Title

Cortical and Striatal Circuits for Learning Adaptive Behaviors and Wireless Ultrasonic Implants for Interfacing with the Nervous System

Author

Neely, Ryan M.

Publication Date

2017

Cortical and Striatal Circuits for Learning Adaptive Behaviors and Wireless Ultrasonic Implants for Interfacing with the Nervous System

By

Ryan M Neely

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

In

Neuroscience

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:

Professor Jose M. Carmena, Chair

Professor Yang Dan

Professor Linda Wilbrecht

Professor Jan M. Rabaey

Summer 2017

Abstract

Cortical and Striatal Circuits for Learning Adaptive Behaviors and Wireless Ultrasonic Implants for Interfacing with the Nervous System

by

Ryan M Neely

Doctor of Philosophy in Neuroscience

University of California, Berkeley

Professor Jose M. Carmena, Chair

Brain and nerve interface systems have shown early promise for alleviating a wide range of debilitating conditions. In the field of brain-machine interfaces (BMI), movement kinematics have been decoded from cortical neurons and used as a control signal for prosthetic devices. In the periphery, recent insights into the connection between nerves and organ systems has sparked new interest in the therapeutic potential of accessing and altering activity in peripheral nerves. Investigating fundamental mechanisms through which networks of neurons coordinate to produce adaptive responses can inform the design of next-generation nervous system interfaces. However, technological challenges must also be addressed before these systems are ready for widespread clinical adoption.

Using a brain-machine interface paradigm, we trained rodents to volitionally modulate the activity of primary visual cortex (V1) neurons. This approach allowed us to observe the instrumental learning process in cortical networks directly, and define the relationship between neural activity and behavioral outcomes. Learning occurred in the absence of visual input, suggesting that modulations were internally driven. Similar to demonstrations of instrumental learning in other cortical areas, learning in V1 engaged and required activity in the striatum, suggesting that cortico-striatal circuits are an essential component for behaviorally-relevant adaptation of cortical outputs. Next, we investigated how factors affecting behavioral choice are represented by striatal neurons as rodents performed a two-alternative probabilistic switching task. We found a rich encoding of task parameters in the dorsomedial striatum both at the level of single neurons and neural populations. We observed activity related to animals' confidence in the current state of the task, and found that confidence levels modulated the strength and timing of signals predicting behavioral choice. Finally, we sought to address the limitations of current methods for interfacing with the nervous system. We designed, built, and tested mm-scale wireless implants for recording electrical activity in peripheral nerves and muscles. This system, called neural dust, utilized ultrasonic backscatter as a scalable means for powering and communicating with miniaturized devices implanted deep in tissue. We showed that this system is capable of recording electroneurogram and electromyogram activity with high fidelity in living animals.

Acknowledgements

Five years ago, I don't believe I could have predicted the path that would eventually lead to this dissertation. When I entered graduate school, I had a neurobiology degree and a laboratory skill set that was best suited for studying clusters of neurons and slices of brain floating in dishes. But ever since grading some 50-odd undergraduate reports on clinical brain-machine interfaces as a teaching assistant, I couldn't ignore the excitement of neurotechnology and its possibilities for neuroscience and beyond. Even so, as a rotation student in my first year at Cal, I was tempted to stick with what I knew. Sitting among electrical and bio-engineers during in-depth discussions of control theory and decoder design was intimidating and felt like a long way away from the dendrites and synapses I was used to. Still, their enthusiasm was infectious and I couldn't help wanting a part of it too. At the other end of this road I am so glad that I took a chance, and also that Jose and the rest of the Carmena lab was willing to give me one. The range of topics discussed below- from the striatum to the sciatic nerve all the way to ultrasonic backscatter- speak to the breadth of opportunities for learning and discovery that I had in my time as a PhD candidate. I have grown a lot as both a researcher and a person and I don't think I could have asked for a better experience.

First and foremost, I have to thank my advisor, Jose. I don't think either of us ever expected to be heatedly discussing the peripheral nerves, and yet Jose was never hesitant to follow new lines of investigation when they had the potential to yield exciting results. "Collaborative" and "interdisciplinary" are popular buzzwords in academic science, but the Carmena lab, both the core group and our many collaborators, truly was a place where people from many different backgrounds came together to pursue common goals. It isn't trivial to foster this kind of environment, but for Jose it seems to come naturally. Combining these and other skills with a palpable excitement for science that follows him everywhere makes Jose an exceptional person to work with.

I also need to thank Michel Maharbiz helping to welcome me into the Berkeley neurotech family. I always looked forward to our meetings because I usually came away dreaming about some crazy new idea that, when you ran the numbers, just might be possible. My PhD would certainly not have been the same if I hadn't been introduced to the Maharbiz group, and I am very glad that I had the opportunity to be part of the team. A thank you is also in order for Rui Costa, whose sage advice and guidance came all the way from Lisbon, and later New York, and magically seemed to clarify concepts that I had been struggling with for days. I don't think I ever left a teleconference that included Rui and Jose without a huge smile on my face.

I had the pleasure of knowing several inspiring Berkeley faculty members, but I especially want to thank the members of my thesis committee, Professors Yang Dan, Linda Wilbrecht, and Jan Rabaey. It was always reassuring to meet with you and I'm glad to have had your guidance over the years.

A major part of what made my graduate experience so memorable was the people I saw every day, and I had the fortune of knowing some pretty awesome ones that deserve a big round of thanks: Aaron, for helping to get me up and running when I first arrived with no idea what I was doing. Preeya and Vivek, my fellow lab “veterans,” who were always available whenever I needed advice or to bounce around on the inflatable ball chairs. DJ and Konlin, for all of the late nights and early mornings we spent staring at oscilloscopes and clipping wires to shelves and tables trying to get neural dust to work. Thanks to postdoc superstars Sam, who always made great science look easy, and Nuria, who was a welcome companion down in the basement rat room. And of course, thank you to all of the many talented lab mates I have had the fortune to know and learn from: Amy, Suraj, Sid, Kelvin, Tanner, Albert, David, Axel, Arda. Keep on truckin’.

I am also lucky to have had some great support systems through the years. My parents, Matt and Denise, and my brother Taylor, were behind me no matter what and always willing to listen. Additionally, I am so glad I chose to become part of the Berkeley neuroscience community, who make Cal an exceptionally inspiring and inclusive place to learn to be a scientist, and I’m happy to know all those who have been part of the HWNI graduate program.

Finally, I can’t forget to thank my partner Kimberly, who has been by my side and helped me see this journey through from start to finish. Through good and bad you’ve always had my back, and I don’t know what I would have done without you.

Contents

Publications related to this work

Journal articles and peer-reviewed conference proceedings

Neely, R.M., Koralek, A.C., Athalye, V, Costa, R.M., Carmena, J.M. (2017). Volitional modulation of primary visual cortex activity requires the basal ganglia. In review.

Diaz-Botia, C., Luna, L., Neely, R., Chamanzar, M., Carraro, C., Carmena, J., Sabes, P., Maboudian, R. and Maharbiz, M. (2017). A silicon carbide electrode technology for the central and the peripheral nervous system. Journal of Neural Engineering.

Seo, D.*, Neely, R.M.*, Shen, K., Singhal, U., Alon, E., Rabaey, J.M., Carmena, J.M. and Maharbiz, M.M. (2016). Wireless recording in the peripheral nervous system with ultrasonic neural dust. Neuron, 91(3), 529-539.

Biederman, W., Yeager, D.J., Narevsky, N., Leverett, J., Neely, R., Carmena, J.M., Alon, E. and Rabaey, J.M. (2015). A 4.78 mm 2 fully-integrated neuromodulation soc combining 64 acquisition channels with digital compression and simultaneous dual stimulation. IEEE Journal of Solid-State Circuits, 50(4), 1038-1047.

Yeager, D., Biederman, W., Narevsky, N., Leverett, J., Neely, R., Carmena, J., Alon, E. and Rabaey, J. (2014, June). A 4.78 mm 2 fully-integrated neuromodulation SoC combining 64 acquisition channels with digital compression and simultaneous dual stimulation. In VLSI Circuits Digest of Technical Papers, 2014 Symposium on (pp. 1-2). IEEE.

Conference Abstracts

Neely, R.M., Seo, D., Shen, K., Singhal, U., Alon, E., Rabaey, J.M., Carmena, J.M. and Maharbiz, M.M. (2016). Neural dust: a wireless, mm-scale device platform for interfacing with the nervous system *In vivo*. Society for Neuroscience annual meeting. San Diego, CA, November 2016.

Neely, R.M., Koralek, A.C., Costa, R.M., Carmena, J.M. (2015). Operant control of primary visual cortex activity using a neuroprosthetic task in rodents. Society for Neuroscience annual meeting. Chicago, IL, October 2015.

Neely, R.M., Koralek, A.C., Carmena, J.M. (2014). Goal directed modulation of neural activity in rodent primary visual cortex via a brain-machine interface. Society for Neuroscience annual meeting. Washington, D.C., November 2014.

***Authors contributed equally**

Chapter 1:

Introduction

From the rhythmic motion of the gut to the remembered melody of a favorite song, the nervous system conducts the tissues of the body in collective harmony. Electrical impulses spark muscles into action, maintain body temperature, and extract meaning from acoustic vibrations in the air. To understand and manipulate the signals carried by nerves and neurons would be to gain access to the mechanisms that keep our bodies and minds in motion. For many years, methods and technologies for interfacing with the nervous system have been developed and refined. Patch-clamp techniques can record voltage fluctuations inside single cells, while functional magnetic resonance imaging has revealed the collective activity of entire brains. In recent decades, advances in brain-machine interface (BMI) technologies have leveraged developments in both technological improvements and biological understanding to create a bridge between the nervous system and electronic devices. Many such technologies seek to attain therapeutic endpoints in humans: cochlear implants to restore hearing and deep brain stimulators for essential tremor represent illustrative examples that have achieved widespread adoption (Chopra et al., 2013; van Schoonhoven et al., 2013). As a proof-of principal therapy to replace lost motor function, the activity of cortical neurons in humans and non-human primates has been decoded and used as a control signal to move artificial effectors (Aflalo et al., 2015; Bouton et al., 2016; Collinger et al., 2013; Gilja et al., 2015; Hochberg et al., 2012). These and other approaches demonstrate the potential application of nervous system interfaces to restore and even augment the normal functions of the human nervous system.

Moving beyond the state-of-the-art in BMI systems requires advances on two main fronts: 1) technological; including hardware design, materials science, and algorithms for decoding neural activity, and 2) biological; which entails a better understanding of how information is encoded in the nervous system, and how that information is translated into thought or action of muscles and organ systems. While progress in technological innovation and biological investigation often occur independently, a concerted approach can often be most effective. New tools can yield major insights into biological functions, while the questions posed by attempts to understand biology can guide and inspire technology development and design.

This dissertation attempts to strike such a collaborative balance between biology and technology in the design and implementation of nervous system interfaces. In this introduction, we first explore how BMI systems can be designed to study volitional control signals in the brain, with a focus on learned volitional modulation of cortical neurons. We then discuss the current limitations of nervous system interfaces imposed by the physical properties of the underlying biology, and improvements to the state-of-the-art that have potential to overcome these limitations. Finally, we will conclude with an overview of chapters to come.

1.1 Closed-loop motor brain-machine interface systems

Brain-machine interface systems can be designed to replace motor function, sensory function, or both. A common goal of BMI systems is to decode movement kinematics or movement intention from neural activity directly, and to translate this control signal into the movement of an artificial effector, such as a computer cursor or robotic arm. Effector control can occur in an open-loop or closed-loop context; here we focus on closed-loop implementations. Closed-loop BMI systems typically involve recording neural activity (such as spike activity, local field potentials (LFP), or electrocorticography (ECoG)), translating this activity into a control signal using a decoding algorithm, and then generating movement in an effector. Finally, to close the loop, the subject must receive some feedback about the movement that took place; for example, visual feedback, by watching the motion in real-time (Fig 1.1). Motor cortices, such as the primary motor cortex (M1) or the dorsal premotor cortex (PMd), are often targeted for this purpose due to the rich encoding of movement parameters that can be found in these regions. However, BMI systems have been implemented in a wide variety of brain regions, including the parietal cortex, temporal lobe, primary sensory cortex, and frontal eye fields, among others (Carmena et al., 2003; Cerf et al., 2010; Clancy et al., 2014; Fetz, 2007; Hwang et al., 2013; Musallam et al., 2004; Schafer and Moore, 2011; Shibata et al., 2011). In theory, precise and accurate decoding of movement kinematics from cortical activity should allow a subject to control an artificial effector using her natural motor repertoire. However, in practice, there are often many differences between how neural activity is translated into movement in a BMI system compared to a natural motor plant. Methods used to record neural activity can only capture a small fraction of the total motor-related activity in the brain, and must estimate a subject's true intention from this sample. Additionally, neural activity is

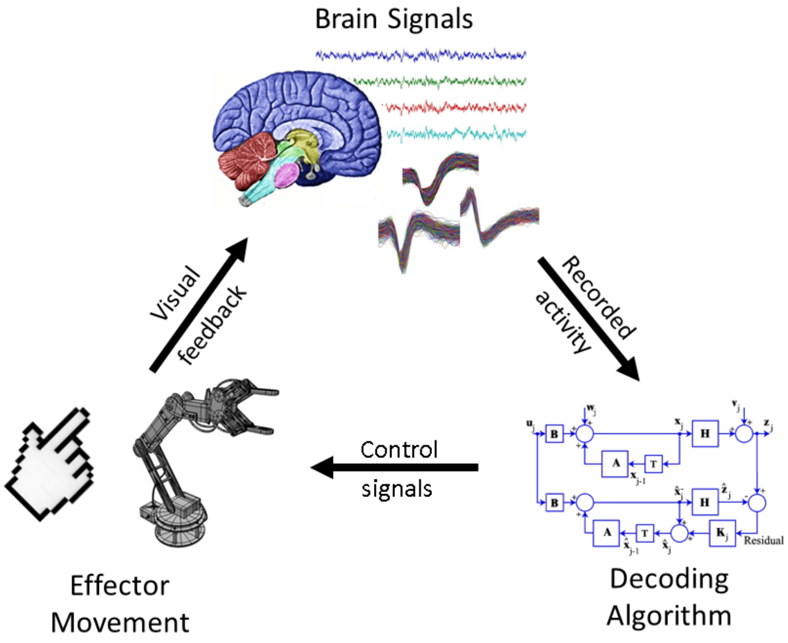


Figure 1.1: *Closed-loop BMI control. Brain activity is recorded from a subject, and then passed through a decoding algorithm, or decoder. The decoder translates neural activity into a control signal used to guide the movement of an artificial effector. Finally, the subject receives feedback about the movement that took place.*

often mapped to the movement of an effector with very different physical properties than an arm or hand. Finally, no BMI studies to date have replicated the full complement of sensory feedback that typically accompanies and guides natural motion, such as proprioceptive feedback or a sense of muscular effort. The result of these factors is that the mapping between a naïve subject's intended movement and the actual effector movement does not always precisely match.

1.2 BMI as a tool to study learning in neural circuits

This potential for mismatch between intended and actual movement in BMI systems creates an intriguing opportunity to study how feedback-based volitional control signals are learned in the brain. Just as the brain can learn new motor skills involving the natural motor plant, learning can also occur in the case of BMI systems when a motivated subject desires to achieve better control of the novel effector (Carmena et al., 2003). In contrast to studying natural motor systems, BMI learning provides a unique window into the cortical dynamics associated with learning for two primary reasons: 1) the final output layer of neurons directly responsible for control is known and observable by the experimenter. This is not always the case in natural motor control, where the mapping between cortical activity and muscle movement involves the activity of many intervening neural circuits. 2) the transform between neural activity and movement is well-defined. This allows both for a clear understanding of how neural activity generates movement, but also provides unique opportunities for perturbations. In the case of motor BMI systems, these properties have been exploited by fixing the decoder across many days and urging subjects to adapt their neural activity to improve control. This can result in new and stable tuning properties for the neurons controlling a cursor (Ganguly and Carmena, 2009). These tuning properties are unique to BMI control and revert back to earlier tuning configurations when a subject returns to manual control using the natural motor plant (Ganguly et al., 2011). Several studies have reported similar observations of neural adaptation as subjects learn in a BMI context (Arduin et al., 2013; Clancy et al., 2014; Jarosiewicz et al., 2008; Koralek et al., 2012, 2013; Prsa et al., 2017; Sadtler et al., 2014). In addition to studying the dynamics of neurons directly involved in learning to control a BMI (and their nearby neighbors), BMI can also be used to ask questions about the larger circuits that support learning in the cortex. Koralek et al (2012) report that the striatum is necessary for learning abstract neuroprosthetic skills using control signals from the primary motor cortex; this result will serve as a starting point for experiments in Chapter 2.

1.3 Limitations of state-of-the-art neural interfaces

1.3.1 Limitations of brain interfaces

Clinical applications of BMI in humans frequently rely on wired, invasive neural probes that penetrate cortical tissue (Aflalo et al., 2015; Bouton et al., 2016; Collinger et al., 2013; Gilja et al., 2015; Hochberg et al., 2012). A major hurdle for reliable implementation of BMI systems in humans is the longevity and reliability of high-quality neural recordings. Non-invasive systems, such as electroencephalogram (EEG) recordings, do not require surgery but are limited to recording low-frequency signals that reflect the activity of millions of neurons due to the attenuation of voltage signals through the skull. However, invasive neural probes, such as penetrating Utah arrays, cause damage to brain tissue that results in a chronic inflammatory response, cell death, and encapsulation of the electrodes by microglia (Biran et al., 2005; Polikov et al., 2005; Turner et al., 1999). Perhaps a related observation is that the ability of implanted electrodes to record spike activity degrades over a variable period of time that can range from weeks to years (Chestek et al., 2011; Suner et al., 2005), although damage of the electrode materials and encapsulation can also contribute to the degradation of signal quality. Wired systems are especially problematic for several reasons; first, these systems typically require tethering of the implant to the skull, which increases the micromotion of the implant relative to the tissue, and therefore exacerbates the tissue response to the electrodes (Biran et al., 2007). Secondly, externalized leads provide opportunities for infection and lead migration that could affect the location of the implant and imperil the health of the patient. Finally, wired systems can be bulky and inconvenient for subjects to use on a daily basis. These factors highlight the need for miniaturized, wireless probes for invasive recording of spike and LFP activity that are optimized for long-term use.

1.3.2 Limitations of peripheral nerve interfaces

In addition to the brain, recording and stimulating from the spinal cord and peripheral nerves has the potential to provide effective therapies for a wide range of conditions. Spinal cord stimulation for chronic pain using implanted neurostimulators has been an established therapy for decades (Cameron, 2004). More recently, closed-loop systems have been approved for the treatment of sleep apnea (Kezirian et al., 2010), and recent discoveries regarding the control of inflammation by cholinergic neurons have sparked new interest in therapies to relieve chronic inflammatory conditions by modulating peripheral nerve activity (Tracey, 2009). Inspired by these developments, an interdisciplinary community ranging from neurophysiologists to materials scientists developed a roadmap to next-generation peripheral nerve interfaces in late 2013 (Birmingham et al., 2014). In addition to a widespread effort to map the function of peripheral nerves, key goals highlighted in the report included improving peripheral nerve interfaces by moving towards miniaturized, implantable, and wireless systems. As discussed above, these features are also desirable for next-generation brain interfaces. The commonalities between features desired for brain as well as peripheral nerve interfaces highlights their importance and serves as the motivation for data presented in Chapter 4.

1.4 Chapter previews

In closed-loop BMI systems, subjects often improve control through the process of instrumental learning. In **Chapter 2**, we explore the possibility that instrumental learning is a common mechanism that tunes a variety of cortical circuits to improve behavioral outcomes. Specifically, we utilize a BMI paradigm to ask whether mice and rats can learn to modulate spike activity in the primary visual cortex (V1) in order to control an auditory cursor and obtain a reward. We explore the neural dynamics that accompany the learning process, and the sensitivity of learned performance to the presence or absence of visual input. Finally, we observe correlated activity in the region of the dorsomedial striatum that receives input from V1, and use an optogenetic strategy to determine whether activity in this region is necessary for learning

In **Chapter 3**, we focus on the encoding of task parameters in the dorsomedial striatum (DMS). We model the behavior of rats as they learn to perform a two-armed bandit task with partially observable states. We then ask how individual neurons in the DMS encode task parameters, including representations of task state. Finally, we use dimensionality reduction strategies to extract task-relevant signals from populations of neurons in DMS, and show that action choice representations are modulated by animals' confidence in the current state of the task.

In **Chapter 4**, we focus on technologies for wireless nervous system interfaces. In order to overcome the size-scaling limitations of systems that rely on electromagnetics to transmit power, we instead demonstrate that ultrasonic backscatter is a scalable means to power and communicate with mm-scale devices implanted deep in tissue. We show *in vivo* electroneurogram and electromyogram recordings from anesthetized rats using a wireless ultrasonic recording system called neural dust.

Chapter 2:

Volitional modulation of primary visual cortex activity requires the basal ganglia

Animals often acquire new behaviors through instrumental conditioning, a process that involves exploring variable behaviors and repeating and refining behaviors that lead to rewarding outcomes. In closed-loop brain-machine interfaces systems, instrumental learning is a process by subjects often learn to control novel effectors. Unlike instrumental conditioning of natural motor behaviors, learning in a BMI context involves the reinforcement of patterns or modulations of neural activity directly. Volitional control of neural activity to obtain future outcomes has been reported in BMI systems that engage frontal and motor cortices, a rich source of signals for voluntary action. However, evidence suggests that activity in primary sensory cortices may also reflect internally driven processes, instead of purely encoding antecedent stimuli. In this chapter, we leverage a BMI paradigm to show that rats and mice can learn to produce arbitrary patterns of neural activity in their primary visual cortex to control an auditory cursor and obtain reward. Furthermore, the ability to acquire the neural patterns that lead to reward was abolished when activity in the region of striatum that receives input from visual cortex was optogenetically inhibited. This inhibition did not affect the ability to produce the reinforced patterns after learning. These data demonstrate that cortico-basal ganglia circuits play a general role in learning to produce cortical activity that leads to desirable outcomes.

2.1 Introduction

Behavioral flexibility is essential for survival in changing and uncertain environments. Task-relevant modification or enhancement of sensory representations can be important to improve behavioral outcomes: for example, attentional resources can be used to amplify activity related to salient stimuli while ignoring distractors. Many sensory areas of the cortex, including primary sensory areas, display activity that reflects task parameters, changing behavioral context, and shifting attention, suggesting that computations in these regions are influenced by ongoing internal processes (Keller et al., 2012; Martinez et al., 1999; Niell and Stryker, 2010; Shuler, 2006; Steinmetz et al., 2000; Zhang et al., 2014). These task-relevant modulations of ongoing sensory representations can emerge and evolve following repeated training or association with a salient stimulus (Makino and Komiyama, 2015). An important question is how modulatory inputs to functionally diverse cortical circuits are tuned such that their outputs contribute positively to the behavioral outcomes of an individual. The basal ganglia, though the striatum, receives input from most cortical areas (Hintiryan et al., 2016; Kemp and Powell, 1970; McGeorge and Faull, 1989; Webster, 1965), feeds back to the cortex via multiple recurrent pathways (Redgrave et al., 2010), and dynamically encodes action-outcome contingencies (Samejima, 2005; Tricomi et al., 2004), making this structure a likely candidate to shape cortical activity based on behavioral experience (Barnes et al., 2005; Graybiel, 2008; Hinterberger et al., 2005; Swanson, 2000). Previous work has demonstrated the importance of the striatum for voluntary behavior and instrumental learning (Hikosaka et al., 1999; Yin et al., 2005, 2006, 2009). Similarly, we have shown that a brain-machine interface controlled by neurons in the primary motor cortex also requires cortico-striatal plasticity in order for animals to learn a novel neuroprosthetic action (Koralek et al., 2012, 2013). However, in addition to overt motor behaviors driven by motor cortices, cortico-striatal circuits have been theorized to also support abstract forms of learning, such as cognitive associations (Graybiel, 1997; Middleton and Strick, 1994). Furthermore, damage to basal ganglia structures in human patients, either through stroke or in diseases like Parkinson's, have been associated with deficits in sensory perception and the control of visual attention (Brown et al., 1997; Husain et al., 1997; Mercuri et al., 1997; Wright et al., 1990; Yamaguchi and Kobayashi, 1998). These data suggest that basal ganglia circuits may be involved in learning modulatory signals that influence many forms of cortical processing based on experience. However, observing and measuring these influences can be difficult, especially when their contributions to overt behavior may not be immediately apparent.

One strategy to overcome this difficulty is to use Brain-Machine Interfaces (BMIs) which directly map a subject's internally generated neural activity to the movement of an artificial effector. By explicitly defining the behavioral relevance of observable patterns of neural activity, BMI can be an important tool for studying how these patterns are generated. In clinical applications, the activity of cortical neurons in humans and non-human primates has been decoded as a proof-of-principle control signal to replace lost motor function by controlling prosthetic devices (Aflalo et al., 2015; Bouton et al., 2016; Collinger et al., 2013; Gilja et al., 2015; Hochberg et al., 2012). However, an important observation is that populations of cortical neurons whose activity is remapped to the movement of an artificial effector can undergo marked learning-related changes, and observing this learning process provides a unique window into how learning proceeds in the cortex (Arduin et al., 2013; Ganguly and Carmena, 2009; Ganguly et al., 2011; Jarosiewicz et al., 2008; Prsa et al., 2017; Sadtler et al., 2014). In order to better facilitate such observations, BMI studies can thus be designed to observe the acquisition and evolution of volitional control signals, rather than to optimize the performance and control of a complex effector.

Here, we asked whether neurons in the primary visual cortex, an area involved in processing low-level visual features, could be instrumentally conditioned to produce arbitrary modulations of ongoing spike activity, and whether this abstract form of learning was dependent on the basal ganglia. To address this question, we trained rats and mice to perform a neuroprosthetic task that virtually re-routed spike activity from the primary visual cortex (V1) into the frequency of an auditory cursor. This allowed us to facilitate and observe learned modulations of V1 activity with a known relationship to behavior. Animals trained on the task successfully learned to produce this novel action by voluntarily modulating spike activity in V1. Then, using the red-shifted inhibitory opsin Jaws (Chuong et al., 2014) to inactivate striatal neurons on a trial-by-trial basis, we then investigated to what degree this instrumental learning process in V1 was also dependent on activity in dorsomedial striatum.

2.2 Methods

2.2.1 Surgery

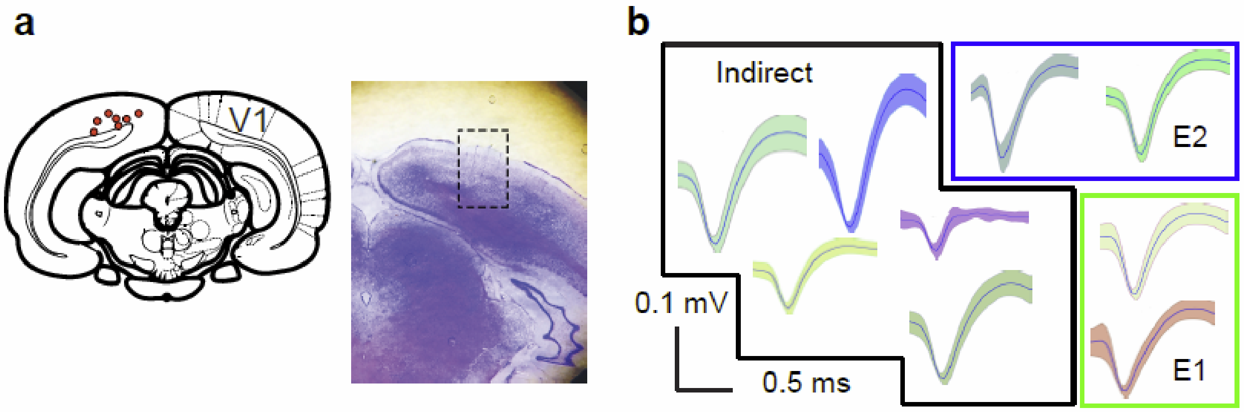


Figure 2.1*: **a:*** *Left: schematic showing the location of V1 electrode arrays, verified by histology for 6 of 10 animals. Right image shows an example section with electrode tracks.* ***b****: Sample V1 waveforms from an example session. Units outlined in green were chosen for Ensemble 1; units outlined in blue were chosen for Ensemble 2; and units outlined in black were recorded but had no direct relationship to cursor control (Indirect units).*

All rat experiments were performed in compliance with the regulations of the Animal Care and Use Committee at the University of California, Berkeley. A total of ten singly-housed, male Long-Evans rats weighing roughly 250 grams were used for the experiments. All rats were chronically implanted with microwire arrays in V1, with 8 rats also receiving implants in the dorsomedial striatum. Each array contained 16 or 32

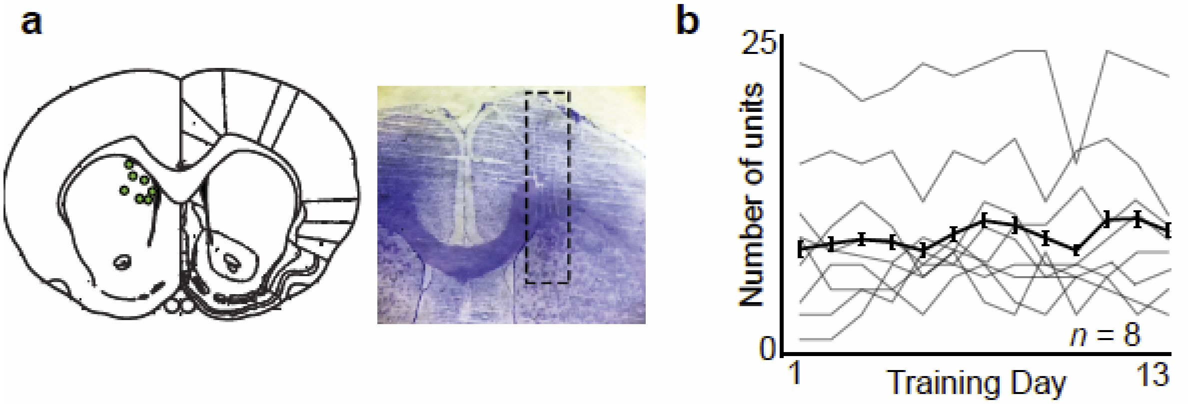


Figure 2.2 a*: Left: schematic showing the location of rat DMS electrode arrays, verified by histology for 6 of 8 animals. Right image shows an example section with electrode tracks.* ***b****: Total number of well-isolated DMS units for 8 animals recorded across several training sessions. Solid line shows the mean and S.E.M; grey lines show data for individual animals.*

tungsten microelectrodes (35 µm diameter, 250 µm electrode spacing, 8x2 or 8x4 configuration; Innovative Neurophysiology, Durham, NC). Stereotactic coordinates relative to bregma and lambda were used to center the arrays (1 mm anterior of lambda, 3.5 mm lateral from the midline, and 1.25 mm ventral from the cortical surface for V1; 1.8 mm anterior of bregma, 2 mm lateral of the midline, and 5.5 mm ventral from the cortical surface for DMS). V1 implants were targeted for layer 5 neurons based on insertion depth, and this was verified histologically at the end of experiments (Fig. 2.1a,b). Rodents were anesthetized with Ketamine (50 mg/kg) and Xylazine (5 mg/kg) with supplemental isoflurance gas as needed. Craniectomies were sealed with cyanoacrylate and rats were allowed to recover for ten days after implantation before behavioral training. Rats were given dexamethasone treatment (0.5 mg/kg) for one week following surgery to minimize tissue damage around the implant (Zhong and Bellamkonda, 2007).

Mouse experiments were performed in accordance with the Champalimaud Centre for the Unknown Ethics Committee guidelines and approved by the Portuguese Veterinary General Board (Direcção Geral de Veterinária, approval 0421/000/000/2014). A total of eight male C57BL mice of 2.5-3.5 months of age were used (four experimental mice and four control mice). Surgeries were performed under isoflurane anesthesia (1-3%). All mice were injected bilaterally with 1 µL of viral solution in the dorsomedial striatum using coordinates relative to Bregma (0.9 mm AP, ±1.5 mm ML, 2 mm below brain surface). Viral solution was injected through a glass pipette by pressure (Nanojet II from Drummond Scientific, 4.6 nl pulses at a rate of 0.2 Hz). For experimental animals, the virus injected was rAAV8/CamKII-Jaws-KGC-GFP-ER2 (University of North Carolina, titer $5.9x10^{12}$). For control animals, the virus injected was rAAV8/CamKII-GFP (University of North Carolina, titer $2.8x10^{12}$). For delivery of red light, mice were implanted bilaterally with optical fibers (250 µm diameter, NA 0.63). All mice were also implanted with 16-channel movable electrode arrays (electrode diameter 23 µm; Innovative Neurophysiology, Durham, NC) in the right primary visual cortex using coordinates relative to Lambda (0 mm AP, 0.3 mm ML, 0.6 mm below brain surface).

2.2.2 Electrophysiology

Single unit activity and local field potentials were simultaneously recorded with a Multichannel Acquisition Processor (MAP in rats, OmniPlex in mice; Plexon Inc., Dallas, TX). Activity was sorted using an online sorting application (Plexon Inc., Dallas, TX) prior to each daily recording session. Only units with a clearly identified waveform and signal-to-noise ratio greater than 2 were used. Sorting templates were further refined using an offline sorting application (Plexon Inc., Dallas, TX). Behavioral timestamps were sent to the MAP recording system through custom Python and C++ programs and synchronized to the neural data for later analyses. Recording arrays were grounded to a screw in the occipital bone, and both arrays were also referenced locally using the online program Ref2 (Plexon Inc., Dallas, TX) to eliminate effects of volume conduction.

For referencing, an electrode on each array was chosen to be subtracted from all other electrodes on that array. This was done independently for both V1 and DMS.

2.2.3 Behavioral Task

After recovering from surgery, rats were trained on the task in a dark behavioral chamber (Lafayette Instrument Company, Lafayette, IL) unless otherwise specified. During training, rats only received access to water during the task unless supplemental water was needed to maintain their body weight at a healthy level. At the start of each session, two ensembles of 2 well-isolated V1 units each were chosen for inclusion in the direct population based on SNR, interspike- interval histograms, and refractory periods. No other selection criteria were used to partition the recorded cells into each ensemble. Although these direct units were consistently well-isolated, we also ensured that many well-isolated units were included in the indirect population to enable a proper comparison. The units assigned to the direct population remained relatively constant throughout training using the stability of spike waveforms, sorting templates, and interspike intervals across sessions as a guide. After sorting and partitioning of direct and indirect units, a 15-minute baseline period was recorded in which animals received a sucrose water reward on a variable-interval schedule. During this time, spike counts were recorded for each ensemble binned into 100 ms bins, and a distribution of state values was calculated by subtracting the counts from E1 from E2 in each time bin. From here, the distribution was fit by a Gaussian mixture model (GMM) comprising between 1 and 10 Gaussian components (the exact number was chosen by finding the Akaike Information Criterion (AIC) value for each possible number of components, and choosing the number with the lowest AIC value). The probability density function (PDF) was then computed for the chosen model. By default, the neural state values (E1 – E2

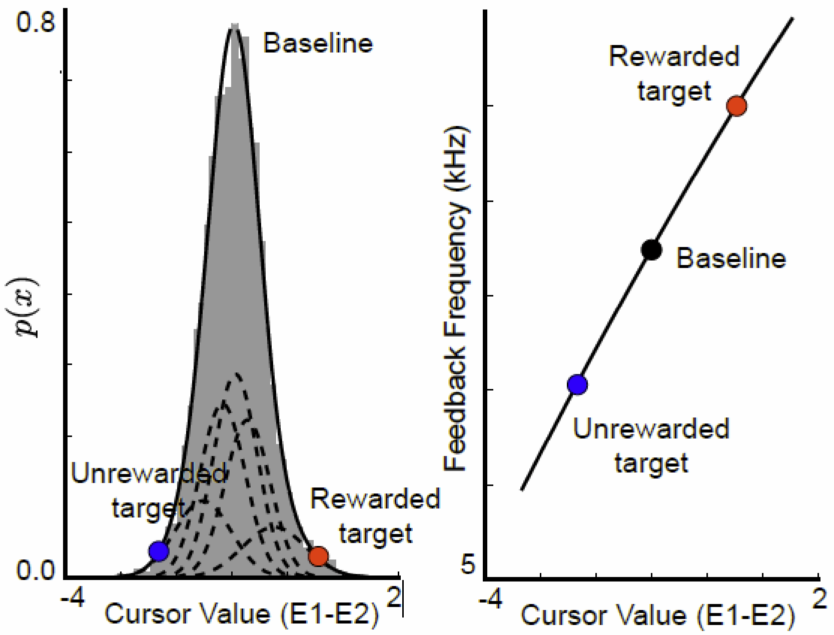


Figure 2.3 ***a****: Calculation of cursor transform. Plot shows a histogram of binned E1-E2 values (i.e., total spike counts for all ensemble 1 neurons minus total spike counts for all ensemble 2 neurons for each 100 ms bin) during a 15-minute baseline period. A mixture of gaussians (dashed lines) are fitted to generate the PDF (solid black line). Based on a simulation of the task using recorded baseline neural activity, thresholds are calculated for a low target (unrewarded in this plot), a high target (rewarded) and a midpoint (Baseline).* ***b:*** *The line fitted to the T2, Baseline, and T1 values that is used to compute a feedback frequency based on the binned E1-E2 value.*

counts for a 100 ms bin) corresponding to the low and high targets were set at the points on the computed PDF where the area under the curve was equivalent to 1.5% and 98.5% of the total area, respectively (Fig. 2.3a). However, these target values were iteratively updated by running a simulation using the data recorded in the baseline period until the probability of hitting each target was approximately 30% Finally, using the lowest and highest target values as well as the mean of the GMM, in combination with the lowest (1000 Hz), highest (15000 Hz) and midpoint (7000 Hz) frequencies, a 2-degree polynomial function was fit to these values in order to map neural state to a frequency value. During online performance, the state value used to compute the instantaneous feedback frequency was smoothed over the most recent 10 bins, updated every 100 ms. The rodents had to then precisely modulate these neuronal ensembles to move the cursor to one of two target frequencies, one which was randomly chosen using a coin flip on a per-animal basis to be associated with a 20% sucrose solution reward (and kept consistent across training sessions). Rodents were free to reach either target, although the cursor value had to return to the middle value for a new trial to begin. A trial was marked as a miss if neither of these target states were achieved within 30 seconds of trial initiation. Recorded neural data was entered in real time to custom routines in Python and C++ that then translated those activity levels into the appropriate feedback frequency as described above and played the frequency on speakers mounted above the behavioral chamber. When a target was hit, a Data Acquisition board (National Instruments, Austin, TX) controlled by a Python script triggered the operant box to supply the appropriate reward to rodents.

2.2.4 Behavioral manipulations

After initial training of at least 8 days, a contingency degradation was performed. This took place over two sessions: during the first session, animals were allowed to achieve stable performance on the task, which took approximately 30 minutes (pre-degradation). Then, reward delivery was decoupled from task performance and delivered on a variable-interval schedule that approximated the rate of reward during training conditions (CD). During this time, animals still received auditory feedback lined to their neural state. The contingency degradation continued for the remainder of the session. The next day, animals were again trained on the task under normal conditions (reinstatement). Similarly, for contingency reversal sessions, we reversed the contingency approximately 30 minutes into a training session. During sessions that involved light manipulation, we again waited for animals to achieve saturating performance in a single session, and then placed a light into the behavioral chamber in an overhead configuration.

2.2.5 Optical inhibition

For optical inhibition experiments, red light was applied to the striatum of both experimental and control groups on 50% of all trials in a session. Light was applied through a fiber-coupled LED system (Prizmatix, Givat-Shmuel, Israel). Power levels tested through the system at the optical fiber tip ranged from 4-6 mW. Red light application consisted of a square pulse that lasted the full duration of the current trial, from trial initiation until either a target or timeout was reached. Both groups were trained in this manner for 7 days. Next, the experimental animals were trained for 5 additional days in the absence of red light. After this initial training, both groups were tested to determine the role of striatal circuits in learning versus performance of the task. On day LED 50, red light was applied on 50% of all trials. On day LED late, no light was applied for the first 45 trials, after which red light was applied on 50% of all remaining trials.

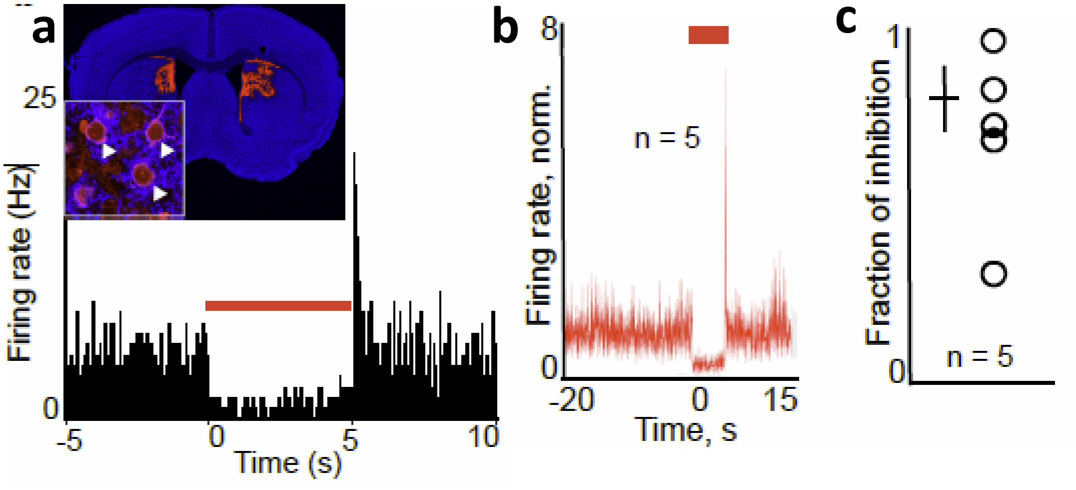


***Figure 2.4 a**: Jaws-mediated inhibition of DMS neurons. Inset shows a coronal section stained for neuronal cell bodies (blue) and Jaws expression (red). Zoomed inset shows a magnification with labeling apparent on single neurons (white arrows). Histogram plot shows suppression of spike activity in one example DMS neuron during Jaws-mediated inhibition (red bar). **b:** Mean firing rates of all recorded DMS units during Jaws-induced inhibition with red light (red bar); N = 5. Firing rates were normalized to baseline firing rate values. Shaded areas indicated SEM. **c**: Fraction of inhibition for recorded DMS units during Jaws-induced inhibition, where a value of 1 indicates complete inhibition of all spike activity. Grey circles show individual units.*

2.2.6 Data Analysis

Analyses were performed in Python with custom- written routines utilizing publicly available software packages. Unit data were first binned in 1 ms time bins and digitized. Firing rate analyses were performed in relation to target achievement as indicated in figures. Unless otherwise specified, firing rates were binned into 50 ms bins for all analyses. Only two-sided statistical tests were used to determine significance. The term “early” indicates that analyses were performed using data taken from animals during the first 3 days of training (inclusive), while “late” specifies data taken from animals during the final 3 days of training. The “first trials” of a session indicates trials in the first 10 minutes, while the “last trials” of a session are defined as occurring in the last 10 minutes, unless otherwise specified in the text. Modulation depths were computed as the difference between the maximum and minimum firing rate values in a 2-second window centered around target achievement. Coherence analyses were performed using algorithms translated to Python from the Chronux toolbox (http://chronux.org) in conjunction with custom routines in Python. A multi-taper method was used to compute spectral estimates of spiking and LFP activity (Jarvis and Mitra, 2001; Thomson, 1982). A total of 5 tapers were used with a time-bandwidth product of 3, and estimates were computed every 50 ms with a window size of 500 ms. Coherence between spiking in LFP activity was calculated and defined as:

$$C_{xy} = \frac{|R_{xy}|}{\sqrt{|R_{xx}|}\sqrt{|R_{yy}}}$$

where R_{xx} and R_{yy} are the power spectra and R_{xy} is the cross-spectrum. Spectral analyses were calculated relative to the delivery of reward and averaged across trials and animals.

Coherence estimates can be affected by firing rate(Lepage et al., 2011) and we therefore performed a thinning procedure to equate firing rates between conditions in which rates differed (Gregoriou et al., 2009). Trial-averaged spike trains in the neuronal populations were smoothed with a moving average of 10 ms. The difference in firing rate between the populations normalized by the maximum firing rate at a given time point determined the probability that a spike would need to be removed from the population with a higher firing rate. Spikes were then removed from the population with a higher firing rate based on this probability in order to eliminate any possible influence of firing rate on coherence estimates.

The signal-to-noise ratio for each recorded waveform was quantified as:

$$SNR = \frac{A}{2 * SD_{noise}}$$

where A is the peak-to-peak voltage of the mean waveform and SD_{noise} is the standard deviation of the residuals from each waveform after the mean waveform has been subtracted (Suner et al., 2005). Units included in the analysis had a minimum SNR of 2.

For logistic regression analyses, we used functions from the publicly available python package scikit-learn (www.scikit-learn.org). Regression was performed using a window of spike activity 500 ms prior to target hits, binned into 50 ms bins. L2 Regularization was done using cross-validation to determine the optimum regularization parameter. 3-fold cross validation was performed 5 times using left out data to compute accuracies, and the average of all 5 results was taken to be the final accuracy value. Chance rates were taken as the accuracy of the analyses using shuffled data. To determine significance values, a permutation test was used that compared the accuracy of the prediction using the original dataset compared to dataset in which target identities for all trials were shuffled. Neural activity was considered to be significantly predictive of target choice if the accuracy of the prediction using the original dataset outperformed the accuracy of the shuffled version on 95% of 500 iterations.

2.3 Results

2.3.1 A V1-driven closed-loop brain-machine interface

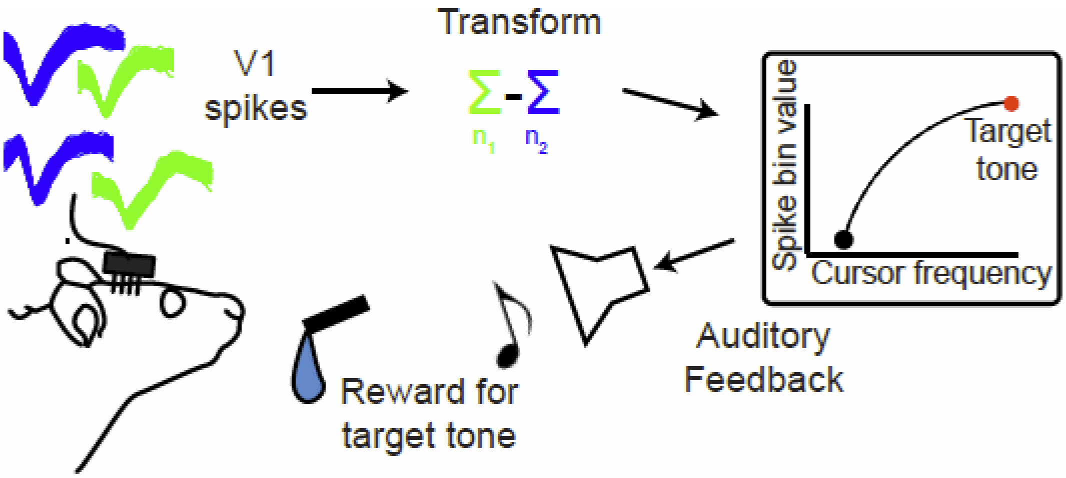


Figure 2.5: *Schematic of V1-BMI paradigm. Activity of well-isolated V1 units (top left) were used to generate auditory tones using a differential transform (top right). Animals were rewarded for producing a target tone (red). A second tone (black) at the opposite end of the frequency range terminated the trial but was not rewarded.*

We adapted our previously developed neuroprosthetic task for rodents (Koralek et al., 2012) in order to directly study abstract reinforcement learning in V1 (Fig. 2.5). Briefly,

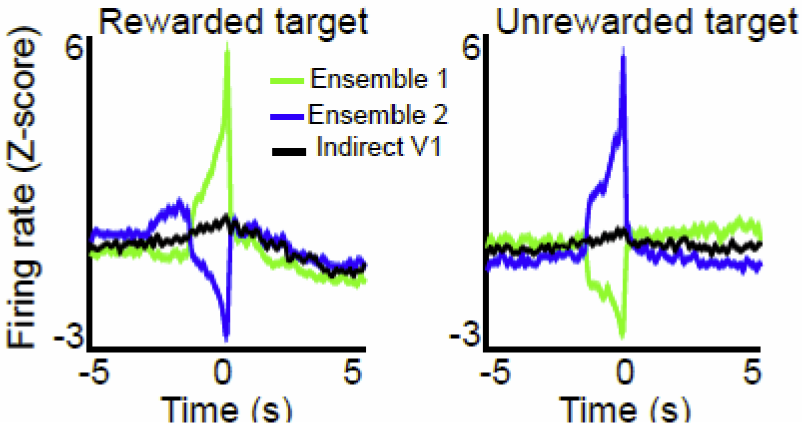


Figure 2.6: ***Left***, *average z-scored firing rates of V1 neurons arbitrarily assigned to ensemble 1 (green), ensemble 2 (blue), or unassigned (indirect; black), time-locked to rewarded targets. Shaded areas show S.E.M.* ***right***, *Same as in left, but time-locked to the unrewarded target. Increased activity in Ensemble 1 units moved the tone frequency in the opposite direction as increased activity in Ensemble 2 units.*

10 rats (and 8 mice, see later) were implanted chronically with microwire electrode arrays positioned in V1 layer 5 (L5), allowing us to isolate and record individual units (Fig. 2.5, Fig. 2.1a,b). In addition to V1, 8 rats were also implanted with electrode arrays in the dorso-medial striatum (DMS) near the projection target of V1 (Hintiryan et al., 2016; Khibnik et al., 2014; McGeorge and Faull, 1989) (Fig. 2.2a,b). During the course of the experiment, animals were placed in a totally dark or lighted behavioral chamber and allowed to move freely while listening to auditory feedback that reported their neural state in real time. Each day, two neural ensembles, consisting of two well-isolated units each, were randomly chosen to directly control the continuous auditory cursor (direct units), while the remaining units recorded in V1 had no defined relationship to cursor control (indirect units, Fig. 2.1b). Activity of the two direct-unit ensembles had an opposing relationship, such that spikes produced by Ensemble 1 (E1) moved the cursor closer to the rewarded frequency, while spikes in Ensemble 2 (E2) moved the cursor away from the rewarded frequency and towards the unrewarded frequency (Fig. 2.6). The highest and lowest possible tones were randomly assigned to be rewarded or unrewarded for each animal, and this association remained constant for the duration of training. Prior the start of every session, a baseline distribution of neural states (binned E1 - E2 spike counts) was used to initialize the target values such that the chance rate of attaining each target (rewarded and unrewarded) within a trial period was approximately 30% (Fig. 2.3; see Methods). Animals were given 30 seconds to reach either target; otherwise the trial was considered a miss and animals received a white noise burst followed by a time-out.

2.3.2 Rodents learn to control V1 activity patterns

Over the course of 7-12 training sessions (average 9.11 sessions), rats learned to perform the task well above chance level (Fig. 2.7). Animals in the late phase of

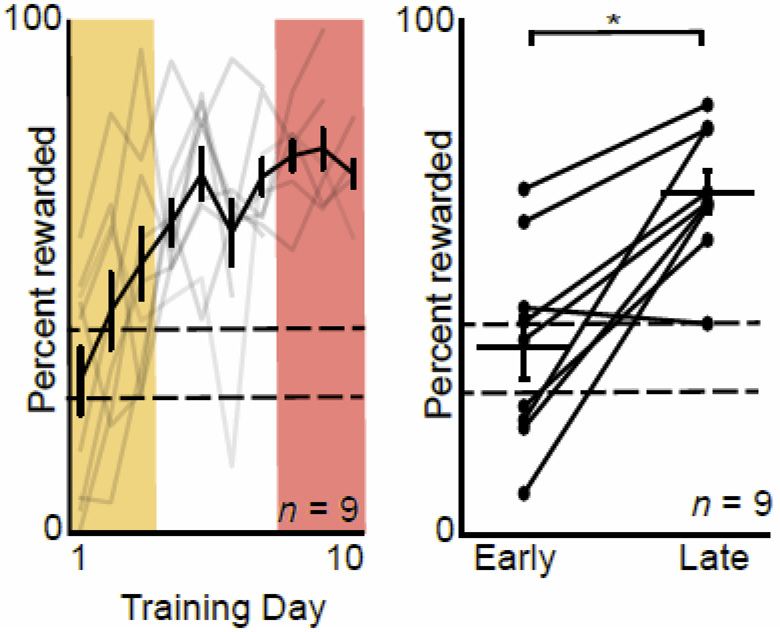


Figure 2.7: ***Left***, *Timecourse of learning across training days. Bold line shows the mean and S.E.M. across 9 rats; grey lines show learning curves for animals individually. One animal only completed 4 sessions; data for this animal has been excluded from this plot. Dashed lines bound the range of chance performance levels. Orange highlighted region shows data range classified as the early learning phase for all animals; red region shows range for late learning phase.* ***Right***, *Comparison of performance between early learning phases and late learning phases. (N = 9 animals, mean for the first 3 days = 39.6% rewarded; mean for the last 3 days 67.2%; P = 0.00162 (paired t-test)).*

learning, considered here the final 3 training sessions, exhibited significant improvements in the percentage of rewarded targets compared to their performance in the early phase (during the first three sessions of training) (Fig. 2.7; mean of the first three days for all animals = 39.7%; mean of the last 3 days = 67.2%; *P* = 0.00162).

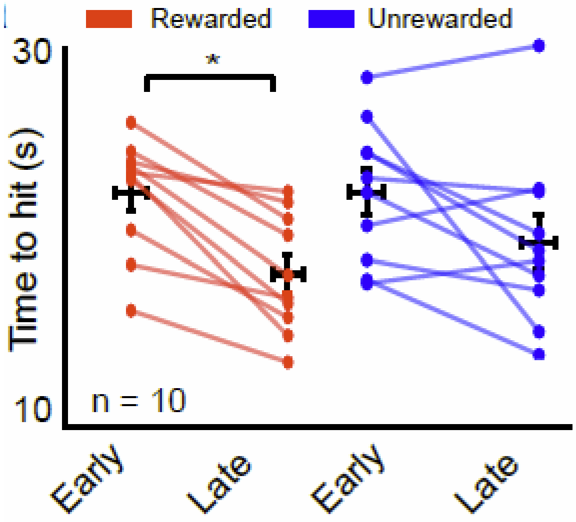


***Figure 2.8**: latencies to rewarded vs unrewarded targets for early training (first three days of training) compared to late training (last three days of training). Mean rewarded targets, early = 21.413 seconds; mean rewarded late = 17.965 seconds; P = 0.0156 (paired t-test). Mean unrewarded targets, early = 20.731 seconds; unrewarded late = 18.131 seconds; P = 0.0684 (paired t-test).*

Additionally, over this same interval, latencies to rewarded targets decreased significantly, while changes in latencies to unrewarded targets was non-significant (Fig 2.8). We observed that simply pairing particular auditory tones with reward was not sufficient to drive V1 activity. After 5 days of performance above chance level, we decoupled auditory tones from neural activity mid-session by playing back the sequence of tones generated in the first part of the session. Although reward was still delivered in tandem with the rewarded tone, modulation of the direct unit ensembles was markedly decreased when animals' neural activity was not driving the cursor (Fig 2.9). This suggests that the learning we observed was not an effect of classical conditioning, and instead resulted from an intentional modulation of V1 activity. Because the chance rate for each target was reset at the start of each session to approximately 30%, increases

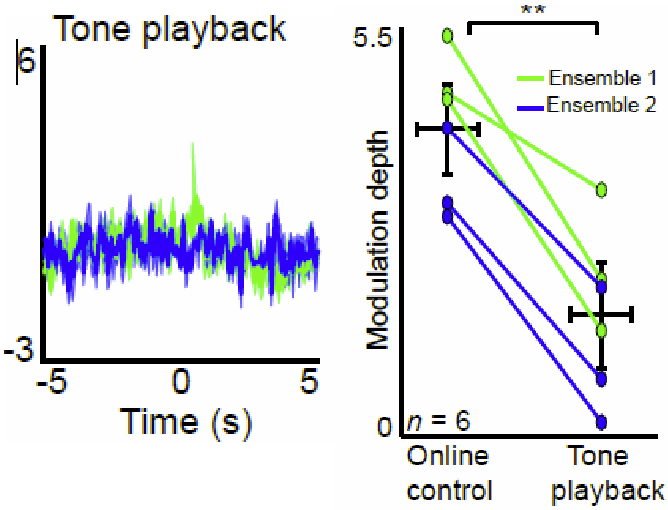


***Figure 2.9**: **Left**: Ensemble 1 (green) and ensemble 2 (blue) modulations during passive playback of tones decoupled from ongoing neural activity, time-locked to the rewarded target tone. **Right**: Modulation depth of Ensemble 1 and Ensemble 2 during online control compared to tone playback. Mean modulation depth online control = 4.341; mean tone playback = 1.739; P = 0.000348 (paired t-test). Black bars show mean and S.E.M*

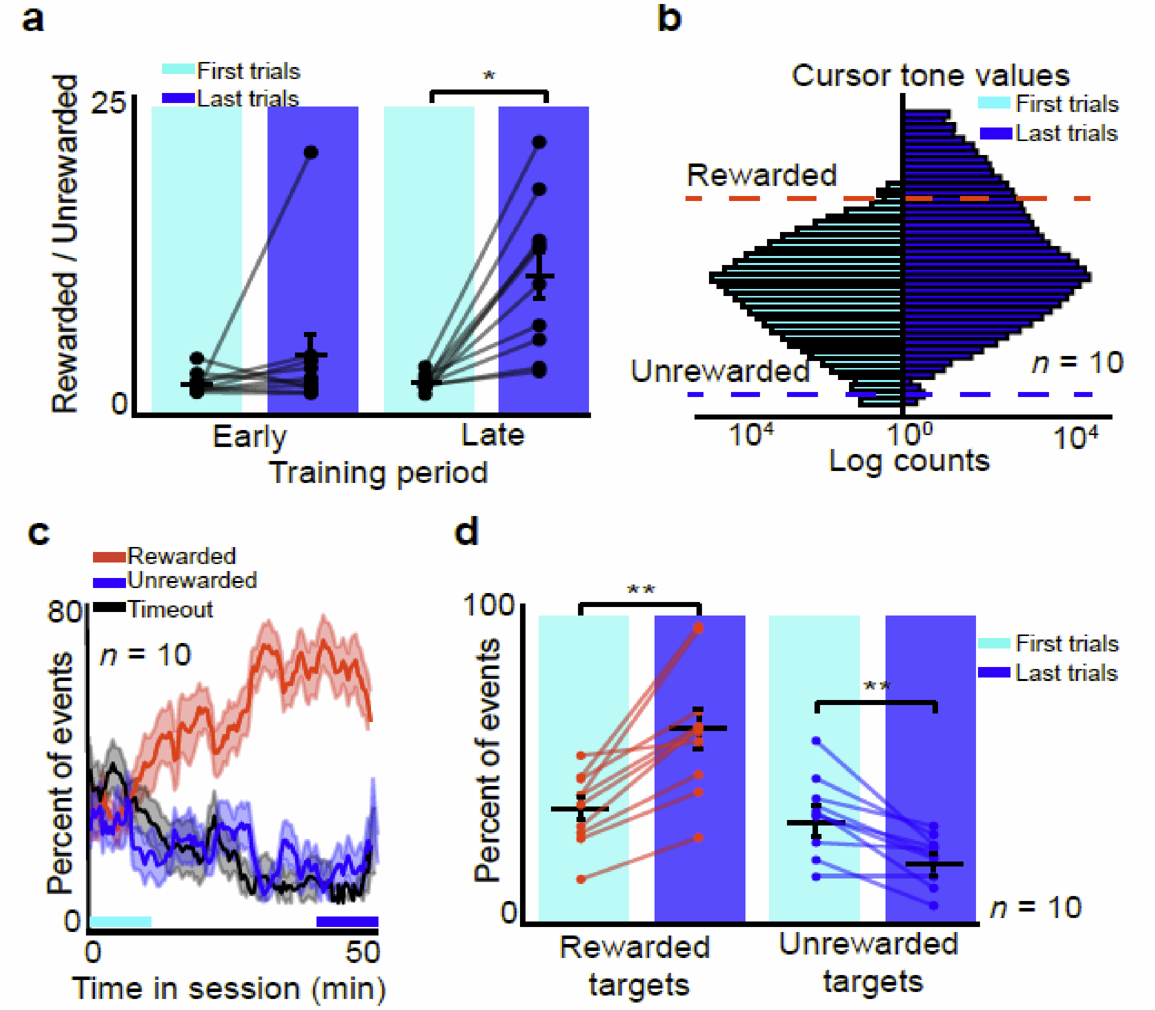


Figure 2.10: ***a***, *comparison of within-session improvements during the early learning period (first 3 days) of training relative to the late learning period (last 3 days) for each animal (N = 9 animals), expressed as the ratio of rewarded to unrewarded targets. Data to compute the ratio for first trials (shaded in cyan) was averaged over the first 10 mins of each session, while data for the last trials (shaded in blue) was averaged over the last 10 mins. During the early learning period, the mean rewarded/unrewarded ratio for first trials was 1.12, last trials session = 3.848; P = 0.275 (paired t-test). For the late period, mean ratio early in session = 1.195, mean late in session = 11.667; P = 0.00130 (paired t-test).* ***b,*** *Mean distribution of cursor values for all animals for the first trials (first 10 mins) compared to the last trials (last 10 mins) of sessions during the late period. Cyan bars show the initial distribution, based on baseline activity, used to set the task parameters, while blue bars show the distribution at the end of the training session for the last trials. Dashed lines show the thresholds corresponding to the rewarded and unrewarded targets.* ***c****, Timecourse of mean within-session task learning during the late period of training for all animals. Shaded areas represent S.E.M. Chance rates for rewarded and unrewarded targets were set at approximately 30% at the start of each training session. Comparisons in performance were made during over first 10 mins of each session (first trials, cyan bar) and over the last 10 minutes (last trials, blue bar) of each session. N = 10 animals, mean of 3 sessions per animal.* ***d****: Quantification of rewarded and unrewarded target hits for the first trials compared to the last trials; data same as c. N = 10 animals, mean of the late period (last 3 sessions) for each animal. Paired t-test between first trials and last trials for rewarded targets: P = 0.00055; mean early = 36.9%, mean late = 59.1%. Paired t-test between first trials and last trials for unrewarded targets: P = 0.00145; mean early = 33.0%, mean late = 20.6%. Red and blue indicate the rewarded and unrewarded targets, respectively (** indicates P < 0.001). Black lines show mean and S.EM.*

in performance seen as animals progressed from early to late phases reflected greater improvements within single training sessions across days. We quantified this by comparing performance in the first 10 minutes of a session (first trials) and the last 10 minutes (last trials) (Fig. 2.10a,). This suggests that the learned ability to control V1 activity was retained between training sessions, even though animals needed to perform some *de novo* learning to adjust to the initial parameters of the transform in any given day. We observed that in late learning, over the course of single training sessions, animals acquired a strong preference for the rewarded target relative to the unrewarded target (Fig. 2.10c,d). This was reflected in the shift of the distribution of auditory cursor values in the direction corresponding to the rewarded target tone (Fig. 2.10b). Additionally, we observed that auditory feedback was necessary for learning: sessions in which the feedback tones were muted resulted in no significant difference between the number of rewarded and unrewarded targets (*P* = 0.738, Fig. 2.11), even though these no-feedback sessions were conducted after several days of successful normal training. These data demonstrate that closed-loop neurofeedback-based reinforcement training can be used to condition the activity of neurons in the primary visual cortex.

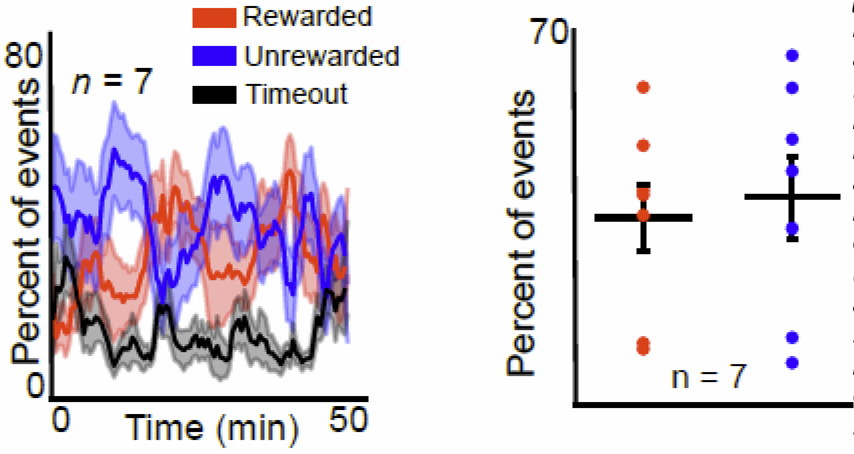


***Figure 2.11**: **Left**: Timecourse of performance when animals were not given auditory feedback about cursor position during training. **Right**: Comparison of rewarded and unrewarded target hits during sessions without auditory feedback Data is the mean over the last 10 minutes of each session. N = 7 animals (mean of 1.5 sessions per animal; range 1-2 sessions). Mean rewarded targets = 38.1%; mean unrewarded = 41.7%; P = 0.738 (unpaired t-test). Horizontal black line shows mean and vertical bars show standard error.*

2.3.3 V1 modulation is sensitive to task contingency

We next investigated the sensitivity of performance to changing task conditions. In late learning animals were able to quickly shift their neural activity to produce the rewarded tone frequency once the auditory feedback began (Fig. 2.12). We asked whether animals' behavior was habitual and therefore insensitive to changes in action-reward contingencies; or goal-directed, in which performance remains sensitive to changing task contingencies (Dias-Ferreira et al., 2009). To test if performance of the V1-controlled task fit either of these regimes, we degraded the task contingency so that animals received randomly timed rewards irrespective of target hits, but at a similar rate.

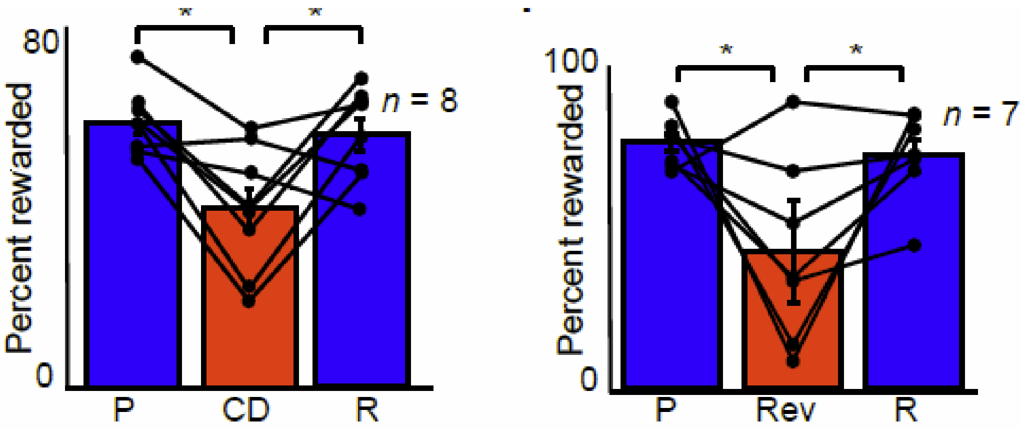


Figure 2.12: Left *Percentage of rewarded trials for contingency degradation sessions. Bars show means, lines show individual animals. N = 8 animals, mean of 2 sessions per animal. P: pre-degradation, mean = 64.5%. CD: peri-degradation (reward decoupled from cursor), mean = 43.1%. R; reinstatement of reward, mean = 61.4%. Paired t-test between pre- (P) and peri- (CD) degradation: P = 0.0038. Paired t-test between CD and reinstatement (R): P = 0.0283. Paired t-test between pre-degradation (P) and reinstatement (R): P = 0.289.* ***Right****, Quantification of performance in contingency reversal sessions; N = 7 animals; average of 1.57 sessions per animal (range 1-2). P: pre-reversal (mean = 85.8%); Rev: peri-reversal (mean = 47.8%); R: recovery of performance with the decoder still reversed (mean = 81.2%). Paired t-test between P and Rev: P = 0.0435; paired t-test between Rev and R: P = 0.447; paired t-test between P and R: P = 0.689.*

Similar to our observations of abstract skill learning in M1(Koralek et al., 2012), task performance during degradation dropped significantly (*N* = 8 animals; 2 sessions per animal. Pre-training = 64.5% rewarded, contingency degradation = 43.1%; *P* = 0.0038) (Fig. 2.12, left), but returned to pre-manipulation levels once the reward contingency was reinstated (reinstatement = 61.4% rewarded; *P* = 0.289). These results suggest that modulation of V1 activity was intentional and goal-directed. To test whether the association between neural activity and rewarded cursor movement could be flexibly adapted to a new task contingency, we reversed the transform after animals had achieved saturating performance. This manipulation caused the cursor to move in the opposite direction for a given spike rate modulation that what animals had initially learned. Although this manipulation caused an initial decrease in performance, animals were able to learn the reversed behavior when given sufficient training time (Fig. 2.12, right), showing that animals could dynamically adapt to changes in the relationship between neural patterns and reward.

2.3.4 Changes in visual context affect performance of a V1-driven task

Neurons in the primary visual cortex are known to respond to visual stimuli. The observation that animals can learn to successfully modulate V1 neurons in total darkness indicates that this activity is at least partially independent of visual input. One possibility is that over the course of learning, E1 and E2 units become decoupled from bottom-up sources of visual input; for example, visually-driven activity from the lateral geniculate nucleus of the thalamus. If this were true, then we can expect trained animals to be able to perform the task under any light condition. To test this, we trained animals both in light and total dark conditions. Interestingly, no significant difference was observed in performance at the end of a training session (last trials) for sessions performed entirely in the dark compared to in the light (train light = 74.0%, train dark = 63.5%, *P* = 0.205, Fig. 2.13a,b), suggesting that learning can occur both in the presence

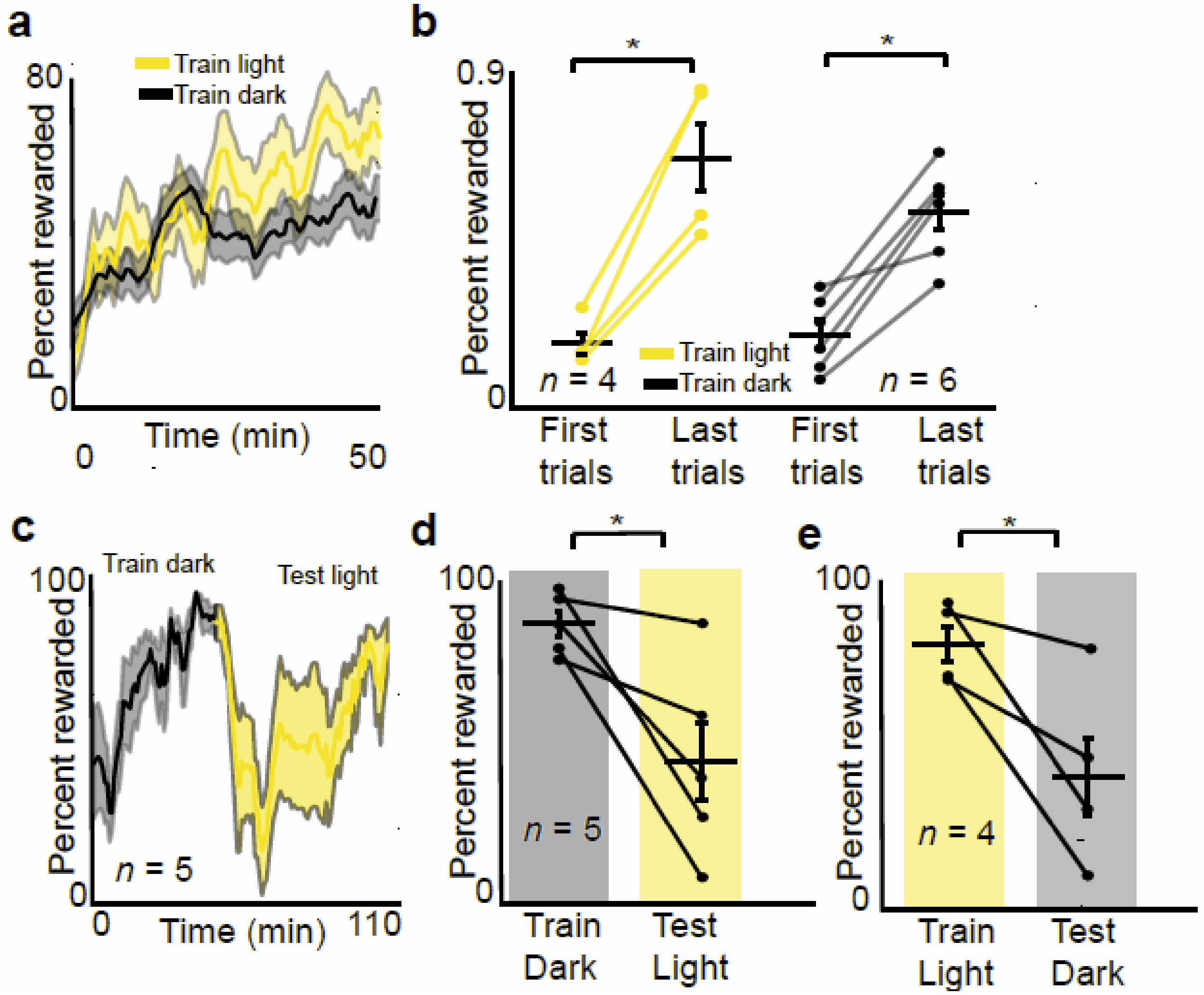


Figure 2.13: ***a****, Timecourse showing the mean percentage of rewarded trials within all sessions under lighted conditions (yellow) compared to dark conditions (black). Shaded areas represent S.E.M.* ***b****, Quantification of data in a, using the first trials and last trials in a session. Train light: N = 4 animals; average of 7.5 sessions per animal (range = 5 to 10 sessions). Paired t-test between first trials and last trials: P = 0.0113; mean first trials = 37.8%; mean last trials = 74.0%. Train dark: N = 6 animals; 8 sessions per animal. Paired t-test between first trials and last trials: P = 0.00167; mean early = 0.396; mean late = 0.635. Unpaired t-test between last trials for light sessions and last trials for dark sessions: P = 0.205 (* indicates P < .05; black crosses show mean and S.E.M.).* ***c****, Timecourse showing the mean percentage of rewarded trials when animals learned a decoder under dark conditions, and were switched to a lighted condition mid-session ("train dark, test light").* ***d****: Mean percentage of rewarded trials when animals were trained in dark, and then tested in the light (same data as c). N = 5 animals, mean of 1.8 sessions per animal (range 1-2). Data taken from last 15 mins of dark training and first 15 mins of light testing. Mean train dark: 87.4%, mean test light 42.6%; P = 0.0309 (paired t-test). Error bars show S.E.M; horizontal lines show mean (* indicates P < 0.05).* ***e****, Mean percentage of rewarded trials when animals were trained in the light and tested in the dark. N = 4 animals, mean of 1.5 sessions per animal (range 1-2). Mean train light: 77.9%. Mean test dark: 35.4%. P = 0.043 (paired t-test).*

and absence of visual stimuli. However, changing the context within a training session, i.e. from dark to light after the animals had learned the task in darkness, or vice-versa, had a significant negative impact on performance (train dark = 87.4%, test light = 42.6%, *P* = 0.0309; train light = 77.9% test dark = 35.4%, *P* = 0.043) (Fig. 2.13c,d,e). Changing the light conditions affected the mean spike rates of all V1 neurons. (Fig. 2.14; mean rate in dark = 4.13 Hz; mean rate in light 6.18 Hz). These results suggest that a modulatory input can learn to generate rewarded patterns of activity in direct units under stable network conditions, but that changing the state of the network (for example, by adding or removing visually-evoked activity) is disruptive and requires compensatory re-learning.

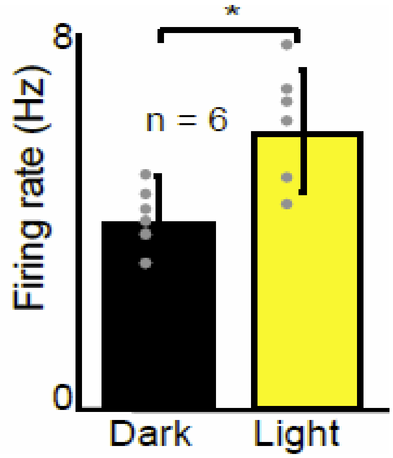


***Figure 2.14**: Mean spike rates for all V1 units when the light conditions were changed mid-session. N = 6 rats. Paired t-test between light and dark conditions: P = 0.314.*

2.3.5 Learning to modulate V1 neurons is associated with changes in neural dynamics

Next, we chose to examine the neural dynamics associated with learning goal-directed modulations in V1. Correlations between E1 units, whose combined positive activity modulations moved the cursor in the rewarded direction, significantly increased over the course of the session (*N* = 10 animals, normalized change 4.42, *P* = 0.042) (Fig. 2.15a,b), suggesting that training resulted in increased coordination between these units. No such change was observed between E2 units or between E1 and E2 units. In 72 out of 102 sessions, the relationship between performance and the E1 unit correlation was positive (mean Pearson correlation coefficient = 0.187). Of these sessions, 55.72% exhibited a significant (P <0.05) correlation. An example session is shown in Fig 2.15c. We also observed that the correlation between E1 units was significantly greater in a 1-second window prior to rewarded target hits, relative to unrewarded targets or timeouts (Fig 2.15d). These data suggest that coordination between E1 units was important for success.

Interestingly, in the late learning phase, we observed an increase in LFP power in V1 prior to rewarded target hits (Fig. 2.15e). Similar changes in ongoing oscillatory activity have previously been associated with top-down processing in visual cortices (Engel et al., 2001), which is one potential mechanism by which animals may be performing the task. To further explore this possibility, we then calculated the coherence between spikes in direct (combined E1 and E2) units and local field potentials (LFP) in V1, time locked to rewarded targets. Previous reports have found that attention alters alpha-band (approx. 5-15 Hz) coherence in the deep layers of visual cortex (Buffalo et al., 2011). We found that the alpha-band spike-field coherence (SFC) of direct units increased from

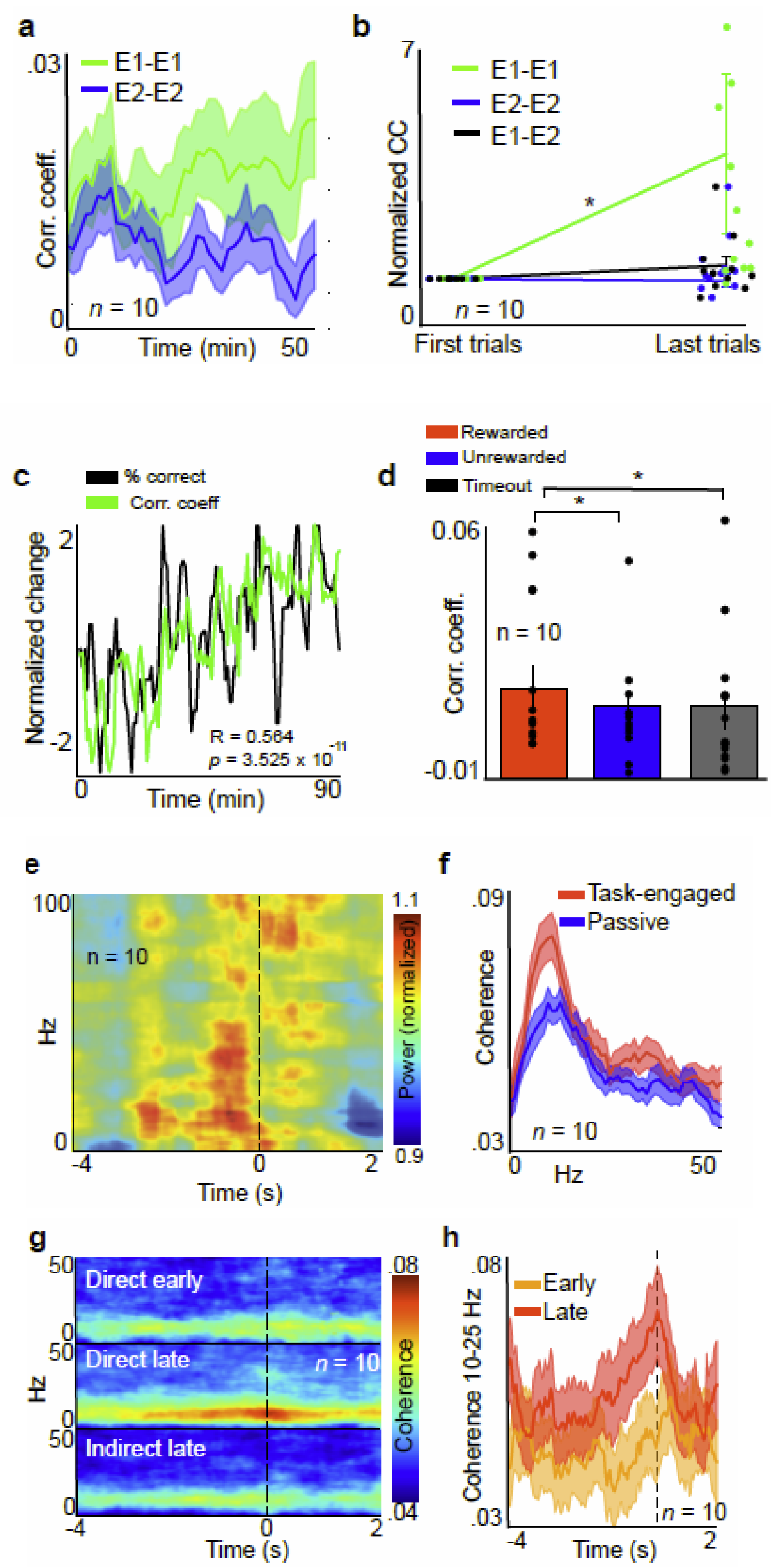


Figure 2.15: ***a***, *Mean pairwise correlations between units within Ensemble 1 (green) or within Ensemble 2 (blue) during training. Correlations were calculated using 1 ms bins.* ***b***, *Change in normalized correlation coefficients (cc) from the first trials to the last trials within sessions. CC calculated between units within Ensemble 1 (mean change = 4.418; P = 0.0417; paired t-test), within Ensemble 2 (mean change = 0.942; P = 0.775; paired t-test)), or between units in ensembles 1 and 2 (mean change = 1.361; P = 0.491; paired t-test). N = 10 rats; mean of 4.4 sessions per rat (range 3-6).* ***c:*** *Example data from one session showing the relationship between changes in correlation of Ensemble 1 units (green) and task performance (percent of correct trials, black).* ***d***: *Mean correlation between Ensemble 1 units during a 2 second window prior to target hits or timeouts. Mean cc prior to rewarded targets = 0.0285; mean cc prior to unrewarded targets = 0.0241 (P = 0.00167; paired t-test). Mean cc prior to timeouts: 0.0240. Comparison between rewarded targets and timeouts: P = 0.0362; comparison between unrewarded targets and timeouts: P = 0.549* ***e***: *Mean spectrogram of V1 LFP time-locked to rewarded target hits.* ***f***, *Spike-field coherence between direct units and V1 LFP for late learning periods during task performance (red) compared to non-engaged passive behavior (blue). Shaded area represents S.E.M.* ***g***, *Spike-field coherograms showing the evolution of coherence over time for LFPs in V1 and spikes from direct (combined Ensemble 1 and Ensemble 2) units in during early training periods (days 1-3; top plot), late training periods (last 3 days, middle plot), and indirect units (no direct relationship to cursor control) for late periods (bottom plot) time-locked to rewarded targets.* ***h***: *Coherence between direct units and V1-LFP in the 10-25 Hz band for early training compared to late training. Solid lines show the mean for 10 animals and shaded areas represent S.E.M.*

early to late phases of learning (Fig. 2.15g,h). This effect was stronger for E1 than for E2 units (Fig 2.16). This increase was only observed during task performance but not when animals were engaged in passive behavior (Fig. 2.15f). Indirect neurons did not show this effect (Fig. 2.15g), suggesting that these learning-related dynamics were specific to units directly involved in cursor control. However, a relatively constant fraction of indirect neurons in each training session did show predictive power for target choice (Figure 2.20e,f), suggesting that there is a broader network of neurons in V1 that may have supported learning in the direct units.

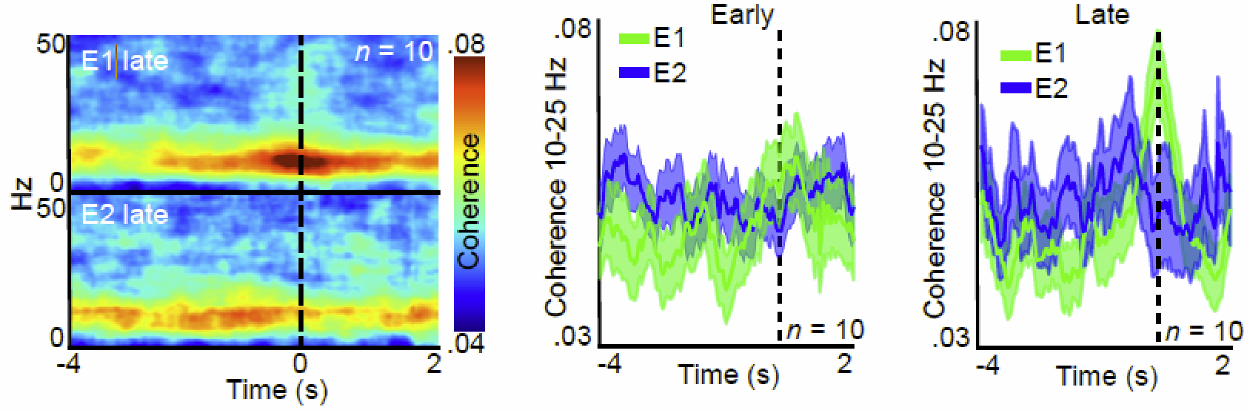


***Figure 2.16**: **Left:** comparison between spike-field coherence between E1 or E2 units and V1 LFP in late training. **Middle**: Spike-field coherence between E1 or E2 units time-locked to rewarded targets and isolated between 10 and 25 Hz for early training periods. **Right**: Coherence as in h for late training periods.*

2.3.6 Striatal activity becomes task-relevant with learning

To address whether the dorso-medial striatum (DMS) plays a role in V1-based reinforcement learning, we next examined whether activity in an area of DMS that receives input from V1 changed with learning. Electrical stimulation of V1 produced a reliable spike response recorded in DMS with a delay of approximately 6 ms, suggesting a direct projection as shown in previous reports (Allen Institute for Brain Science, 2015; Faull et al., 1986; Hintiryan et al., 2016; Khibnik et al., 2014; McGeorge and Faull, 1989) (Fig 2.17a,b). Conversely, stimulation of DMS while recording in V1 did not produce an observable response (not shown). In late learning, DMS units exhibited a strong modulation time-locked to the rewarded target that was not present in the early phase (Fig. 2.17c,d). This was accompanied by increases in beta and gamma LFP power over the same time interval (Fig 2.17f). A linear regression analysis revealed that over the course of training days, population activity of recorded DMS units became increasingly predictive of direct unit activity in V1 (Fig. 2.17e, Fig 2.20e), suggesting that extended training of V1 activity increasingly recruited the striatum. In accordance with this observation, field-field coherence between V1 LFP and DMS LFP was also increased in late learning around the time of target achievement (Fig. 2.17g,h), suggesting a possible role for the striatum in learning to produce rewarded modulations in V1.

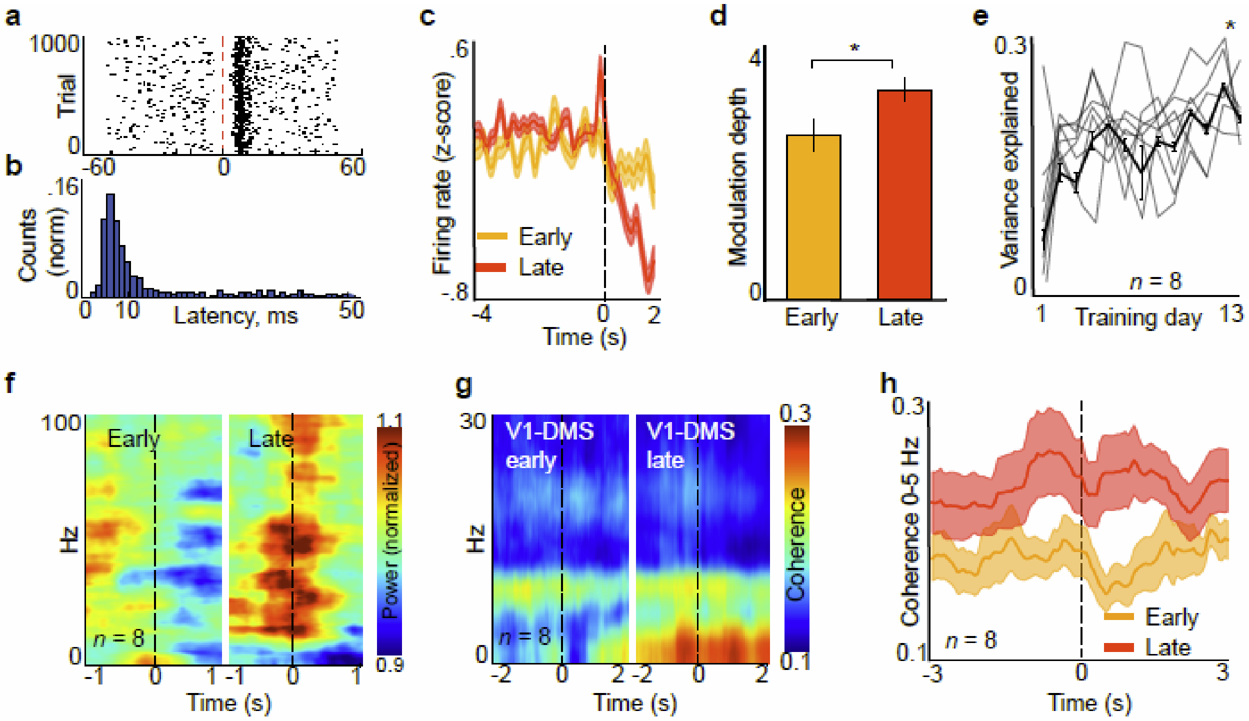


***Figure 2.17**: **a:** Raster plot of an example DMS unit time-locked to V1 ICMS. **b:** histogram of evoked spike counts from unit shown in c. **c**, Mean z-scored firing rate of units recorded in the dorsal medial striatum (DMS) time-locked to rewarded targets for early compared to late learning. **d**: Modulation depth of DMS units in a 1-second window surrounding rewarded target hits. Mean for early periods = 2.872; mean for late periods = 3.691; P = 0.00015. **e**: Proportion of variance of direct unit spikes in V1 explained by DMS unit spikes in a 500 ms window prior to rewarded target hits, computed using a linear regression analysis on individual training days. Solid black lines show mean and S.E.M; grey lines show data for individual animals. Mean of first two days =0.0788; mean of last 2 days = 0.276; P = 0.0196 (paired t-test). **f**: Mean spectrogram of DMS LFP time-locked to rewarded target hits for all animals during the early period (sessions 1-3), left), compared to the late period (last 3 sessions, right). **g**: Field-field coherograms between V1 LFP and DMS LFP for early (left) compared to late training (right) sessions time-locked to rewarded targets. **h**: field-field coherence between V1 and DMS in the 0-5 Hz band in for early and late training, time locked to rewarded targets. Solid lines show mean for all animals; shaded areas represent S.E.M.*

2.3.7 Striatal activity is critical for learning to modulate V1 activity

Next, we asked whether DMS activity was necessary for learning to volitionally modulate V1 neural activity. Mice were injected bilaterally with the red-shifted inhibitory opsin Jaws (rAAV8/CamKII-Jaws-KGC-GFP-ER2) (Chuong et al., 2014) into the area of DMS which receives input from V1, and implanted chronically with optical fibers targeting DMS, and with recording electrodes in L5 of V1 (Fig. 2.18a). Red light stimulation through the optical fiber decreased activity in optogentically-identified DMS units (Fig. 2.4), but had no effect on spike rates in V1 (not shown). Mice were trained on the same task as rats as described above and in Fig. 2.5. During the first 7 days of training, Jaws-injected mice and GFP-injected controls randomly received red light

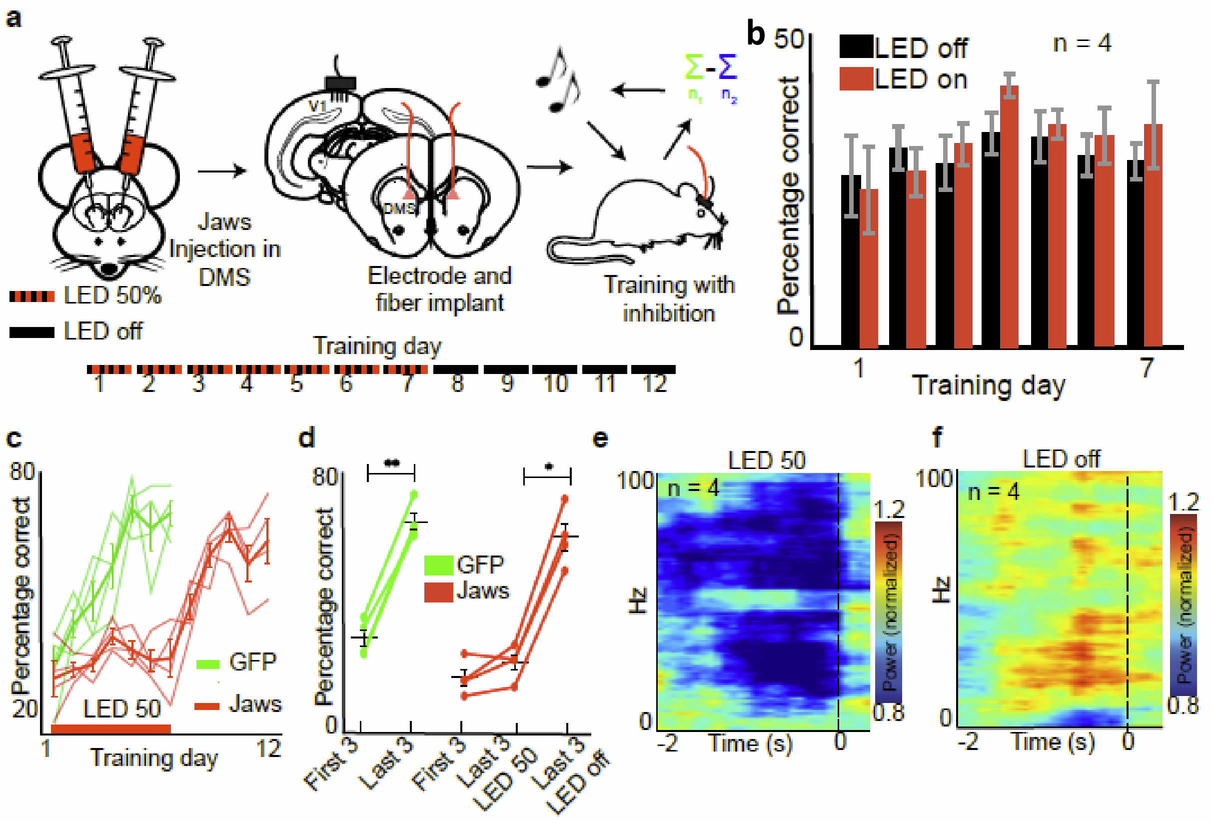


Figure 2.18: ***a****, Schematic showing the replication of rat experiments in Jaws-injected mice. Left: mice were injected bilaterally with rAAV8/CamKII-Jaws-KGC-GFP-ER2. Middle: After Jaws expression stabilized, mice were implanted unilaterally with an electrode array in V1 L5 and bilaterally in DMS. Right: animals were trained on the same task as rats (see Figure 1a) with the addition of optogenetic inhibition. Bottom: timecourse of experiments. LED 50% indicates that Jaws was activated via red LED light in 50% of trials.* ***b****, Success rates for LED-on VS LED-off trials in Jaws-injected animals across training sessions.* ***c****, Percentage of rewarded trials for Jaws compared to control GFP mice over the course of several days of training. Red bar indicates sessions where the LED was active on 50% of trials for both groups (LED 50).* ***d****, Quantification of performance across days for Jaws-injected animals (red; N = 4) and GFP-injected animals (green, N = 4) Each time period is a mean over 3 sessions. Mean GFP, first 3 sessions = 39.7%, Mean GFP, last 3 sessions = 69.7%; P = 0.00072 (paired t-test). Mean Jaws, first 3 sessions = 29.3%, mean Jaws, last 3 sessions LED 50 = 33.3%; P = 0.183. Mean jaws, last 3 sessions LED off = 62.1%. Paired t-test between Jaws last 3 sessions LED 50 and Jaws LED off: P = 0.00131.* ***e and f****: Mean spectrogram of V1 LFP in Jaws animals time-locked to rewarded targets for LED-50 sessions (e) and LED-off sessions (f). LED-50 sessions include both trials with and without Jaws activation.*

stimulation on 50% of trials for the whole duration of the trial (see Methods; LED 50, Fig. 2.18a, bottom). Control GFP animals were able to learn the task and improved with several days of training (*N* = 4 animals, mean of first 3 days = 39.7%, mean of last 3 days = 69.7%, *P* = 0.00072) while Jaws animals did not learn (*N* = 4 animals, mean first 3 days LED 50 = 29.3%, mean of last 3 days LED 50 = 33.3%, *P* = 0.183) (Fig. 2.18c,d). This appeared to be a deficit in acquisition of the task, and not just poor performance in LED on trials, as no difference in success rate was observed between performance during LED on trials compared to LED off trials in any of these sessions (Fig. 2.18b). However, Jaws animals were able to learn the task and improve over the course of 5 subsequent training days if no LED stimulation was provided and activity DMS was unimpaired (mean performance after LED off training = 62.1%, *P* = 0.00131), (Fig. 2.18c,d). Interestingly, V1 LFP power in the gamma band (25-60Hz) during sessions with Jaws inhibition was markedly reduced around the time of target hits, while gamma power during LED-off sessions in which animals successfully learned the task was similar to that observed in rats after learning (Figure 2.18e,f; 2.19a). Together, these results suggest that inhibition of DMS activity prevents animals from learning to generate arbitrary patterns of V1 activity in order to obtain a desirable outcome.

2.3.8 Optogenetic inhibition of DMS does not impair learned performance

In our task, the parameters used to translate neural activity to auditory tones were re-calculated each day such that at the start of every session animals were required to undergo some *de novo* learning in order to adapt to the new calibration. Despite this, animals were able to retain some memory of training from previous days to perform better over the course of each session as training progressed (Fig. 2.10a; comparison between early and late phases). Interestingly, we observed that returning Jaws animals to an LED on 50% condition after several successful days of LED off training impaired their ability to learn during that session, even though they performed well above chance during the previous session and during following session when the LED was not active (*N* = 4 animals, performance with LED off, pre = 63.3%, LED 50 = 38.9%, LED off, post = 59.6%, Fig. 2.19b). This could suggest that striatal activity is required for animals to learn the initial parameters set at the beginning of the session; however, it could also

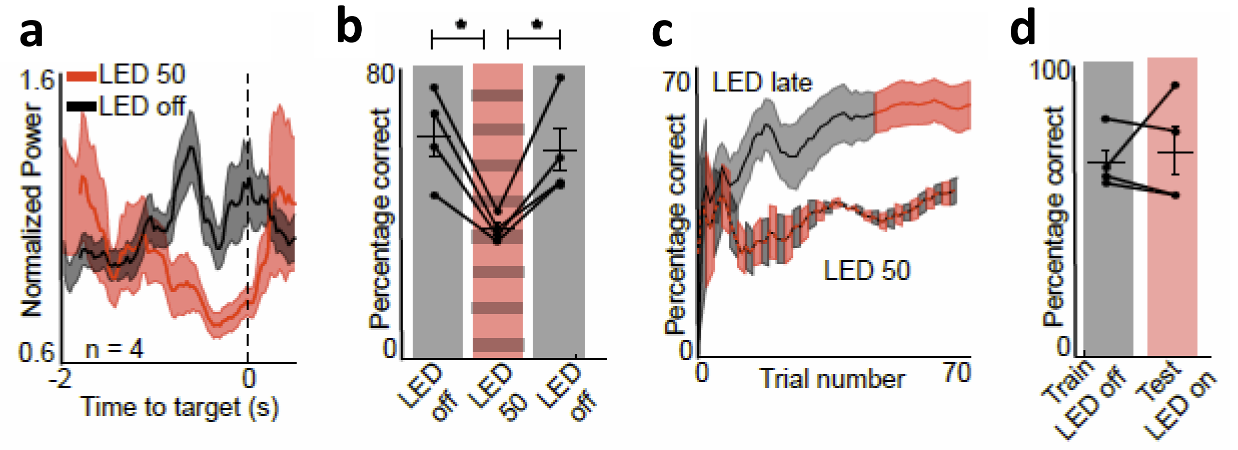


Figure 2.19: ***a***: *Mean gamma power (25-60 Hz, solid lines) in V1 time-locked to rewarded targets during LED-50 vs LED-off sessions. Shaded areas show S.E.M.* ***b***, *Mean performance of trained Jaws-injected animals after several days of training. Data is plotted in the order that the training sessions occurred. Black bars indicate sessions without LED activation. Striped red bars indicate sessions where the red LED was active on 50% of trials. N = 4 animals. Mean LED off, first session = 63.3%; mean LED 50 % (second session) = 38.9%; mean LED off, last session = 59.6%. Paired t-test between LED off, first session and LED 50%: P = 0.0122. Paired t-test between LED 50% and LED off, second last session: P = 0.0259. Paired t-test between LED off, first and last sessions: P = 0.502.* ***c***, *Mean performance within a session for trained Jaws animals with late-session LED only activation compared to full-session LED 50% sessions.* ***d,*** *Quantification of performance when animals were trained with LED off and tested with LED on in the same session; data same as in f. N = 4 animals; mean of train LED off = 59.6%; mean of test LED on = 64.4%; P = 0.657 (paired t-test).*

indicate that striatal inhibition was interfering with task performance rather than acquisition. To disambiguate these two possibilities, we allowed animals to achieve saturating performance on the task within a single session (Train LED off, Fig. 2.19d), and then continued to train animals with the LED turned on in every trial for the remainder of the session (Test LED on, Fig. 2.19d). Interestingly, inhibition of the striatum after within-session learning had taken place did not impair animals' ability to perform the task, and they continued to perform well above chance level (Fig. 2.19c,d), and in an indistinguishable manner from the LED off trials (training with LED off = 59.6%, testing with LED on = 64.4%; $P = 0.657$). These data suggest that striatal inhibition prevents animals from learning to modulate neural patterns in V1 through instrumental conditioning, but not from executing these patterns after learning has taken place.

2.4 Discussion

Here we demonstrate that animals can learn to modulate spike activity in the primary visual cortex in a goal-directed manner using an abstract virtual task. These data demonstrate that feedback-based reinforcement learning can modulate activity on the scale of a few neurons, even in a primary sensory region which is strongly driven by external sensory input. Because we observed that successful performance occurred in both the presence and absence of light, it would appear that the learned modulation of V1 units in our task is an internally-driven process. Taken along with a body of previous work describing brain-machine interface learning in other diverse cortical areas (Cerf et al., 2010; Clancy et al., 2014; Fetz, 2007; Prsa et al., 2017; Schafer and Moore, 2011; Shibata et al., 2011), these results suggest that this type of instrumental learning ability may be a common feature that tunes the activity of cortical circuits more generally.

In the realm of motor control, the cortico-basal ganglia circuit has been hypothesized to perform a selection function in which competing cortical motor programs are either maintained or released from inhibitory control (Costa, 2011; Redgrave et al., 2011). A similar function has also been postulated to operate in the realm of abstract cognition, by which various cognitive patterns may be selected that are appropriate for the current behavioral context, and have previously led to positive outcomes (Graybiel, 1997). These models propose an inhibitory feed-forward projection from basal ganglia output nuclei (globus pallidus internal (GPi) and substantia nigra pars reticulata (SNr) that can activate cortical programs when inhibition is transiently released from the thalamus. Interestingly, although basal ganglia outputs are known to target frontal cortical areas and even higher-order visual areas like area TE in the primate (Middleton and Strick, 1994, 1996), we are not aware of any direct projections from the basal ganglia that target V1-projecting thalamic nuclei.

Despite this, we observed that activity in the striatum was necessary for instrumental learning of neural patterns in the primary visual cortex. This result may be due to an induced learning deficit in a cortical region other than V1 whose input modulates

Ensemble 1 and Ensemble 2 activity in the absence of visual stimulation. Frontal cortical areas, such as the cingulate cortex (Cg) in the rodent, are known to powerfully and directly influence processing in V1 to select and amplify representations of behaviorally relevant stimuli (Zhang et al., 2014). Thus, disrupting cortico-basal ganglia circuit function might impair learning of a top-down modulatory signal that generates rewarded activity in V1. Perhaps analogously, Parkinson's patients with abnormal basal ganglia function have been shown to be impaired in voluntary and sustained control of visual attention in the absence of eye movements (Wright et al., 1990; Yamaguchi and Kobayashi, 1998). Alternatively, the learning process we observed could have recruited structures upstream of V1, such as the superior colliculus, that also form connections with the basal ganglia (McHaffie et al., 2005). Our results do not rule out these or other possibilities; rather, they simply demonstrate that the basal ganglia are necessary to learn to modulate activity in V1, and that the input for this circuit is the striatum.

In our experiments, we observed that animals performed in a goal-directed manner: performance remained sensitive to changing task contingencies, even after many days of training (Figure 2.12). The projection of the primary visual cortex to the striatum lies along the most medial-dorsal aspect, adjacent to the ventricle wall (Khibnik et al., 2014). This lies well within the dorsomedial division of the striatum, which is known to be necessary for and to facilitate goal-directed behaviors (Yin et al., 2005, 2009), as opposed to the dorsolateral division which is required for habitual action (Redgrave et al., 2010; Wickens et al., 2007; Yin et al., 2006). One possibility is that segregation of V1 projections in the dorsomedial division of the striatum favors goal-directed learning and behavior in V1. However, it is also possible that the daily recalibration of task parameters or simply insufficient training time prevented behavior from becoming habitual. Previous work utilizing a similar task design but controlled by neurons in M1 also observed that animals behaved in a goal-directed manner (Clancy et al., 2014; Koralek et al., 2012).

From our analyses, we observed that learning-related changes in neural dynamics, such as changes in correlations and spike-field coherence (Figure 2.15) were largely restricted to the direct population consisting of units from Ensemble 1 and Ensemble 2. Absolute modulation depth of indirect (non- E1 or E2) neurons in V1 remained low relative to direct units (Figure 2.20a,b), suggesting that the learning we observed operated primarily on the small scale of a few neurons. Furthermore, the modulation depth of task-irrelevant indirect neurons declined over the course of training (Figure 2.20c), echoing similar results observed across mice and monkeys using calcium imaging and electrophysiology techniques (Clancy et al., 2014; Ganguly et al., 2011; Prsa et al., 2017). It has been reported that the activity of single cells in sensory cortex is sufficient to drive a percept (Houweling and Brecht, 2008), suggesting that cortical circuits may be optimized to operate on these microscales. However, a closer analysis of indirect unit activity showed that many single units as well as the full population of indirect cells contained activity that was predictive of target choice (Figure 2.20e,f).

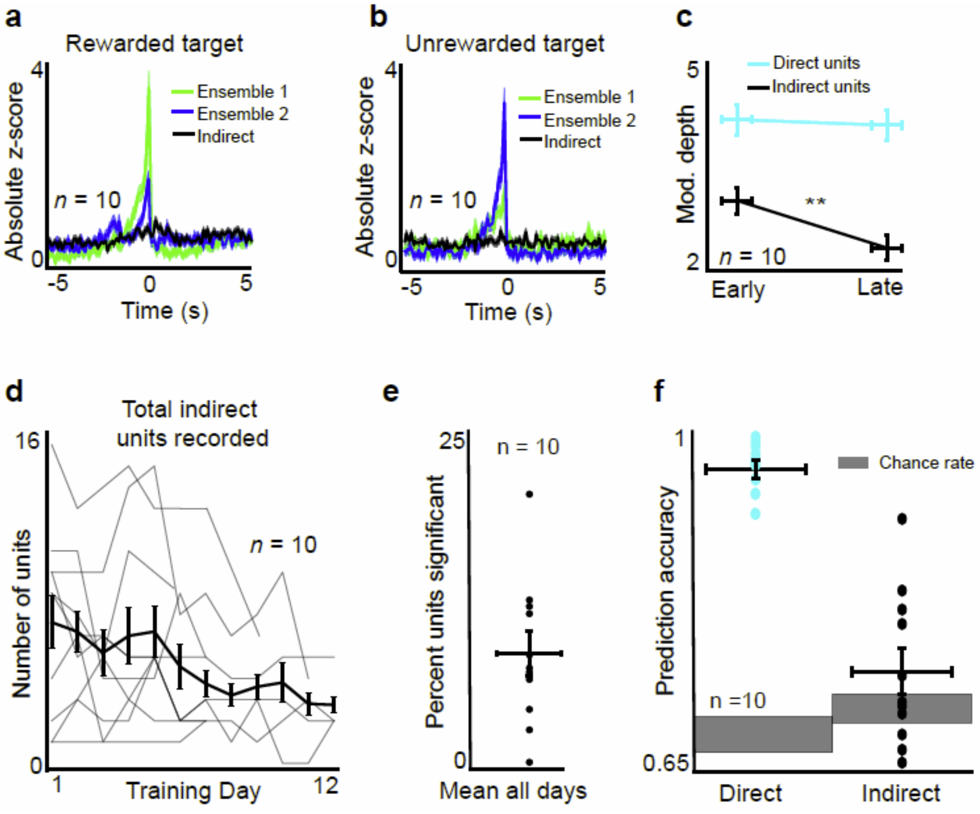


Figure 2.20:* a and b:** *absolute z-scored firing rates for E1, E2 and indirect units, time-locked to rewarded (a) and unrewarded (b) target hits.* ***c*: Mean modulation depth of direct (pooled E1 and E2) units and indirect units in early (sessions 1-3) and late (sessions 8-10) training. Data taken during a 3-second window prior to rewarded target hits. Mean direct early = 3.988, mean direct late = 3.942, P = 0.743 (unpaired t-test). Mean indirect early = 3.135, mean indirect late = 2.718, P = 0.0218 (unpaired t-test).* ***d****: total number of indirect units recorded across training sessions. Black lines show mean and S.E.M; grey lines show values for individual animals.* ***e****: Mean percent of single indirect units with significant predictability for target choice over all training days. Significance determined by permutation test with P < 0.05. Mean = 10.782%. Black lines show mean and S.E.M. for all animals; dots show mean for each animal individually.* ***f****: Prediction of target choice using population activity. Direct (E1 and E2) units or indirect units were used to predict target choice using a logistic regression analysis. Spike data was taken in a 500 ms window prior to target hits. Grey bars show chance accuracy rates calculated by permutation test. Mean direct =0.964 ; mean indirect = 0.729*

These results suggest a subtle role for this population in the learning and execution of the task; however, as we only capture a small slice of the neural population using our technique, it is difficult to conclusively say much about the relative contributions of this population as a whole. Future projects will no doubt involve closer study of how nearby neural populations support learning in a particular subset of cells, perhaps by using new approaches (such as calcium imaging) that allow for stable recording of large neural populations.

An important goal of BMI research is often to decode movement parameters with high accuracy in order to translate a subjects' existing motor control repertoire into the movement of a complex artificial effector. In these cases, using high channel count recordings in motor cortices is an effective strategy due to the rich encoding of movement parameters in areas such as M1 and PMd; however, BMI tasks have successfully been implemented in other cortical regions as well (Carmena et al., 2003; Cerf et al., 2010; Clancy et al., 2014; Fetz, 2007; Hwang et al., 2013; Musallam et al., 2004; Schafer and Moore, 2011; Shibata et al., 2011). In this study, our goal was not to optimize performance or control of an effector, but rather develop a task that would enable us to study learning. The learning rates we observed for a 1-dimensional BMI task using low numbers (~1-10) of neurons in rodents were consistent with what we and others have previously reported (Arduin et al., 2013; Clancy et al., 2014; Koralek et al., 2012; Prsa et al., 2017). It is important to note our task is different in several respects from BMI paradigms designed with the goal of maximizing performance and control in a therapeutic context, and we believe that comparing learning rates or performance metrics observed in this study with those reported across the BMI field as a whole would be misleading. Instead, our goal was to use a BMI paradigm as a method of operantly conditioning neural activity in V1 directly in order to study the learning process and the neural dynamics associated with it. This enabled us to define the final output layer of neurons directly responsible for controlling a virtual action as well as their relationship to task performance, and allowed us to observe their activity relative to each other, other V1 neurons, and activity in the dorsomedial striatum.

Although neurons in the primary visual cortex are thought to represent low-level visual features early in the visual processing stream, we observed that V1 neurons could learn to produce rewarded activity patterns that were independent of visual stimulation when spike activity was used as a control signal for a closed-loop brain-machine interface task. While here we focus on learning in the primary visual cortex, the dynamics of striatal activation, cortico-striatal dynamics over the course of learning, and the necessity of the striatum in the learning process is similar to what has been observed in a variety of tasks that engage diverse cortical regions (Barnes et al., 2005; Corbit and Janak, 2010; Koralek et al., 2012; Pasupathy and Miller, 2005; Shohamy et al., 2004; Yin et al., 2009). These results suggest that cortico-striatal projections, as part of larger cortico-basal ganglia circuits, play a generalizable role in shaping cortical activity based on ongoing experience and behavioral outcomes.

Chapter 3:

Representation of task-relevant parameters by population activity in the dorsomedial striatum

Selecting the appropriate behavior for a given time and place can be a complicated decision. Evidence suggests that cortico-basal ganglia circuits shape behavior by learning the value of different actions, and selecting actions that are most likely to lead to positive outcomes. In **Chapter 2**, we demonstrated that the dorsomedial striatum (DMS) was necessary for learning to modulate neural activity in the primary visual cortex (V1). These data suggest that patterns of neural activity can be learned instrumentally, similar to overt actions, and that cortico-basal ganglia circuits may be necessary to assigning value to these patterns. However, the value of a given action may not be static and can fluctuate according to the current state of an animal's environment, creating uncertainty about which action to take. How cortico-basal ganglia circuits facilitate action selection under different environmental conditions remains an active area of investigation. We trained rats on a two-alternative probabilistic switching task while recording neural activity from the dorsomedial region of the striatum, the main input nucleus of the basal ganglia. We found that the encoding of task parameters was mixed at the level of single neurons, but latent signals corresponding to individual parameters could be robustly decoded from population activity. Furthermore, we found that the magnitude of animals' uncertainty about the correct action was encoded by neural populations on a trial-by-trial basis, and that uncertainty influenced the timing and strength of signals encoding action choice.

3.1 Introduction

Instrumental learning is a common mechanism by which animals acquire new behaviors, including natural motor behaviors and brain-machine interface-driven behaviors that condition neural activity directly. In **Chapter 2**, we showed that instrumental learning was a common mechanism by which animals learned to control neural activity to obtain rewarding outcomes. In many models of instrumental learning, there is an early period of exploration, during which an agent will sample many possible actions and acquire information about their relative values. This is often followed by a period of exploitation, when the agent repeats the action with the highest estimated value to accumulate rewarding outcomes. In real-life scenarios, intelligent prediction about the value of possible actions is essential. Cues about state of an environment can signal that previously estimated action values may no longer be valid, requiring a change of strategy or further exploration. Conversely, failing to exploit valuable options when they are available can result in suboptimal outcomes. A deficit in the ability to respond flexibly to changing conditions is feature of several neuropsychiatric disorders, such as obsessive-compulsive disorder (Gillan et al., 2011) and addiction (Everitt and Robbins, 2005), that can severely impact an individual's quality of life.

In the brain, deliberative decisions about behavioral choices have been attributed to regions of the prefrontal cortex, such as the orbitofrontal and anterior cingulate cortices (Izquierdo et al., 2016; Quilodran et al., 2008; Saez et al., 2015; Wilson et al., 2014). However, accumulating evidence suggests that these prefrontal areas are likely part of larger circuits involved in action selection that include the basal ganglia (Clarke et al., 2008; Daw et al., 2005, 2006; Gremel and Costa, 2013; Kimchi and Laubach, 2009; McClure et al., 2003; Owen, 1997; Tanaka et al., 2016). The striatum, which is the main input nucleus of the basal ganglia, is known to encode action-outcome contingencies (Samejima, 2005; Tricomi et al., 2004) and has a well-documented role in instrumental learning (Dang et al., 2006; Hikosaka et al., 2002; Yin et al., 2005, 2006, 2009). Distinct roles in the acquisition of new behaviors have been attributed to different sub-regions of the striatum (O'Doherty et al., 2004). The ventral compartment is thought to encode the difference between expected and actual reward (Gläscher et al., 2010; O'Doherty et al., 2004; Schultz, 2000), whereas the dorsal region is thought to be involved in action selection (Costa, 2011; Kimchi and Laubach, 2009; Redgrave et al., 2011). A further distinction may be drawn between the medial and lateral compartments of the dorsal striatum. The dorsolateral part (DLS), which is recipient of sensorimotor inputs from the cortex, is thought to be involved during the later phases of skill learning when actions become habitual (Yin et al., 2006, 2009). The dorsomedial part (DMS) receives input primarily from associative areas of the cortex, and is thought to be engaged during early periods of skill learning, or when actions are goal-directed and sensitive to action-outcome contingency (Redgrave et al., 2010; Yin et al., 2005, 2009).

In addition to accumulating experience through trial and error, goal-directed actions can be guided by acquired knowledge or predictions about an environment. For example, an individual with confidence in the current state of the environment may choose to exclusively and repeatedly execute a single action, while the same individual in a more uncertain environment might choose to sample many different options. Behaviors that rely on abstract knowledge about environmental state have been shown to engage the frontal cortex (Saez et al., 2015; Samejima and Doya, 2007; Wilson et al., 2014; Yoshida and Ishii, 2006)(Wolter, 2014), and lesions of frontal cortices are known to cause deficits in these behaviors (Boulougouris et al., 2007; Brigman and Rothblat, 2008; Cordova et al., 2014; McAlonan and Brown, 2003). However, similar deficits have been observed following lesions of the medial striatum (Castañé Anna et al., 2010; Clarke et al., 2008). This evidence, as well as the anatomical convergence of frontal cortical inputs to the dorsomedial striatum, suggests that striatal circuits play an important role in flexible decision making.

Here, we asked how the activity of neurons in the dorsomedial striatum encoded task parameters as rats learned to perform a two-alternative probabilistic switching task. Rats exploited actions when the state of the environment was clear, and explored alternatives when the state was uncertain. We observed that a fraction of single units displayed activity that significantly predicted one or more task parameters, including the degree of uncertainty about the best action to take in a given trial. We then used dimensionality reduction techniques to decompose the activity of neural populations into latent factors that encoded individual parameters, and found factors that robustly encoded action, choice, and uncertainty from trial-to-trial. Furthermore, we found that the magnitude of choice signals was affected by animals' uncertainty levels, suggesting that uncertainty may promote exploration by weakening striatal signals that favor selection of a particular action.

3.2 Methods

3.2.1 Surgery

All rat experiments were performed in compliance with the regulations of the Animal Care and Use Committee at the University of California, Berkeley. A total of three singly-housed, male Long-Evans rats weighing roughly 250 grams were used for the experiments. After a total of 17 days of behavioral training, all rats were chronically implanted bilaterally with microwire arrays the dorsomedial striatum. Each array contained 16 tungsten microelectrodes (35 µm diameter, 250 µm electrode spacing, 8x2 or 4x4 configuration; Innovative Neurophysiology, Durham, NC). Stereotactic coordinates relative to bregma and lambda were used to center the arrays (1.8 mm anterior of bregma, 2 mm lateral of the midline, and 5.5 mm ventral from the cortical surface for DMS). Rodents were anesthetized isoflurance gas as needed. Craniectomies were sealed with cyanoacrylate and rats were allowed to recover for ten days after implantation before behavioral training. Rats were given dexamethasone treatment (0.5 mg/kg) for one week following surgery to minimize tissue damage around the implant (Zhong and Bellamkonda, 2007).

3.2.2 Electrophysiology

Single unit activity and local field potentials were simultaneously recorded with a Multichannel Acquisition Processor (MAP; Plexon Inc., Dallas, TX). A commutator (Plexon Inc., Dallas, TX) was mounted above the behavioral chamber and used to route and detangle headstage cables connected to the animals' implants while they moved freely. Activity was sorted using an online sorting application (Plexon Inc., Dallas, TX) prior to each daily recording session. Only units with a clearly identified waveform and signal-to-noise ratio greater than 2 were used. Sorting templates were further refined using an offline sorting application (Plexon Inc., Dallas, TX). Recording arrays were grounded to a screw in the occipital bone, and both arrays were also referenced locally using the online program Ref2 (Plexon Inc., Dallas, TX) to eliminate effects of volume conduction. For referencing, an electrode on each array was chosen to be subtracted from all other electrodes on that array. This was done independently for arrays in each hemisphere.

3.2.3 Behavioral Training

Rats were trained on the task in a lit behavioral chamber (Lafayette Instrument Company, Lafayette, IL). One wall of the chamber was configured with two levers in the center of the wall, positioned one on top of the other. The bottom lever was 2 cm from the floor of the chamber, and the other was 10 cm above the first. Animals needed to stand on their hind legs in order to press the top lever. The opposite wall of the chamber was configured with a nose poke port connected to a IR beam break detector, as well as a syringe pump that delivered a sucrose water reward. (Fig. 3.1). Food pellets were freely available to the animal during the task. A speaker was mounted at the top of the cage to signal behavioral events: a single beep indicated that a lever had been fully depressed, while two beeps indicated that a new trial had started and the levers were reset. A programmable Raspberri Pi computer (Raspberri Pi Foundation, Cambridge, UK) along with custom-written Python routines was used to control the behavioral task, log behavioral events, and record video using USB webcams. During the first two days of training, animals were familiarized with the behavioral chamber and rewarded for nose pokes into the reward port. In the next three days of training, animals were trained to first press the lower lever exclusively (with a 100% reward rate), followed by the top lever exclusively (again with 100% reward rate). Pressing the correct lever primed the reward port such that when animals poked their nose inside and broke the IR beam, a sucrose water reward was delivered. Subsequent days of training then utilized both levers. In the two-lever task, the first lever press in the trial was considered to be the chosen lever for that trial, meaning that any subsequent lever presses had no effect on the outcome for that trial. At any point in the two-lever task, one lever was associated with an 85% probability of reward, while pressing the other lever did not deliver any reward. At a randomized and un-cued number of trials between 25 and 100, the identities of the rewarded lever and unrewarded lever were reversed. Training sessions lasted approximately 1 hour each day, and animals trained for a total of 30 days on the two-lever task.

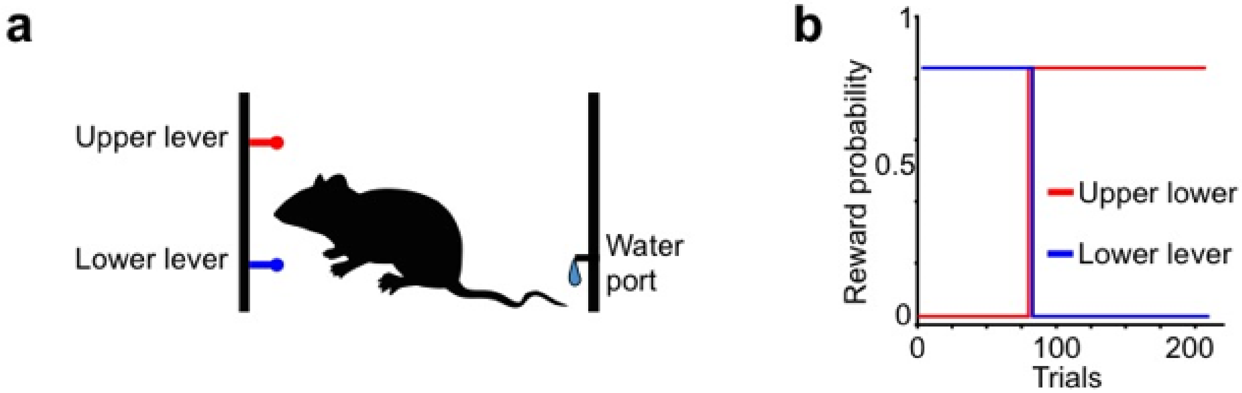


Figure 3.1: ***a:*** *side view of behavioral chamber configuration.* ***b****: example of hidden task structure for a single session. In this example, the identity of the correct lever switches at approximately 90 trials. The correct lever had an 85% chance of delivering a sucrose water reward.*

3.2.4 Data Analysis

Analyses were performed in Python with custom- written routines utilizing publicly available software packages. Unit data were first binned in 1 ms time bins and digitized. Analyses were performed in relation to different behavioral events as indicated in figures. For all analyses, firing rates were first smoothed with a Gaussian window of 100-ms width, and then binned into 50-ms bins. Only two-sided statistical tests were used to determine significance. Any trials in which the interval between lever press and nose poke was greater than 5 seconds was excluded from analyses.

RL model. To compute estimates of action value, we adapted a previously described method (Samejima, 2005) based on a Rescorla-Wagner rule. In the model, a sucrose reward on the $i-$th trial following nose poke was considered to have a value of $r(i) = 1$, while an unrewarded poke had a value of $r(i) = 0$. The action value for an action a is defined as $Qa = E[r(i) \mid a(i) = a]$. After taking an action $a(i)$ and receiving a reward $r(i)$, the on-line estimate $Qa(i)$ is updated by:

$Qa\,(i + 1) = (1 - \alpha\,)Qa\,(i) + \alpha r(i)\; if\; a = a(i)$, or

$Qa\,(i + 1) = Qa\,(i)\; if\; a \neq a(i)$,where α is the learning rate parameter.

The choice of lever a is chosen according to a Boltzmann distribution:

$$P_a(i) = \frac{e^{\beta Q_a(i)}}{\sum_{a\prime \in A} e^{\beta Q_{a\prime}(i)}}$$

where β is the inverse temperature parameter that regulates the randomness of action selection. Because our task has only two possible actions, the probability of choosing one lever is a sigmoid function of the difference between the two estimated action values:

$$P_a(i) = \frac{1}{1 + \, e(^{-\beta\{Q_a(i) - Q_{a\prime}(i)\}})}$$

Hidden Markov Model. We adapted equations from previous reports (Hampton et al., 2006; Schlagenhauf et al., 2014) in order to construct a Bayesian Hidden State Markov Model (HMM) that accounts for the partially observable structure of the probabilistic switching task. In this model, s_t is the abstract hidden state (upper lever rewarded or lower lever rewarded) that subjects have to infer at time $t.y_t$ is the result of a trial at time t; that is the action-outcome pair of the trial y_t= a_t, o_t; where a is the lever choice and o is whether the trial is rewarded or unrewarded. Subjects therefore need to estimate their belief in the state of the task at the next trial, $t+1$. Here, the belief over states at time $t+1$, based on all observations (action-reward pairs) up to time t be $p(s_{t+1}|o^t) = p(s_{t+1}|a^t, r^t)$, where $y^t = \{y_\tau\}_{\tau=1}^t = \{a_\tau, r_\tau\}_{\tau=1}^t$. We write:

$$p(s_{t+1}|y^t) = \int ds_t \; p(s_{t+1}|s_t) \frac{p(y_t|s_t)p(s_t|y^{t-1})}{\int ds'_t \;\; p(y_t|s_t)p(s_t|y^{t-1})}$$

The probability of observing a given action-outcome pair can be estimated by the reward probabilities of different actions in different states. A reward tells the subject with probability c that the choice was correct and the state corresponds to action a; the absence of reward tells us with probability d that the state is the one not corresponding to action a:

$$\begin{aligned} p(a=1 \,\&\, r=+1 \mid s=1) &= p(a=2 \,\&\, r=+1 \mid s=2) = 0.5+0.5c \\ p(a=2 \,\&\, r=-1 \mid s=1) &= p(a=1 \,\&\, r=-1 \mid s=2) = 0.5+0.5d \end{aligned}$$

Next, we consider the probability of transitioning from one state to another. Here, the probability of staying in a state is γ. We can write:

$$p(s_{t+1}|s_t) = \begin{cases} \gamma & if \; s_{t+1} = s_t \\ 1-\gamma & if \; s_{t+1} \neq s_t \end{cases}$$

Finally, once the belief in state $b = p(s_t = 1|y^{t-1})$ is computed based on past observations, this belief is mapped onto action probabilities using a sigmoidal function as in the RL model, but using a fixed steepness:

$$p(a_t = 1|b) = \frac{1}{1+e^{-20(b-0.5)}}$$

Model fitting. Parameters for both RL models and HMM models were fit with the Sequential Monte Carlo (SMC) method (Samejima, 2005), which acts as a particle filter that represents and updates parameter distributions using a large number of samples or "particles." We randomly seeded 100,000 particles around reasonable means for each of the parameters. At each time step, particles whose parameter values provided the best estimates of behavior were assigned the strongest weight and used to update the distributions, along with some added Gaussian jitter. Parameter values evolved at each time step to better fit the subject's true behavior. At the end of the evolution, the converged parameter means were used as the final model values. Goodness-of-fit was computed using log-liklihood estimation as well as accuracy of performance by comparing to subjects' actual action choices.

Linear regression. To analyze single-unit encoding of task parameters, we used smoothed spike data from individual trials that began 800 ms window prior to lever press. and ended 800 ms after nose poke. Linear regression models were computed using activity for each unit in all trials within one session. We used the open-source statsmodels (www.statsmodels.org) package to fit regression models to the data using 3-fold cross validation. 10 regressors were used to fit the model: action choice, outcome, task state (upper or lower rewarded), estimated uncertainty, and the pairwise interactions between each of these variables. Statistical significance was computed using parametric t-testing available in statsmodels. Permutation tests yielded similar significance values.

Logistic regression. To compute the population encoding of decision variables, we used the logit models available from the statsmodels package. We computed the log-likelihood of upper or lower lever pressing at each time bin using bin values for all recorded neurons in a single session. To decode decision variables as they evolved over time, we computed the mean coefficient values over all time bins and then used these as a fixed value in a linear decoder built from the logistic regression model to predict the log-likelihood of upper or lower lever choice at each time point for individual trials. To compare trials with high or low uncertainty, we ordered trials by the estimated uncertainty for each session. Then, we averaged the decision variables for the top 10 percent of trials with the highest uncertainty, and the top 10 percent of trials with the lowest uncertainty.

De-mixed principal components analysis. dPCA models were fit by following the procedures in Kobak et al, 2016. Briefly, we used action choice (upper or lower) and task state (upper rewarded or lower rewarded) to define the dPCA marginalizations. Data matrices were constructed by separating out trial types according to combinations of these marginalizations; for example, one trial type was (state = upper rewarded; action = lower lever). Data from all animals and all sessions were pooled into a single data matrix. Because the data were not recorded sequentially and the number of trials differed between various trial types, a re-balancing procedure was used (Kobak et al., 2016). We fit models using 15 components using code available at https://github.com/machenslab/dPCA. Significance was determined by using each component as a linear decoder to classify the condition it was associated computing the accuracy of decoding on a cross-validation procedure, and comparing this accuracy to performance of the decoder using shuffled labels.

Tensor analysis. Data from each session individually was used to create a third-order tensor of dimensions units x bins x trial number, such that each entry was the binned spike rate of a single unit at a single time point within a single trial. Tensors were then decomposed using canonical polyadic tensor decomposition using the methods described elsewhere and 10 factors (Williams et al.). To compute the correlation between trial factors and uncertainty estimates, we computed the Pearson correlation coefficient between the top three trial factors accounting for the most variance and the uncertainty estimate. To compute significance, we compared the number of sessions per animal where the time-varying uncertainty value showed a significant correlation with at least one of the trial factors, and then compared this to an analysis of correlation where the trial factors and uncertainty estimates were shuffled randomly between trials.

3.3 Results

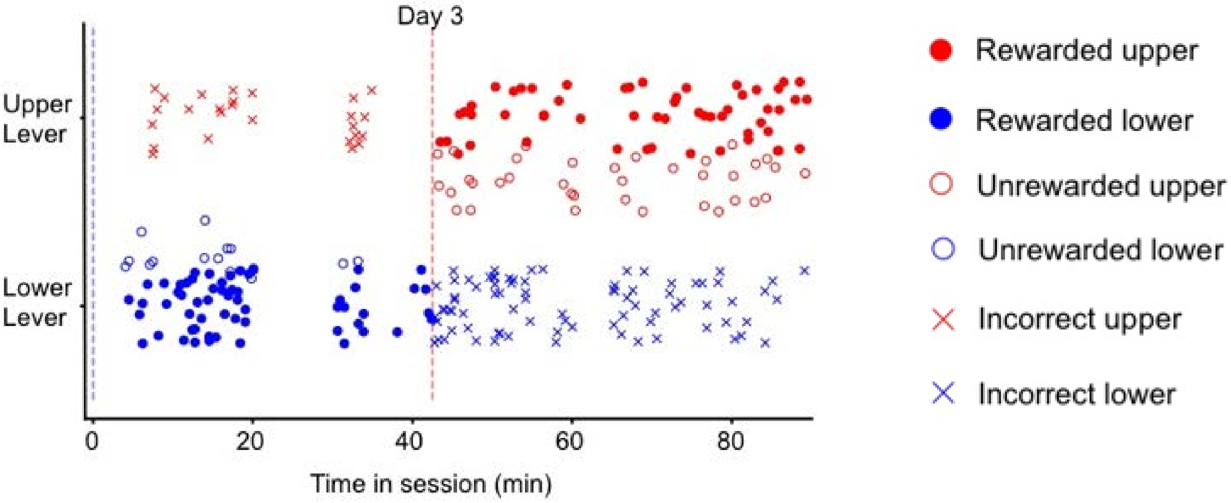


Figure 3.2: *Actions of one animal on the third day of training. Blue and red dashed lines indicate the start of lower = correct and upper = correct task states, respectively.*

3.3.1 A probabilistic switching task To observe possible state-dependent signals in the dorsomedial striatum, we trained rats on a two-alternative probabilistic switching task. Rats were placed in an operant chamber and allowed to move freely. On one side of the chamber were two levers in the center of the wall, one higher than the other. On the opposite side of the chamber was a nose poke port that dispensed a sucrose water solution (Fig. 3.1a). At the beginning of each session, one lever was designated as correct, and the other lever was considered incorrect. Animals were given an auditory cue that signaled the start of a trial, at which point they were free to press either lever. After choosing to press one lever, animals were trained to move to the opposite side of the cage and nose poke into the reward port to terminate the trial. If the correct lever was chosen, the reward port delivered a sucrose water reward with an 85% probability, while the incorrect lever was not associated with any reward. After a randomized number of trials between 25 and 50, the identities of the correct and incorrect levers were reversed without providing any cue to the animal (Fig. 3.1b). Thus, animals were forced to sample both levers to determine which was associated with reward at a given time during a session. Animals trained on this task for a total of 30 days, completing one session per day. After 9 days of training, animals were implanted bilaterally with 16-channel microwire arrays in the dorsomedial striatum (DMS). After implantation, we recorded a mean of 31.7 (range 22-52) well-isolated units across both arrays in each session.

3.3.2 Animals learn to balance exploration and exploitation

Animals' performance, measured as the fraction of correct lever presses over total lever presses, significantly improved across days (Fig. 3.2, 3.3, 3.4a; mean percent correct for animals during the first 4 days = 0.5553, mean for animals in the last 4 days = 0.8467; *P* = 0.000665, T = -38.770). Training was also associated with shorter trial durations (Fig. 3.4b, mean trial duration for first 4 days = 2.36 sec; mean for last 4 days = 1.48 seconds; P = 1.933×10^{-62}, T = 16.9066). As animals learned to perform the task, they adapted their behavior to fit the hidden task structure. We observed that when the identities of the correct and incorrect levers were switched, animals in the early phase of training persisted in pressing the previously rewarded lever. In contrast, animals in later

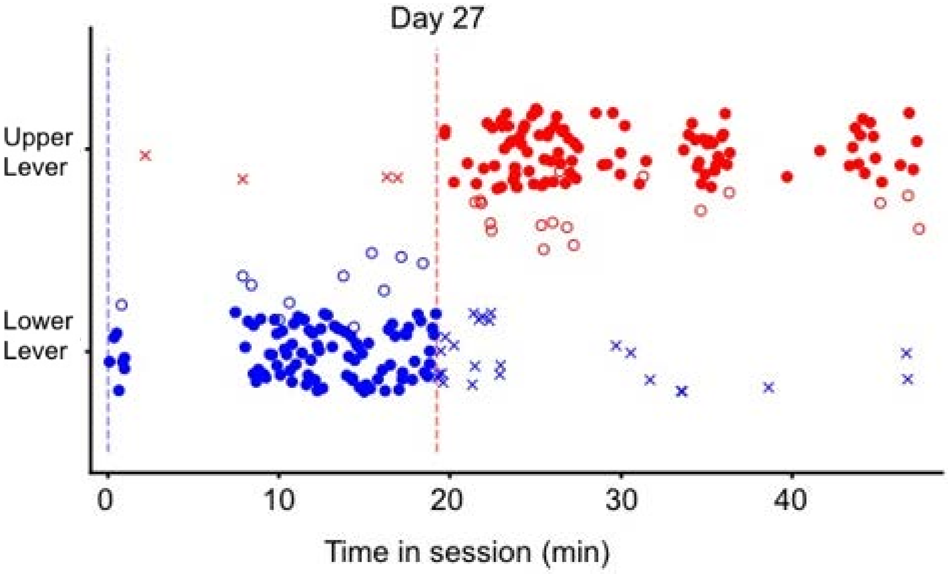


Figure 3.3: *Same as in Figure 3.2, but for Day 27 of training for the same animal.*

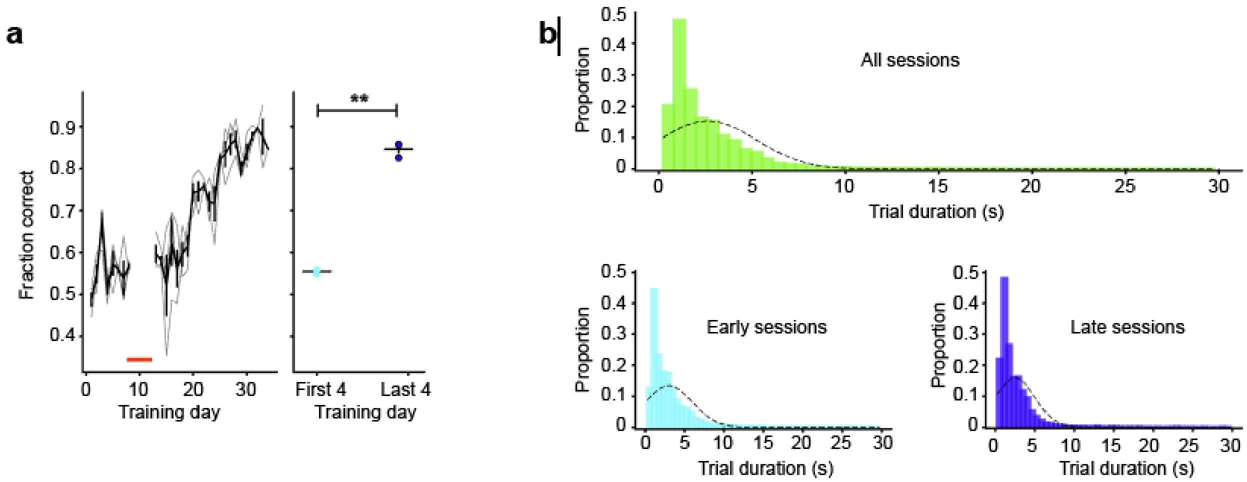


Figure 3.4: ***a**, left panel: fraction of correct trials out of all trials for each day. Black bars show mean and SEM; grey lines show trajectories for each animal individually. Red line indicates days when animals underwent electrode implants in DMS. Right panel: quantification of mean performance on the first 4 and last 4 training days. Black bars show mean and SEM across animals; dots show data for each animal separately. Double asterisk indicates P < 0.01. **b**: Histograms showing the distribution of trial durations for all sessions (top), early sessions (first 4 days, left bottom) and late sessions (last 4 days, bottom right.)*

training sessions quickly recognized the state change and adjusted their behavior (compare Figs. 3.2 and 3.3; Fig. 3.5a). To quantify this effect, we computed the number of trials to reach a criterion performance of 70% correct following a lever switch. In the early period, animals needed a mean of 26.632 trials before reaching criterion, while in the late period, animals needed only 12.319 trials (Figure 3.5b,c; $P = 0.0301$; T = 5.630). In order to obtain more rewards, trained animals may have adopted a strategy that better balanced exploration and exploitation. In this task, receiving a reward means that the chosen lever was correct. However, not receiving a reward does not provide definitive information on the identity of the chosen lever, as both correct and incorrect levers could be unrewarded. We thus asked whether animals' behavioral strategy reflected this relationship. We observed that compared to choice behavior during the first 4 training sessions, animals in the last 4 sessions were significantly more likely to exploit rewarded lever presses. After receiving a reward, animals in the early phase returned to the same lever on 89.9% of trials on average, while animals in the late phase returned on 97.24% of trials (Figure 3.5d,e, P = 0.0322; T = -5.435). Conversely, after unrewarded choices, animals in all training phases exhibited more exploratory behavior, returning to the same lever on only 65.20% of trials and 61.89% of trials for early and late training, respectively (Figure 3.5f,g; $P = 0.562$, T = 0.690). These data suggest that increasingly with training, animals adopted a win-stay, lose-shift strategy that accounted for task structure.

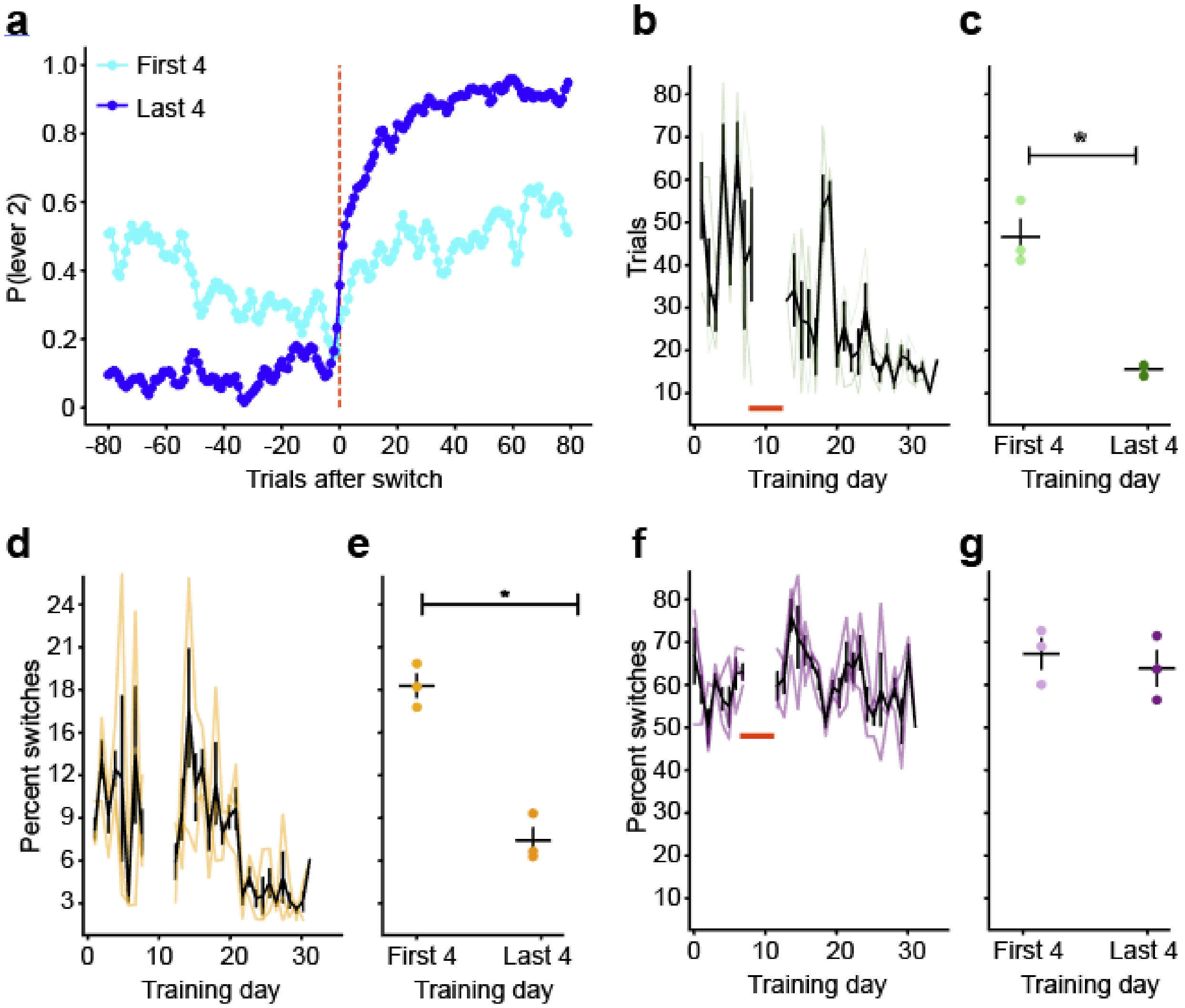


Figure 3.5: ***a****: mean switch curves for all animals during the first 4 and last 4 training sessions. In this plot, lever 1 is the lever designated as correct prior to trial 0 (red line), where lever 2 is correct after trial 0. Lines show the probability of an animal pressing lever 2 as a function of trials.* ***b****: number of trials to reach criterion (70% correct) after a state switch as a function of training day. Red line represents a break in training for array implantation.* ***c:*** *quantification of number of trials to criterion by pooling across the first and last 4 training days. Black lines show mean and SEM for all animals, and dots show data for animals individually.* ***d****: Exploitative behavior, expressed as the mean percentage trials in which animals switched to the second lever after being rewarded for pressing the first, as a function of training day. Black lines show mean and SEM for all animals; orange lines show data for animals individually.* ***e****: Quantification of exploitive behavior in d comparing the first and last 4 training days. Black lines show mean and SEM; dots show data for animals individually. Asterisk indicates P < 0.05.* ***f****: Exploratory behavior, expressed as the mean percentage of trials in which animals switched to a new lever after not being rewarded for pressing the other lever.* ***g****: Quantification of exploratory behavior in f comparing the first and last 4 training days. Black lines show mean and SEM; dots show data for animals individually.*

3.3.3 Animals' behavior reflects the hidden state of the environment

We next chose to model animals' behavior in order to make predictions about behavioral choice. First, we fit a simple reinforcement learning (RL) model (also known as a Q-learning model, Fig. 3.6a) (Samejima, 2005). This model estimates the value of possible actions at each trial, and updates those estimates by computing the difference between expected and received rewards. Importantly, in this model, only the value of the chosen action is updated on each trial. Additionally, we fit a Hidden Markov Model (HMM, Fig. 3.6b). This model estimates the hidden state of the environment using observations about actions and outcomes on each trial. Unlike the Q-learning model, the estimate of the hidden state can be used to update the expected values of both actions on each trial. For early training periods, both the HMM model and the RL model performed similarly (mean RL accuracy = 0.580, mean HMM accuracy = 0.584; *P* = 0.915, T = 0.114). However, in late training periods after animals had learned behavioral strategies that reflected the task structure, the HMM significantly outperformed the RL model (Fig. 3.6c,d; mean RL accuracy = 0.616, mean HMM accuracy = 0.800; *P* = 0.0207, T = 3.706). An examination of the behavior of the fitted models provides a likely explanation: following a lever identity switch, trained animals quickly begin exploiting the newly assigned correct lever after experiencing a rewarded trial. The HMM behaves similarly; however, the RL model relies on experience with an action in order to update

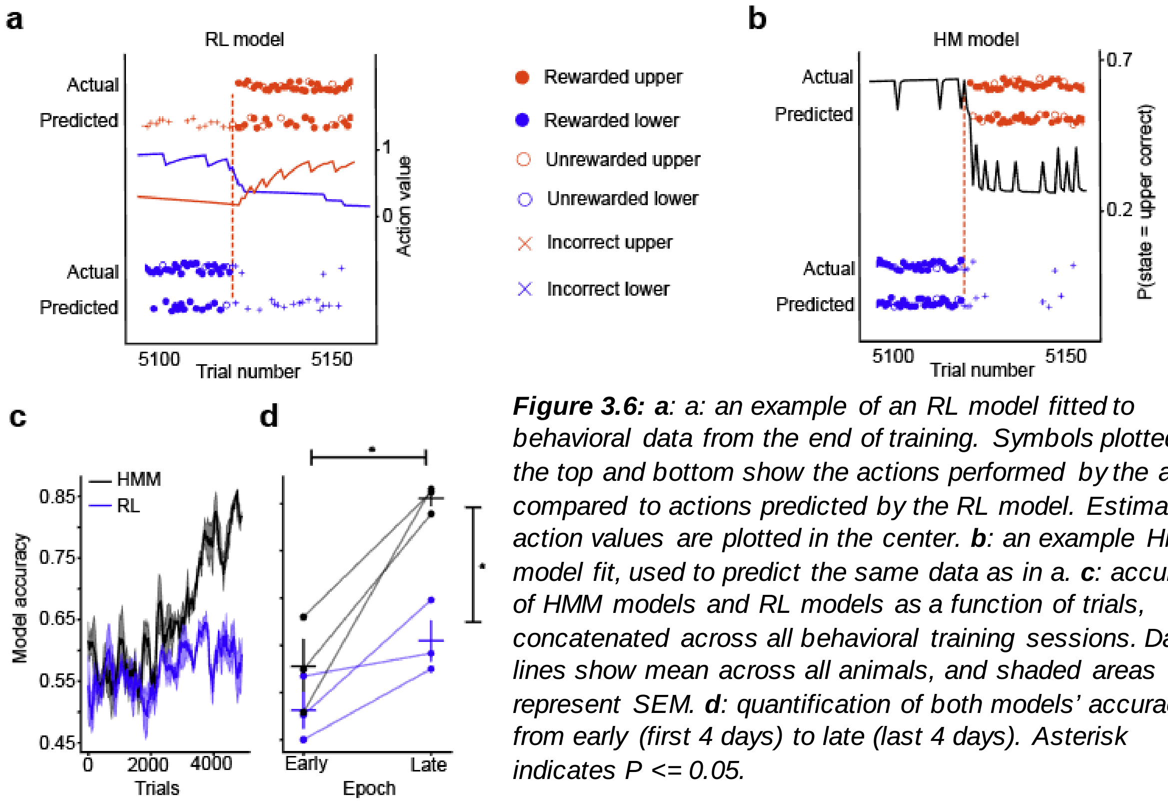


Figure 3.6: ***a****: a: an example of an RL model fitted to behavioral data from the end of training. Symbols plotted on the top and bottom show the actions performed by the animal compared to actions predicted by the RL model. Estimated action values are plotted in the center.* ***b****: an example HMM model fit, used to predict the same data as in a.* ***c****: accuracy of HMM models and RL models as a function of trials, concatenated across all behavioral training sessions. Dark lines show mean across all animals, and shaded areas represent SEM.* ***d****: quantification of both models' accuracy from early (first 4 days) to late (last 4 days). Asterisk indicates P <= 0.05.*

its value and thus the value of the newly assigned incorrect lever remains high (Fig. 3.6a,b). Thus, the RL model persists in choosing the unrewarded lever even after the animal has switched its strategy and thus performs less accurately. These data suggest that well-trained animals acquire an abstract representation of the task structure that influences their behavior.

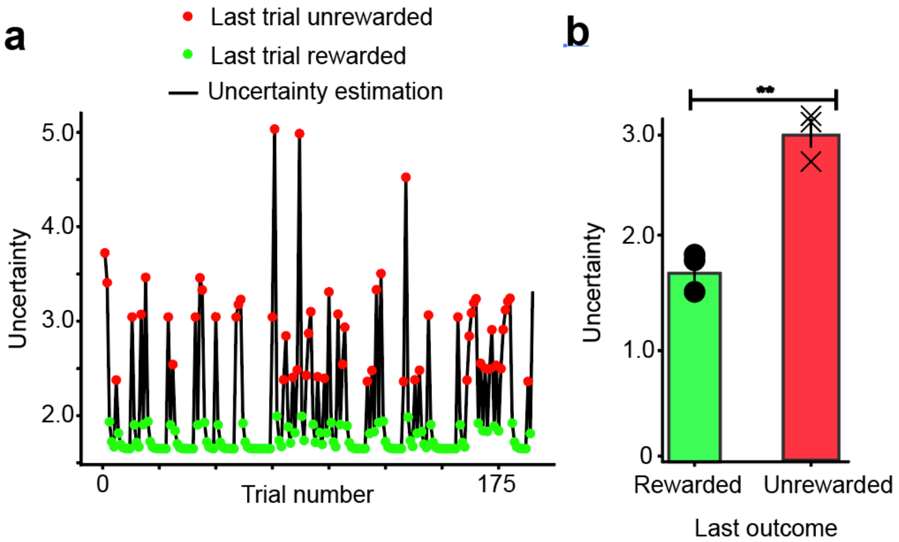


Figure 3.7: ***a**: an example of an uncertainty estimation for one behavioral session. Green dots indicate trials following a rewarded trial, while red dots indicate trials following unrewarded trials. Uncertainty about action values is influenced by trial history.* ***b**: Mean uncertainty scores for trials following a rewarded trial compared to scores following unrewarded trials. Lines show SEM; symbols show scores for animals individually. Asterisk indicates P < 0.01.*

Given that the HMM can provide an accurate description of behavior, we used the model parameters to estimate animals' uncertainty about their behavioral choice on any given trial. The HMM computes a probability estimate that the task is in either state (upper lever correct or lower lever correct) at each time step, which equates to the relative value of either action. Thus, we estimated the uncertainty on one trial as the the inverse of the difference between these two state estimates. For example, when P(upper lever = correct) was high and P(lower lever = correct) was low, the relative difference between action values was large and thus the uncertainty was low. Conversely, when the probabilities of each state were similar, the uncertainty value was high. We predicted that animals should behave in a more exploratory manner when uncertainty was high, and be exploitative when uncertainty was low. To test this, we computed the mean uncertainty values for trials following rewarded trials, after which animals were more likely to exploit, and mean uncertainty values following unrewarded trials when animals were more likely to explore. We found that for all animals, the mean uncertainty value was significantly higher following unrewarded trials compared to rewarded trials (Fig. 3.7a,b; mean uncertainty after rewarded trials = 1.758, mean uncertainty after unrewarded trials = 2.992; P = 0.001, T = -30.027). Thus, this uncertainty value provides a meaningful estimate of animals' decision to explore or exploit on a trial-by-trial basis.

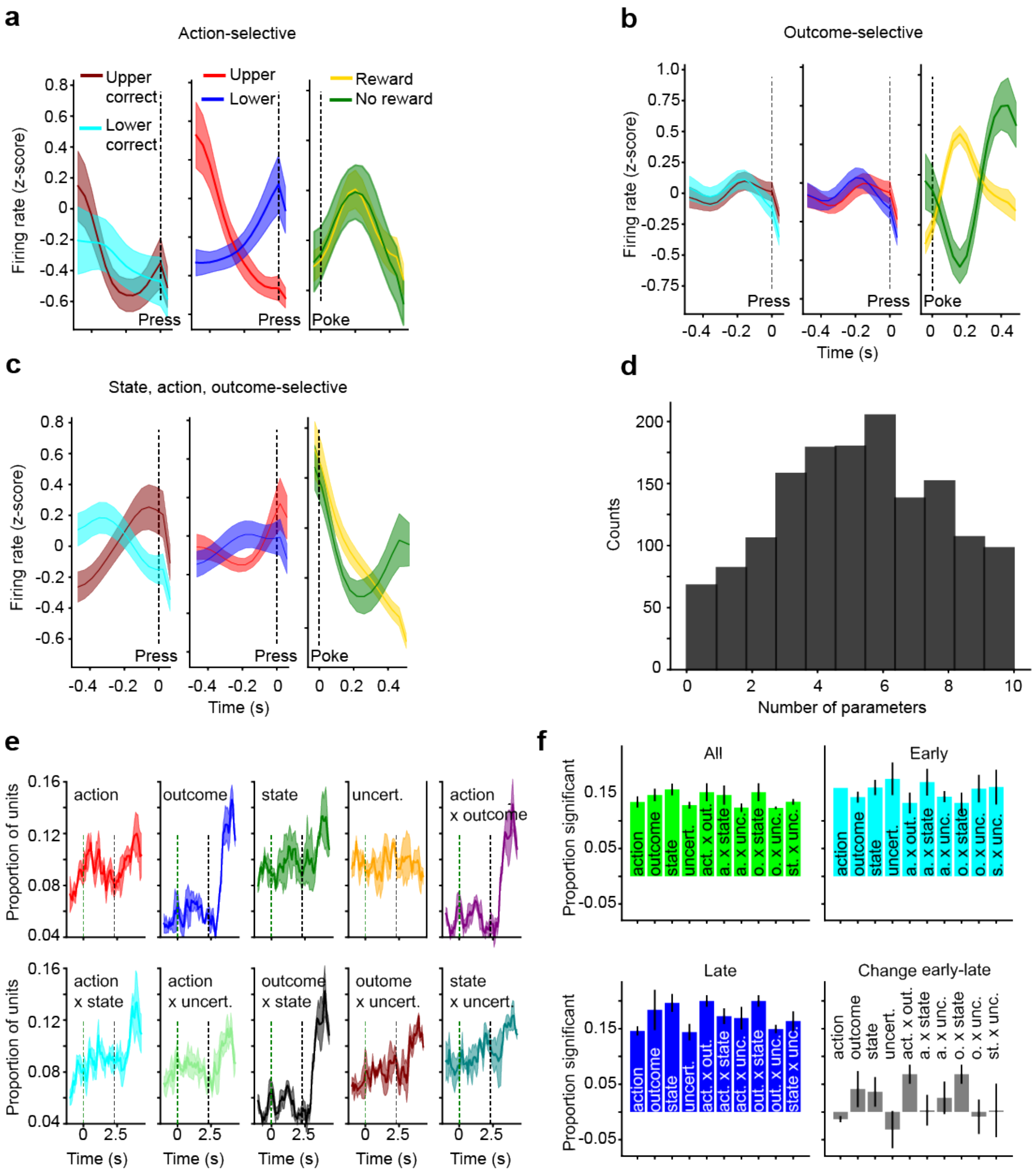


***Figure 3.8 a**: An example of an action-selective neuron with significant predictive value for lever choice. The left panel shows mean activity averaged over states; the middle shows mean activity averaged over actions; the right shows activity averaged over outcomes. State and action epochs are aligned to lever press (dashed line), while the outcome epoch is aligned to time of reward delivery. **b** and **c**: examples of different individual single neurons encoding different task parameters. Legends same as **a**. **d**: distribution of the number of parameters significantly encoded by single units, not including units that show no significant encoding. **e**: Change in the proportion of units significantly encoding each task paramter over the course of a trial. Solid lines show mean, and shaded areas represent S.E.M. for all animals. **f**: proportions of units with significant predictability for different task parameters in across training epochs, as well as the difference between early and late epochs (bottom right).*

3.3.4 Single neurons encode one or more task parameters

Previous reports have demonstrated that single striatum neurons can encode task parameters, including action choice, reward, and action value (Kim et al., 2009; Lau and Glimcher, 2008; Samejima, 2005). Using a linear regression analysis, we asked whether single DMS neurons encoded the parameters of this task. Specifically, we asked whether the activity of single neurons was predictive of action choice (upper lever or lower lever), task state (upper rewarded or lower rewarded), trial outcome (rewarded or unrewarded poke), the estimated uncertainty value, and interactions between each of these terms. Data used in this analysis was taken from a window that began 800 ms before a lever press, and ended 800 ms following a nose poke. We found that single neurons could be significantly predictive of one or more task parameters (Fig. 3.8a,b,c), with the number of parameters encoded by each neuron roughly following a normal distribution (Fig 3.8d). On average, single neurons encoded a mean of 5.25 parameters, suggesting that task encoding was mixed at the level of single units. We observed that the fraction of neurons significantly encoding individual parameters fluctuated over the course of the trial (Fig 3.8e). We next asked whether the proportion of units displaying selectivity for different parameters changed over the course of training. Although several parameters did show different encoding levels between early and late training, these differences were non-significant (Fig 3.8f).

3.3.5 Task parameters can be decoded from population activity

Given that single neurons displayed mixed selectivity for task parameters, we next asked how task-relevant activity of the recorded population was encoded at the population level. To extract relevant signals simultaneously from many recorded neurons, we used a directed dimensionality reduction technique known as de-mixed Principal Components Analysis (dPCA) (Kobak et al., 2016). Similar to principal components analysis, this technique aims to decompose high-dimensional neural activity into a few low-dimensional latent components that capture a large proportion of the variance. However, in addition, dPCA takes into account information about task parameters such that each component is explicitly related single task variables, de-mixing task-relevant signals from the population activity. Combining data from all animals and all trials, we performed a dPCA analysis that decomposed population activity into signals related to time, task state, action choice, and their interaction. We fitted dPCA models using neural data beginning 2 seconds prior to lever press and ending 2 second after receipt of reward. Models were fit with 15 components, which explained over 75% of the variance in the neural data (Fig. 3.9).

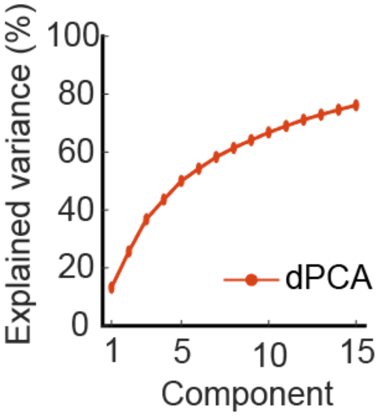


Figure 3.9: *Variance explained by dPCA models fit with 15 components.*

Using this method, latent signals encoding individual task parameters were clearly

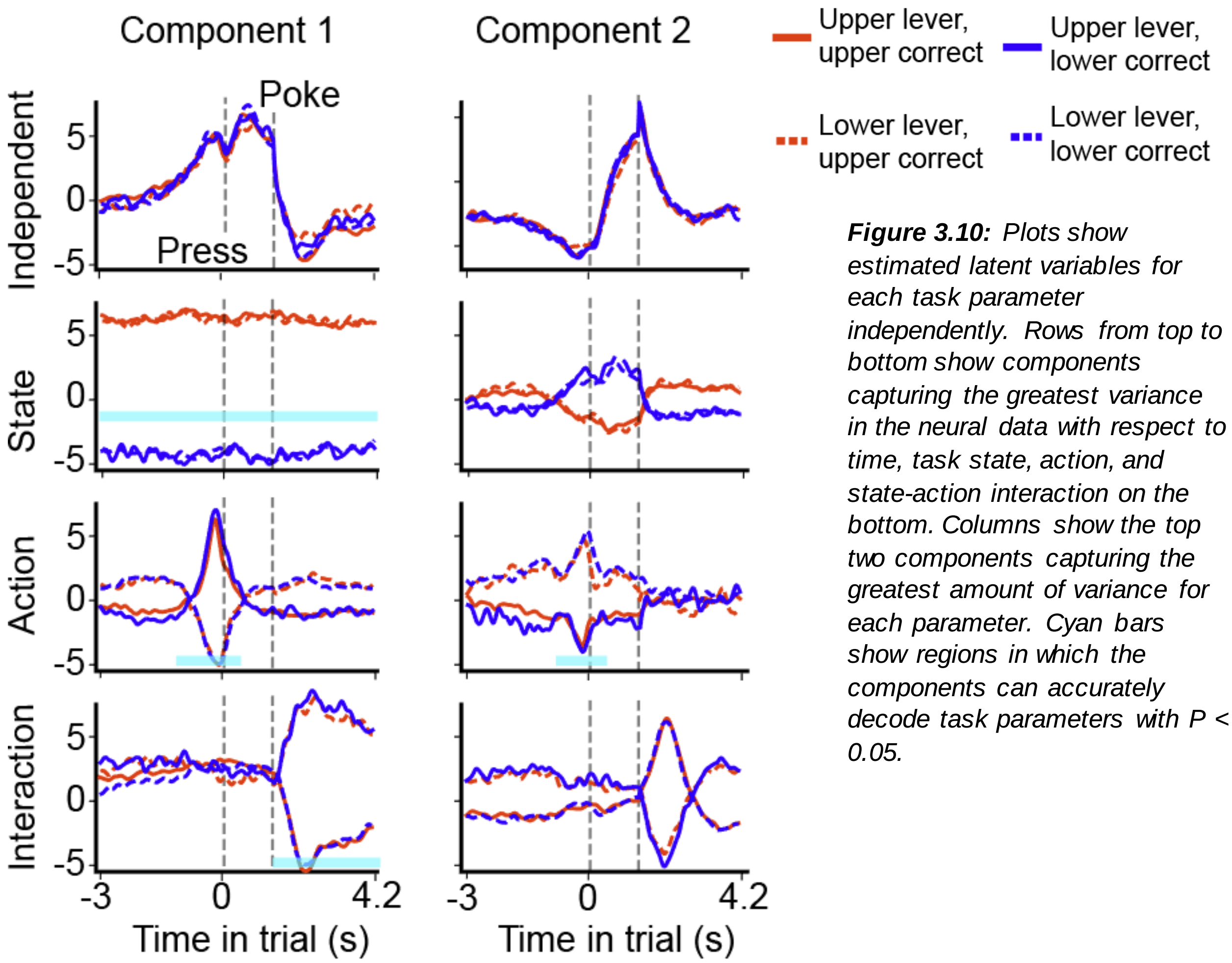


Figure 3.10: *Plots show estimated latent variables for each task parameter independently. Rows from top to bottom show components capturing the greatest variance in the neural data with respect to time, task state, action, and state-action interaction on the bottom. Columns show the top two components capturing the greatest amount of variance for each parameter. Cyan bars show regions in which the components can accurately decode task parameters with P < 0.05.*

separable (Fig. 3.10). Some components displayed sustained activity related to particular task states or lever choices, while others displayed activity that peaked around the time of lever press. The interaction components represented the interaction between task states (upper or lower correct) and action choices (upper or lower lever) and were therefore loosely related to task outcomes. Activity in these components peaked around the time of nose poke in the reward port, at which time the outcome of the trial was revealed to the animal. To test whether de-mixed components could accurately classify task variables, we measured the classification accuracy of individual components at each time bin for the component's related task variable using a cross-validation

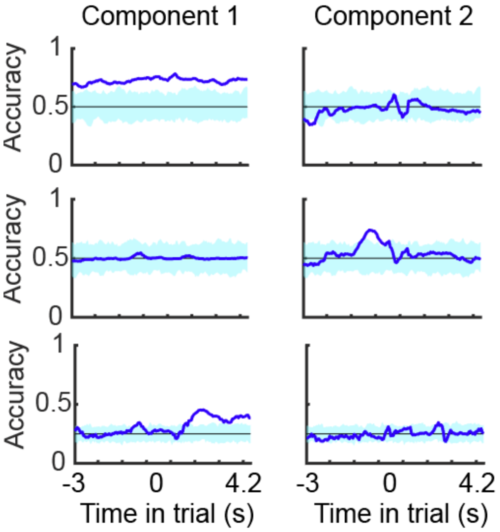


Figure 3.11 *Classification performance for each component in b. Solid line shows cross-validated accuracy; cyan bars show performance with shuffled data.*

procedure (see Methods). Then, to determine the significance of classification, we used a permutation test on shuffled data. In addition to appearing visually separable, the first component capturing the most variance for each task parameter (task state, action choice, and state-choice interaction), as well as the second action component, contained time regions in which they were predictive of the related task parameter (Fig. 3.10, cyan bars, Fig. 3.11). These results demonstrate that the population activity of dorsal striatum neurons can be decomposed into low-dimensional latent factors that encode task parameters, including those related to the state of a probabilistic switching task. Interestingly, the first latent component encoding task state showed significant classification accuracy across the entire duration of the trial, suggesting that task state is persistently encoded in DMS across many behavioral epochs within a trial. The 2-second pre-action and post-reward time windows we used in this analysis likely overlapped between trials, suggesting that the encoding of task state is likely persistent across epochs of several trials as well.

3.3.6 Population activity in DMS reflects trial-by-trial estimates of uncertainty about behavioral choice

Given that signals encoding the state of the task and thus action value are strongly encoded in DMS, we wondered whether uncertainty was also represented in population activity, and whether this representation fluctuated from trial to trial based on evidence from the sequential sequence of actions and outcomes. We showed previously that a hidden markov model could be used to accurately predict animals' behavior as well as uncertainty, and that this uncertainty value fluctuated depending on whether a trial was

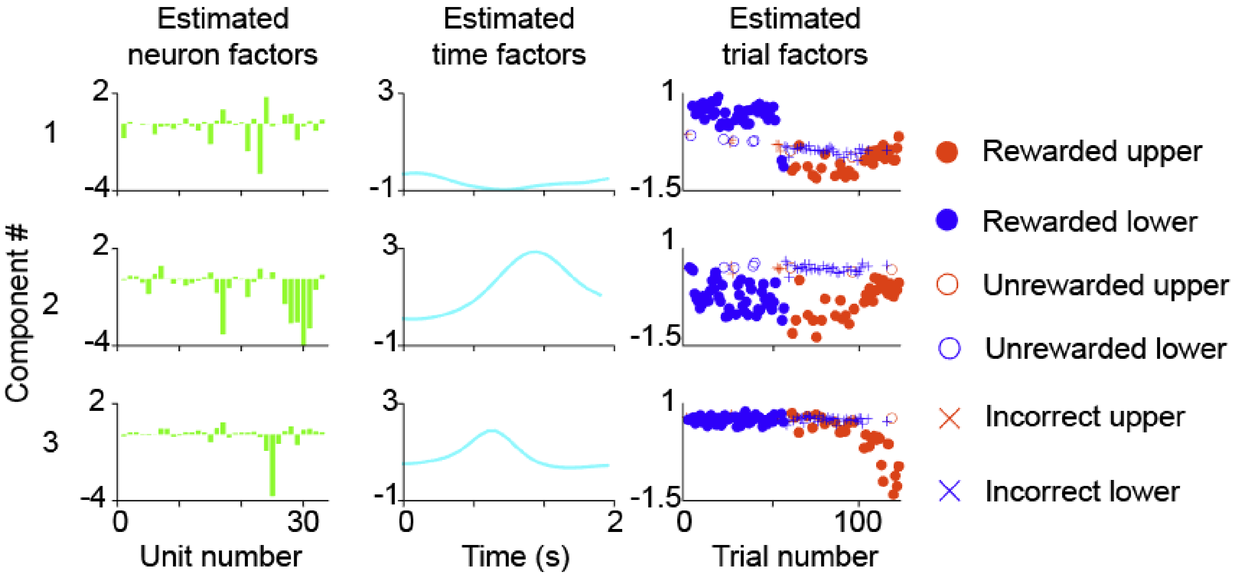


Figure 3.12: *Example latent factors decomposed from DMS activity recorded one training session. Rows show top 3 factors explaining the greatest proportion of variance. First column shows the weights assigned to units for the first 3 factors. Middle column shows the estimate time factors that capture variance within trials. Third column shows the estimated trial factors, color-coded for various trial types.*

rewarded or not. We thus sought to detect whether this time-varying uncertainty parameter was also encoded by population activity in DMS. Many techniques for dimensionality reduction in neural data, including dPCA used here, rely on computation of peri-stimulus time histograms constructed by averaging activity over many trials. However, uncertainty can fluctuate widely on a trial-to-trial basis as new observations are made (Fig. 3.7). To extract time-varying uncertainty signals from DMS population activity as training progressed, we utilized canonical polyadic (CP) tensor decomposition (Williams et al.). For each training session, we constructed a third-order tensor in which each element represents the binned spike activity of one neuron at a particular time point in a single trial. This tensor decomposition allowed us to decompose this spike activity into low-dimensional latent factors that capture how neural activity evolves across time within trials, such as in the dPCA analysis above, but also how it evolves in time across many trials (Fig. 3.12). CP tensor decomposition was fit to neural data from each session independently using 10 latent factors. Scrutinizing these models for each session individually, we observed that in many sessions, at least one of the first three trial factors accounting for the largest proportions of variance appeared to track our estimates of uncertainty as they evolved across trials (Fig. 3.13a). To better quantify this, we computed the Pearson correlation coefficient for each of the first three trial factors compared to the uncertainty values estimated by the HMM in each session. We found that across all animals, an average of 84.51% of all training sessions contained at least one trial factor that correlated significantly with the estimated uncertainty value on that trial (Fig. 3.13b). This relationship did not hold when uncertainty estimations and trial factors were shuffled for different days. Performing the same analysis with the shuffled data yielded a significant correlation on only 27.5% of

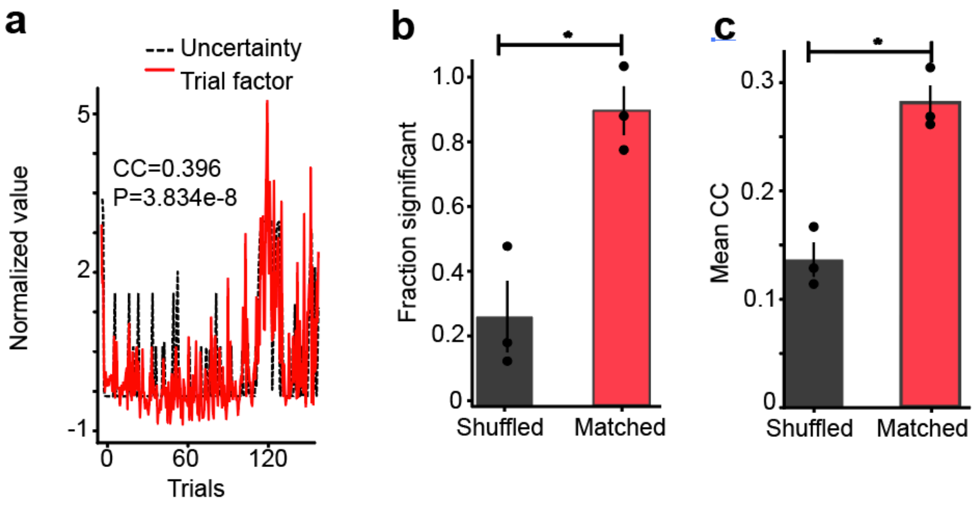


Figure 3.13: ***a****: An example of the correlation between an animal's estimated uncertainty in task state across one session, estimated from behavioral observations, and a trial factor extracted from DMS neuron population activity.* ***b****: fraction of trials for all animals in which one of the top three trial factors correlated significantly with the estimated uncertainty values. Black bar shows data computed for shuffled data; red bar shows matched data. Lines show SEM across all animals; dots show data for animals individually. Asterisk shows P < 0.05.* ***c****: Mean Pearson correlation coefficient between trial factors and uncertainty estimations for all animals. Legend same as in b.*

sessions for all animals ($P = 0.023$; T = 6.483), comparison between percent of significant days using real vs shuffled data). Additionally, the mean correlation coefficient for the intact dataset was 0.246, while the mean correlation coefficient for the shuffled data was 0.119 (Fig. 3.13c; $P = 0.027$; T = 5.963). These data demonstrate that in addition to selection of actions and evaluation of outcomes as has been previously reported, DMS neurons also encode signals that track an animal's uncertainty about the best course of action.

3.3.7 Uncertainty influences the magnitude and timing of action choice signals in DMS

Given that population activity of striatal neurons encodes actions related to action choice as well as uncertainty about action value, we asked whether this uncertainty value also influenced choice-related activity. For each training session, we used a logistic regression model to construct a linear decoder of population activity that computed the log odds of choosing the upper or lower lever. On average for each animal, the decoder was able to predict lever choice with a peak accuracy of 77.1 percent, with decoders tested on left-out data. This peak occurred approximately 250 ms before lever press (Fig. 3.14a). Using a decoder fit on all trials in one session, the log odds of pressing the upper or lower lever could then be computed at each time bin for each trial individually (see Methods). Here, we refer to these odds computed across time as decision variables. Fig. 3.14b shows the timecourse of the mean decision variable for upper lever presses and lower lever presses individually. Next, we chose to separate decision variables based on animals' uncertainty about the current choice. From the pool of all trials in one session, we sorted trials according to the estimated uncertainty magnitude. From here, we chose the top 10 percent of trials with high uncertainty as well as the lowest 10 percent of trials. We then averaged these decision

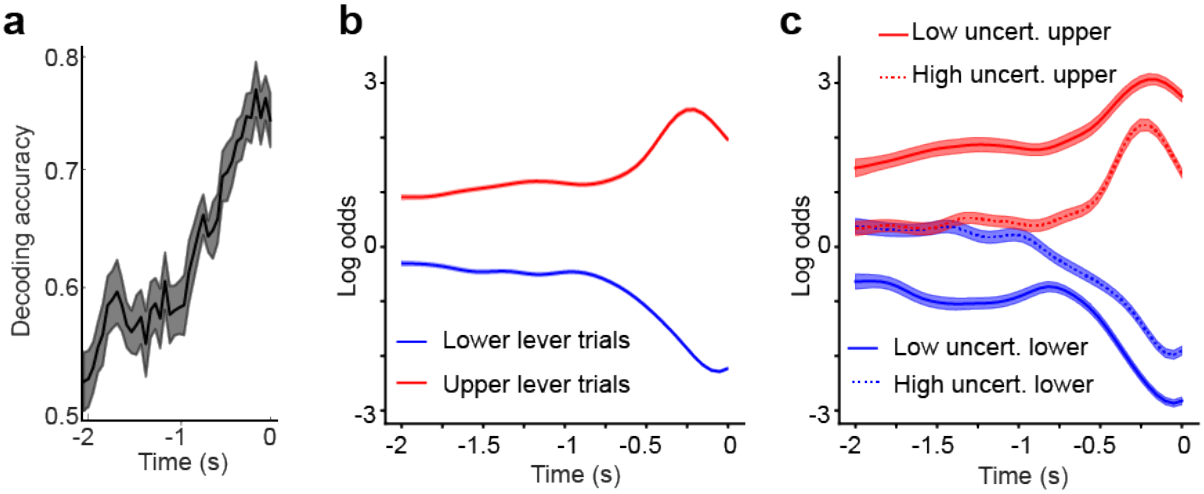


Figure 3.14: ***a***: *Mean decoding accuracy of lever choice using cross-validation. Solid lines show mean across all animals; shaded areas represent SEM.* ***b***: *Decision variables represented by the decoded log odds of choosing either lever prior to lever press.* ***c***: *Mean decision variables separated by trials with high estimated uncertainty about action value and low estimated uncertainty.*

variables across all animals. As shown in Fig. 3.14c, decision variables for upper or lower lever presses on low uncertainty trials were separable throughout the length of the analysis window, suggesting that signals for action choice persisted from seconds before action until the action took place. However, for high uncertainty trials, animals' choice only became clear approximately 800 ms prior to lever press. Additionally, the magnitude of decision variables was greater for low uncertainty trials than for high uncertainty trials. These data suggest that signals encoding action choice in the striatum are modulated by an animal's uncertainty about whether a particular decision is correct.

3.4 Discussion

The striatum is a site of convergence for efferent activity from most cortical regions (Hintiryan et al., 2016; Kemp and Powell, 1970; McGeorge and Faull, 1989; Webster, 1965), and thus is recipient of task-related information of many different types. One popular hypothesis is that the cortico-basal ganglia circuitry performs a selection function, in which action values learned through experience are used to select between many possible competing actions (Redgrave et al., 2011). The anatomy of cortico-basal ganglia loops is well-suited for such a function: glutmatergic collaterals from the cortex pass to the striatum, which sends inhibitory GABA-ergic output to the SNr or GPi, the output nuclei of the basal ganglia. Both of these nuclei in turn tonically inhibit the thalamus which projects back to the cortex. Therefore, the net effect of activating the striatum is to release inhibition over thalamic inputs to the cortex. Although initial reports proposed that these recurrent anatomical loops were parallel and segregated into different circuits processing different types of information (Alexander et al., 1986), more recent work has suggested that there is greater convergence between striatal areas receiving different types of information (Joel and Weiner, 1994).

Despite receiving input from many cortical areas involved in higher-order information processing, basal ganglia circuits have long been associated with basic sensorimotor instrumental learning and the encoding of reward prediction error (Oyama et al., 2015). However, more recently, human neuroimaging studies have demonstrated that the dorsomedial striatum is also active during tasks involving internal models of task states (Daw et al., 2011; Gläscher et al., 2010; Li and Daw, 2011). Here, we show that spike activity in medial striatum neurons robustly encoded signals not only related to actions and outcomes, but also related to the current state of a switching task, and the uncertainty about that state estimate. Previous reports have shown that lesions of the frontal cortex and DMS produce similar deficits of perseverative responding in reversal learning tasks (Boulougouris et al., 2007; Clarke et al., 2008; Izquierdo et al., 2016). Combined with anatomical reports of frontal cortical projections to the striatum (Hintiryan et al., 2016; Kemp and Powell, 1970), our results seem to support the idea that signals relating to model-based estimates of the environment pass from the frontal cortex through the striatum before influencing behavior.

Furthermore, we found that animals' uncertainty in the state of the task influenced the magnitude and timing of decision signals in encoded by the population. If the striatum is indeed performing an action selection function, this result suggests that reliable knowledge about the correct action to perform in a given context may bias decision-making processes in the striatum. Conversely, uncertainty may result in weak choice signals, the result of which may direct animals to be more exploratory in their behavior. Given the diversity of convergent information in striatal circuits, our results support the notion that cortico-basal ganglia circuits play an integral role in integrating relevant information, both from external and internal sources, in order to guide adaptive behavior.

Chapter 4:

Wireless recording in the peripheral nervous system with ultrasonic neural dust

The experiments described in **Chapters 2 and 3** relied upon chronically implanted microwire arrays to obtain neural recordings. Although this approach is feasible in laboratory setting, state-of-the-art invasive BMI systems that rely on cabled electronics create challenges for widespread clinical adoption and daily use in human subjects. As discussed in **Chapter 1**, the need for miniaturized, wireless, implantable neural recording systems extends beyond the field of motor BMI systems. Recent technological innovations and fundamental discoveries have renewed interest in strategies to treat diseases by altering activity in the peripheral nervous system with electronic devices (Birmingham et al., 2014; Famm et al., 2013). These emerging technologies, known as bioelectronic medicines or electroceuticals, seek to decipher and modulate electrophysiological activity in the body to attain therapeutic effects at target organs. Similar to the technologies used to record and stimulate the brain, current approaches to interfacing with peripheral nerves and muscles rely heavily on wires, creating problems for chronic use. Simultaneously, while emerging wireless approaches exist, they lack the size scalability necessary to interrogate small-diameter nerves. Furthermore, conventional electrode-based technologies lack the capability to record from nerves with high spatial resolution or to record independently from many discrete sites within a nerve bundle. In this chapter, we demonstrate neural dust, a wireless and scalable ultrasonic backscatter system for powering and communicating with implanted bioelectronics. We show that ultrasound is effective at delivering power to mm-scale devices in tissue; likewise, passive, battery-less communication using backscatter enables high-fidelity transmission of electromyogram (EMG) and electroneurogram (ENG) signals from anesthetized rats. These results highlight the potential for an ultrasound-based neural interface system for advancing future bioelectronics-based therapies.

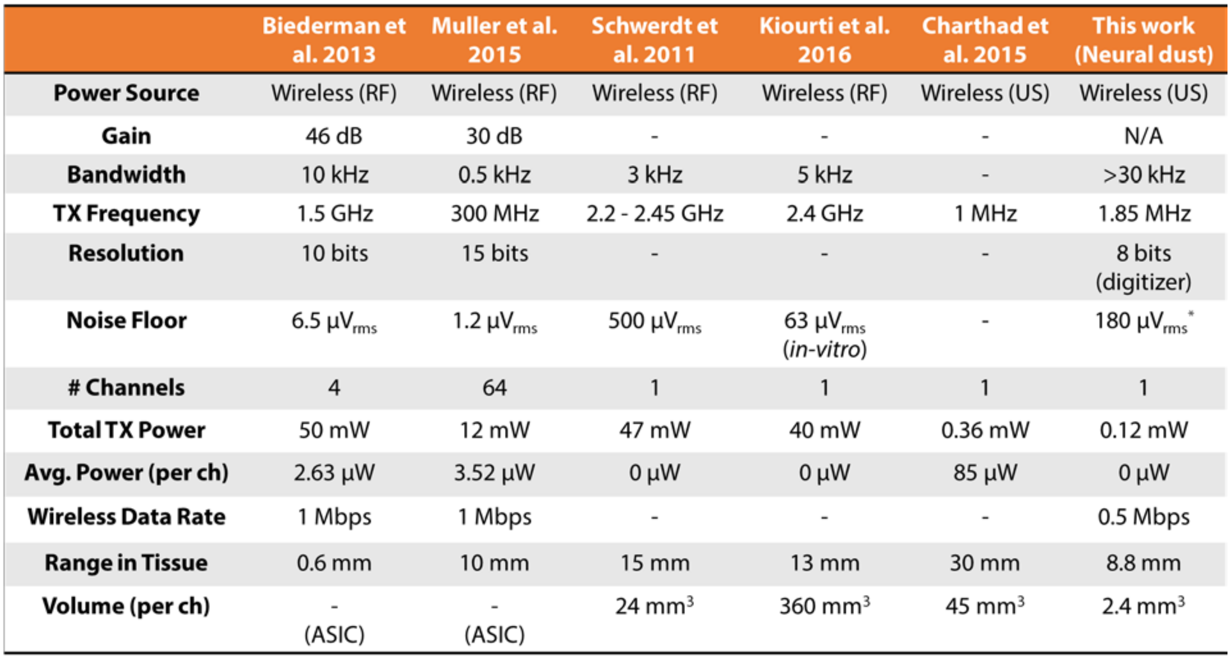
4.1 Introduction

Recent technological advances (Boretius et al., 2010; Delivopoulos et al., 2012) and fundamental discoveries (Bhadra and Kilgore, 2005; Pavlov and Tracey, 2012; Rosas-Ballina et al., 2011)have renewed interest in implantable systems for interfacing with the peripheral nervous system. Early clinical successes with peripheral neurostimulation devices, such as those used to treat sleep apnea (Strollo et al., 2014) or control bladder function in paraplegics (Creasey et al., 2001) have led clinicians and researchers to propose new disease targets ranging from diabetes to rheumatoid arthritis (Famm et al., 2013). A recently proposed roadmap for the field of bioelectronic medicines highlights the need for new electrode-based recording technologies that can detect abnormalities in physiological signals and be used to update stimulation parameters in real-time. Key features of such technologies include high-density, stable recordings of up to 100 channels in single nerves, wireless and implantable modules to enable characterization of functionally specific neural and electromyographic signals, and scalable device platforms that can interface with small nerves of 100 µm diameter or less (Birmingham et al., 2014) as well as specific muscle fibers. Current approaches to recording peripheral nerve activity fall short of this goal; for example, cuff electrodes provide stable chronic performance, but are limited to recording compound activity from the entire nerve. Single-lead intrafascicular electrodes can record from multiple sites within a single fascicle, but do not enable high-density recording from discrete sites in multiple fascicles (Lefurge et al., 1991). Similarly, surface EMG arrays allow for very high density recording (Lapatki et al., 2004; Martinez-Valdes et al., 2016) but do not capture fine details of deep or small muscles. Recently, wireless devices to enable untethered recording in rodents (Lee et al., 2013; Szuts et al., 2011) and nonhuman primates (Foster et al., 2014; Schwarz et al., 2014; Yin et al., 2014), as well as mm-scale integrated circuits for neurosensing applications have been developed (Biederman et al., 2015; Denison et al., 2007; Muller et al., 2015). However, most wireless systems use electromagnetic (EM) energy coupling and communication, which becomes extremely inefficient in systems smaller than ~5 mm due to the inefficiency of coupling radio waves at these scales within tissue (Rabaey et al., 2011; Seo et al., 2013); see also *Size Scaling and Electromagnetics* in the Discussion section, below. Further miniaturization of wireless electronics platforms that can effectively interface with small-diameter nerves will require new approaches.

In contrast to EM, ultrasound offers an attractive alternative for wirelessly powering and communicating with sub-mm implantable (Charthad et al., 2015; Larson and Towe, 2011; Meng and Sahin, 2013; Ozeri and Shmilovitz, 2010; Seo et al., 2015). Ultrasound has two advantages. First, the speed of sound is 10^5 x lower than the speed of light in water, leading to much smaller wavelengths at similar frequencies; this yields excellent spatial resolution at these lower frequencies as compared to radio waves. Second, ultrasonic energy attenuates far less in tissue than EM radiation; this results not only in much higher penetration depths for a given power, but also significantly decreases the amount of unwanted power introduced into tissue due to scattering or absorption. In fact, for most frequencies and power levels, ultrasound is safe in the human body. These limits are well-defined and ultrasound technologies have long been used for diagnostic and therapeutic purposes. As a rough guide, about 72x more power is allowable into the human body when using ultrasound as compared to radio waves (Lin, 2006; US Food and Drug Administration, 2008)

We previously introduced the *neural dust* ultrasonic backscattering concept to harness the potential advantages of ultrasound and showed that, theoretically, such a system could be scaled well below the mm-scale when used for wireless electrophysiological neural recording (Seo et al., 2013, 2015). Here, we present the first experimental validation of a neural dust system *in-vivo* in the rat peripheral nervous system (PNS) and skeletal muscle, reporting both electroneurogram (ENG) recordings from the sciatic nerve and electromyographic (EMG) recordings from the gastrocnemius muscle. The neural dust system consists of an external ultrasonic transceiver board which powers and communicates with a millimeter-scale sensor implanted into either a nerve or muscle (Fig. 4.6a). The implanted mote consists of a piezoelectric crystal, a single custom transistor, and a pair of recording electrodes (Fig. 4.6b,c and Fig. 4.1).

During operation, the external transducer alternates between a) emitting a series of six 540 ns pulses every 100 µs and b) listening for any reflected pulses. The entire sequence of transmit, receive and reconstruction events are detailed in Fig. 4.8; this sequence (steps a – h) is repeated every 100 µs during operation. Briefly, pulses of ultrasonic energy emitted by the external transducer impinge on the piezocrystal and are, in part, reflected back towards the external transducer. In addition, some of the ultrasonic energy causes the piezocrystal to vibrate; as this occurs, the piezocrystal converts the mechanical power of the ultrasound wave into electrical power, which is supplied to the transistor. Any extracellular voltage change across the two recording electrodes modulates the transistor's gate, changing the amount of current flowing between the terminals of the crystal. These changes in current, in turn, alter the vibration of the crystal and the intensity of the reflected ultrasonic energy. Thus, the shape of the reflected *ultrasonic* pulses encodes the *electrophysiological voltage* signal seen by the implanted electrodes and this electrophysiological signal can be reconstructed externally. The performance specifications of neural dust in comparison to other state-of-the-art systems are summarized the Table below.



	Biederman et al. 2013	Muller et al. 2015	Schwerdt et al. 2011	Kiourti et al. 2016	Charthad et al. 2015	This work (Neural dust)
Power Source	Wireless (RF)	Wireless (RF)	Wireless (RF)	Wireless (RF)	Wireless (US)	Wireless (US)
Gain	46 dB	30 dB	-	-	-	N/A
Bandwidth	10 kHz	0.5 kHz	3 kHz	5 kHz	-	>30 kHz
TX Frequency	1.5 GHz	300 MHz	2.2 - 2.45 GHz	2.4 GHz	1 MHz	1.85 MHz
Resolution	10 bits	15 bits	-	-	-	8 bits (digitizer)
Noise Floor	6.5 μV_{rms}	1.2 μV_{rms}	500 μV_{rms}	63 μV_{rms} (*in-vitro*)	-	180 μV_{rms}^{*}
# Channels	4	64	1	1	1	1
Total TX Power	50 mW	12 mW	47 mW	40 mW	0.36 mW	0.12 mW
Avg. Power (per ch)	2.63 µW	3.52 µW	0 µW	0 µW	85 µW	0 µW
Wireless Data Rate	1 Mbps	1 Mbps	-	-	-	0.5 Mbps
Range in Tissue	0.6 mm	10 mm	15 mm	13 mm	30 mm	8.8 mm
Volume (per ch)	- (ASIC)	- (ASIC)	24 mm^3	360 mm^3	45 mm^3	2.4 mm^3

* In a stationary, water tank setup

4.2 Methods

4.2.1 Neural dust mote assembly

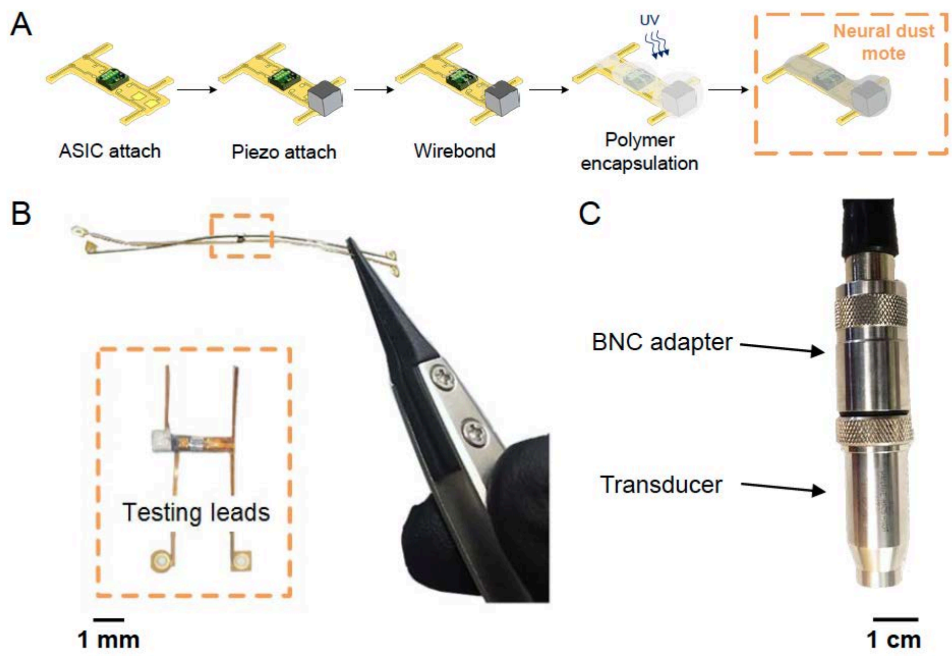


Figure 4.1: ***a:*** *Robust, high-yield fabrication steps for the mote, which is encapsulated with medical grade, UV-curable epoxy.* ***b.*** *Close-up of neural dust mote on a flexible PCB with calibration leads to measure electrophysiological signals (ground truth) and voltages harvested on the piezocrystal. During in-vivo experiments, calibration leads are removed.* ***c.*** *Single-element Olympus transducer used in this manuscript.*

Lead zirconate titanate (PZT) sheets (841, APC Int., Mackeyvile, PA) with ~12 µm of fired on silver were diced to desired dimensions using a dicing saw (DAD3240, Disco, Santa Clara, CA) with a ceramic blade (PN CX-010-270-080-H). The diced PZT coupon, along with the custom transistor, were attached to a 50 µm thick polyimide flexible PCB with immersion gold (Altaflex, Santa Clara, CA) using a thin layer of two-part silver epoxy with 1:1 mix ratio (H20E, Epotek, Billerica, MA). The board was cured at 150°C, which is far below the melting temperature of polyimide and the Curie temperature of the PZT, for 10 minutes. The custom transistor was wirebonded using an aluminum ultrasonic wirebonder (7400B, West Bond, Scotts Valley, CA) to pre-patterned targets. In order to prevent charge build-up on the PZT from the wedge contact, top and bottom contacts of the PZT were discharged to a thin metal sheet prior to wirebonding the top contact of the PZT to close the circuits. Medical-grade, UV-curable epoxy (OG116-31, Epotek) was used to protect the wirebond and provide insulation. The platform was then cured in UV chamber (Flash, Asiga, Anaheim Hills, CA) with 92 mW/cm^2 @ 365 nm for 3 minutes.

4.2.2 Electrical and ultrasonic characterization of the assembly in water

The custom transistor was electrically tested with a precision current meter (2400-LV, Keithley) and a DC-power supply (3631A, Agilent, Santa Clara, CA). To characterize the piezocrystal prior to assembly, an impedance plot was obtained with an impedance analyzer (4285A, Agilent) using two-terminal measurements with open/short/load calibration scheme. The impedance of exposed gold recording pads (0.2 mm x 0.2 mm), separated by 1.8 mm on the bottom of the PCB, was measured in Phosphate Buffered Solution (PBS 1X) with an electrochemical impedance spectroscope (nanoZ, White Matter LLC, Mercer Island, WA). The device formed the active electrode and a silver wire formed the reference electrode.

Ultrasonic characterization of the transducer was performed in a custom-built water tank. A capsule hydrophone (HGL-0400, Onda Corp., Sunnyvale, CA) with 20 dB preamplification (AH-2020, Onda Corp.) was mounted on a computer-controlled 2D translating stage (XSlide, VelMex Inc., Bloomfield, NY) and was used to calibrate the output pressure and characterize beam patterns of a 2.25 MHz single element transducer (V323-SU, Olympus). Verification of ultrasonic power transfer and communication sensitivity was performed in a smaller water-tank with the transducer mounted on manual translational and rotational stages (Thorlabs Inc.). The outline of the neural dust mote was patterned on an extruded acrylic piece with UV-laser and the mote was clamped to the acrylic stage with nylon screws. The position and angle of the transducer with relative to the mote were manually adjusted until the maximum voltage was measured across the piezocrystal. Cable capacitances and parasitics were carefully calibrated by adjusting the series capacitance in the high-impedance probes (N2863B, Agilent). An electric field in the water tank was generated with a current source (2400-LV, Keithley) forcing electrical current at varying current densities through two 0.127 mm thick platinum wires (773000, A-M systems) immersed in the tank.

The transceiver board consisted of a custom integrated circuit (IC) in a QFN-64 package that achieved an on-chip 1.8V to 32V charge pump efficiency of 33% and system latency of 20 ns and consumed 16.5 µJ per each transmit cycle (Tang et al., 2015). During the receive mode, the high voltage switch was closed and the signal was amplified by 28 dB; both operations were performed on-chip. The output signal from the chip was digitized by an off-chip 10-bit, 100 MHz analog-to-digital converter (ADC) (LTC2261-12, Linear Technology, Milpitas, CA). The outputs of the ADC were fed back into the field-programmable gate array (FPGA) and USB 3.0 integration module (XEM6310-LX45, Opal Kelly, Portland, OR) and transferred to the laptop. The FPGA-USB module was also used to serially program the IC.

4.2.3 Experiment setup and surgical procedures

All animal procedures were performed in accordance with University of California Berkeley Animal Care and Use Committee regulations. Adult male Long-Evans rats were used for all experiments. Prior to the start of surgery, animals were anesthetized with a mixture of ketamine (50 mg/kg) and xylazine (5 mg/kg) IP. The fur surrounding the surgical site was shaved and cleaned. For EMG recordings, a patch of gastrocnemius muscle roughly 10 mm x 5 mm in size was exposed by removing the overlying skin and fascia. The neural dust mote was then placed on the exposed muscle, and the skin and fascia were replaced and the wound was closed with 5/0 surgical suture. For ENG recordings, the sciatic nerve was exposed by making an incision from the sciatic notch to the knee, and separating the hamstring muscles. The mote was then placed in contact with the epineurium of the main branch of the sciatic nerve bundle, and sutured to the nerve using 10/0 microsurgical suture. Animals were euthanized at the conclusion of the experiments.

Constant-current stimulation was delivered using an isolated pulse stimulator (2100, A-M Systems). Single biphasic pulses with a 2 ms pulse width were used to deliver stimulation at various current amplitudes. For each experiment, electrophysiological responses from 10 stimulations (i.e., samples) were recorded. The FPGA-USB module generated a trigger for the stimulator every 6 seconds. For EMG experiments, bipolar Ag-AgCl hook electrodes placed around the trunk of the sciatic nerve were used for stimulation. To evoke ENG activity, 28G stainless steel needle electrodes were placed in the foot with an inter-electrode distance of approximately 5mm.

The wired signals were amplified (100x) by a battery-powered differential amplifier with a built-in bandpass filter (DAM50, WPI, Sarasota, FL) set at 10 Hz – 1 kHz. The ground reference for the amplifier was a 28G stainless steel needle electrode placed in the contralateral foot relative to the recording setup. The output of the amplifier was connected to a multi-channel digitizer, sampled at 100 kHz, and recorded on computer.

4.2.4 Transmit pulse waveforms and data acquisition

The neural dust mote was placed one Rayleigh distance from the transducer (8.9 mm), which corresponded to 5.9 µs transit time, assuming an acoustic velocity of ~1500 m/s in water. 6-cycles of square waves at 1.85 MHz with peak voltage of 5 V were launched every 100 µs (pulse repetition frequency (PRF) of 10 kHz). The total transmit pulse width was approximately, 3.3 µs, which was sufficiently small to prevent any overlaps with the first harvested voltage measurement at 5.9 µs. Given that the first reflection back to the transducer (e.g., backscatter) occurred at approximately 11.8 µs (twice the transit time) and persisted until for 3.3 µs, the maximum PRF (e.g., in this context, the sampling rate) was ~66 kHz. Given that the bulk peripheral nerve responses occurred below 1 kHz (Mezzarane et al., 2013), a PRF of 10 kHz was chosen to sufficiently capture the dynamics.

In order to sample the backscatter waveform at 1.85 MHz without losing signal fidelity, the off-chip ADC on the transceiver board was heavily oversampled at 50 MHz. This resulted in ~8 Mbits of data in a 10 ms neural recording, which was stored in a 128 MByte, 16-bit wide, synchronous DDR2 DRAM (MT47H64M16HR-3, Micron Technology, Boise, ID). The raw waveforms were transferred to the laptop via the USB interface post-recording. The raw waveforms were simultaneously recorded using an 8-bit digitizer (USB-5133, National Instruments, Santa Clara, CA) for comparison.

4.2.5 Backscatter data processing

Raw backscatter waveforms, sampled at 50 MHz, from each experiment were sliced and time-aligned to be averaged over samples. The averaged signals were bandpass-filtered with a symmetric 4th order Butterworth filter from 10 Hz to 1 kHz. The distinct characteristics of the backscatter waveform (Fig. 4.7e) were used as a template to locate the region of interest. The signals were then rectified and the integral of the region was computed to estimate the input voltage signal, which exhibited a linear response (Fig. 4.7f). Multiplication factor for the signal was extracted from the ground truth measurement.

4.2.6 Piezocrystal model

There are a number of equivalent circuit models to describe the electromechanical operation of a piezoelectric crystal. In this study, a combination of KLM model (Krimholtz et al., 1970) and resonance model (Kino, 1987) are used. Across a wide range of frequencies, KLM model (Fig. 4.2a) can be used; the frequency-dependent three-port network consists of one electrical port (where electric power is applied or collected) and two acoustical ports (where mechanical waves are produced or sensed from the front and back faces of the transducer). The parallel-plate capacitance due to the electrodes and the frequency-dependent acoustic capacitance are modeled as C_o and X_i, respectively, and the transduction between electrical and mechanical domains is modeled as an ideal electromechanical transformer with a turn ratio of ϕ, connected to the middle of a transmission line of length λ/2 (Fig. 4.2a). Near the resonant frequency of a piezocrystal, KLM model can be simplified to the resonance model (Fig. 4.2b). However, both models are derived under the assumption of one-dimensional operation, and therefore can only provide a valid representation for a piezoelectric transducer with an aspect ratio (width/thickness) greater than 10 or less than 1/10. Therefore, piezocrystals are also simulated using a 3D finite element package (COMSOL Multiphysics, Acoustic Module) to model anisotropies and mode coupling between several resonant modes.

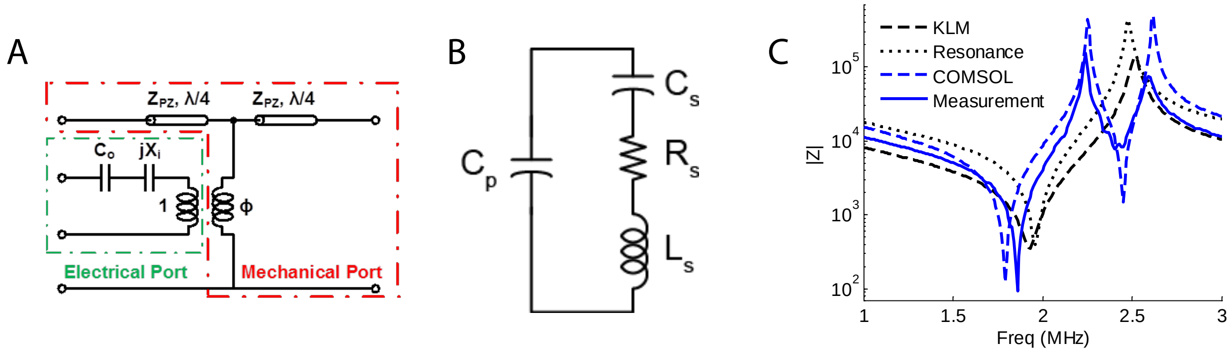


Figure 4.2: ***a:*** *KLM model or* ***b****. simplified near resonance model.* ***c****. Measured impedance spectrum of a 0.75 mm x 0.75 mm x 0.75 mm piezocrystal cube used in this manuscript compared to impedances estimated by the resonance, KLM, and COMSOL models.*

Simulated impedance spectrums using various models of the piezocrystal matched the measured resonant frequency at 1.85 MHz with the impedance magnitude of ~100 Ohm (Fig. 4.2c). Given the aspect ratio of 1, the measurement of 0.75 mm x 0.75 mm x 0.75 mm piezocrystal used in this study showed a splitting of the anti-resonant peak due to mode coupling, at 2.25 MHz and 2.6 MHz, which was captured by the COMSOL model. Both KLM and the resonance model, however, did not capture this effect. Although the COMSOL model of the piezocrystal was more accurate, in order to reduce the computational complexity, KLM model was used to simulate the link behavior near the operation frequency of 1.85 MHz.

4.2.7 *In-vivo* ultrasonic transmission

A 2.25 MHz single element transducer (V323-SU, Olympus NDT, Waltham, MA) was used to generate 6 pulses at 1.85 MHz. The transducer had a measured half-power bandwidth (HPBW) of more than 2.5 MHz (Fig. 4.3a). In order to measure the transmission loss through the tissue, various thicknesses of skin found near the gastrocnemius muscle of a male Long-Evans rat was placed in between the transducer and the neural dust prototype. The harvested voltage on the piezocrystal with and without tissue was obtained and 8.9 mm-thick tissue resulted in 10 dB of tissue attenuation (Fig. 4.3b).

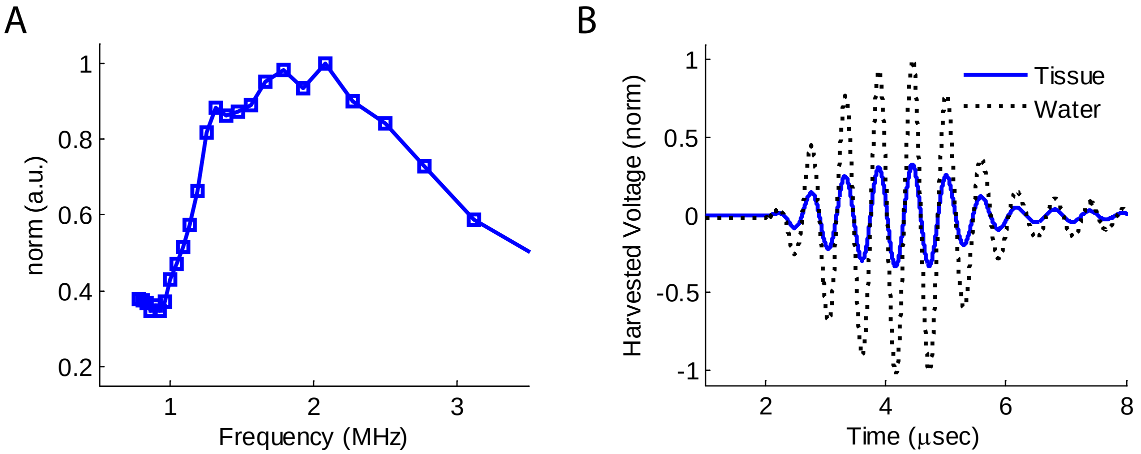
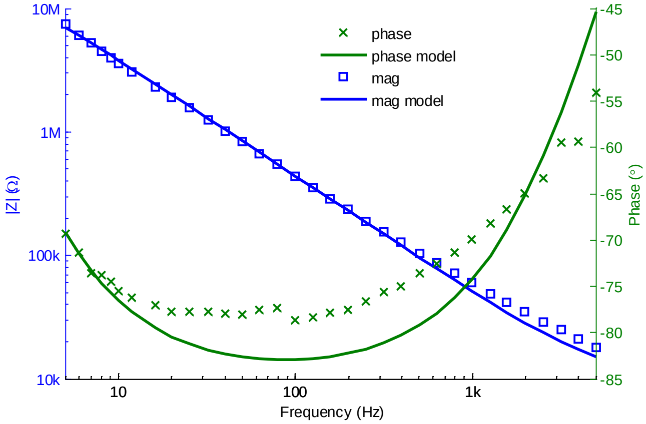


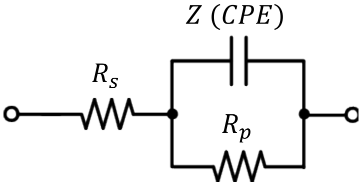
Figure 4.3: ***a***: Frequency response of the transducer. **b:** Ultrasonic attenuation in 8.9 mm of tissue.

4.2.8 Electrode characterization

A recording electrode pair was made of immersion gold by a flexible PCB vendor (Altaflex, Santa Clara, CA) and measured 0.2 mm x 0.2 mm. We characterized the electrical properties of the surface electrode by measuring the recording site impedances in Phosphate Buffered Solution (PBS 1X) with an electrochemical impedance spectroscope (nanoZ, White Matter LLC, Mercer Island, WA). The device formed the active electrode and a silver wire formed the reference electrode.

The electrode/electrolyte interface can be fitted to a Randles Cell model (Randles, 1947) to extract the line resistance (R_s = 9.73 $k\Omega$), charge transfer resistance (R_p = 25.6 $M\Omega$), and the parameters of the constant phase element (CPE: n = 0.94, P=5.1 nF·s^n); the MATLAB package Zfit (https://www.mathworks.com/matlabcentral/fileexchange/19460-zfit) was used to fit these parameters (Fig. 4.4).





$$Z\ (CPE) = \frac{1}{P(j\omega)^n}$$

$$Z_{total} = R_s + \frac{R_p}{1 + R_p P(j\omega)^n}$$

R_s	$9.73\ k\Omega$
R_p	$25.6\ M\Omega$
n	0.94
P	$5.1\ nF$

Figure 4.4: *Impedance spectroscopy of the electrode and the fit using Randles Cell model.*

4.2.9 ENG recording with different electrode spacing

Recording electrodes with various spacing were fabricated on a 50 µm thick polyimide flexible printed circuit board (PCB). There were a total of 5 electrodes, each measuring 0.2 mm x 0.2 mm, and one of them was used as the reference electrode. Other electrodes were spaced 0.3 mm, 0.8 mm, 1.3 mm, and 1.8 mm, respectively, apart from the reference electrode.

The spacing board was placed in contact with the epineurium of the main branch of the sciatic nerve bundle (distal) and sutured to the nerve. Bipolar Ag-AgCl hook electrodes placed around the trunk of the sciatic nerve (proximal) were used for stimulation. Constant-current simulation of a single biphasic pulse with a duration of 0.5 ms every 1 second was delivered using an isolated pulse stimulator (2100, A-M Systems, Sequim, WA).

The recorded signals with various spacing between the electrodes were amplified (100x) by a battery-powered differential amplifier with a built-in bandpass filter (DAM50, WPI, Sarasota, FL) set at 10 Hz – 1 kHz (Fig. 4.5a). As expected, the peak-to-peak voltage recorded on the electrode increased with the spacing at least quadratically. The amplitude saturated after the spacing of 1.3 mm, confirming that the electrode spacing of 1.8 mm on the recording sensor was sufficient to capture the maximum, saturated ENG response (Fig. 4.5b).

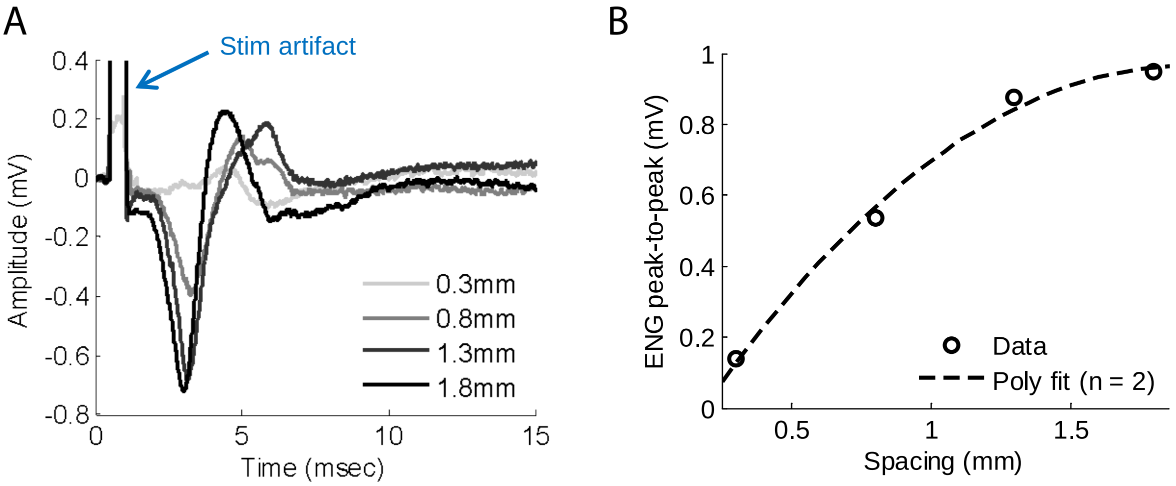


Figure 4.5: a: Recorded time-domain ENG responses. **b:** Peak-to-peak ENG with varying electrode spacing.

4.2.10 Calculation of acoustic intensity

Several parameters are established by the American Institute for Ultrasound in Medicine and National Electronics Manufacturers Administration to assess the safety of an ultrasonic system (American Institute of Ultrasound in Medicine, 1998). The acoustic power output of diagnostic ultrasonic system is limited by the de-rated values of spatial-peak pulse-average intensity (I_{SPPA}), spatial-peak temporal average intensity (I_{SPTA}), and mechanical index (MI). These de-rated values are computed by multiplying the measured values in water by an attenuation factor of 0.3 dB/cm/MHz to simulate the effects on tissue (US Food and Drug Administration, 2008).

A capsule hydrophone (HGL-0400, Onda Corp) with 20 dB preamplification (AH-2020, Onda Corp., Sunnyvale, CA) was mounted on a computer-controlled 2D translating stage (XSlide, VelMex Inc., Newton, NJ) and immersed in a custom-built water tank to calibrate the output pressure of a 2.25 MHz single element transducer (V323-SU, Olympus NDT). 6-cycles of square waves at 1.85 MHz with peak input voltage of 5 V were launched every 1 ms (pulse repetition frequency (PRF) of 10 kHz) to the transducer. The hydrophone was placed one Rayleigh distance from the transducer (8.9 mm).

The pulse intensity integral (PII) is defined as: $PII = \int \frac{p^2(t)}{Z_0} dt$ where p is the instantaneous peak pressure, Z_0 is the characteristic acoustic impedance of the medium. In the case of water, Z_0 is estimated to be 1.5 MRayl.

The I_{SPPA} is defined as: $I_{SPPA} = \frac{PII}{PD}$, where PD is the pulse duration defined as $(t)(0.9 \cdot PII - 0.1 \cdot PII) \cdot 1.25$ as outlined by the standards established by NEMA (American Institute of Ultrasound in Medicine, 1998).

The I_{SPTA} is defined as: $I_{SPTA} = PII \cdot PRF$ where PRF is the pulse repetition frequency.

The MI is defined as: $MI = \frac{p_r}{\sqrt{f}}$ where p_r is the peak rarefaction pressure and f is the acoustic frequency.

4.3 Results

4.3.1 Commercially-available components can be assembled into mm-scale recording implants

The assembly process (Fig. 4.1) shows a neural dust implant mote integrated on a 50 µm thick polyimide flexible printed circuit board (PCB) where both the piezocrystal (0.75 mm x 0.75 mm x 0.75 mm) and the custom transistor (0.5 mm x 0.45 mm) are attached to the topside of the board with a conductive silver paste. Electrical connections between the components are made using aluminum wirebonds and conductive gold traces. Exposed gold recording pads on the bottom of the board (0.2 mm x 0.2 mm) are separated by 1.8 mm and make contact on the nerve or muscle to record

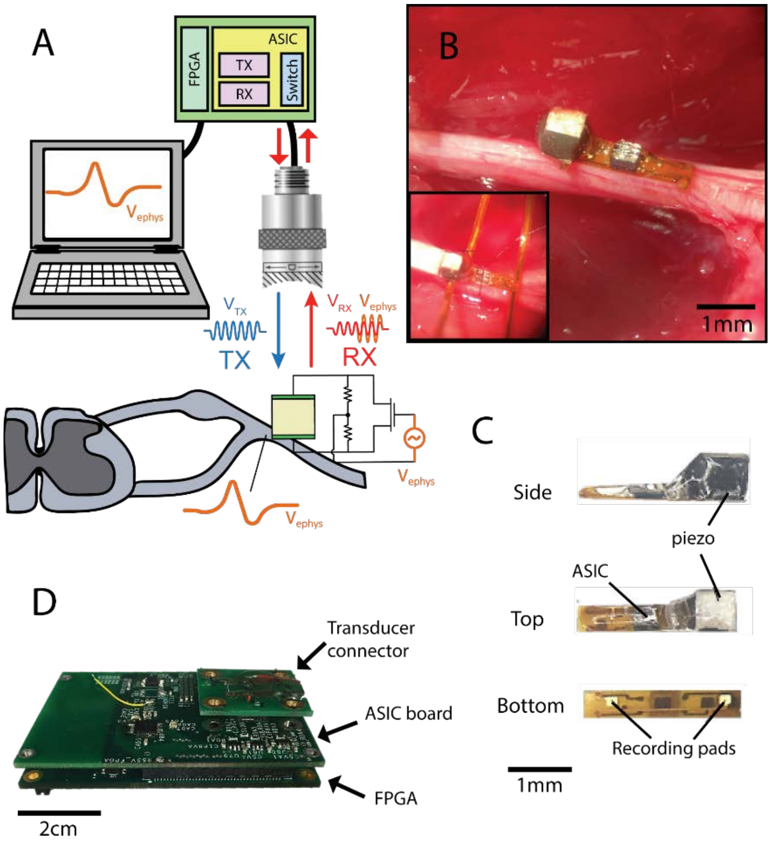


Figure 4.6:* *a: *An external transducer powers and communicates with a neural dust mote placed remotely in the body. Driven by a custom transceiver board, the transducer alternates between transmitting a series of pulses that power the device and listening for reflected pulses that are modulated by electrophysiological signals.* ***b:*** *A neural dust mote anchored to the sciatic nerve in an anesthetized rat. Inset shows neural dust mote with optional testing leads.* ***c:*** *Components of a neural dust mote. The devices were assembled on a flexible PCB and consist of a piezoelectric crystal, a single custom transistor, and a pair of recording electrodes.* ***d:*** *The transceiver board consisted of Opal Kelly FPGA board, ASIC board (Seo et al., 2015; Tang et al., 2015), and the transducer connector board.*

electrophysiological signals (Fig. 4.6c). Recorded signals are sent to the transistor's input through micro-vias. Additionally, some implants were equipped with 0.35 mm-wide, 25 mm-long, flexible, compliant leads (Fig. 4.1, bottom) with test points for simultaneous measurement of both the voltage across the piezocrystal and direct wired measurement of the extracellular potential across the electrode pair used by the mote (we refer to this direct, wired recording of extracellular potential as the *ground truth measurement* below, which is used as a control for the ultrasonically reconstructed data). The entire implant is encapsulated in a medical grade UV-curable epoxy to protect wirebonds and provide insulation. A single neural dust mote implant measures roughly 0.8 mm x 3 mm x 1 mm (Fig. 4.6c and 4.1). The size of the implants presented here is limited only by our use of commercial polyimide backplane technology, which is commercially accessible to anyone; relying on more aggressive assembly techniques with in-house polymer patterning would produce implants not much larger than the piezocrystal dimensions (yielding a ~1 mm^3 implant).

4.3.2 A custom integrated circuit operates the external transceiver board and enables low-noise interrogation

An external, ultrasonic transceiver board (Fig. 4.6d) interfaces with neural dust motes by both supplying power (transmit (TX) mode) and receiving reflected signals (receive (RX) mode). This system is a low-power, programmable, and portable transceiver board that drives a commercially available external ultrasonic transducer (V323-SU, Olympus, Waltham, MA). Details of the custom integrated circuit (IC) that drove the external ultrasonic transducer with high energy-efficiency were presented elsewhere (Seo et al., 2015; Tang et al., 2015). The transceiver board exhibited a de-rated focus at ~8.9 mm (Fig. 4.7a). The XY cross-sectional beam-pattern clearly demonstrated the transition from the near-field to far-field propagation of the beam, with the narrowest beam at the Rayleigh distance (Fig. 4.7b). The transducer was driven with a 5 V peak-to-peak voltage signal at 1.85 MHz. The measured de-rated peak rarefaction pressure was 14 kPa, resulting in a mechanical index (MI) of 0.01. De-rated spatial pulse peak average (I_{SPPA}) and spatial peak time average (I_{SPTA}) of 6.37 mW/cm^2 and 0.21 mW/cm^2 at 10 kHz pulse repetition were 0.0034% and 0.03% of the FDA regulatory limit, respectively (US Food and Drug Administration, 2008). The transceiver board was capable of outputting up to 32 V peak-to-peak and the output pressure increased linearly with the input voltage (Fig. 4.7c).

4.3.3 Reflections from non-piezocrystal interfaces provide a built-in reference for movement artifacts and temperature drift

The entire system was submerged and characterized in a custom-built water tank with manual 6 degrees-of-freedom (DOF) linear translational and rotational stages (Thorlabs Inc., Newton, NJ). Distilled water was used as a propagation medium, which exhibits similar acoustic impedance as tissue, at 1.5 MRayls (Kino, 1987). For initial calibration of the system, a current source (2400-LV, Keithley, Cleveland, OH) was used to mimic

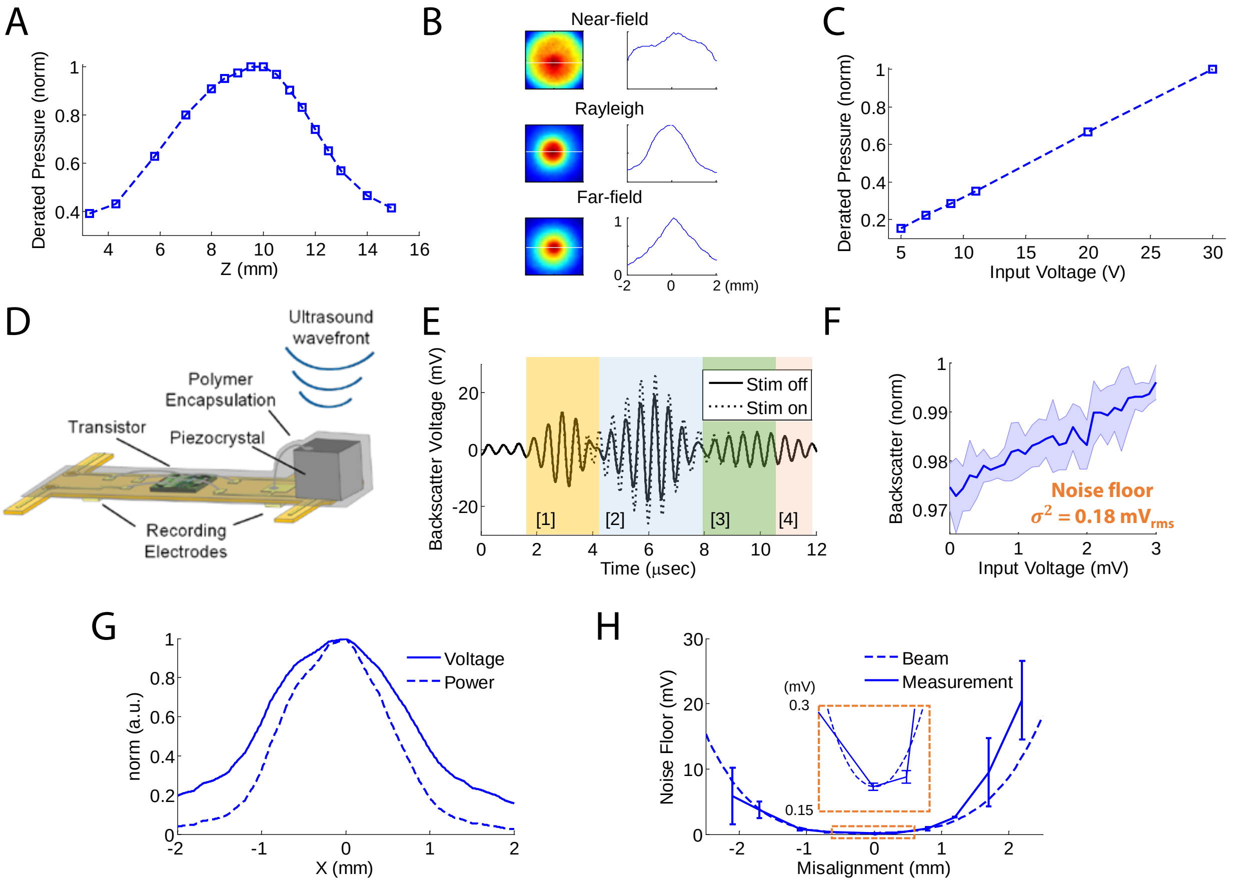


Figure 4.7: ***a***: *A de-rated, normalized peak pressure as a function of distance from the surface of the transducer showed a de-rated focus at ~8.9 mm at 1.85 MHz.* ***b***: *The XY cross-sectional beampatterns and the corresponding 1-D voltage plot at y = 0 at near-field, Rayleigh distance, and far-field showed beam focusing at the Rayleigh distance.* ***c:*** *The transducer's output pressure was a linear function of input voltage (up to 32 V peak-to-peak).* ***d:*** *Cross-section of the neural dust mote.* ***e:*** *Example backscatter waveform showing different regions of backscatter. The backscatter waveform is found flanked (in time) by regions which correspond to reflections arising from non-responsive regions; these correspond to reflected pulses from other device components shown in D. The measurement from the non-responsive regions, which do not encode biological data) can be used as a reference. As a result of taking this differential measurement, any movements of the entire structure relative to the external transducer during the experiment can be subtracted out.* ***f:*** *Calibration curve obtained in the custom water tank setup showed the noise floor of 0.18 mVrms.* ***g:*** *1-D plot of the transducer's off-axis voltage and power drop-off at y = 0 at Rayleigh distance.* ***h:*** *The effect of noise floor as a function of lateral misalignment followed the beampattern power fall-off.*

extracellular signals by forcing electrical current at varying current densities through 0.127 mm thick platinum wires (773000, A-M Systems, Sequim, WA) immersed in the tank. The neural dust mote was submerged in the current path between the electrodes. As current was applied between the wires, a potential difference arose across the implant electrodes. This potential difference was used to mimic extracellular electrophysiological signals during tank testing.

To interrogate the neural dust mote, six 540 ns pulses every 100 µs were emitted by the external transducer (Fig. 4.8). These emitted pulses reflect off the neural dust mote and produce backscatter pulses back towards the external transducer. Reflected backscatter pulses were recorded by the same transceiver board (Fig. 4.6a and d). The received backscatter waveform exhibits four regions of interest; these are pulses reflecting from four distinct interfaces (Fig. 4.7d,e): 1) the water-polymer encapsulation boundary, 2) the top surface of the piezoelectric crystal, 3) the piezo-PCB boundary, and 4) the back of the PCB. As expected, the backscatter amplitude of the signals reflected from the piezoelectric crystal (second region) changed as a function of changes in potential at the recording electrodes. Reflected pulses from other interfaces did not respond to changes in potential at the recording electrodes. Importantly, pulses

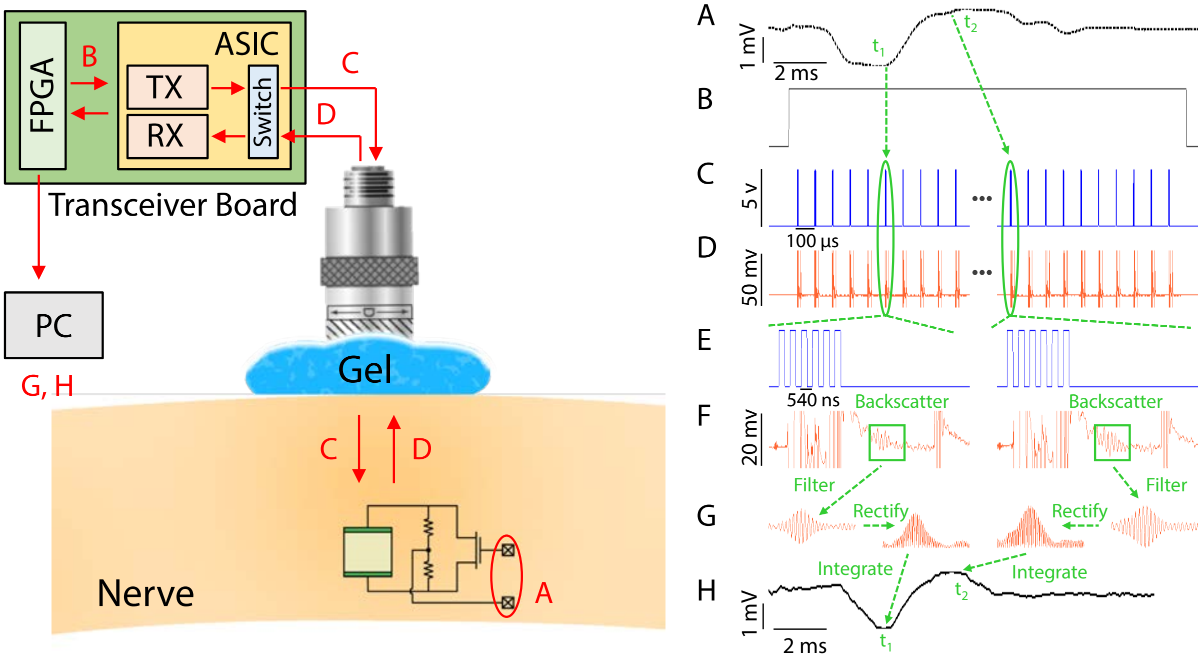


Figure 4.8: (left) schematic flow of information; (right) representative time traces of signals at each step (referenced to the left diagram). The sequence is shown on the right for two time points, t_1 and t_2.

a: An extracellular, electrophysiological potential is presented to the recording electrodes on a neural dust mote. **b:** The FPGA generates a trigger signal to initiate recording. **c:** Upon receiving the trigger from the FPGA, the transceiver board generates a series of transmit pulses. At the end of the transmit cycle, the switch on the ASIC disconnects the transmit module and connects the receive module. **d:** Backscatter from the neural dust mote reaches the transducer approximately $2t_{Rayleigh}$. **e:** Zoomed-in transmit pulse sequence, showing 6 pulses at 1.85 MHz. **f:** Zoomed-in backscatter waveforms, aligned in time with **e.** Note the large, saturating signal which overlaps with the transmitted pulses is electrical feedthrough and is ignored. The returning, backscattered pulses can be seen subsequent to the transmission window (green box). A close-up of the backscatter pulses is shown in **Fig. 4.7e** and discussed in the text. **g:** These backscattered pulses are filtered, rectified, and the area under the curve is computed in order to produce reconstructed waveforms. **h:** Reconstructed waveform is sampled at 10 kHz. Each point of the reconstructed waveform is computed by calculating the area under the curve of the appropriate reflected pulses, received every 100 µs.

from the other non-responsive regions were used as a signal level reference, making the system robust to motion or heat-induced artifacts (since pulses reflected from all interfaces change with physical or thermal disturbances of the neural dust mote but only pulses from the second region change as a function of electrophysiological signals). In a water tank, the system showed a linear response to changes in recording electrode potential and a noise floor of ~0.18 mV_{rms} (Fig. 4.7f). The overall dynamic range of the system is limited by the input range of the transistor and is greater than >500 mV (i.e., there is only an incremental change in the current once the transistor is fully on (input exceeds its threshold voltage) or fully off). The noise floor increased with the measured power drop-off of the beam; 0.7 mm of misalignment degraded it by a factor of two (N = 5 devices, Fig. 4.7h). This lateral mis-alignment-induced increase in the noise floor constitutes the most significant challenge to neural recordings without a beam-steering system (that is, without the use of an external transducer array that can keep the ultrasonic beam focused on the implanted dust mote and, thus, on-axis). On axis, the neural dust mote converted incident acoustic power to electrical power across the load resistance of the piezo with ~25% efficiency. Fig. 4.7g plots the off-axis drop-off of voltage and power at one Rayleigh distance for the transducer used in this manuscript. Likewise, Fig. 4.7h plots the change in effective noise floor as a function of angular misalignment.

4.3.4 EMG and ENG can be recorded tetherlessly *in-vivo* in rodents

We recorded evoked EMG responses from the gastrocnemius muscle of adult Long-Evans rats under anesthesia using the neural dust system. The mote was placed on the exposed muscle surface, the skin and surrounding connective tissue were then replaced, and the wound was closed with surgical suture (Fig. 4.9a).

The ultrasonic transducer was positioned 8.9 mm away from the implant (one Rayleigh distance of the external transducer) and commercial ultrasound gel (Aquasonic 100, Parker Labs, Fairfield, NJ) was used to enhance coupling. The system was aligned using a manual manipulator by maximizing the harvested voltage on the piezocrystal measured from the flexible leads. Ag/AgCl wire hook electrodes were placed approximately 2 cm distally on the trunk of the sciatic nerve for the bulk stimulation of muscle fiber responses. Stimulation pulses of 200 µs duration were applied every 6 seconds and data was recorded for 20 ms around the stimulation window (Fig. 4.9b). The power spectral density (PSD) of the reconstructed data with several harmonics due to edges in the waveform is shown in Fig. 4.9c. This process could be continued indefinitely, within the limit of the anesthesia protocol; a comparison of data taken after 30 minutes of continuous recording showed no appreciable degradation in recording quality (Fig. 4.9d).

We obtained EMG recruitment curves with both ground truth and wireless dust backscatter by varying stimulation amplitude (Fig. 4.10a and 4.10b). Reconstruction of the EMG signal from the wireless backscatter data was sampled at 10 kHz, while the

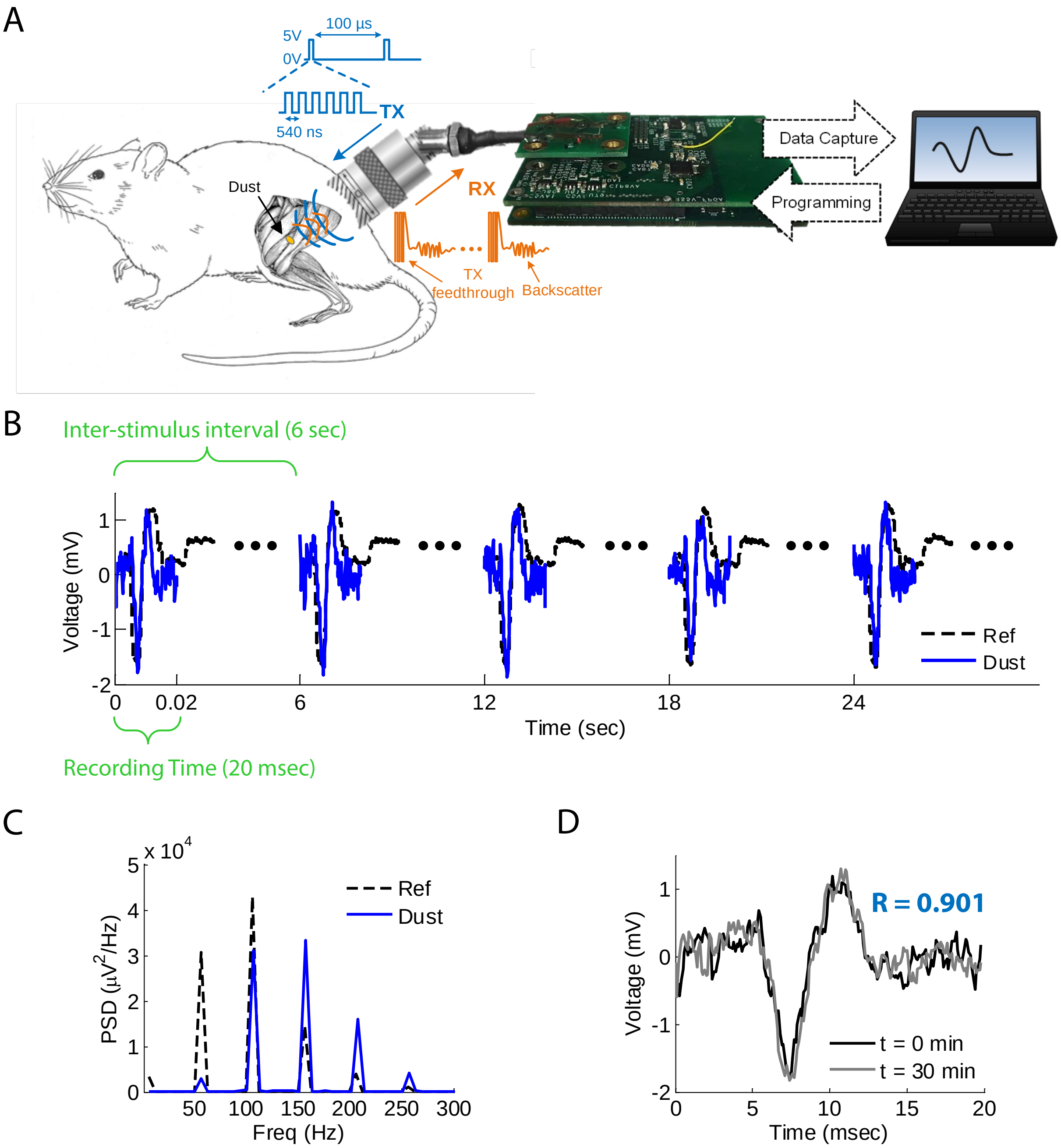


Figure 4.9: a: *In-vivo experimental setup for EMG recording from gastrocnemius muscle in rats; the neural dust mote was placed on the exposed muscle surface and the wound was closed with surgical suture. The external transducer couples ultrasound to the mote and the wireless data is recorded and displayed on the laptop.* ***b:*** *Comparison between ground truth measurement and the reconstructed EMG signals over a number of trials. 20 msec samples were recorded and the inter-stimulus interval was 6 sec.* ***c:*** *Power spectral density (PSD) of the recorded EMG signal showed 4.29e4 µV2/Hz and 3.11e4 µV2/Hz at 107 Hz for ground truth and the reconstructed dust data, respectively, and several harmonics due to edges in the waveform.* ***d:*** *The wireless backscatter data recorded at t = 0 min and t = 30 min matched with R = 0.901.*

wired, ground truth measurement was sampled at 100 kHz with a noise floor of 0.02 mV. The two signals at response-saturating stimulation amplitude (100%) matched with R = 0.795 (Fig. 4.10c). The difference between the wireless and wired data was within ± 0.4 mV (Fig. 4.10d). The salient feature of the neural dust mote EMG response was approximately 1 ms narrower than the ground truth, which caused the largest error in the difference plot (Fig. 4.10c and 4.10d). The responses from skeletal muscle fibers

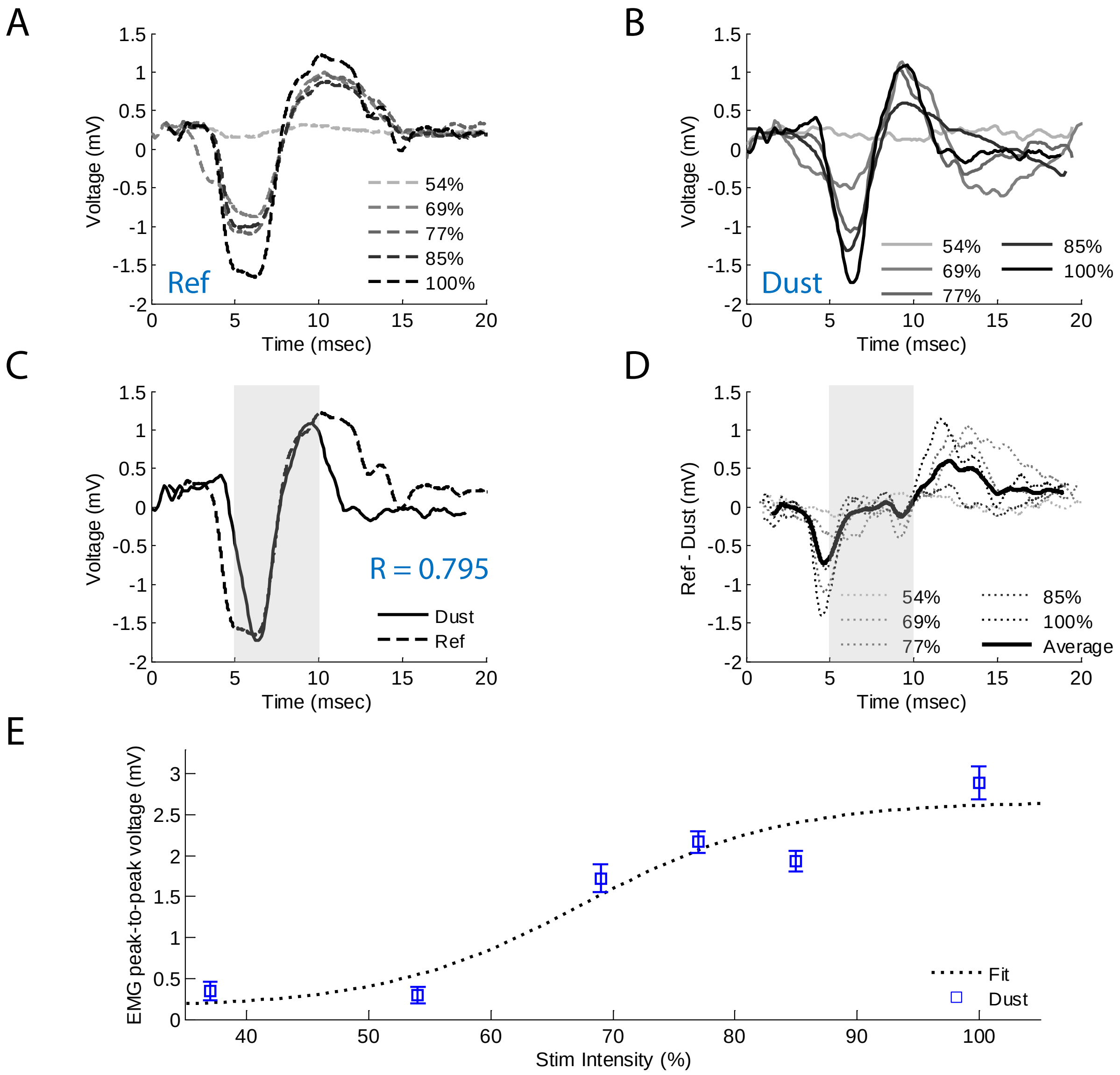


Figure 4.10: ***a:*** *Different intensities of EMG signals were recorded in-vivo with the electrodes on the PCB with varying stimulation intensities.* ***b:*** *Similar gradient EMG responses were recorded wirelessly with the mote.* ***c:*** *Ground truth and reconstruction of EMG signal from the wireless backscatter data at response-saturating stimulation amplitude (100%) matched with R = 0.795 (R = 0.60, 0.64, 0.67, 0.92 for 54%, 69%, 77%, 89%, respectively).* ***d:*** *Quantitative comparison showed < 0.4 mV match of the salient feature (shaded regions). E. EMG peak-to-peak voltage showed an expected sigmoidal relationship with the stimulation intensity.*

occurred 5 ms post-stimulation and persisted for 5 ms. The peak-to-peak voltage of the EMG shows a sigmoidal response as a function of stimulation intensity (Fig. 4.10e) as expected (Gruner and Mason, 1989). The error bars indicate the measurement uncertainties from two rats and 10 samples each per stimulation amplitude. The minimum signal detected by the neural dust mote is approximately 0.25 mV, which is in good agreement with the noise floor measurement made in a water tank.

A similar setup was prepared to measure the electroneurogram (ENG) response from the main branch of the sciatic nerve in anesthetized rats. The sciatic nerve was exposed

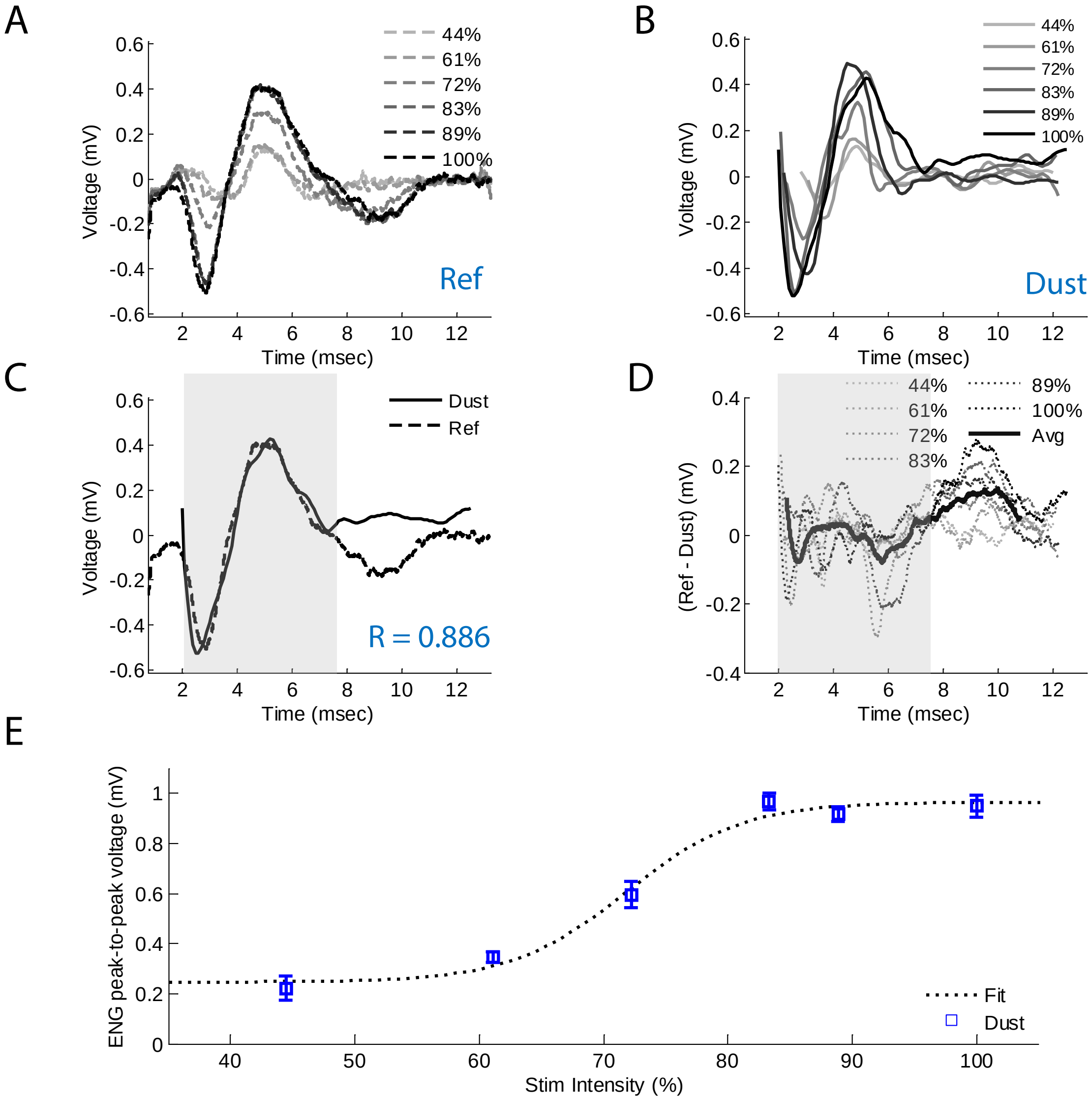


Figure 4.11: ***a:*** *A. Different intensities of ENG signals were recorded in-vivo with the electrodes on the PCB with varying stimulation intensities.* ***b:*** *Similar gradient ENG responses were recorded wirelessly with the mote.* ***c:*** *Ground truth and reconstruction of ENG signal from the wireless backscatter data at response-saturating stimulation amplitude (100%) matched with R = 0.886 (R = 0.822, 0.821, 0.69, 0.918, 0.87 for 44%, 61%, 72%, 83%, 89%, respectively).* ***d:*** *Quantitative comparison showed < 0.2 mV match of the salient feature (shaded regions).* ***e:*** *ENG peak-to-peak voltage showed an expected sigmoidal relationship with the stimulation intensity.*

by separating the hamstring muscles and the neural dust mote was placed and sutured to the nerve, with the recording electrodes making contact with the epineurium (Fig. 4.6b). We measured a similar graded response on both ground truth and wireless dust backscatter by varying stimulation current amplitude delivered to bipolar stainless steel electrodes placed in the foot (Fig. 4.11a and 4.11b). The two signals at response-saturating stimulation amplitude (100%) matched with R = 0.886 (Fig. 4.11c); the average error was within ± 0.2 mV (Fig. 4.11d). The peak-to-peak ENG voltage showed a sigmoidal response with the error bars indicating uncertainties from two rats and 10 samples each per stimulation amplitude. The minimum signal detected by the neural dust mote was again at 0.25 mV (Fig. 4.11e).

4.4 Discussion

In recent years, there has been growing interest in the use of neural recording technologies to improve neurostimulation-based treatments as well as to develop new closed-loop neuromodulation therapies for disorders in the central (Krook-Magnuson et al., 2015) and peripheral (Famm et al., 2013) nervous systems. Because nerves carry both efferent and afferent signals to a variety of target organs, effective recording technologies will need high spatiotemporal resolution to record from multiple discrete sites within a single nerve. In order for these technologies to become clinically viable they will need to be tether-less to avoid potential infections and adverse biological responses due to micro-motion of the implant within the tissue.

To address this looming issue, we designed, built, and implanted a wireless, ultrasonic neural sensor and communication system that enables neural recordings in the peripheral nervous system. *In-vivo*, acute recordings in a stationary, anaesthetized rat model were used to collect compound action potentials from the main branch of the sciatic nerve as well as evoked EMG from the gastrocnemius muscle. The performance of the neural dust system was equivalent to conventional electrophysiological recordings employing microelectrodes and cabled electronics.

One of the principal strengths of the demonstrated technology is that, unlike conventional radio frequency technology, ultrasound-based systems appear scalable down to <100 µm sizes (see *Size Scaling and Electromagnetics* below), opening the door to a new technological path in implantable electronics. A complete analysis of this scaling can be found in (Seo et al., 2013, 2015). In brief, physics limits how small a good radio frequency receiver can be due to the long wavelengths of radio frequency energy (millimeters to centimeters) and the high degree of absorption of radio frequency energy into tissue (which heats up the tissue and limits the total power than can be sent to an implant). Ultrasonic systems fare much better in both areas, allowing for the design of extremely small receiver devices. In addition, the extreme miniaturization of lower power electronics allows for useful recording electronics to be incorporated into such small packages.

A number of technical challenges remain open. The power levels used in this study were limited by the specifications of commercially-available transducers; custom transducers will reduce the overall external device footprint, lower the noise floor (by producing higher power densities at the focal spot), and allow for selection of the focal depth to suit specific applications. For example, a flat, low-profile piezo-transducer with proper impedance matching would enable a wearable neural dust transceiver board small enough for awake, behaving rodent neurophysiology. Additionally, the development of wearable, battery-powered multi-element arrays would allow for beam-steering of the ultrasonic beam, with several advantages: 1) motes could be maintained on-axis even in the face of relative motion between mote and external transducer, which is the most significant challenge of the present work; 2) multiple motes could potentially be interrogated by sweeping the focused beam electronically; 3) post-surgical tuning of mote location would be made easier. Additional de-noising of the transceiver drive electronics should also help decrease the noise floor (see *Methods*). The modifications above are all well-within current state of the art; with others, we have recently shown theoretical and experimental advantages to using beam-forming systems (Bertrand et al., 2014; Seo et al., 2015).

In addition, the calculated scaling predictions suggest that <500 µm scale motes are feasible. To do this, a number of material and microfabrication challenges exist, including the use of microfabricated backplanes, solder microbumping assembly of components (instead of the conventional wirebonding approach used here) and the use of thin film encapsulants (instead of medical grade epoxy) such as parylene. Transitioning away from PZT piezocrystals to biocompatible $BaTiO_3$ single crystal transducers is also underway; taken together, these developments would open the way for chronic studies of neural dust recording.

Lastly, as this platform presents a generalized power delivery system, the design and fabrication of neural stimulation systems based on charge-delivery through electrodes on the dust motes is also underway.

4.4.1 Size Scaling and Electromagnetics

The most popular existing wireless transcutaneous energy transfer technique relies on electromagnetics (EM) as the energy modality. An external transmitter generates and transfers information through purely electric or magnetic near field or electromagnetic far field coupling; this energy can be harvested by the implanted device and converted into a stable DC supply voltage. Energy transmission via magnetic near field has been used in a wide variety of medical applications (e.g. cochlear implants). As EM requires no moving parts or the need for chemical processing or temperature gradients, it is considered more robust and stable than other forms energy scavenging. When used in-body, however, EM coupling power density is restricted by the potential adverse health effects associated with excess tissue heating in the vicinity of the human body due to electromagnetic fields. This is regulated by the well-known FCC and IEEErecommended levels. Roughly, the upper limit for EM power density transiting through tissue is set by the minimum required to heat a model sample of human tissue by 1°C. For electromagnetic waves, the output power density is frequency dependent and cannot exceed a maximum of 10 mW/cm^2. Consider, in this context, the problem of transmitting EM power to (and information from) very small CMOS chiplets embedded in tissue; does this approach scale to allow high density neural recordings? Regardless of the specific implementation, any such chiplet will contain a resonant component that couples to the EM waves; such a system can be modeled as a series/parallel RLC (for the purposes of this exercise, one may presume that a suitable method exists for modulating the quality factor or mutual coupling of the RLC as a function of neural activity for wireless communication). Given this, the performance of electromagnetic power transfer suffers from two fundamental issues. First, the extreme constraint on the size of the node limits the maximum achievable values of the passives. Assuming a 3-turn planar square loop inductor with 1 oz. copper wire and capacitor density of 1 fF/µm^2, calculations to maximize the link efficiency using equations derived in (Harrison, 2007; Jow and Ghovanloo, 2007) predict the resonant frequency of a 100 µm neural dust would be ~10 GHz.

Fig. 4.12a plots the modeled channel loss, or the attenuation of the EM signal as it propagates through 8.9 mm of tissue, due to tissue absorption and beam spreading, as a function of frequency. We observe that there is an exponential relationship between the propagation loss and the frequency, and at 10 GHz – the total combined loss for one-way transmission is approximately 50 dB. Moreover, at these very small footprints (compared to the wavelength, which is in the millimeter range), the receive antenna efficiency becomes quite small, thereby easily adding roughly 20 dB of additional loss. The tissue absorption loss penalty incurred by operating at high frequency can be reduced by increasing the capacitance density using 3D inter-digitized capacitor layouts, but even then, as shown in Fig. 4.12b, scaling down the dimensions of the motes increases the resonant frequency of the link, causing an exponential increase in the tissue absorption loss and the overall channel loss, and the efficiency of EM transmission becomes miniscule.

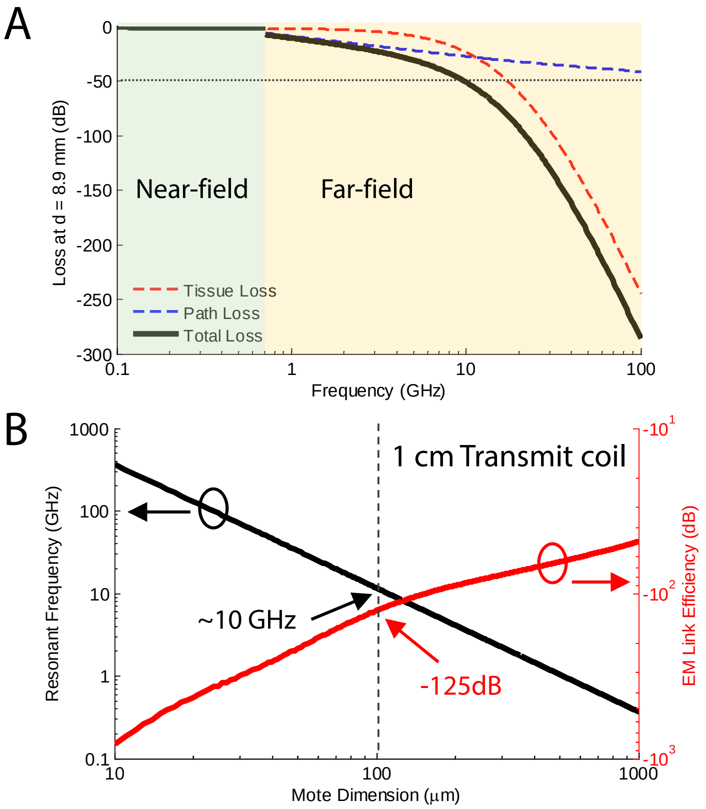


***Figure 4.12:* *a:* Total propagation loss through 8.9 mm of tissue, due to both tissue and path loss, increases exponentially with frequency, resulting in approximately 50 dB of loss at 10 GHz. *b:* The mutual coupling, and therefore link efficiency, also reduces dramatically with the scaling of mote dimensions. Note that the transmit coil is assumed to be 1 cm in diameter, which is equivalent to the diameter of the ultrasonic transducer surface used in the manuscript.*

Ultrasound transmission performs much better at the same size scale. Our experimental data confirms this, as can be seen from the data plotted in Fig. 8 of Seo et al., 2014. There, dust motes were manufactured with crystals ranging from ~1 mm down to 125 µm and the power delivery efficiency and backscatter sensitivity were measured in a water tank. Comparing data from Fig. 8 of that paper with Fig. 4.12b, we note that for a 100 µm mote for transmission through 8.9 mm of tissue, the single path loss of EM is 125 dB (or ~3e-13 efficiency) vs. ~1e-5 for ultrasound (or 7 orders of magnitude better).

In addition, we may ask just how small an ultrasonically-powered implant could be made and still reasonably be expected to function. Our earlier theoretical work (Seo et al., 2013) matches that above results closely and, in that work, we show that a dust mote embedded 2 mm into the brain tissue would not scale well below 50 µm in size. This is because unlike conventional penetrating recording shanks (which measure electrical potential at each recording site in relation to a common electrode which is placed relatively far away – centimeters – from the recording sites), with neural dust motes both the recording and the common electrode must be placed within the same (very small) dust mote. Thus, the distance between electrodes and therefore, the maximum differential signal between the electrodes are inherently limited by the neural dust footprint size, and follow the dipole-dipole voltage characteristic that decreases at least quadratically (unless very near a cell body, in which case it appears to scale exponentially; see (Gold et al., 2007) for a more thorough review) with increasing separation distance. Since the power available to the implant has a fixed upper bound, the reduction of extracellular potential amplitude as the neural dust dimensions are scaled down in the presence of biological, thermal, electronic, and mechanical noise (which do not scale), causes the signal-to-noise (SNR) ratio to degrade significantly. At a separation distance of 100 µm between recording electrodes, we expect a 10 µV AP amplitude (data derived from (Du et al., 2011)), with the amplitude further reducing quadratically as the separation is reduced.

Since the power available to the neural dust is limited, the design goal of a front-end architecture is to minimize the input-referred noise within this power budget. The power efficiency factor (NEF^2 x Vdd) quantifies the tradeoff between power and noise (Bjorninen et al., 2012) and extrapolating from the measurement result of a recent CMOS neural front-end design (NEF^2 x Vdd of 11.3 (Muller et al., 2015), we can estimate the relationship between the input-referred noise level and the DC power consumption of an optimally designed front-end architecture as we scale. The fundamental limit to the NEF^2 x Vdd occurs at a supply voltage of at least ~4 kBT/q or 100 mV, in order to reliably operate the FET, and by definition, the NEF of 1 for a single BJT amplifier (Steyaert and Sansen, 1987). Fixing the minimum input SNR to 0 dB for extracting neural signals, we can evaluate the scaling capability of neural dust as shown in Fig 4.13. The point of intersection in Fig. 4.13 denotes the minimum size of neural dust that enables the operation of the complete link. For transmission through 8.9 mm of tissue (as is the case in this manuscript), this occurs approximately at 90 µm. This effectively means that, staying within FDA-approved ultrasound power limits, assuming an SNR of 0 dB is required, neural dust motes smaller than 90 µm cannot receive enough power to distinguish neural activity from noise.

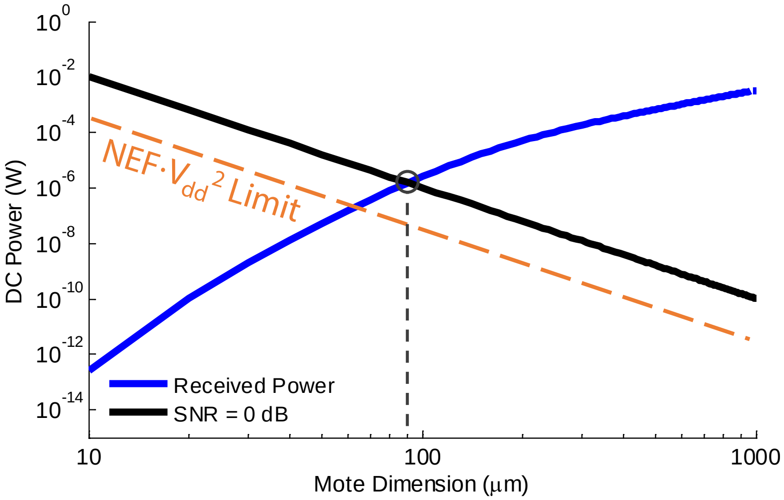


Figure 4.13: As we scale down the neural dust size, more power is needed to keep the noise floor down to maintain SNR while less power is captured. The intersection of these two trends is the smallest node that will still operate. Scaling with an SNR of 0 dB shows operation down to 90 µm for transmission distance of 8.9 mm.

Chapter 5:

Conclusions and open questions

As our understanding of the nervous system expands in tandem with the increasing pace of technological innovation, the capacity to seamlessly merge the brain and nerves with electronic devices will only increase. Technological innovation can provide new methods and devices to improve reliability, resolution, and efficacy of BMI systems and bioelectronic medicines. However, intelligent neural interface design requires a strong, foundational understanding of neurobiology. Efforts to better map the contributions of the nervous system to voluntary behaviors, such as speech and motion, as well as involuntary activity, such as gut motility and organ function, will inform the design of future devices to alter or correct nervous system function. For example, understanding the circuits and mechanisms behind how the brain learns to control novel neuroprosthetics can inform the design of adaptive decoders that learn in concert with the brain (Orsborn et al., 2014). These “top-down” approaches help to focus the direction and specificity of neural interface design to ensure that the resulting devices are best suited for the target application. On the other hand, “bottom-up” approaches to designing neural interfaces that begin by setting aggressive performance targets based on fundamental limitations can provide new tools that enable new insights into biological function. The point at which these two approaches meet will likely be the cutting edge of new innovation.

This dissertation was written with the goal of balancing both investigation of neurobiological functions, using brain-machine interfaces as a tool to study learning in neural circuits, as well as technological innovation to improve the tools that make such investigations possible. A summary of the contributions to the field appear below in Chapter 5.1. Finally, although this work has yielded insights into both neural function and possibilities for the design of next-generation neural interfaces, many open questions still remain; these are summarized in Chapter 5.2.

5.1 Summary of contributions

One of the primary motivations of this work was to better understand how the brain learns to control its own neural activity in behaviorally constructive ways. In **Chapter 2**, we investigated a possible circuit-level mechanism by which cortical neurons learn to modulate their activity to contribute to rewarding behavioral outcomes. Using a BMI paradigm that transformed neural activity in the pitch of an auditory tone, we showed that rats could learn to produce arbitrary modulations of primary visual cortex (V1) neurons, both in the presence or absence of light. We demonstrated that performance on this task was goal-directed and intentional using contingency degradation. Then, we found that this learning process was accompanied by cell-specific changes in neural dynamics, including spike-field coherence and correlations between neurons directly responsible for cursor control. These results mirror observations of learning in the primary motor cortex reported elsewhere (Clancy et al., 2014; Koralek et al., 2012, 2013).

In the second half of Chapter 2, we reported that with training, activity in the dorsomedial striatum (DMS) became increasingly modulated and coherent with learned activity in V1. To test the relevance of this observation, we injected mice with a viral vector encoding the red-shifted inhibitory opsin, Jaws. Injections were targeted to the region of DMS that receives projections from V1. We demonstrated that GFP-injected (control) mice were able to learn the same BMI task as was implemented in rats. Then, using the Jaws-infected animals, we silenced activity in DMS on approximately 50% of all trials during training sessions. We observed that inhibition of DMS prevented learning to modulate activity in V1, but that DMS inhibition after learning had taken place did not affect performance. Given that the striatum is the primary input nucleus of the basal ganglia, these results argue for a necessary role for cortico-basal ganglia circuits in learning to modulate cortical activity in behaviorally relevant ways.

After documenting an important role for the striatum in goal-directed instrumental behavior, in **Chapter 3** we focused on the representation of task parameters in the dorsomedial striatum. We trained rats to perform a probabilistic switching task that involved choosing between two levers. We quantified animals' behavior and observed that with learning, animals adopted patterns of responding that allowed them to exploit the hidden structure of the task. Then, we showed that behavior was best fit by a hidden markov model (HMM), and used models of behavior to compute a measure of confidence. This measurement, which was intended to estimate animals' confidence in the current state of the task, closely tracked behavioral outcomes from trial-to-trial, as well as the likelihood of an animal to explore alternative actions.

Then, we analyzed the task-relevant components of neural activity as animals learned and performed the switching task. Many single neurons in DMS encoded one or more task parameters, including task state, action choice, and the outcome of a trial. We then utilized de-mixed principal components analysis (dPCA) to extract task-relevant neural dynamics from the full population of recorded DMS units. We observed that population activity robustly captured task-relevant information, and could accurately classify different aspects of animals' choice. In order to achieve trial-by-trial resolution of the encoding of confidence by population activity, we employed canonical polyadic tensor decomposition on simultaneously recorded data within single training sessions. We observed the existence of trial factors that robustly correlated with fluctuations in animals' confidence level on a trial-by-trial basis. Given that confidence signals could be decoded from neural activity, we then investigated the influence of confidence on action-predictive decision variables that could also be decoded from population activity. We found that high confidence levels biased choice signals in DMS to be larger in magnitude and appear earlier in a trial. Conversely, weak confidence levels resulted in lower-magnitude choice signals that only appeared immediately prior to an animals' choice. These data demonstrate a rich encoding of task parameters in DMS that may influence the selection of appropriate behaviors according to abstract predictions about the environment.

In **Chapter 4**, we demonstrated a novel approach for recording neural activity *in vivo*. We designed, built, and tested an implantable, mm-scale wireless ultrasound system, known as neural dust, to record electroneurogram (ENG) and electromyogram (EMG) activity in anesthetized rats. These devices utilized ultrasonic backscatter to encode the neural activity of nerves and muscles. We used an off-the-shelf ultrasonic transducer to communicate and power neural dust motes and demonstrated that the system could operate efficiently at power levels well within established safety limits for ultrasonic devices. Then, we demonstrated that implanted neural dust motes could reliably record ENG and EMG activity relative to conventional wired electronics. We show that ultrasonic backscatter is a safe, scalable method for powering and communicating with miniaturized devices implanted deep in tissue. These results demonstrate the potential for ultrasonic implantables to overcome many of the limitations facing state-of-the-art neural interfaces.

5.1.1 Additional engineering collaborations

Beyond the results presented in this dissertation, the work reported here also resulted in additional fruitful engineering collaborations. Although they are not covered in detail in the preceding chapters, they will be briefly reported below.

Prior to beginning work on in vivo testing of the neural dust system, we validated the performance of a wireless, low-power system-on-chip (SoC) for neural recording and stimulation. This 65 nm CMOS SoC had a 4.78mm^2 footprint, and consumed 348 µA from an unregulated 1.2 V to 1.8 V power source. The system operated 64 data acquisition channels with compression and was capable of simultaneously engaging two stimulation channels. At the time of initial reporting, this device was the lowest area and power for the highest integration complexity to date.

As mentioned in Chapter 1, a common failure mode of brain-machine interfaces is breakdown of encapsulation material or the foreign body response caused by implantation. Silicon carbide is extremely stable and biocompatible, and therefore a strong candidate for fabrication of neural interfaces. We presented *in vivo* validation of a novel method for fabricating electrode arrays that integrated polycrystalline conducting silicon carbide with insulating silicon carbide. The result of this process was a seamless transition between doped and amorphous silicon carbide at recording sites, avoiding heterogenous interfaces that are a common site of degradation. We successfully demonstrated high-fidelity recording of sciatic nerve ENG and electrocorticography in the visual cortex from anesthetized rats. This novel fabrication process represents an exciting new avenue for neural interfaces designed to last years to decades.

5.2 Open questions and future directions

5.2.1 The source of voluntary control signals

In Chapter 2, we observed that animals could learn to volitionally modulate the activity of V1 neurons in total darkness. Where, then, is the source of modulatory input? Previous reports have demonstrated motor-related activity in V1 (Niell and Stryker, 2010). In addition, attention is known to modulate neural activity in V1 (Gregoriou et al., 2009; Martinez et al., 1999; Zhang et al., 2014), suggesting that the learning we observe might be a result of top-down inputs from other areas of cortex. Previous reports have shown that learning enhances top-down cortical influences in V1 while reducing the impact of visually-driven activity (Makino and Komiyama, 2015). It seems possible that these same circuit mechanisms are in play. A broader understanding of this process would contribute greatly to knowledge of how instrumental learning proceeds in the cortex.

5.2.2 The role of DMS in learned V1 modulations

The primary visual cortex has a well-established connection to the dorsomedial striatum (Khibnik et al., 2014). However, we know of no reports of direct projections from the output nuclei of the basal ganglia that return to neurons in the primary visual cortex. How the basal ganglia, which consists of multiple unique nuclei and several converging pathways, contributes to this learning process remains unknown. Dissecting the multi-synaptic pathways through the basal ganglia that enable instrumental learning in the cortex could yield important mechanistic insights.

5.2.3 The role of convergent inputs in the striatum for action selection

In Chapter 3, we show that the striatum contains several simultaneous representations of task variables, including those that relate to confidence in the current state of a switching task. The striatum is recipient of inputs from most cortical areas, and is implicated in the selection of actions and motor programs. These data seem to suggest that the striatum is important for selecting a course of action based on many different sources of information from a variety of cortical areas. However, many cortical regions are also known to be critical for action planning and execution, which leaves open the question of how different components of cortico-basal ganglia circuits contribute to the selection of appropriate actions, and how they may work in concert to give rise to adaptive behavior.

5.2.4 Digital neural dust

In Chapter 4, we demonstrated a wireless neural recording system that operated using acoustic energy for power and communication. The devices reported here were analog, such that changes in voltage recorded by the electrodes directly modulated the gate of a single transistor, which in turn modulated the reflectivity of a piezoelectric crystal. One drawback of this design is that only a small fraction of the piezocrystal's dynamic range was utilized, resulting in the fairly high noise floor value of 0.18 mV. A future direction that is already in progress is to develop a digital version of recording neural dust motes that utilizes the full dynamic range of the crystal, allowing for more robust operation under instances of misalignment, as well as a lower noise floor.

5.2.5 Stimulation neural dust

In addition to recording, closed-loop BMI systems and bioelectronic medicines often require stimulation of nerves or neurons to achieve therapeutic endpoints. Another future direction of neural dust technology, again already in progress, is to develop an implantable device that can harvest ultrasonic energy, and then use it to stimulate neural tissue on demand. These devices would an essential component of a fully closed-loop ultrasonic neural interface.

5.2.6 Neural dust in the brain

In Chapter 4, we described the application of neural dust for recording the activity of peripheral nerves. Theoretically, further miniaturization of neural dust motes could allow for wireless recording from many sites deep in brain tissue. This possibility has been described in detail elsewhere (Seo et al., 2013). However, several important challenges remain. Acoustic waves are attenuated by bone to a degree that makes transmission of ultrasonic backscatter through the skull infeasible. Other strategies could be employed to alleviate this issue: an intermediate interrogator placed on the cortical surface could communicate with motes implanted in tissue and then relay data through the skull using electromagnetic coupling. Additionally, the noise floor of analog motes is likely too high to record single unit activity; however, increasing the complexity of the devices may also increase their size, which could create challenges for implantation. The electrode spacing of cortical neural dust motes is also an important consideration if devices reach the order of 10s of microns, due to a much smaller potential difference that would be measured across the electrodes relative to a mm-scale device. Finally, an implantation method for delivering motes precisely in brain tissue remains to be determined.

5.3 Conclusion

Direct interfaces between brains, nerves and machines represent a compelling direction for future investigation. Outlining the neural mechanisms that underlie the voluntary and involuntary motions of the body open up new possibilities for therapeutic interventions and can inspire new technologies. These same technologies can also spur greater understanding of the biological tissues with which they connect. However, many important avenues of investigation must still be explored in order to fully grasp the complex mechanisms by which the nervous system directs and adapts the behavior of the body. In this dissertation, we demonstrated both the potential of technology to elucidate learning mechanisms in the brain, as well as developed novel systems to improve the way we record neural activity. Although many questions remain unanswered, identifying fundamental properties of the learning process as well as scalable tools for interfacing with nervous tissue present exciting new experimental and clinical possibilities.

Bibliography

Aflalo, T., Kellis, S., Klaes, C., Lee, B., Shi, Y., Pejsa, K., Shanfield, K., Hayes-Jackson, S., Aisen, M., Heck, C., et al. (2015). Decoding motor imagery from the posterior parietal cortex of a tetraplegic human. Science *348*, 906–910.

Alexander, G.E., DeLong, M.R., and Strick, P.L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu. Rev. Neurosci. *9*, 357–381.

Allen Institute for Brain Science (2015). Allen Mouse Brain Atlas. Allen Mouse Brain Atlas.

American Institute of Ultrasound in Medicine (1998). Acoustic output measurement standard for diagnostic ultrasound equipment.

Arduin, P.-J., Frégnac, Y., Shulz, D.E., and Ego-Stengel, V. (2013). “Master” neurons induced by operant conditioning in rat motor cortex during a brain-machine interface task. J. Neurosci. *33*, 8308–8320.

Barnes, T.T.D., Kubota, Y., Hu, D., Jin, D.D.Z., and Graybiel, A.M.A. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature *437*, 1158–1161.

Bertrand, A., Seo, D., Maksimovic, F., Carmena, J.M., Maharbiz, M.M., Alon, E., and Rabaey, J.M. (2014). Beamforming approaches for untethered, ultrasonic neural dust motes for cortical recording: a simulation study. Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf. *2014*, 2625–2628.

Bhadra, N., and Kilgore, K.L. (2005). High-frequency electrical conduction block of mammalian peripheral motor nerve. Muscle Nerve *32*, 782–790.

Biederman, W., Yeager, D.J., Narevsky, N., Member, S., Leverett, J., Neely, R., Carmena, J.M., Member, S., Alon, E., and Rabaey, J.M. (2015). Combining 64 Acquisition Channels With Digital Compression and Simultaneous Dual Stimulation. JSSC *50*, 1038–1047.

Biran, R., Martin, D.C., and Tresco, P.A. (2005). Neuronal cell loss accompanies the brain tissue response to chronically implanted silicon microelectrode arrays. Exp. Neurol. *195*, 115–126.

Biran, R., Martin, D.C., and Tresco, P.A. (2007). The brain tissue response to implanted silicon microelectrode arrays is increased when the device is tethered to the skull. J. Biomed. Mater. Res. Part A *82A*, 169–178.

Birmingham, K., Gradinaru, V., Anikeeva, P., Grill, W.M., Pikov, V., McLaughlin, B., Pasricha, P., Weber, D., Ludwig, K., and Famm, K. (2014). Bioelectronic medicines: a research roadmap. Nat. Rev. Drug Discov. *13*, 399–400.

Bjorninen, T., Muller, R., Ledochowitsch, P., Sydanheimo, L., Ukkonen, L., Maharbiz, M.M., and Rabaey, J.M. (2012). Design of Wireless Links to Implanted Brain–Machine

Interface Microelectronic Systems. IEEE Antennas Wirel. Propag. Lett. *11*, 1663–1666.

Boretius, T., Badia, J., Pascual-Font, A., Schuettler, M., Navarro, X., Yoshida, K., and Stieglitz, T. (2010). A transverse intrafascicular multichannel electrode (TIME) to interface with the peripheral nerve. Biosens. Bioelectron. *26*, 62–69.

Boulougouris, V., Dalley, J.W., and Robbins, T.W. (2007). Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. Behav. Brain Res. *179*, 219–228.

Bouton, C.E., Shaikhouni, A., Annetta, N. V., Bockbrader, M.A., Friedenberg, D.A., Nielson, D.M., Sharma, G., Sederberg, P.B., Glenn, B.C., Mysiw, W.J., et al. (2016). Restoring cortical control of functional movement in a human with quadriplegia. Nature *0*, 1–13.

Brigman, J.L., and Rothblat, L.A. (2008). Stimulus specific deficit on visual reversal learning after lesions of medial prefrontal cortex in the mouse. Behav. Brain Res. *187*, 405–410.

Brown, L.L., Schneider, J.S., and Lidsky, T.I. (1997). Sensory and cognitive functions of the basal ganglia. Curr. Opin. Neurobiol. *7*, 157–163.

Buffalo, E.A., Fries, P., Landman, R., Buschman, T.J., and Desimone, R. (2011). Laminar differences in gamma and alpha coherence in the ventral stream. Proc. Natl. Acad. Sci. U. S. A. *108*, 11262–11267.

Cameron, T. (2004). Safety and efficacy of spinal cord stimulation for the treatment of chronic pain: a 20-year literature review. J. Neurosurg. Spine *100*, 254–267.

Carmena, J.M., Lebedev, M.A., Crist, R.E., O'Doherty, J.E., Santucci, D.M., Dimitrov, D.F., Patil, P.G., Henriquez, C.S., and Nicolelis, M.A.L. (2003). Learning to Control a Brain-Machine Interface for Reaching and Grasping by Primates. PLoS Biol. *1*, e42.

Castañé Anna, A., Theobald, D.E.H., and Robbins, T.W. (2010). Selective lesions of the dorsomedial striatum impair serial spatial reversal learning in rats. Behav. Brain Res. *210*, 74–83.

Cerf, M., Thiruvengadam, N., Mormann, F., Kraskov, A., Quiroga, R.Q., Koch, C., and Fried, I. (2010). On-line, voluntary control of human temporal lobe neurons. Nature *467*, 1104–1108.

Charthad, J., Weber, M.J., Chang, T.C., and Arbabian, A. (2015). A mm-Sized Implantable Medical Device (IMD) With Ultrasonic Power Transfer and a Hybrid Bi-Directional Data Link. IEEE J. Solid-State Circuits *50*, 1741–1753.

Chestek, C., Gilja, V., and Nuyujukian, P. (2011). Long-term stability of neural prosthetic control signals from silicon cortical arrays in rhesus macaque motor cortex. J. Neural.

Chopra, A., Klassen, B., and Stead, S. (Matt) (2013). Current clinical application of deep-brain stimulation for essential tremor. Neuropsychiatr. Dis. Treat. *Volume 9*, 1859.

Chuong, A.S., Miri, M.L., Busskamp, V., Matthews, G.A.C., Acker, L.C., Sørensen, A.T., Young, A., Klapoetke, N.C., Henninger, M.A., Kodandaramaiah, S.B., et al. (2014). Noninvasive optical inhibition with a red-shifted microbial rhodopsin. Nat. Neurosci. *17*, 1123–1129.

Clancy, K.K.B., Koralek, A.C.A., Costa, R.R.M., Feldman, D.E. DE, and Carmena, J.M. (2014). Volitional modulation of optically recorded calcium signals during neuroprosthetic learning. Nat. Neurosci. *17*, 807–809.

Clarke, H.F., Robbins, T.W., and Roberts, A.C. (2008). Lesions of the Medial Striatum in Monkeys Produce Perseverative Impairments during Reversal Learning Similar to Those Produced by Lesions of the Orbitofrontal Cortex. J. Neurosci. *28*, 10972–10982.

Collinger, J., Wodlinger, B., Downey, J., and Wang, W. (2013). High-performance neuroprosthetic control by an individual with tetraplegia. Lancet.

Corbit, L.H., and Janak, P.H. (2010). Posterior dorsomedial striatum is critical for both selective instrumental and Pavlovian reward learning. Eur. J. Neurosci. *31*, 1312–1321.

Cordova, C.A., Jackson, D., Langdon, K.D., Hewlett, K.A., and Corbett, D. (2014). Impaired executive function following ischemic stroke in the rat medial prefrontal cortex. Behav. Brain Res. *258*, 106–111.

Costa, R.M. (2011). A selectionist account of de novo action learning. Curr. Opin. Neurobiol. *21*, 579–586.

Creasey, G.H., Grill, J.H., Korsten, M., U, H.S., Betz, R., Anderson, R., Walter, J., and Implanted Neuroprosthesis Research Group (2001). An implantable neuroprosthesis for restoring bladder and bowel control to patients with spinal cord injuries: a multicenter trial. Arch. Phys. Med. Rehabil. *82*, 1512–1519.

Dang, M.T., Yokoi, F., Yin, H.H., Lovinger, D.M., Wang, Y., and Li, Y. (2006). Disrupted motor learning and long-term synaptic plasticity in mice lacking NMDAR1 in the striatum. Proc. Natl. Acad. Sci. U. S. A. *103*, 15254–15259.

Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. *8*, 1704–1711.

Daw, N.D., O'Doherty, J.P., Dayan, P., Dolan, R.J., and Seymour, B. (2006). Cortical substrates for exploratory decisions in humans. Nature *441*, 876–879.

Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., and Dolan, R.J. (2011). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. Neuron *69*, 1204–1215.

Delivopoulos, E., Chew, D.J., Minev, I.R., Fawcett, J.W., and Lacour, S.P. (2012). Concurrent recordings of bladder afferents from multiple nerves using a microfabricated PDMS microchannel electrode array. Lab Chip *12*, 2540–2551.

Denison, T., Consoer, K., Santa, W., Molnar, G., and Mieser, K. (2007). A 2 μW, 95nV/rtHz, chopper-stabilized instrumentation amplifier for chronic measurement of biopotentials. 2007 IEEE Instrum. Meas. Technol. Conf. IMTC 2007 *42*, 3–8.

Dias-Ferreira, E., Sousa, J.C., Melo, I., Morgado, P., Mesquita, A.R., Cerqueira, J.J., Costa, R.M., and Sousa, N. (2009). Chronic Stress Causes Frontostriatal Reorganization and Affects Decision-Making. Science (80-.). *325*, 621–625.

Du, J., Blanche, T.J., Harrison, R.R., Lester, H.A., and Masmanidis, S.C. (2011). Multiplexed, high density electrophysiology with nanofabricated neural probes. PLoS One *6*.

Engel, A.K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. Nat. Rev. Neurosci. *2*, 704–716.

Everitt, B.J., and Robbins, T.W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. Nat. Neurosci. *8*, 1481–1489.

Famm, K., Litt, B., Tracey, K.J., Boyden, E.S., and Slaoui, M. (2013). Drug discovery: a jump-start for electroceuticals. Nature *496*, 159–161.

Faull, R.L.M., Nauta, W.J.H., and Domesick, V.B. (1986). The visual cortico-striato-nigral pathway in the rat. Neuroscience *19*, 1119–1132.

Fetz, E.E. (2007). Volitional control of neural activity: implications for brain–computer interfaces. J. Physiol. *579*, 571–579.

Foster, J.D., Nuyujukian, P., Freifeld, O., Gao, H., Walker, R., I Ryu, S., H Meng, T., Murmann, B., J Black, M., and Shenoy, K. V (2014). A freely-moving monkey treadmill model. J. Neural Eng. *11*, 46020.

Ganguly, K., and Carmena, J.M.J. (2009). Emergence of a stable cortical map for neuroprosthetic control. PLoS Biol. *7*, e1000153.

Ganguly, K., Dimitrov, D.D.F., Wallis, J.J.D., and Carmena, J.M.J. (2011). Reversible large-scale modification of cortical networks during neuroprosthetic control. Nat. Neurosci. *14*, 662–667.

Gilja, V., Pandarinath, C., Blabe, C.H., Nuyujukian, P., Simeral, J.D., Sarma, A. a, Sorice, B.L., Perge, J. a, Jarosiewicz, B., Hochberg, L.R., et al. (2015). Clinical translation of a high-performance neural prosthesis. Nat. Med. *21*, 6–8.

Gillan, C.M., Papmeyer, M., Morein-Zamir, S., Sahakian, B.J., Fineberg, N.A., Robbins, T.W., and De Wit, S. (2011). Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. Am. J. Psychiatry *168*, 718–726.

Gläscher, J., Daw, N., Dayan, P., and O'Doherty, J.P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron *66*, 585–595.

Gold, C., Henze, D.A., and Koch, C. (2007). Using extracellular action potential recordings to constrain compartmental models. J. Comput. Neurosci. *23*, 39–58.

Graybiel, A.M. (1997). The basal ganglia and cognitive pattern generators. Schizophr.

Bull. *23*, 459–469.

Graybiel, A.M. (2008). Habits, rituals, and the evaluative brain. Annu. Rev. Neurosci. *31*, 359–387.

Gregoriou, G.G., Gotts, S.J., Zhou, H., and Desimone, R. (2009). High-frequency, long-range coupling between prefrontal and visual cortex during attention. Science *324*, 1207–1210.

Gremel, C.C.M., and Costa, R.R.M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nat. Commun. *4*, 2264.

Gruner, J.A., and Mason, C.P. (1989). Nonlinear muscle recruitment during intramuscular and nerve stimulation. J. Rehabil. Res. Dev. *26*, 1–16.

Hampton, A.N., Bossaerts, P., and O'Doherty, J.P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. J. Neurosci. *26*, 8360–8367.

Harrison, R.R. (2007). Designing Efficient Inductive Power Links for Implantable Devices. 2007 IEEE Int. Symp. Circuits Syst. *1*, 2080–2083.

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

Hikosaka, O., Nakamura, K., Sakai, K., and Nakahara, H. (2002). Central mechanisms of motor skill learning. Curr. Opin. Neurobiol. *12*, 217–222.

Hinterberger, T., Veit, R., Wilhelm, B., Weiskopf, N., Vatine, J.J., and Birbaumer, N. (2005). Neuronal mechanisms underlying control of a brain-computer interface. Eur. J. Neurosci. *21*, 3169–3181.

Hintiryan, H., Foster, N.N., Bowman, I., Bay, M., Song, M.Y., Gou, L., Yamashita, S., Bienkowski, M.S., Zingg, B., Zhu, M., et al. (2016). The mouse cortico-striatal projectome. Nat. Neurosci. *19*, 1100.

Hochberg, L.R., Bacher, D., Jarosiewicz, B., Masse, N.Y., Simeral, J.D., Vogel, J., Haddadin, S., Liu, J., Cash, S.S., van der Smagt, P., et al. (2012). Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. Nature *485*, 372–375.

Houweling, A.R., and Brecht, M. (2008). Behavioural report of single neuron stimulation in somatosensory cortex. Nature *451*, 65–68.

Husain, M., Shapiro, K., Martin, J., and Kennard, C. (1997). Abnormal temporal dynamics of visual attention in spatial neglect patients. Nature *385*, 154–156.

Hwang, E.J., Bailey, P.M., and Andersen, R.A. (2013). Volitional control of neural activity relies on the natural motor repertoire. Curr. Biol. *23*, 353–361.

Izquierdo, A., Brigman, J.L., Radke, A.K., Rudebeck, P.H., and Holmes, A. (2016). The neural basis of reversal learning: An updated perspective. Neuroscience *345*, 12–26.

Jarosiewicz, B., Chase, S.M., Fraser, G.W., Velliste, M., Kass, R.E., and Schwartz, A.B. (2008). Functional network reorganization during learning in a brain-computer interface paradigm. Proc. Natl. Acad. Sci. U. S. A. *105*, 19486–19491.

Jarvis, M.R., and Mitra, P.P. (2001). Sampling Properties of the Spectrum and Coherency of Sequences of Action Potentials. Neural Comput. *13*, 717–749.

Joel, D., and Weiner, I. (1994). The organization of the basal ganglia-thalamocortical circuits: Open interconnected rather than closed segregated. Neuroscience *63*, 363–379.

Jow, U.M., and Ghovanloo, M. (2007). Design and optimization of printed spiral coils for efficient inductive power transmission. In Proceedings of the IEEE International Conference on Electronics, Circuits, and Systems, pp. 70–73.

Keller, G.B., Bonhoeffer, T., and Hübener, M. (2012). Sensorimotor Mismatch Signals in Primary Visual Cortex of the Behaving Mouse. Neuron *74*, 809–815.

Kemp, J.M., and Powell, T.P.S. (1970). The cortico-striate projection in the monkey. Brain *93*.

Kezirian, E.J., Boudewyns, A., Eisele, D.W., Schwartz, A.R., Smith, P.L., Van de Heyning, P.H., and De Backer, W.A. (2010). Electrical stimulation of the hypoglossal nerve in the treatment of obstructive sleep apnea. Sleep Med. Rev. *14*, 299–305.

Khibnik, L.A., Tritsch, N.X., and Sabatini, B.L. (2014). A direct projection from mouse primary visual cortex to dorsomedial striatum. PLoS One *9*, e104501.

Kim, H., Sul, J.H., Huh, N., Lee, D., and Jung, M.W. (2009). Role of striatum in updating values of chosen actions. J Neurosci *29*, 14701–14712.

Kimchi, E.Y., and Laubach, M. (2009). Dynamic encoding of action selection by the medial striatum. J. Neurosci. *29*, 3148–3159.

Kino, G.S. (1987). Acoustic Waves: Devices, Imaging, and Analog Signal Processing.

Kobak, D., Brendel, W., Constantinidis, C., Feierstein, C.E., Kepecs, A., Mainen, Z.F., Qi, X.L., Romo, R., Uchida, N., and Machens, C.K. (2016). Demixed principal component analysis of neural population data. Elife *5*, 614–635.

Koralek, A.C.A., Jin, X., II, J.L., Costa, R.R.M., Carmena, J.M.J., Long, J.D., Costa, R.R.M., and Carmena, J.M.J. (2012). Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. Nature *483*, 331–335.

Koralek, A.C.A., Costa, R.R.M., and Carmena, J.M.J. (2013). Temporally precise cell-specific coherence develops in corticostriatal networks during learning. Neuron *79*, 865–872.

Krimholtz, R., Leedom, D.A., and Matthaei, G.L. (1970). New Equivalent Circuits for Elementary Piezoelectric Transducers. Electron. Lett. *6*, 398–399.

Krook-Magnuson, E., Gelinas, J.N., Soltesz, I., and Buzsáki, G. (2015). Neuroelectronics and Biooptics: Closed-Loop Technologies in Neurological Disorders. JAMA Neurol. *72*, 1–7.

Lapatki, B.G., Van Dijk, J.P., Jonas, I.E., Zwarts, M.J., and Stegeman, D.F. (2004). A thin, flexible multielectrode grid for high-density surface EMG. J. Appl. Physiol. *96*, 327–336.

Larson, P.J., and Towe, B.C. (2011). Miniature ultrasonically powered wireless nerve cuff stimulator. In 2011 5th International IEEE/EMBS Conference on Neural Engineering, NER 2011, pp. 265–268.

Lau, B., and Glimcher, P.W. (2008). Value Representations in the Primate Striatum during??Matching Behavior. Neuron *58*, 451–463.

Lee, S.B., Yin, M., Manns, J.R., and Ghovanloo, M. (2013). A wideband dual-antenna receiver for wireless recording from animals behaving in large arenas. IEEE Trans. Biomed. Eng. *60*, 1993–2004.

Lefurge, T., Goodall, E., Horch, K., Stensaas, L., and Schoenberg, A. (1991). Chronically implanted intrafascicular recording electrodes. Ann. Biomed. Eng. *19*, 197–207.

Lepage, K.Q., Kramer, M.A., and Eden, U.T. (2011). The dependence of spike field coherence on expected intensity. Neural Comput. *23*, 2209–2241.

Li, J., and Daw, N.D. (2011). Signals in Human Striatum Are Appropriate for Policy Update Rather than Value Prediction. J. Neurosci. *31*, 5504–5511.

Lin, J.C. (2006). A new IEEE standard for safety levels with respect to human exposure to radio-frequency radiation. IEEE Antennas Propag. Mag. *48*, 157–159.

Makino, H., and Komiyama, T. (2015). Learning enhances the relative impact of top-down processing in the visual cortex. Nat. Neurosci. *18*, 1116–1122.

Martinez, A., Anllo-Vento, L., and Sereno, M. (1999). Involvement of striate and extrastriate visual cortical areas in spatial attention. Nat. Neurosci. *2*, 364–369.

Martinez-Valdes, E., Laine, C.M., Falla, D., Mayer, F., and Farina, D. (2016). High-density surface electromyography provides reliable estimates of motor unit behavior. Clin. Neurophysiol. *127*, 2534–2541.

McAlonan, K., and Brown, V.J. (2003). Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. Behav. Brain Res. *146*, 97–103.

McClure, S.M., Berns, G.S., and Montague, P.R. (2003). Temporal prediction errors in a passive learning task activate human striatum. Neuron *38*, 339–346.

McGeorge, A.J., and Faull, R.L.M. (1989). The organization of the projection from the cerebral cortex to the striatum in the rat. Neuroscience *29*, 503–537.

McHaffie, J.G., Stanford, T.R., Stein, B.E., Coizet, V., and Redgrave, P. (2005).

Subcortical loops through the basal ganglia. Trends Neurosci. *28*, 401–407.

Meng, H., and Sahin, M. (2013). An electroacoustic recording device for wireless sensing of neural signals. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, pp. 3086–3088.

Mercuri, E., Atkinson, J., Braddick, O., Anker, S., Cowan, F., Rutherford, M., Pennock, J., and Dubowitz, L. (1997). Basal ganglia damage and impaired visual function in the newborn infant. Arch Dis Child Fetal Neonatal Ed *77*, F111-4.

Mezzarane, R., Elias, L., and Magalhães, F. (2013). Experimental and simulated EMG responses in the study of the human spinal cord. New Front. ….

Middleton, F.A., and Strick, P.L. (1994). Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. Science (80-.). *266*, 458–461.

Middleton, F.A., and Strick, P.L. (1996). The temporal lobe is a target of output from the basal ganglia. Proc Natl Acad Sci U S A *93*, 8683–8687.

Muller, R., Le, H.-P., Li, W., Ledochowitsch, P., Gambini, S., Bjorninen, T., Koralek, A., Carmena, J.M., Maharbiz, M.M., Alon, E., et al. (2015). A Minimally Invasive 64-Channel Wireless uECoG Implant. IEEE J. Solid-State Circuits *50*, 344–359.

Musallam, S., Corneil, B.D., Greger, B., Scherberger, H., and Andersen, R.A. (2004). Cognitive control signals for neural prosthetics. Science *305*, 258–262.

Niell, C.M.C.C.M., and Stryker, M.P.M. (2010). Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. Neuron *65*, 472–479.

O'Doherty, J., Becker-Hapak, S.M.S., Mcallister, S.F., Dowdy, H.P., Grill, J.L., Zweier, P., Kuppusamy, M.L., Weisfeldt, J.T., Flaherty 14, D., and Morgan (2004). Dissociable Roles of Ventral and Dorsal Striatum in Instrumental. Science (80-.). *304*, 1604–1563.

Orsborn, A.L., Moorman, H.G., Overduin, S.A., Shanechi, M.M., Dimitrov, D.F., and Carmena, J.M. (2014). Closed-Loop Decoder Adaptation Shapes Neural Plasticity for Skillful Neuroprosthetic Control. Neuron *82*, 1380–1393.

Owen, A.M. (1997). Cognitive planning in humans: Neuropsychological, neuroanatomical and neuropharmacological perspectives. Prog. Neurobiol. *53*, 431–450.

Oyama, K., Tateyama, Y., Hernádi, I., Tobler, P.N., Iijima, T., and Tsutsui, K.-I. (2015). Discrete coding of stimulus value, reward expectation, and reward prediction error in the dorsal striatum. J. Neurophysiol. *114*, 2600–2615.

Ozeri, S., and Shmilovitz, D. (2010). Ultrasonic transcutaneous energy transfer for powering implanted devices. Ultrasonics *50*, 556–566.

Pasupathy, A., and Miller, E.K. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. Nature *433*, 873–876.

Pavlov, V.A., and Tracey, K.J. (2012). The vagus nerve and the inflammatory reflex--

linking immunity and metabolism. Nat. Rev. Endocrinol. *8*, 743–754.

Polikov, V.V.S., Tresco, P.A., and Reichert, W.W.M.W. (2005). Response of brain tissue to chronically implanted neural electrodes. J. Neurosci. Methods *148*, 1–18.

Prsa, M., Galinanes, G.L., and Huber, D. (2017). Rapid Integration of Artificial Sensory Feedback during Operant Conditioning of Motor Cortex Neurons. Neuron *93*, 929–939.e6.

Quilodran, R., Rothé, M., and Procyk, E. (2008). Behavioral Shifts and Action Valuation in the Anterior Cingulate Cortex. Neuron *57*, 314–325.

Rabaey, J.M., Mark, M., Chen, D., Sutardja, C., Tang, C., Gowda, S., Wagner, M., and Werthimer, D. (2011). Powering and communicating with mm-size implants. Des. Autom. Test Eur. Conf. Exhib. 1–6.

Randles, J. (1947). Kinetics of rapid electrode reactions. Discuss. Faraday Soc.

Redgrave, P., Rodriguez, M., Smith, Y., Rodriguez-Oroz, M.C., Lehericy, S., Bergman, H., Agid, Y., DeLong, M.R., and Obeso, J.A. (2010). Goal-directed and habitual control in the basal ganglia: implications for Parkinson’s disease. Nat. Rev. Neurosci. *11*, 760–772.

Redgrave, P., Vautrelle, N., and Reynolds, J.N.J. (2011). Functional properties of the basal ganglia’s re-entrant loop architecture: selection and reinforcement. Neuroscience *198*, 138–151.

Rosas-Ballina, M., Olofsson, P.S., Ochani, M., Valdés-Ferrer, S.I., Levine, Y.A., Reardon, C., Tusche, M.W., Pavlov, V.A., Andersson, U., Chavan, S., et al. (2011). Acetylcholine-synthesizing T cells relay neural signals in a vagus nerve circuit. Science *334*, 98–101.

Sadtler, P.T., Quick, K.M., Golub, M.D., Chase, S.M., Ryu, S.I., Tyler-Kabara, E.C., Yu, B.M., and Batista, A.P. (2014). Neural constraints on learning. Nature *512*, 423–426.

Saez, A., Rigotti, M., Ostojic, S., Fusi, S., and Salzman, C.D. (2015). Abstract Context Representations in Primate Amygdala and Prefrontal Cortex. Neuron *87*, 869–881.

Samejima, K. (2005). Representation of Action-Specific Reward Values in the Striatum. Science (80-.). *310*, 1337–1340.

Samejima, K., and Doya, K. (2007). Multiple representations of belief states and action values in corticobasal ganglia loops. In Annals of the New York Academy of Sciences, pp. 213–228.

Schafer, R.J., and Moore, T. (2011). Selective Attention from Voluntary Control of Neurons in Prefrontal Cortex. Science (80-.). *332*.

Schlagenhauf, F., Huys, Q.J.M., Deserno, L., Rapp, M.A., Beck, A., Heinze, H.J., Dolan, R., and Heinz, A. (2014). Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. Neuroimage *89*, 171–180.

van Schoonhoven, J., Sparreboom, M., van Zanten, B.G. a, Scholten, R.J.P.M., Mylanus, E. a M., Dreschler, W. a, Grolman, W., and Maat, B. (2013). The Effectiveness of Bilateral Cochlear Implants for Severe-to-Profound Deafness in Adults : A Systematic Review. Otol. Neurotol. *34*, 190–198.

Schultz, W. (2000). Multiple reward signals in the brain. Nat. Rev. *1*, 199–207.

Schwarz, D.A., Lebedev, M.A., Hanson, T.L., Dimitrov, D.F., Lehew, G., Meloy, J., Rajangam, S., Subramanian, V., Ifft, P.J., Li, Z., et al. (2014). Chronic, wireless recordings of large-scale brain activity in freely moving rhesus monkeys. Nat. Methods *11*, 670–676.

Seo, D., Carmena, J.M., Rabaey, J.M., Alon, E., and Maharbiz, M.M. (2013). Neural dust: an ultrasonic, low powersSolution for chronic brain-machine interfaces. Cornell Univ. Libr. 1–11.

Seo, D., Carmena, J.M., Rabaey, J.M., Maharbiz, M.M., and Alon, E. (2015). Model validation of untethered, ultrasonic neural dust motes for cortical recording. J. Neurosci. Methods *244*, 114–122.

Shibata, K., Watanabe, T., Sasaki, Y., and Kawato, M. (2011). Perceptual Learning Incepted by Decoded fMRI Neurofeedback Without Stimulus Presentation. Science (80- .). *334*, 1413 LP-1415.

Shohamy, D., Myers, C.E., Grossman, S., Sage, J., Gluck, M.A., and Poldrack, R.A. (2004). Cortico-striatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. Brain *127*.

Shuler, M.G. (2006). Reward Timing in the Primary Visual Cortex. Science (80-.). *311*, 1606–1609.

Steinmetz, P.N., Roy, A., Fitzgerald, P.J., Hsiao, S.S., Johnson, K.O., and Niebur, E. (2000). Attention modulates synchronized neuronal firing in primate somatosensory cortex. Nature *404*, 187–190.

Steyaert, M., and Sansen, W. (1987). A micropower low-noise monolithic instrumentation amplifier for medical purposes. IEEE J. Solid-State Circuits.

Strollo, P.J., Soose, R.J., Maurer, J.T., de Vries, N., Cornelius, J., Froymovich, O., Hanson, R.D., Padhya, T. a, Steward, D.L., Gillespie, M.B., et al. (2014). Upper-airway stimulation for obstructive sleep apnea. N. Engl. J. Med. *370*, 139–149.

Suner, S., Fellows, M.R., Vargas-Irwin, C., Nakata, G.K., and Donoghue, J.P. (2005). Reliability of signals from a chronically implanted, silicon-based electrode array in non-human primate primary motor cortex. IEEE Trans. Neural Syst. Rehabil. Eng. *13*, 524–541.

Swanson, L.W. (2000). Cerebral hemisphere regulation of motivated behavior. Brain Res. *886*, 113–164.

Szuts, T.A., Fadeyev, V., Kachiguine, S., Sher, A., Grivich, M. V, Agrochão, M., Hottowy, P., Dabrowski, W., Lubenov, E. V, Siapas, A.G., et al. (2011). A wireless multi-channel neural amplifier for freely moving animals. Nat. Neurosci. *14*, 263–269.

Tanaka, S.C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., and Yamawaki, S. (2016). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. In Behavioral Economics of Preferences, Choices, and Happiness, pp. 593–616.

Tang, H.H.-Y., Seo, D., Singhal, U., Li, X., Maharbiz, M.M., Alon, E., and Boser, B.E. (2015). Miniaturizing Ultrasonic System for Portable Health Care and Fitness. IEEE Trans. Biomed. Circuits Syst. *9*, 767–776.

Thomson, D.J. (1982). Spectrum estimation and harmonic analysis. Proc. IEEE *70*, 1055–1096.

Tracey, K.J. (2009). Reflex control of immunity. Nat. Rev. Immunol. *9*, 418–428.

Tricomi, E.M., Delgado, M.R., and Fiez, J.A. (2004). Modulation of Caudate Activity by Action Contingency. Neuron *41*, 281–292.

Turner, J., Shain, W., Szarowski, D., and Andersen, M. (1999). Cerebral astrocyte response to micromachined silicon implants. Experimental.

US Food and Drug Administration (2008). Information for Manufacturers Seeking Marketing Clearance of Diagnostic Ultrasound Systems and Transducers. Ultrasound 1–64.

Webster, K.E. (1965). The cortico-striatal projection in the cat. J. Anat. *99*, 329–337.

Wickens, J.R., Horvitz, J.C., Costa, R.M., and Killcross, S. (2007). Dopaminergic mechanisms in actions and habits. J. Neurosci. *27*, 8181–8183.

Williams, A., Kim, H., Wang, F., Vyas, S., Shenoy, K. V., Schnitzer, M., Kolda, T., and Ganguli, S. Dimensionality reduction of multi-trial neural data by canonical polyadic tensor decomposition. In COSYNE, p.

Wilson, R.C., Takahashi, Y.K., Schoenbaum, G., and Niv, Y. (2014). Orbitofrontal cortex as a cognitive map of task space. Neuron *81*, 267–278.

Wolter, T. (2014). Spinal cord stimulation for neuropathic pain: Current perspectives. J. Pain Res. *7*, 651–663.

Wright, M.J., Burns, R.J., Geffen, G.M., and Geffen, L.B. (1990). Covert orientation of visual attention in Parkinson’s disease: An impairment in the maintenance of attention. Neuropsychologia *28*, 151–159.

Yamaguchi, S., and Kobayashi, S. (1998). Contributions of the dopaminergic system to voluntary and automatic orienting of visuospatial attention. J. Neurosci. *18*, 1869–1878.

Yin, H.H., Ostlund, S.B., Knowlton, B.J., and Balleine, B.W. (2005). The role of the dorsomedial striatum in instrumental conditioning. Eur. J. Neurosci. *22*, 513–523.

Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. Behav. Brain Res. *166*, 189–196.

Yin, H.H., Mulcare, S.P., Hilário, M.R.F., Clouse, E., Holloway, T., Davis, M.I., Hansson, A.C., Lovinger, D.M., and Costa, R.M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. Nat. Neurosci. *12*, 333–341.

Yin, M., Borton, D.A., Komar, J., Agha, N., Lu, Y., Li, H., Laurens, J., Lang, Y., Li, Q., Bull, C., et al. (2014). Wireless neurosensor for full-spectrum electrophysiology recordings during free behavior. Neuron *84*, 1170–1182.

Yoshida, W., and Ishii, S. (2006). Resolution of Uncertainty in Prefrontal Cortex. Neuron *50*, 781–789.

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J.P., Chang, W.-C., Jenvay, S., Miyamichi, K., Luo, L., and Dan, Y. (2014). Selective attention. Long-range and local circuits for top-down modulation of visual cortex processing. Science *345*, 660–665.

Zhong, Y., and Bellamkonda, R. V (2007). Dexamethasone-coated neural probes elicit attenuated inflammatory response and neuronal loss compared to uncoated neural probes. Brain Res. *1148*, 15–27.