a socially monogamous songbird. *Proceedings of the Royal Society B: Biological Sciences 280*(1750), 20122396.

Kleim, J. A., E. Lussnig, E. R. Schwarz, T. A. Comery, and W. T. Greenough (1996). Synaptogenesis and fos expression in the motor cortex of the adult rat after motor skill learning. *Journal of Neuroscience 16*(14), 4529–4535.

Klintsova, A. Y. and W. T. Greenough (1999). Synaptic plasticity in cortical systems. *Current opinion in neurobiology 9*(2), 203–208.

Konishi, M. (1964). Effects of deafening on song development in two species of juncos. *The Condor 66*(2), 85–102.

Konishi, M. (1965). The role of auditory feedback in the control of vocalization in the white-crowned sparrow 1. *Zeitschrift für Tierpsychologie 22*(7), 770–783.

Konopka, G. and T. F. Roberts (2016). Insights into the neural and genetic basis of vocal communication. *Cell 164*(6), 1269–1276.

Krakauer, J. W. (2009). Motor learning and consolidation: the case of visuomotor rotation. *Progress in motor control*, 405–421.

Krakauer, J. W., A. A. Ghazanfar, A. Gomez-Marin, M. A. MacIver, and D. Poeppel (2017). Neuroscience needs behavior: correcting a reductionist bias. *Neuron 93*(3), 480–490.

Krakauer, J. W. and P. Mazzoni (2011). Human sensorimotor learning: adaptation, skill, and beyond. *Current opinion in neurobiology 21*(4), 636–644.

Krakauer, J. W., Z. M. Pine, M.-F. Ghilardi, and C. Ghez (2000). Learning of visuomotor transformations for vectorial planning of reaching trajectories. *Journal of Neuroscience 20*(23), 8916–8924.

Kravitz, A. V., B. S. Freeze, P. R. Parker, K. Kay, M. T. Thwin, K. Deisseroth, and A. C. Kreitzer (2010). Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature 466*(7306), 622–626.

Kröner, S. and O. Güntürkün (1999). Afferent and efferent connections of the caudolateral neostriatum in the pigeon (columba livia): a retro-and anterograde pathway tracing study. *Journal of Comparative Neurology 407*(2), 228–260.

Krützfeldt, N. O. and J. M. Wild (2004). Definition and connections of the entopallium in the zebra finch (taeniopygia guttata). *Journal of Comparative Neurology 468*(3), 452–465.

Kubikova, L., K. Wada, and E. D. Jarvis (2010). Dopamine receptors in a songbird brain. *Journal of Comparative Neurology 518*(6), 741–769.

Kuebrich, B. and S. Sober (2015). Variations on a theme: Songbirds, variability, and sensorimotor error correction. *Neuroscience 296*, 48–54.

Kuhl, P. K. (2007). Is speech learning 'gated'by the social brain? *Developmental science 10*(1), 110–120.

Kuhl, P. K. and A. N. Meltzoff (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *The journal of the Acoustical Society of America 100*(4), 2425–2438.

Kuhl, P. K., F.-M. Tsao, and H.-M. Liu (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences 100*(15), 9096–9101.

Lackner, J. R. and P. Dizio (1994). Rapid adaptation to coriolis force perturbations of arm trajectory. *Journal of neurophysiology 72*(1), 299–313.

Lai, C. S., D. Gerrelli, A. P. Monaco, S. E. Fisher, and A. J. Copp (2003). Foxp2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain 126*(11), 2455–2462.

Lametti, D. R., S. M. Nasir, and D. J. Ostry (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *Journal of Neuroscience 32*(27), 9351–9358.

Lammel, S., B. K. Lim, C. Ran, K. W. Huang, M. J. Betley, K. M. Tye, K. Deisseroth, and R. C. Malenka (2012). Input-specific control of reward and aversion in the ventral tegmental area. *Nature 491*(7423), 212–217.

Leblois, A., B. J. Wendel, and D. J. Perkel (2010). Striatal dopamine modulates basal ganglia output and regulates social context-dependent behavioral variability through d1 receptors. *Journal of Neuroscience 30*(16), 5730–5743.

LeDoux, J. E., P. Cicchetti, A. Xagoraris, and L. M. Romanski (1990). The lateral

amygdaloid nucleus: sensory interface of the amygdala in fear conditioning. *Journal of neuroscience 10*(4), 1062–1069.

Lee, B. S. (1950). Effects of delayed speech feedback. *The Journal of the Acoustical Society of America 22*(6), 824–826.

Leonardo, A. and M. Konishi (1999). Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature 399*(6735), 466–470.

Li, X. and E. Jarvis (2001). Sensory-and motor-driven bdnf expression in a vocal communication system. *Soc Neurosci Meet 31*(538.8).

Lieberman, P., E. Kako, J. Friedman, G. Tajchman, L. S. Feldman, and E. B. Jiminez (1992). Speech production, syntax comprehension, and cognitive deficits in parkinson's disease. *Brain and language 43*(2), 169–189.

Liotti, M. and L. Ramig (2003). Improvement of voicing in patients with parkinson's disease by speech therapy. *Neurology 61*(9), 1316–1317.

Lipkind, D., G. F. Marcus, D. K. Bemis, K. Sasahara, N. Jacoby, M. Takahasi, K. Suzuki, O. Feher, P. Ravbar, K. Okanoya, et al. (2013). Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature 498*(7452), 104–108.

Liu, H., E. Q. Wang, L. V. Metman, and C. R. Larson (2012). Vocal responses to perturbations in voice auditory feedback in individuals with parkinson's disease. *PloS one 7*(3), e33629.

Liu, W.-c., J. Kohn, S. K. Szwed, E. Pariser, S. Sepe, B. Haripal, N. Oshimori, M. Marsala, A. Miyanohara, and R. Lee (2015). Human mutant huntingtin disrupts vocal learning in transgenic songbirds. *Nature neuroscience 18*(11), 1617.

Locke, J. L. and C. Snow (1997). Social influences on vocal learning in human and nonhuman primates.

Long, M. A. and M. S. Fee (2008). Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature 456*(7219), 189–194.

Lotharius, J. and P. Brundin (2002). Pathogenesis of parkinson's disease: dopamine, vesicles and alpha-synuclein. *Nature reviews neuroscience 3*(12), 932–942.

Luft, A. R., M. M. Buitrago, T. Ringer, J. Dichgans, and J. B. Schulz (2004). Motor skill learning depends on protein synthesis in motor cortex after training. *Journal of Neuroscience 24*(29), 6515–6520.

Luo, M. and D. J. Perkel (1999a). A gabaergic, strongly inhibitory projection to a thalamic nucleus in the zebra finch song system. *Journal of Neuroscience 19*(15), 6700–6711.

Luo, M. and D. J. Perkel (1999b). Long-range gabaergic projection in a circuit essential for vocal learning. *Journal of Comparative Neurology 403*(1), 68–84.

Machado, A. S., D. M. Darmohray, J. Fayad, H. G. Marques, and M. R. Carey (2015). A quantitative framework for whole-body coordination reveals specific deficits in freely walking ataxic mice. *Elife 4*, e07892.

Malenka, R. C. and R. A. Nicoll (1999). Long-term potentiation–a decade of progress? *Science 285*(5435), 1870–1874.

Mann, N., P. Slater, L. Eales, and C. Richards (1991). The influence of visual stimuli on song tutor choice in the zebra finch, taeniopygia guttata. *Animal Behaviour 42*(2), 285–293.

Marler, P. (1970). Birdsong and speech development: Could there be parallels?

there may be basic rules governing vocal learning to which many species conform, including man. *American scientist 58*(6), 669–673.

Martin, T., J. Keating, H. Goodkin, A. Bastian, and W. Thach (1996). Throwing while looking through prisms: I. focal olivocerebellar lesions impair adaptation. *Brain 119*(4), 1183–1198.

Matheson, L. E., H. Sun, and J. T. Sakata (2016). Forebrain circuits underlying the social modulation of vocal communication signals. *Developmental neurobiology 76*(1), 47–63.

Mathis, M. W., A. Mathis, and N. Uchida (2017). Somatosensory cortex plays an essential role in forelimb motor adaptation in mice. *Neuron 93*(6), 1493–1503.

Matsumoto, N., T. Hanakawa, S. Maki, A. M. Graybiel, and M. Kimura (1999). Nigrostriatal dopamine system in learning to perform sequential motor tasks in a predictive manner. *Journal of neurophysiology 82*(2), 978–998.

Mayer, U., O. Rosa-Salva, J. L. Loveland, and G. Vallortigara (2019). Selective response of the nucleus taeniae of the amygdala to a naturalistic social stimulus in visually naive domestic chicks. *Scientific reports 9*(1), 1–10.

McGurk, H. and J. MacDonald (1976). Hearing lips and seeing voices. *Nature 264*(5588), 746–748.

Medina, L. and A. Reiner (1995). Neurotransmitter organization and connectivity of the basal ganglia in vertebrates: implications for the evolution of basal ganglia (part 1 of 2). *Brain, behavior and evolution 46*(4-5), 235–246.

Medina, L. and A. Reiner (2000). Do birds possess homologues of mammalian primary visual, somatosensory and motor cortices? *Trends in neurosciences 23*(1), 1–12.

Mello, C. V. and D. F. Clayton (1994). Song-induced zenk gene expression in auditory pathways of songbird brain and its relation to the song control system. *Journal of Neuroscience 14*(11), 6652–6666.

Mello, C. V., D. S. Vicario, and D. F. Clayton (1992). Song presentation induces gene expression in the songbird forebrain. *Proceedings of the National Academy of Sciences 89*(15), 6818–6822.

Meltzoff, A. N. and M. K. Moore (1983). Newborn infants imitate adult facial gestures. *Child development*, 702–709.

Miyachi, S., O. Hikosaka, K. Miyashita, Z. Kárádi, and M. K. Rand (1997). Differential roles of monkey striatum in learning of sequential hand movement. *Experimental brain research 115*(1), 1–5.

Mohebi, A., J. R. Pettibone, A. A. Hamid, J.-M. T. Wong, L. T. Vinson, T. Patriarchi, L. Tian, R. T. Kennedy, and J. D. Berke (2019). Dissociable dopamine dynamics for learning and motivation. *Nature 570*(7759), 65–70.

Molina-Luna, K., A. Pekanovic, S. Röhrich, B. Hertler, M. Schubring-Giese, M.-S. Rioult-Pedotti, and A. R. Luft (2009). Dopamine in motor cortex is necessary for skill learning and synaptic plasticity. *PloS one 4*(9), e7082.

Mollaei, F., D. M. Shiller, and V. L. Gracco (2013). Sensorimotor adaptation of speech in parkinson's disease. *Movement Disorders 28*(12), 1668–1674.

Monfils, M.-H. and G. Teskey (2004). Skilled-learning-induced potentiation in rat sensorimotor cortex: a transient form of behavioural long-term potentiation. *Neuroscience 125*(2), 329–336.

Mongeon, D., P. Blanchet, and J. Messier (2013). Impact of parkinson's disease

and dopaminergic medication on adaptation to explicit and implicit visuomotor perturbations. *Brain and cognition 81*(2), 271–282.

Mooney, R. (2009a). Neural mechanisms for learned birdsong. *Learning & memory 16*(11), 655–669.

Mooney, R. (2009b). Neurobiology of song learning. *Current opinion in neurobiology 19*(6), 654–660.

Morris, D. (1954). The reproductive behaviour of the zebra finch (poephila guttata), with special reference to pseudofemale behaviour and displacement activities. *Behaviour 6*(1), 271–322.

Morton, S. M. and A. J. Bastian (2006). Cerebellar contributions to locomotor adaptations during splitbelt treadmill walking. *Journal of Neuroscience 26*(36), 9107–9116.

Murdoch, D., R. Chen, and J. H. Goldberg (2018). Place preference and vocal learning rely on distinct reinforcers in songbirds. *Scientific reports 8*(1), 1–9.

Murugan, M., S. Harward, C. Scharff, and R. Mooney (2013). Diminished foxp2 levels affect dopaminergic modulation of corticostriatal signaling important to song variability. *Neuron 80*(6), 1464–1476.

Nagel, K. I. and A. J. Doupe (2006). Temporal processing and adaptation in the songbird auditory forebrain. *Neuron 51*(6), 845–859.

Nakamura, K., K. Sakai, and O. Hikosaka (1999). Effects of local inactivation of monkey medial frontal cortex in learning of sequential procedures. *Journal of Neurophysiology 82*(2), 1063–1068.

Nasir, S. M. and D. J. Ostry (2008). Speech motor learning in profoundly deaf adults. *Nature neuroscience 11*(10), 1217–1222.

Nicholson, D. A., T. F. Roberts, and S. J. Sober (2018). Thalamostriatal and cerebellothalamic pathways in a songbird, the bengalese finch. *Journal of Comparative Neurology 526*(9), 1550–1570.

Nordeen, K. W. and E. J. Nordeen (1992). Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behavioral and neural biology 57*(1), 58–66.

Nordeen, K. W. and E. J. Nordeen (1993). Long-term maintenance of song in adult zebra finches is not affected by lesions of a forebrain region involved in song learning. *Behavioral and neural biology 59*(1), 79–82.

Norman-Haignere, S., N. G. Kanwisher, and J. H. McDermott (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron 88*(6), 1281–1296.

Nottebohm, F. and A. P. Arnold (1976). Sexual dimorphism in vocal control areas of the songbird brain. *Science 194*(4261), 211–213.

Nottebohm, F., J. A. Paton, and D. B. Kelley (1982). Connections of vocal control nuclei in the canary telencephalon. *Journal of Comparative Neurology 207*(4), 344–357.

Nottebohm, F., T. M. Stokes, and C. M. Leonard (1976). Central control of song in the canary, serinus canarius. *Journal of Comparative Neurology 165*(4), 457–486.

O'Donovan, K. J., W. G. Tourtellotte, J. Millbrandt, and J. M. Baraban (1999). The egr family of transcription-regulatory factors: progress at the interface of molecular and systems neuroscience. *Trends in neurosciences 22*(4), 167–173.

Okanoya, K. and A. Yamaguchi (1997). Adult bengalese finches (lonchura striata

var. domestica) require real-time auditory feedback to produce normal song syntax. *Journal of neurobiology 33*(4), 343–356.

Okuhata, S. and N. Saito (1987). Synaptic connections of thalamo-cerebral vocal nuclei of the canary. *Brain research bulletin 18*(1), 35–44.

Oller, D. K. and R. E. Eilers (1988). The role of audition in infant babbling. *Child development*, 441–449.

Ölveczky, B. P., A. S. Andalman, and M. S. Fee (2005). Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol 3*(5), e153.

Ota, N., M. Gahr, and M. Soma (2015). Tap dancing birds: the multimodal mutual courtship display of males and females in a socially monogamous songbird. *Scientific reports 5*(1), 1–6.

Ota, N., M. Gahr, and M. Soma (2018). Couples showing off: Audience promotes both male and female multimodal courtship display in a songbird. *Science advances 4*(10), eaat4779.

Paquet, F., M. Bedard, M. Levesque, P. Tremblay, M. Lemay, P. Blanchet, P. Scherzer, S. Chouinard, and J. Filion (2008). Sensorimotor adaptation in parkinson's disease: evidence for a dopamine dependent remapping disturbance. *Experimental brain research 185*(2), 227–236.

Parent, A. and L.-N. Hazrati (1995). Functional anatomy of the basal ganglia. i. the cortico-basal ganglia-thalamo-cortical loop. *Brain research reviews 20*(1), 91–127.

Paterson, A. K. and S. W. Bottjer (2017). Cortical inter-hemispheric circuits for multimodal vocal learning in songbirds. *Journal of Comparative Neurology 525*(15), 3312–3340.

Patriarchi, T., J. R. Cho, K. Merten, M. W. Howe, A. Marley, W.-H. Xiong, R. W. Folk, G. J. Broussard, R. Liang, M. J. Jang, et al. (2018). Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. *Science 360*(6396).

Perkel, D. J., M. A. Farries, M. Luo, and L. Ding (2002). Electrophysiological analysis of a songbird basal ganglia circuit essential for vocal plasticity. *Brain research bulletin 57*(3-4), 529–532.

Person, A. L., S. D. Gale, M. A. Farries, and D. J. Perkel (2008). Organization of the songbird basal ganglia, including area x. *Journal of Comparative Neurology 508*(5), 840–866.

Petersen, S. E., H. Van Mier, J. A. Fiez, and M. E. Raichle (1998). The effects of practice on the functional anatomy of task performance. *Proceedings of the National Academy of Sciences 95*(3), 853–860.

Petitto, L. A. (1993). On the ontogenetic requirements for early language acquisition. In *Developmental neurocognition: Speech and face processing in the first year of life*, pp. 365–383. Springer.

Petkov, C. I. and E. Jarvis (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Frontiers in evolutionary neuroscience 4*, 12.

Phillips, R. and J. LeDoux (1992). Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behavioral neuroscience 106*(2), 274.

Picardo, M. A., J. Merel, K. A. Katlowitz, D. Vallentin, D. E. Okobi, S. E. Benezra, R. C. Clary, E. A. Pnevmatikakis, L. Paninski, and M. A. Long (2016). Population-

level representation of a temporal sequence underlying song production in the zebra finch. *Neuron 90*(4), 866–876.

Plant, G. and B. Hammarberg (1983). Acoustic and perceptual analysis of the speech of the deafened. *Speech Trans. Lab. Q. Prog. Stat. Rep 2*(3), 85–107.

Price, P. H. (1979). Developmental determinants of structure in zebra finch song. *journal of comparative and physiological psychology 93*(2), 260.

Rasmussen, K. and B. L. Jacobs (1986). Single unit activity of locus coeruleus neurons in the freely moving cat: Ii. conditioning and pharmacologic studies. *Brain research 371*(2), 335–344.

Rasmussen, K., D. A. Morilak, and B. L. Jacobs (1986). Single unit activity of locus coeruleus neurons in the freely moving cat: I. during naturalistic behaviors and in response to simple and complex stimuli. *Brain research 371*(2), 324–334.

Redding, G. M., Y. Rossetti, and B. Wallace (2005). Applications of prism adaptation: a tutorial in theory and method. *Neuroscience & Biobehavioral Reviews 29*(3), 431–444.

Reiner, A., D. J. Perkel, L. L. Bruce, A. B. Butler, A. Csillag, W. Kuenzel, L. Medina, G. Paxinos, T. Shimizu, G. Striedter, et al. (2004). Revised nomenclature for avian telencephalon and some related brainstem nuclei. *Journal of Comparative Neurology 473*(3), 377–414.

Riede, T. and F. Goller (2010). Peripheral mechanisms for vocal production in birds–differences and similarities to human speech and singing. *Brain and language 115*(1), 69–80.

Rioult-Pedotti, M.-S., D. Friedman, and J. P. Donoghue (2000). Learning-induced ltp in neocortex. *science 290*(5491), 533–536.

Rodriguez-Oroz, M. C., J. Obeso, A. Lang, J.-L. Houeto, P. Pollak, S. Rehncrona, J. Kulisevsky, A. Albanese, J. Volkmann, M. Hariz, et al. (2005). Bilateral deep brain stimulation in parkinson's disease: a multicentre study with 4 years follow-up. *Brain 128*(10), 2240–2249.

Rogan, M. T., U. V. Stäubli, and J. E. LeDoux (1997). Fear conditioning induces associative long-term potentiation in the amygdala. *Nature 390*(6660), 604–607.

Rosenbaum, D. A., S. B. Kenny, and M. A. Derr (1983). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology: Human Perception and Performance 9*(1), 86.

Saravanan, V., G. J. Berman, and S. J. Sober (2020). Application of the hierarchical bootstrap to multi-level data in neuroscience. *arXiv preprint arXiv:2007.07797*.

Saravanan, V., L. A. Hoffmann, A. L. Jacob, G. J. Berman, and S. J. Sober (2019). Dopamine depletion affects vocal acoustics and disrupts sensorimotor adaptation in songbirds. *Eneuro 6*(3).

Sasaki, A., T. D. Sotnikova, R. R. Gainetdinov, and E. D. Jarvis (2006). Social context-dependent singing-regulated dopamine. *Journal of Neuroscience 26*(35), 9010–9014.

Scharff, C. and F. Nottebohm (1991). A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *Journal of Neuroscience 11*(9), 2896–2913.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of neurophysiology 80*(1), 1–27.

Schultz, W. (2017). Reward prediction error. *Current Biology 27*(10), R369–R371.

Schultz, W., P. Apicella, E. Scarnati, and T. Ljungberg (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. *Journal of neuroscience 12*(12), 4595–4610.

Schultz, W., P. Dayan, and P. R. Montague (1997). A neural substrate of prediction and reward. *Science 275*(5306), 1593–1599.

Schultz, W. and A. Dickinson (2000). Neuronal coding of prediction errors. *Annual review of neuroscience 23*(1), 473–500.

Scott, S. and F. Caird (1983). Speech therapy for parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry 46*(2), 140–144.

Shadmehr, R. and F. A. Mussa-Ivaldi (1994). Adaptive representation of dynamics during learning of a motor task. *Journal of neuroscience 14*(5), 3208–3224.

Shadmehr, R., M. A. Smith, and J. W. Krakauer (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual review of neuroscience 33*, 89–108.

Shima, K. and J. Tanji (2000). Neuronal activity in the supplementary and presupplementary motor areas for temporal organization of multiple movements. *Journal of neurophysiology 84*(4), 2148–2160.

Shimizu, T., K. Cox, and H. J. Karten (1995). Intratelencephalic projections of the visual wulst in pigeons (columba livia). *Journal of Comparative Neurology 359*(4), 551–572.

Simonyan, K., B. Horwitz, and E. D. Jarvis (2012). Dopamine regulation of human speech and bird song: a critical review. *Brain and language 122*(3), 142–150.

Smith, M. A. and R. Shadmehr (2005). Intact ability to learn internal models of

arm dynamics in huntington's disease but not cerebellar degeneration. *Journal of neurophysiology 93*(5), 2809–2821.

Sober, S. J. and M. S. Brainard (2009). Adult birdsong is actively maintained by error correction. *Nature neuroscience 12*(7), 927.

Sober, S. J. and P. N. Sabes (2003). Multisensory integration during motor planning. *Journal of Neuroscience 23*(18), 6982–6992.

Sober, S. J. and P. N. Sabes (2005). Flexible strategies for sensory integration during motor planning. *Nature neuroscience 8*(4), 490–497.

Sober, S. J., M. J. Wohlgemuth, and M. S. Brainard (2008). Central contributions to acoustic variation in birdsong. *Journal of Neuroscience 28*(41), 10370–10379.

Sohrabji, F., E. J. Nordeen, and K. W. Nordeen (1990). Selective impairment of song learning following lesions of a forebrain nucleus in the juvenile zebra finch. *Behavioral and neural biology 53*(1), 51–63.

Sossinka, R. and J. Böhner (1980). Song types in the zebra finch poephila guttata castanotis 1. *Zeitschrift für Tierpsychologie 53*(2), 123–132.

Steinberg, E. E., R. Keiflin, J. R. Boivin, I. B. Witten, K. Deisseroth, and P. H. Janak (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature neuroscience 16*(7), 966–973.

Stern, Y., R. Mayeux, A. Hermann, and J. Rosen (1988). Prism adaptation in parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry 51*(12), 1584–1587.

Stoel-Gammon, C. and K. Otomo (1986). Babbling development of hearing-impaired and normally hearing subjects. *Journal of Speech and Hearing Disorders 51*(1), 33–41.

Stratton, G. M. (1897). Vision without inversion of the retinal image. *Psychological review 4*(4), 341.

Stripling, R., S. F. Volman, and D. F. Clayton (1997). Response modulation in the zebra finch neostriatum: relationship to nuclear gene regulation. *Journal of Neuroscience 17*(10), 3883–3893.

Suthers, R. A., F. Goller, and J. M. Wild (2002). Somatosensory feedback modulates the respiratory motor program of crystallized birdsong. *Proceedings of the National Academy of Sciences 99*(8), 5680–5685.

Teramitsu, I., L. C. Kudo, S. E. London, D. H. Geschwind, and S. A. White (2004). Parallel foxp1 and foxp2 expression in songbird and human brain predicts functional interaction. *Journal of Neuroscience 24*(13), 3152–3163.

Thorndike, E. L. (1970). *Animal intelligence: Experimental studies*. Transaction Publishers.

Thoroughman, K. A. and R. Shadmehr (2000). Learning of action through adaptive combination of motor primitives. *Nature 407*(6805), 742–747.

Tobler, P. N., C. D. Fiorillo, and W. Schultz (2005). Adaptive coding of reward value by dopamine neurons. *Science 307*(5715), 1642–1645.

Tremblay, S., D. M. Shiller, and D. J. Ostry (2003). Somatosensory basis of speech production. *Nature 423*(6942), 866–869.

Tsai, H.-C., F. Zhang, A. Adamantidis, G. D. Stuber, A. Bonci, L. De Lecea, and K. Deisseroth (2009). Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science 324*(5930), 1080–1084.

Tsanas, A., M. A. Little, P. E. McSharry, J. Spielman, and L. O. Ramig (2012). Novel

speech signal processing algorithms for high-accuracy classification of parkinson's disease. *IEEE transactions on biomedical engineering 59*(5), 1264–1271.

Tseng, Y.-w., J. Diedrichsen, J. W. Krakauer, R. Shadmehr, and A. J. Bastian (2007). Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *Journal of neurophysiology 98*(1), 54–62.

Tumer, E. C. and M. S. Brainard (2007). Performance variability enables adaptive plasticity of 'crystallized'adult birdsong. *Nature 450*(7173), 1240–1244.

Vates, G. E., B. M. Broome, C. V. Mello, and F. Nottebohm (1996). Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches (taenopygia guttata). *Journal of Comparative Neurology 366*(4), 613–642.

Vicario, D. S. (1991). Organization of the zebra finch song control system: functional organization of outputs from nucleus robustus archistriatalis. *Journal of Comparative Neurology 309*(4), 486–494.

Von Helmholtz, H. (1925). *Helmholtz's treatise on physiological optics*, Volume 3. Optical Society of America.

Vonsattel, J.-P., R. H. Myers, T. J. Stevens, R. J. Ferrante, E. D. Bird, and E. P. Richardson (1985). Neuropathological classification of huntington's disease. *Journal of Neuropathology & Experimental Neurology 44*(6), 559–577.

Wada, K., H. Sakaguchi, E. D. Jarvis, and M. Hagiwara (2004). Differential expression of glutamate receptors in avian neural pathways for learned vocalization. *Journal of Comparative Neurology 476*(1), 44–64.

Warren, T. L., J. D. Charlesworth, E. C. Tumer, and M. S. Brainard (2012). Vari-

able sequencing is actively maintained in a well learned motor skill. *Journal of neuroscience 32*(44), 15414–15425.

Warren, T. L., E. C. Tumer, J. D. Charlesworth, and M. S. Brainard (2011). Mechanisms and time course of vocal learning and consolidation in the adult songbird. *Journal of neurophysiology 106*(4), 1806–1821.

Watabe-Uchida, M., L. Zhu, S. K. Ogawa, A. Vamanrao, and N. Uchida (2012). Whole-brain mapping of direct inputs to midbrain dopamine neurons. *Neuron 74*(5), 858–873.

West, M. J. and A. P. King (1988). Female visual displays affect the development of male song in the cowbird. *Nature 334*(6179), 244–246.

Whitlock, J. R., A. J. Heynen, M. G. Shuler, and M. F. Bear (2006). Learning induces long-term potentiation in the hippocampus. *science 313*(5790), 1093–1097.

Wild, J. and M. Williams (1999). Rostral wulst of passerine birds: Ii. intratelencephalic projections to nuclei associated with the auditory and song systems. *Journal of Comparative Neurology 413*(4), 520–534.

Wild, J. M. (1993). Descending projections of the songbird nucleus robustus archistriatalis. *Journal of Comparative Neurology 338*(2), 225–241.

Wild, J. M. (1994). Visual and somatosensory inputs to the avian song system via nucleus uvaeformis (uva) and a comparison with the projections of a similar thalamic nucleus in a nonsongbird, columbia livia. *Journal of Comparative Neurology 349*(4), 512–535.

Wild, J. M. (1995). Convergence of somatosensory and auditory projections in the avian torus semicircularis, including the central auditory nucleus. *Journal of Comparative Neurology 358*(4), 465–486.

Wild, J. M. (1997a). The avian somatosensory system: the pathway from wing to wulst in a passerine (chloris chloris). *Brain research 759*(1), 122–134.

Wild, J. M. (1997b). Neural pathways for the control of birdsong production. *Journal of neurobiology 33*(5), 653–670.

Wild, J. M. and S. M. Farabaugh (1996). Organization of afferent and efferent projections of the nucleus basalis prosencephali in a passerine, taeniopygia guttata. *Journal of Comparative Neurology 365*(2), 306–328.

Wild, J. M. and A. H. Gaede (2016). Second tectofugal pathway in a songbird (taeniopygia guttata) revisited: tectal and lateral pontine projections to the posterior thalamus, thence to the intermediate nidopallium. *Journal of Comparative Neurology 524*(5), 963–985.

Williams, H. (2001). Choreography of song, dance and beak movements in the zebra finch (taeniopygia guttata). *Journal of Experimental Biology 204*(20), 3497–3506.

Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological review 105*(3), 558.

Wolpert, D. M., J. Diedrichsen, and J. R. Flanagan (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience 12*(12), 739–751.

Woolley, S. and M. Kao (2015). Variability in action: Contributions of a songbird cortical-basal ganglia circuit to vocal motor learning and control. *Neuroscience 296*, 39–47.

Woolley, S. C. and A. J. Doupe (2008). Social context–induced song variation affects female behavior and gene expression. *PLoS Biol 6*(3), e62.

Woolley, S. M. and E. W. Rubel (1997). Bengalese finches lonchura striata domestica

depend upon auditory feedback for the maintenance of adult song. *Journal of Neuroscience 17*(16), 6380–6390.

Xiao, L., D. P. Merullo, M. Cao, A. Kulkarni, G. Konopka, T. F. Roberts, et al. (2020). Functional identification of specialized basal ganglia circuits that regulate vocal motor sequences. *bioRxiv*.

Xu, T., X. Yu, A. J. Perlik, W. F. Tobin, J. A. Zweig, K. Tennant, T. Jones, and Y. Zuo (2009). Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature 462*(7275), 915–919.

Yang, G., F. Pan, and W.-B. Gan (2009). Stably maintained dendritic spines are associated with lifelong memories. *Nature 462*(7275), 920–924.

Yizhar, O., L. E. Fenno, T. J. Davidson, M. Mogri, and K. Deisseroth (2011). Optogenetics in neural systems. *Neuron 71*(1), 9–34.

Yuan, R. C. and S. W. Bottjer (2020). Multidimensional tuning in motor cortical neurons during active behavior. *Eneuro 7*(4).

Zann, R. A. (1996). *The zebra finch: a synthesis of field and laboratory studies*, Volume 5. Oxford University Press.

Zeier, H. and H. J. Karten (1971). The archistriatum of the pigeon: organization of afferent and efferent connections. *Brain research 31*(2), 313–326.

Zhang, F., L.-P. Wang, M. Brauner, J. F. Liewald, K. Kay, N. Watzke, P. G. Wood, E. Bamberg, G. Nagel, A. Gottschalk, et al. (2007). Multimodal fast optical interrogation of neural circuitry. *Nature 446*(7136), 633–639.