Structure and time course of neural population activity during learning

Jay A. Hennig
March 2021

Center for the Neural Basis of Cognition, Neuroscience Institute
Machine Learning Department, School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee
Steven Chase (co-chair)
Byron Yu (co-chair)
Aaron Batista
Robert Kass
Eric Shea-Brown (University of Washington)

Submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Neural Computation and Machine Learning

This work was supported by the Richard King Mellon Presidential Fellowship and the Carnegie Prize Fellowship in Mind and Brain Sciences.

Abstract

Our ability to perform a variety of difficult tasks—everything from reasoning about the best chess move, to shooting a free throw, or finely dicing an onion—is due to the coordinated activity of populations of neurons throughout the nervous system. And yet, we lack an understanding of how the brain generates the activity appropriate for achieving something as simple as pressing an elevator button. In part, this is because we do not know which neural activity patterns the brain is capable of generating, nor how that activity will change with experience. By exploring the structure and constraints on the activity patterns the brain can express, we move closer to understanding how the brain can generate the activity supportive of such a rich variety of behaviors and adaptations.

Presently, in studies of arm or eye movements, we typically don't know the causal relationship between neural activity and behavior. Here we use a brain-computer interface (BCI) paradigm to study learning, because the exact relationship between neural activity and behavior is controlled by the experimenter. To generate proficient behavior, the animal must change the activity of the neurons currently being recorded. This provides us with the means to causally relate any observed structure in neural population activity with animals' performance at the task.

The focus of this thesis is to characterize the structure and time course of neural population activity during learning. In the first part of this thesis, we note that just as there is more than one way to win a game of chess, the brain has many different patterns of neural activity it can produce to drive the same behavior. Which of these redundant options does the brain prefer? We find that the frequency with which animals used different patterns of neural population activity was remarkably similar before and after learning. This suggests that the brain's ability to take advantage of redundancy may be somewhat limited, at least within the span of a few hours.

In the second part of this thesis, we asked how internal states such as our arousal, attention, and motivation interact with how we learn new tasks. We identified large, abrupt fluctuations in neural population activity in motor cortex indicative of arousal-like internal state changes, which we term "neural engagement." We find that stereotyped changes in neural engagement during learning were unrelated to goal-seeking behavior, but nevertheless influenced how quickly different tasks were learned. Overall, this thesis characterizes a variety of different constraints and influences on how populations of neurons change their activity during learning.

Acknowledgments

I would like to thank my advisors Byron Yu and Steve Chase for all of their time, advice, and feedback over the years. I have learned so much from their distinct approaches to doing research. I admire Byron's optimism about what it means to do good science, his thoroughness, and his ability to pinpoint questions that everyone wants answers to. And I admire Steve for his incredibly quick mind, sharp memory, and eagerness to dive into the details.

I thank Aaron Batista for all of his support, and also for his excitement about what felt like every result I ever presented. I would also like to thank my other committee members, Rob Kass and Eric Shea-Brown, for their feedback on this thesis. Thanks to Emily Oby for all of her amazing feedback on analyses, figures, colors, and words over the years. I am grateful for my collaborators, Matt Golub, Lindsay Bahureksa, Patrick Sadtler, Kristin Quick, Stephen Ryu, Elizabeth Tyler-Kabara, for designing and collecting such an amazing set of experiments. Thank you to Scabby, as well as my lab mates in the Yu and Chase lab, for an endless supply of feedback and criticism. I thank Larry Wasserman and Ryan Tibshirani for their incredible lectures that continually reminded me why I love math and statistics. And thanks to Ryan Williamson and Akash Umakantha for their company and teamwork throughout our time as graduate students.

I feel so lucky to have two parents and two sisters who have supported me in everything I do. I also feel extremely grateful for my friends in Pittsburgh, who welcomed us into their circle when we first moved here. Thank you to my two boys, Robin and Jordan, for always making me laugh. And to Jess, for improving everything I do.

This thesis is dedicated to my mom, for supporting my education, and cultivating my love for math and baked goods.

Contents

List of Figures

1 Introduction

Our ability to learn and perform a variety of difficult tasks—everything from reasoning about the best chess move, to shooting a free throw, to finely dicing an onion—is due to the coordinated activity of populations of neurons throughout the nervous system. And yet we lack an understanding of how the brain generates the activity appropriate for achieving something as simple as pressing an elevator button.

How *does* our brain generate the activity necessary to guide our finger towards an elevator button? And how does it learn to generate this activity in the first place? There are three key challenges to answering questions like these, which the work in this thesis will address. First, we typically do not know the causal relationship between neural activity and behavior (e.g., the movements of our arm and finger). To get around this difficulty, we can use a brain-computer interface (BCI) paradigm (Taylor et al., 2002; Carmena et al., 2003; Hochberg et al., 2006; Ganguly and Carmena, 2009; Gilja et al., 2012; Hauschild et al., 2012; Sadtler et al., 2014), where the exact relationship, or "mapping," between neural activity and behavior is controlled by the experimenter. In this thesis, we will discuss a series of BCI learning experiments in which monkeys controlled a computer cursor on a screen by modulating the activity of hundreds of neurons recorded in motor cortex (Figure 2.1) (Sadtler et al., 2014). To achieve the desired behavior (in this case, moving the cursor to acquire a target—not unlike moving a finger to touch an elevator button), the monkey must generate a particular pattern of neural activity. Importantly, the BCI paradigm provides us with the means to causally relate the neural activity we record in the experiment with animals' performance at the task (Golub et al., 2016).

The second challenge is to understand which neural activity patterns the brain is capable of generating. For example, when we move our arms, millions of neurons in primary motor cortex (M1) fire up to hundreds of action potentials, also known as "spikes," per second. These neurons do not fire independently, but are rather highly correlated with one another. In the work presented in this thesis, we use a dimensionality reduction technique known as factor analysis (FA) to identify the dominant patterns of co-modulation across the recorded neural units (Churchland et al., 2010; Harvey et al., 2012; Sadtler et al., 2014), termed the *intrinsic manifold* (Sadtler et al., 2014). Previous work found that, over the course of a few hours, monkeys could not readily learn BCI mappings that required neural activity to leave the intrinsic manifold, suggesting that the intrinsic manifold characterizes which patterns of neural population activity monkeys could easily generate (Sadtler et al.,

2014). Later work demonstrated that the set of neural population activity monkeys could generate was further constrained by a fixed *neural repertoire* within the intrinsic manifold (Golub et al., 2018). In this thesis, we characterize constraints on the *distribution* of neural population activity within the neural repertoire during short-term learning, and also find that changes to the neural repertoire occur given weeks of practice.

The third and final challenge is that we do not know how neural population activity changes with experience. While reaching out to touch an elevator button may seem effortless to you now, it is a skill that takes infants months and months of practice to perfect. Learning requires using feedback about our performance (e.g., when our finger misses the elevator button to the right) to change which neural activity we will generate on future attempts. At the same time, we know that neural activity changes throughout the brain according to changes in our internal state such as our arousal, attention, and motivation. How do these two sources of changes in neural population activity during learning—i.e., those that are learning-driven and those related to changes in our internal state–interact when we learn something new? We will present results indicating that even changes in our internal state unrelated to goal-seeking behavior can nevertheless impact how we learn new tasks.

The contributions of this thesis are organized as follows.

- In Chapter 2, I will describe the brain-computer interface (BCI) paradigm, including details of the experiments analyzed in this thesis.
- In Chapter 3, I will present results on how the brain selects among the redundant patterns of neural activity that would all achieve the same goal.
- In Chapter 4, I will show how internal state changes such as changes in our arousal can interact with how quickly we learn different tasks.
- In Chapter 5, I will present work I contributed to regarding the emergence of new neural activity patterns during long-term learning, as well as a textbook chapter on intracortical brain-machine interfaces.
- In Chapter 6 I will place our results in a broader context, reviewing what the field has learned about learning in the brain from studies of neural population activity.
- Finally, in Chapter 7, I will outline future directions for understanding how neural population activity changes during learning.

2 BRAIN-COMPUTER INTERFACES

A brain-computer interface, or BCI, directly connects the brain to the external world. A BCI consists of four basic components: neural recording, a decoding algorithm, an output device, and sensory feedback. Intracortical BCIs begin by recording neural signals from electrodes implanted in the cortex. Next, the salient features of the neural signal useful for control are extracted with a decoding algorithm. This algorithm translates the neural signal into an intended action which is executed by the output device, which can be a robotic limb, the person's own muscles, or in the experiments considered in this thesis, a computer interface. Finally, the user receives sensory feedback about the action, allowing them to make corrections if they move off course, and also allowing them to improve over time with learning. We discuss these components in more detail in the following paragraphs.

Building an effective BCI depends on choosing a brain area for the neural recordings, a decoding algorithm, an output device, and feedback for the desired use. These choices are interrelated. The necessary control signal will depend on the goal of the task and the device being controlled. In turn, these choices influence which motor cortical area is most appropriate to record from, and what type of signals to record. For example, when the goal of a BCI is to provide continuous control of a limb, primary motor cortex (M1) has long been thought to be an ideal location for recording BMI control signals because it is involved in generating voluntary movements. A number of groups have reported a correlation between M1 firing rates and various kinematic variables, including direction and distance of targets, and direction, speed, and spatial path of hand displacement (Georgopoulos et al., 1982). Other groups have found M1 firing rates to be related to forces, and even to muscle activity (Evarts, 1968). It appears that M1 includes a heterogeneous representation of both the kinematics and kinetics of limb movements. The good news, from the perspective of designing a BCI, is that either representation can be exploited as a BCI control signal, depending on the intended function of the device.

Once the neural signals have been recorded, a decoding algorithm (or "decoder") translates the user's movement intentions into a control signal suitable for guiding the output device. There are two classes of BCI decoders: discrete and continuous. A discrete decoder estimates one of several possible movement goals by solving a classification problem. The most common use for this is a communication device, where patients use their BCI to type letters, much as one would if one were composing a text or an email. Communication BCIs focus on the speed and accuracy with which

a desired key on a keyboard can be selected. A continuous decoder estimates the moment-by-moment details of a movement trajectory. This is needed for guiding a computer cursor or robotic limb along a desired path. For example, a person may wish to guide a robotic limb to pick up a glass of milk without knocking over the milk carton. The decoder produces a control signal that is then fed into an output device. There are a variety of output devices for BCIs that depend on the needs of the user. One common device is a computer cursor, where a person controls the cursor using their thoughts much as they would control a computer mouse. Other common devices being developed to help people with paralysis interact with the world around them include robotic arms and hands and motorized wheelchairs.

The final element of the BCI control loop is sensory feedback. The most common sensory feedback is visual: a user can look at the device they are controlling and see how it is responding, which allows them to make corrective movements and learn to better control the device. Feedback has been shown to dramatically improve BCI performance. Some tasks, however, require more than just visual feedback to be performed dexterously. Consider tying a shoelace or fastening buttons. We rely on touch to perform these tasks correctly, which is why it is difficult to tie one's shoes while wearing gloves. This has motivated the inclusion of proprioceptive and somatosensory feedback into BMI systems.

In the experiments considered in this thesis, we use a continuous decoder to control the 2D velocity of a computer cursor on a screen (Fig. 2.1). At each moment in time, the cursor's velocity, $\mathbf{v}_t \in \mathbb{R}^2$, is determined by the simultaneous firing activity, $\mathbf{u}_t \in \mathbb{R}^d$, of $d \approx 90$ neural units recorded in primary motor cortex using a Utah array:

$$\mathbf{v}_t = A\mathbf{v}_{t-1} + M\mathbf{z}_t + \mathbf{c}$$

where $A \in \mathbb{R}^{2\times 2}$, $M \in \mathbb{R}^{2\times 10}$, and $\mathbf{c} \in \mathbb{R}^{2\times 1}$ are the parameters of the BCI decoder, or "mapping." More details can be found in Chapter A.

While BCI technologies can serve as a neural engineering solution to replace or restore motor or sensory function to patients with neurological injury or disease, they can also serve as a general tool for understanding the relationship between neural activity and behavior. This is because the BCI paradigm provides us with the means to causally relate the neural activity we record in the experiment with animals' performance at the task (Golub et al., 2016).

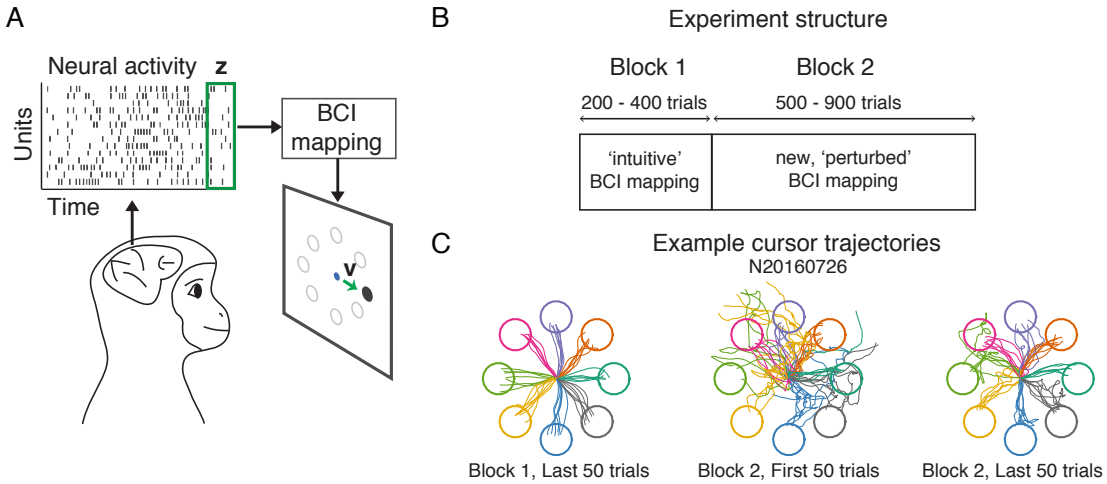


Figure 2.1. Studying learning using a brain-computer interface (BCI). A. Monkeys controlled the velocity of a visual cursor ($\mathbf{v}$) by modulating the spike counts ($\mathbf{z}$) recorded from a multi-electrode array implanted in primary motor cortex. The relationship between spike counts and the cursor's velocity was determined by the BCI mapping. Monkeys were rewarded for successfully guiding the cursor to hit the visually instructed target. **B.** Each experiment consisted of two blocks of trials. In Block 1, monkeys used an intuitive BCI mapping, during which they exhibited proficient cursor control. Later, in Block 2, monkeys learned to control a new BCI mapping they had not used before. **C.** A snapshot of the paths taken by the cursor for trials to each target throughout an example experiment. Early in learning ("Block 2, First 50 trials"), the monkey's control of the cursor was poor. But after a few hundred trials of practice ("Last 50 trials"), monkeys had learned to modify the neural activity they produced, allowing them to achieve more proficient cursor control.

3 Constraints on neural redundancy

Millions of neurons drive the activity of hundreds of muscles, meaning many different neural population activity patterns could generate the same movement. Studies have suggested that these redundant (i.e., behaviorally equivalent) activity patterns may be beneficial for neural computation. However, it is unknown what constraints may limit the selection of different redundant activity patterns. We leveraged a brain-computer interface, allowing us to define precisely which neural activity patterns were redundant. Rhesus monkeys made cursor movements by modulating neural activity in primary motor cortex. We attempted to predict the observed distribution of redundant neural activity. Principles inspired by work on muscular redundancy did not accurately predict these distributions. Surprisingly, the distributions of redundant neural activity and task-relevant activity were coupled, which enabled accurate predictions of the distributions of redundant activity. This suggests limits on the extent to which redundancy may be exploited by the brain for computation.

Published as: Hennig, J.A., Golub, M.D., Lund, P.J., Sadtler, P.T., Oby, E.R., Quick, K.M., Ryu, S.I., Tyler-Kabara, E.C., Batista, A.P., Byron, M.Y. and Chase, S.M., 2018. "Constraints on neural redundancy" *Elife*, 7, p.e36774

3.1 Introduction

Neural circuits relay information from one population of neurons to another. This relay involves successive stages of downstream neurons reading out the activity of upstream neurons. In many cases, the same activity in the downstream population can be produced by different population activity patterns in the upstream population, a phenomenon termed *neural redundancy*. Redundancy is ubiquitous in neural computation, from sensory input to motor output. For example, during a task where subjects need to discriminate the color of a stimulus while ignoring its orientation (Mante et al., 2013), population activity patterns corresponding to the same color but different orientations are read out equivalently, and are therefore redundant. There is mounting evidence that redundancy in readouts may provide various computational benefits. For example, neural redundancy may allow us to prepare movements without executing them (Kaufman et al., 2014; Elsayed et al., 2016), enable stable

computation despite unstable neural dynamics (Druckmann and Chklovskii, 2012; Murray et al., 2017; Driscoll et al., 2017), and allow the central nervous system to filter out unwanted noise (Moreno-Bote et al., 2014).

To fully utilize the proposed benefits of neural redundancy, the population activity should be allowed to freely vary, as long as the readout of this activity remains consistent with task demands. This would allow the population activity to perform computations that are not reflected in the readout. However, a commonly held assumption is that neural activity might also be constrained by energetics: All things being equal, if two population activity patterns are read out equivalently, the brain should prefer the pattern that requires less energy to produce (Laughlin et al., 1998; Barlow, 1969; Levy and Baxter, 1996). These two lines of reasoning raise the following questions: What principles guide the production of redundant neural activity patterns? Are there constraints on which redundant activity patterns can be produced? If so, this may limit the extent to which neural circuits can exploit the proposed computational benefits of redundancy.

Redundancy has been studied extensively in motor control (Lashley, 1933; Bernstein, 1967), albeit in terms of muscular redundancy rather than neural redundancy. During arm movements, different combinations of muscle activity can lead to the same arm kinematics, meaning these different muscle activity patterns are redundant. Previous work on this muscle redundancy problem has identified two principles guiding the selection of redundant muscle activity. First, because muscle contraction requires energy in the form of ATP, the selected muscle activity should require minimum energy relative to the other redundant options (Thoroughman and Shadmehr, 1999; Huang et al., 2012; Fagg et al., 2002). Second, a minimal intervention strategy has been proposed in which subjects control only the aspects of muscle activity that influence the task outcome, and allow for variability in the aspects of muscle activity that do not influence the task outcome (Scholz and Schöner, 1999; Todorov and Jordan, 2002; Valero-Cuevas et al., 2009). To generate movements, the brain not only needs to deal with muscle redundancy, but also *neural* redundancy, which has been less studied.

One way in which neural redundancy can arise is when there are more elements (neurons or muscles) upstream than downstream. During arm movements, the activity of around thirty muscles in the arm and hand is controlled by tens of thousands of neurons in the spinal cord (Gray, 1918; Feinstein et al., 1955). Those neurons are in turn influenced by millions of neurons in the primary motor cortex and other motor areas (Ettema et al., 1998; Lemon, 2008). Thus, the neural control of arm movement is redundant (Fig. 3.1A), in that different population activity patterns can generate the same movement (Li et al., 2001; Rokni et al., 2007; Ajemian et al., 2013). Can the principles of muscular redundancy inform our understanding of neural redundancy?

A common challenge in studying neural redundancy is that it is typically not known which neural activity patterns are redundant, because we do not know how downstream neurons or muscles read out information. In this study we overcome this problem by leveraging a brain-computer interface (BCI), in which the activity of dozens of neurons is read out as movements of a cursor on a computer screen (Fig.

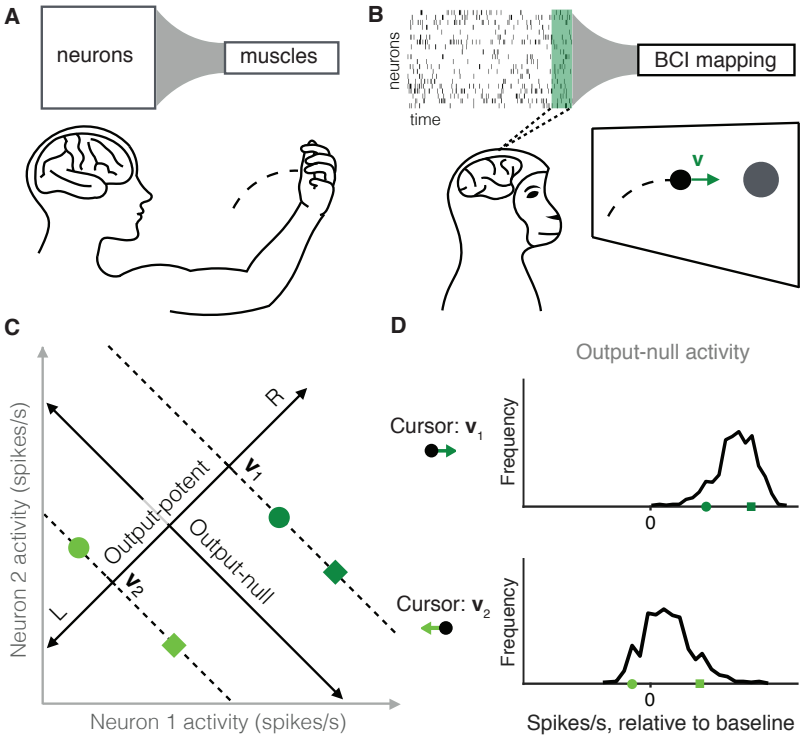


Figure 3.1. Studying the selection of redundant neural activity. (A) Millions of neurons in motor cortex drive tens of muscles to move our arms. Thus, different population activity patterns can be redundant, meaning they produce the same muscle activations and movement. (B) In a BCI, the mapping between neural activity and movement is defined by the experimenter. A subject modulates the spiking activity of tens of neurons (green rectangle) to control the 2D velocity ($\mathbf{v}$) of a cursor on a screen. (C) Example of redundant neural activity in a simplified example where the activity of two neurons (horizontal and vertical axes) drives a 1D cursor velocity (left, L, or right, R). For each of the population activity patterns shown (green squares and circles), the component of the activity along the "Output-potent" axis determines the cursor velocity (e.g., $\mathbf{v}_1$ or $\mathbf{v}_2$), while the position of this activity along the orthogonal axis ("Output-null" axis) has no effect on the cursor's movement. Activity patterns on the same dotted line (e.g., the two dark green patterns) are redundant, because these patterns have the same output-potent activity and produce the same cursor velocity (e.g., $\mathbf{v}_1$). (D) Example distributions of neural activity along the output-null dimension (corresponding to dotted lines in (C)). Each black trace depicts the density of output-null activity observed over the course of an experiment when the cursor velocity was $\mathbf{v}_1$ (top) or $\mathbf{v}_2$ (bottom). The output-null activity of the green symbols from (C) are marked for reference. In the actual experiments, there were two output-potent dimensions and eight output-null dimensions. Output-null activity has units of spikes/s, presented relative to the vector of mean activity for each neuron ("baseline").

3.1B) (Taylor et al., 2002; Carmena et al., 2003; Hochberg et al., 2006; Ganguly and Carmena, 2009; Gilja et al., 2012; Hauschild et al., 2012; Sadtler et al., 2014). A key advantage of a BCI is that the readout of the population activity (termed the BCI mapping) is fully known and defined by the experimenter (Golub et al., 2016). This allows us to determine precisely the redundant set of population activity patterns, i.e., those that move the cursor in exactly the same way. To illustrate this, consider a simplified example where the activity of two neurons controls a 1D cursor velocity (Fig. 3.1C). The two dark green activity patterns produce the same cursor movement ($\mathbf{v}_1$), and the two light green patterns produce a different movement ($\mathbf{v}_2$). We can decompose any population activity pattern into two orthogonal components: output-potent activity and output-null activity (Fig. 3.1c, black axes) (Law et al., 2014; Kaufman et al., 2014). The output-potent component determines the cursor's movement, whereas the output-null component has no effect on the cursor. Two population activity patterns are redundant if they have the same output-potent activity, but different output-null activity (e.g., the dark green square and circle on the "$\mathbf{v}_1$" dotted line in Fig. 3.1C). The question we address here is, which redundant population activity patterns were preferred by the nervous system? To answer this, we assessed the distribution of output-null activity produced during each cursor movement (Fig. 3.1D), and compared it to what we would expect to observe under each of several candidate hypotheses for explaining neural redundancy.

We trained three Rhesus macaques to perform a brain-computer interface task in which they controlled the velocity of a cursor on a computer screen by volitionally modulating neural activity in primary motor cortex. To understand the principles guiding the selection of redundant neural activity, we compared the observed distributions of output-null activity to those predicted by three different hypotheses. The first two hypotheses we considered were inspired by studies of muscle redundancy. First, by analogy to minimum energy principles (Thoroughman and Shadmehr, 1999; Huang et al., 2012; Fagg et al., 2002), neural activity may minimize unnecessary spiking (Barlow, 1969; Levy and Baxter, 1996). Second, by analogy to the minimal intervention strategy (Scholz and Schöner, 1999; Todorov and Jordan, 2002; Valero-Cuevas et al., 2009), output-null activity might be uncontrolled (i.e., output-potent activity is modified independently of output-null activity) because neural variability in this space has no effect on cursor movement. Third, we considered the possibility that the distribution of redundant activity may be coupled with the task-relevant activity, so that producing particular activity patterns in output-potent dimensions requires changing the distribution of activity in output-null dimensions.

We tested all hypotheses in terms of their ability to predict the distribution of output-null activity, given the output-potent activity. Hypotheses were tested within the space in which the population activity naturally resides, termed the *intrinsic manifold* (Sadtler et al., 2014). The results of Sadtler et al. (2014) indicate that neural activity cannot readily leave this manifold, and more recent results demonstrate that neural activity is further constrained by a *neural repertoire* within the intrinsic manifold (Golub et al., 2018). However, a repertoire defines only a set of population activity patterns, and not how often different activity patterns within the repertoire

are produced. Therefore, to understand the principles governing the selection among redundant population activity patterns, we focused on predicting the *distribution* of redundant population activity within the intrinsic manifold and neural repertoire.

We found strong evidence for the third hypothesis, that redundant activity is coupled with task-relevant activity. This indicates that neural redundancy is resolved differently than muscular redundancy. Furthermore, the output-null space should not be thought of as a space in which neural activity can freely vary to carry out computations without regard to the output-potent activity. Instead, the distribution of output-null activity is constrained by the corresponding output-potent activity. If the required output-potent activity is defined by the task demands, this can constrain how the output-null activity can vary, and correspondingly the computations that can be carried out in the output-null space.

3.2 Results

To study the selection of redundant neural activity, we used a BCI based on 85-94 neural units recorded using a Utah array in the primary motor cortex in each of three Rhesus macaques. Animals modulated their neural activity to move a computer cursor in a 2D center-out task (see Methods; Supplemental Fig. A.1). At the beginning of each experiment, we identified the 10 dimensions of the population activity that described the largest activity modulations shared among the neural units, termed the *intrinsic manifold* (Sadtler et al., 2014). A two-dimensional subspace of the 10-dimensional intrinsic manifold was mapped to horizontal and vertical cursor velocity and was therefore output-potent, while the eight orthogonal dimensions were output-null. Our goal was to predict the joint distribution of the observed neural activity in this eight-dimensional output-null space.

We tested several hypotheses for the selection of redundant neural activity using the following logic. First, we predicted the distributions of output-null activity expected under each hypothesis. All hypotheses' predictions were consistent with the observed behavior (i.e., the output-potent activity), and we ensured that none of these predictions required unrealistic firing rates when combined with the output-potent activity. Next, we compared the predicted distributions to the observed distributions of output-null activity to determine which hypothesis provided the best match to the observed distributions. We built the observed distributions of output-null activity as follows: At each time step during the BCI task, we assigned the recorded population activity pattern to one of eight bins corresponding to the direction of cursor movement (0°, 45°, 90°, etc.) produced by that neural activity. We binned by the cursor movement because we are studying the population activity that is redundant for a given cursor movement direction. For each bin, we projected the corresponding population activity patterns onto the eight output-null dimensions of the intrinsic manifold. The black histograms in Fig. 3.2, Fig. 3.3, and Fig. 3.4 show the marginal distributions in the first three output-null dimensions (ordered by variance accounted for). The colored histograms in Fig. 3.2, Fig. 3.3, and Fig.

3.4 are the predicted output-null distributions built under each hypothesis, which we compared to the observed distributions. The ensuing three subsections describe each hypothesis, and compare how well the corresponding predicted distributions matched the observed distributions.

During each experiment, animals controlled two different BCI mappings (i.e., the two mappings had different output-potent subspaces). The first mapping was an "intuitive" one that required no learning for proficient control. The second mapping was a within-manifold perturbation (see Methods). For the second mapping, we analyzed the trials after the behavioral performance had asymptoted. Each hypothesis predicted the distribution of output-null activity that the animal would produce under the second mapping. To form its prediction, a hypothesis could utilize the output-potent activity observed during the second mapping, as well as all neural activity recorded under control of the first mapping. This technique allowed us to avoid circularity in our results because we built the hypothesized distributions using the first behavioral context and evaluated those predictions in the second. Additionally, because animals learned to use the BCI mappings through trial and error, it is possible that the animals' assumptions about the output-null dimensions do not align perfectly with the actual output-null dimensions of the BCI mapping. To control for this, we estimated the animal's internal model of the BCI mapping (Golub et al., 2015). The results in the main text are based on this internal model, and we show in supplemental figures that all results still hold when using the actual BCI mapping.

Minimal firing hypotheses do not accurately predict output-null activity

Previous work in motor control has found that subjects select muscle activations that minimize energy use, i.e., subjects tend not to make movements with more stiffness or muscular co-contraction than necessary to complete the task (Thoroughman and Shadmehr, 1999; Fagg et al., 2002; Huang et al., 2012). We tested whether an analogous principle might hold true at the level of neurons (Fig. 3.2A, *Minimal Firing* hypothesis). Because spiking incurs a metabolic cost (Laughlin, 2001; Laughlin et al., 1998), we first considered the hypothesis that among all the population activity patterns that produce the same cursor movement, the subject will select the one requiring the fewest spikes (Barlow, 1969; Softky and Kammen, 1991; Levy and Baxter, 1996).

To predict the distribution of output-null activity under this hypothesis, at each time step we found the population activity pattern that would produce the observed cursor movement with the fewest spikes across all recorded neurons (see Methods). This means population activity will have minimal variability in output-null dimensions, because spiking in these dimensions does not affect cursor movement. In Fig. 3.2A, the orange square depicts the activity pattern nearest zero spikes/s (gray square) among all activity patterns that would produce the same cursor movement (black dotted line). This would produce a delta distribution of output-null activity, where the delta would be located at the predicted value (orange square). To make

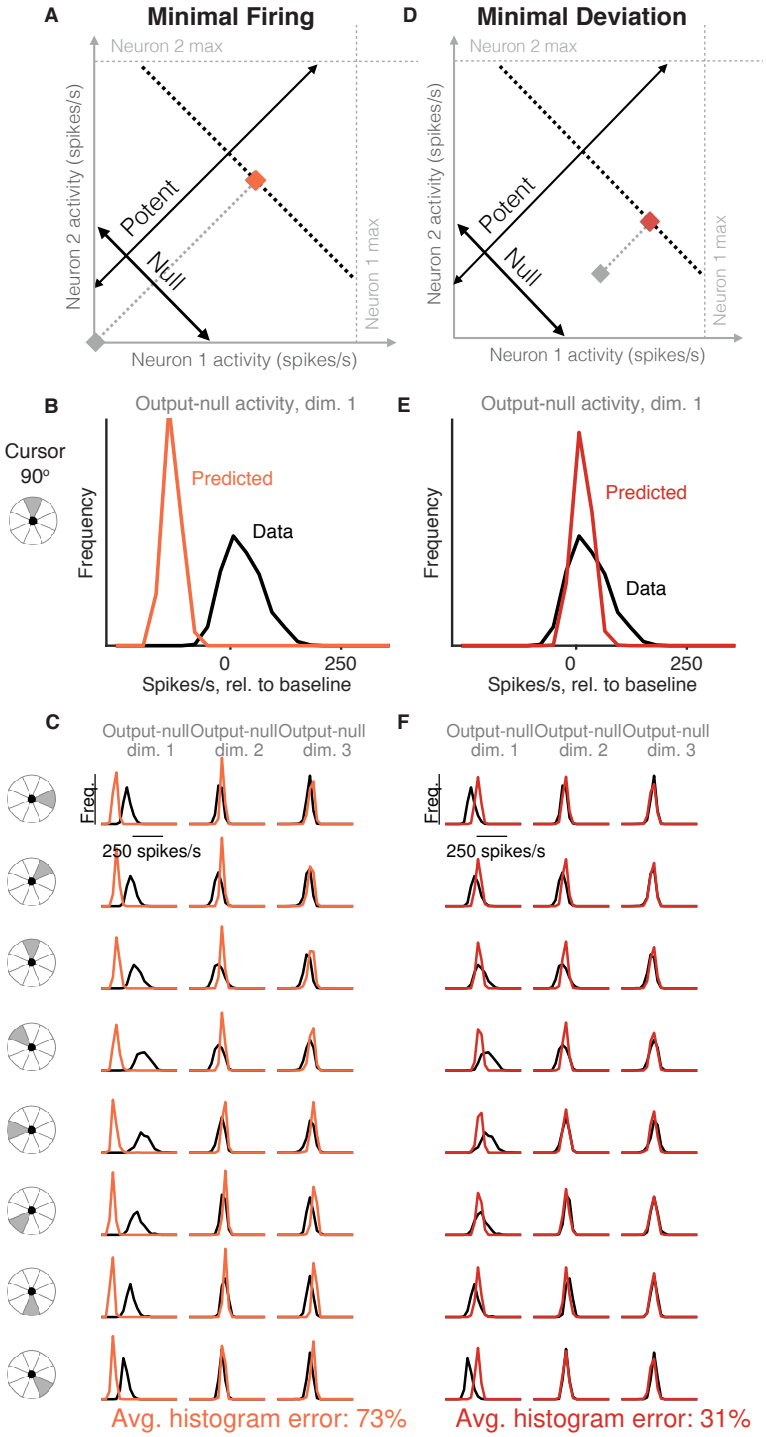


Figure 3.2. Minimal firing hypotheses. (A) Minimal Firing hypothesis: Given a particular output-potent activity (i.e., activity is constrained to black dotted line), subject selects the activity pattern (orange square) that requires the fewest spikes (i.e., nearest the gray square). (B) Distribution of observed output-null activity ("Data", in black) and activity predicted by the Minimal Firing hypothesis ("Predicted", in orange), in the first output-null dimension for upwards cursor movements. For this visualization, we applied PCA to the observed output-null activity to display the dimensions ordered by the amount of shared variance, with only the first of those dimensions shown here. The range of activity (e.g., ± 150 spikes/s) appears larger than that expected for a single neuron because the range tends to increase with the number of neural units contributing to that dimension. Session L20131218. (C) Distributions of observed and predicted output-null activity as in (B), for time steps when the cursor was moving in eight different directions (rows), in three (of eight) output-null dimensions explaining the most output-null variance (columns). (D) Minimal Deviation hypothesis: Given a particular output-potent activity, subject selects the activity pattern (red square) nearest a fixed population activity pattern chosen for each session by cross-validation (gray square). (E-F) Same conventions as in (B-C) for the Minimal Deviation hypothesis.

this prediction more realistic, we incorporated Poisson spiking noise. In addition, for this hypothesis and those following, we ensured that all predictions were physiologically plausible (i.e., firing rates were between zero and the maximum rates observed in the experiment; see Methods).

We constructed histograms of the output-null activity predicted by the Minimal Firing hypothesis by pooling over all time steps in which the cursor moved in a similar direction (e.g., 0°, 45°, etc.) (Fig. 3.2B, orange). We compared these predicted distributions to the observed distributions of output-null activity measured for that movement direction during the experiment (Fig. 3.2B, black). Fig. 3.2C depicts these histograms for the same session across eight different cursor directions (rows), in three of the eight output-null dimensions (columns). For visualization, we applied principal components analysis (PCA) to display the output-null dimensions ordered by the amount of shared variance in the output-null activity. To assess how well the Minimal Firing hypothesis predicted the observed output-null activity, we computed the absolute error between the predicted and observed histograms. These errors were averaged across histograms for all eight cursor directions and eight output-null dimensions in a given session. We normalized the errors so that a perfect match between the observed and predicted histograms would result in an error of 0%, while complete mismatch between the predicted and observed histograms would yield an error of 100% (see Methods). We found that the predictions of the Minimal Firing hypothesis differed from the observed activity by 73.2% ± 1.3% (mean ± SE) across sessions.

One possible explanation as to why these predictions were so different from the observed activity is that minimal energy principles in the brain may not equate to minimal spiking. Perhaps a more relevant constraint is not how far the activity is away from zero firing, but rather how far the activity is from a different level of activity, such as the mean firing rate for each neuron. This alternative version of a minimal energy hypothesis (Fig. 3.2D, *Minimal Deviation* hypothesis) predicts that among all the population activity patterns that produce the same cursor movement, subjects select the one with the smallest deviation from some baseline population activity pattern. For each session, we identified the population activity pattern that would minimize the output-null prediction error across cursor directions in a cross-validated fashion (see Methods) (Fig. 3.2E). This hypothesis yielded an average histogram error of 30.9% ± 1.2% (mean ± SE) across sessions. While this represents a substantial improvement over the Minimal Firing hypothesis (paired t-test of histogram errors in each session, $p < 0.001$), the predicted distributions of output-null activity still show clear discrepancies from the observed distributions (Fig. 3.2F). Thus, we sought a hypothesis that could better predict the observed distributions of output-null activity.

Uncontrolled hypotheses do not accurately predict output-null activity

It has been shown that muscle activity exhibits more variability in output-null dimensions than in output-potent dimensions (Scholz and Schöner, 1999; Todorov and Jordan, 2002; Valero-Cuevas et al., 2009). An explanation of this variability asymmetry is the "minimal intervention" principle (Todorov and Jordan, 2002; Valero-Cuevas et al., 2009; Diedrichsen et al., 2010b), which states that while variability in output-potent dimensions should be corrected to ensure task success, variability in output-null dimensions can be left uncorrected because it does not lead to deficits in task performance. While this principle has been used to explain muscle activity, here we investigate whether it also explains neural activity. This hypothesis, that output-null activity will be "uncontrolled" and have high variability, is in contrast to the minimal firing hypotheses, which predict that output-null activity will have low variability.

The idea that neural activity may be selected according to a minimal intervention principle does not, by itself, specify the form of the distribution in output-null dimensions. We therefore considered two specific forms of uncontrolled hypotheses. First, we supposed that if all values of output-null activity are equally likely, then output-null activity would have a uniform distribution with bounds determined by each neuron's physiological range (Fig. 3.3A, *Uncontrolled-uniform*). We emphasize that the minimal intervention principle does not specify a candidate distribution, and so we consider this particular hypothesis as a limiting case, where output-null activity has maximum entropy within bounds on minimum and maximum activity. At each time step, we sampled the output-null activity from a uniform distribution within ranges observed experimentally (see Methods). This procedure predicts that the output-null activity is selected independently of the current output-potent activity, reflecting the minimal intervention principle. However, note that the extent of the uniform distribution depends on the physiological range of each neuron, and so the predicted distributions of output-null activity vary slightly with the cursor direction (Fig. 3.3B-C) (e.g., the length of the green bar in Fig. 3.3A depends on the output-potent activity). As before, for visualization we ordered the eight output-null dimensions by the amount of shared variance explained in the recorded activity, and displayed the first three of these output-null dimensions (Fig. 3.3C). Because these three dimensions were rotated along the dimensions of highest variance, the predicted histograms are mound-shaped rather than uniformly distributed (see Methods). The predictions of the Uncontrolled-uniform hypothesis differed from the observed output-null activity by 56.6% $\pm$ 1.1% (mean $\pm$ SE) across sessions.

In the second variant of this hypothesis, we considered a non-uniform distribution of output-null activity. If the natural variability of output-null activity is truly unmodified, then the distribution of activity observed in the same dimensions when a subject was controlling a different (previous) BCI mapping should have the same distribution under the current mapping (Fig. 3.3D, *Uncontrolled-empirical*). Thus, under this hypothesis we construct an empirical distribution of output-null activity,

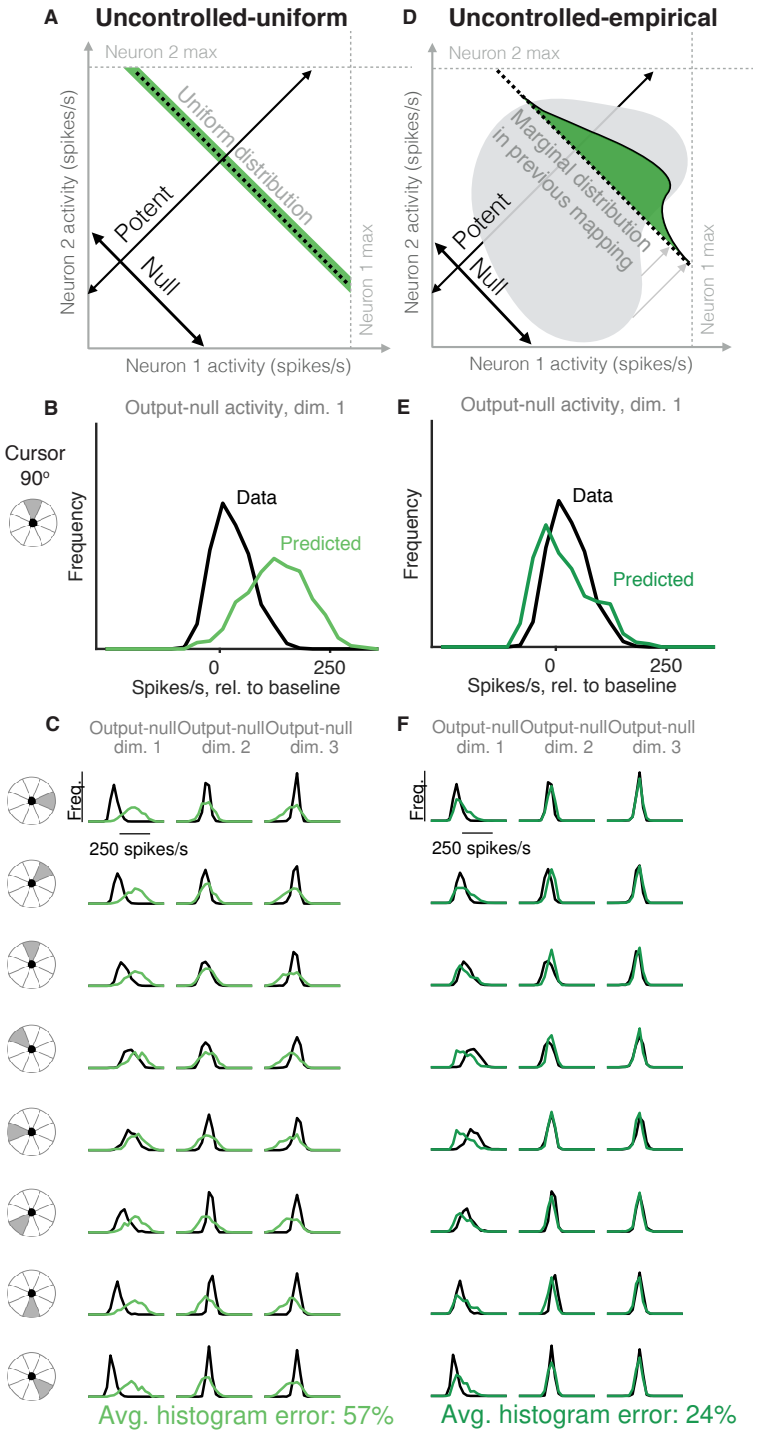


Figure 3.3. Uncontrolled hypotheses. (A) Uncontrolled-uniform hypothesis: Given a particular output-potent activity, subject selects any activity within the physiological range (dark green), sampled from a uniform distribution. (B-C) Distributions of output-null activity observed and predicted by the Uncontrolled-uniform hypothesis; same conventions as in Fig. 3.2. The predicted distributions appear mound-shaped rather than uniform because we applied PCA to display the dimensions of output-null activity with the most shared variance (see Methods). The range of activity increases with the number of neural units. Session L20131218. (D) Uncontrolled-empirical hypothesis: Subject selects output-null activity from the distribution of all output-null activity produced at any time while subjects used a different BCI mapping. (E-F) Same conventions as in (B-C) for the Uncontrolled-empirical hypothesis.

which we form by projecting all of the population activity that the subject produced under the *previous* mapping onto the output-null dimensions of the *current* BCI mapping (see Methods). At each time step, we sampled from this empirical distribution of output-null activity independently of the output-potent activity, again reflecting the minimal intervention principle (Fig. 3.3D). We checked that combining the output-null and output-potent activity resulted in physiologically plausible population activity (see Methods). If it did not, then we re-sampled a different output-null activity pattern until the combination resulted in physiologically plausible population activity. Due to this resampling, the predicted distributions of output-null activity vary slightly with the cursor direction (Fig. 3.3E-F). The histograms of the predictions differed from the observed data by only 23.8% ± 0.8% (mean ± SE) across sessions, which is the lowest error of all hypotheses considered so far. This suggests that previously observed population activity (in this case, recorded during use of a different BCI mapping) offers greater predictive power of the selection of output-null activity than a priori predictions such as those of the Minimal Firing, Minimal Deviation, and Uncontrolled-uniform hypotheses.

Task-transfer hypotheses accurately predict output-null activity

Thus far, the hypothesis that best predicts the observed output-null activity is the one that uses previously observed activity to generate its predictions (Uncontrolled-empirical). This motivated us to consider more refined hypotheses that make use of this previously observed activity to generate predictions.

We first considered the hypothesis that in order to produce a desired movement, the subject selects neural activity as if he were still using the previous mapping, and corrects this activity only to ensure task success (Fig. 3.4A, *Persistent Strategy*). Conceptually, when the subject wants to move the cursor in a particular direction using the current BCI mapping, he starts with the population activity patterns that he used to move the cursor in that direction under an earlier mapping (Fig. 3.4A, light blue shading). Because this activity will not move the cursor in the same way that it did under the previous mapping, this activity is modified along the output-potent dimensions of the current mapping (Fig. 3.4A, red arrows), reflecting the minimal intervention principle (Todorov and Jordan, 2002; Valero-Cuevas et al., 2009; Diedrichsen et al., 2010b). This is similar to the Uncontrolled-empirical hypothesis in that we assume activity in output-null dimensions can be corrected independently of the activity in output-potent dimensions. However, instead of sampling from the entire distribution of previously observed output-null activity at each time step, here we only sample from the subset of this activity observed when subjects needed to move the cursor in the same direction as for the current time step. The predictions of this hypothesis (Fig. 3.4B-C) differed from the observed output-null activity by 17.4% ± 0.7% (mean ± SE) across sessions.

The principle of minimal intervention posits that output-null activity can change independently from output-potent activity. Here we examine this assumption in de-

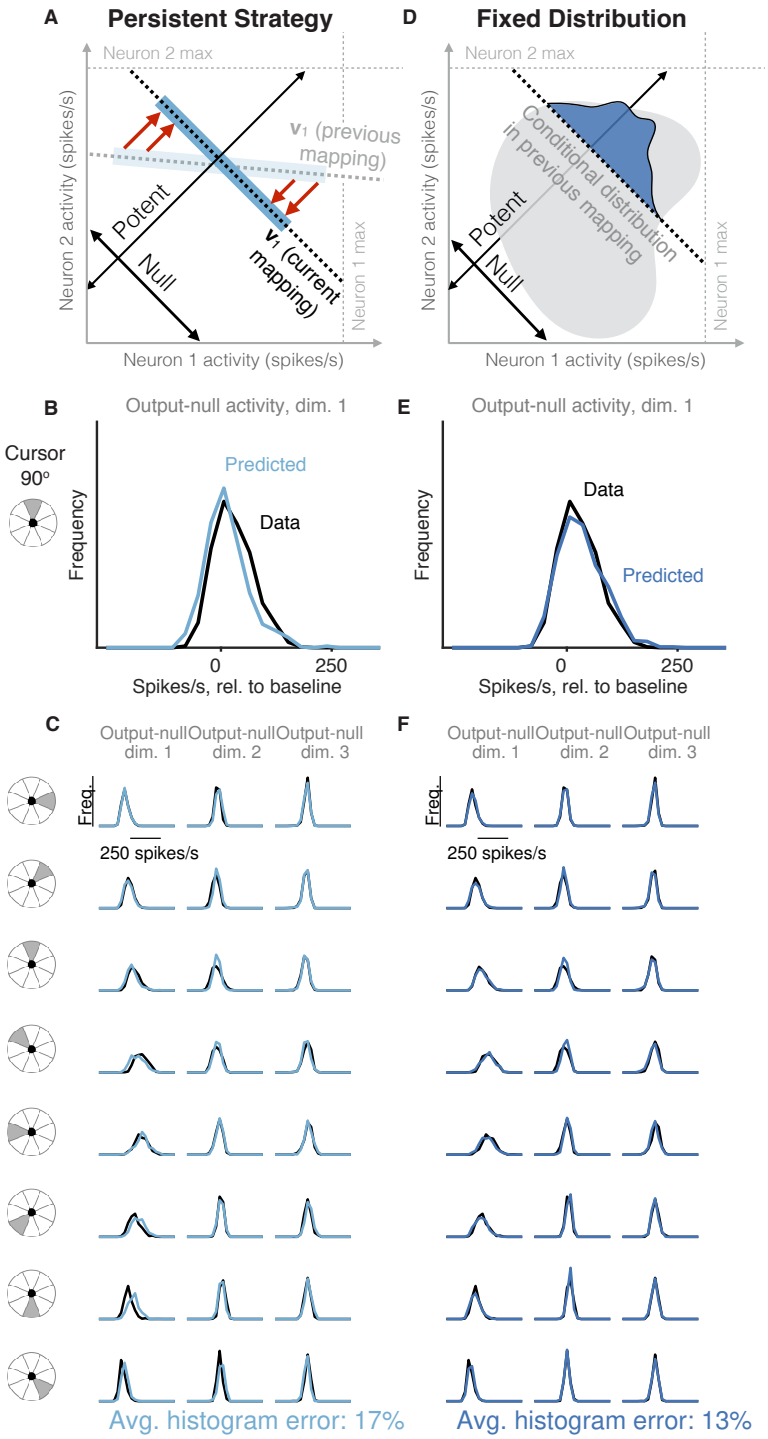


Figure 3.4. Task-transfer hypotheses. (A) Persistent Strategy hypothesis: Given a particular output-potent activity, subject selects an activity pattern appropriate under a different mapping (light blue rectangle), and corrects its output-potent component (red arrows) so as to produce the desired output-potent value under the current mapping (darker blue rectangle). (B-C) Distributions of output-null activity observed and predicted by the Persistent Strategy hypothesis; same conventions as in Fig. 3.2. The range of activity increases with the number of neural units. Session L20131218. (D) Fixed Distribution hypothesis: Given a particular output-potent activity, subject selects from the output-null activity patterns that were observed concurrently with this output-potent activity while controlling a different mapping. Different patterns are selected with the same frequencies as they were under the previous mapping. (E-F) Same conventions as in (B-C) for the Fixed Distribution hypothesis.

tail. Previous work has found that the characteristic ways in which neurons covary (i.e., the dimensions of the intrinsic manifold) persist even under different BCI mappings, perhaps owing to underlying network constraints (Sadtler et al., 2014). All hypotheses we consider here are evaluated within the intrinsic manifold, and thus respect these constraints on population covariability. Because the dimensions of the intrinsic manifold capture the covariability among the neurons, it is plausible that the activity along different dimensions of the intrinsic manifold can vary independently, consistent with the minimal intervention principle. By contrast, in the next hypothesis we consider the possibility that activity along different dimensions exhibit dependencies.

We considered the hypothesis that the distribution of activity in output-null dimensions would be predictably coupled with the activity in output-potent dimensions, even under a different BCI mapping when those dimensions were not necessarily potent and null. Under this hypothesis (Fig. 3.4D, *Fixed Distribution*), given the output-potent activity, the distribution of the corresponding output-null activity remains the same as it was under a different BCI mapping (Fig. 3.4D, blue frequency distribution), even if this activity was not output-null under the other mapping. This hypothesis predicts that neural activity patterns are "yoked" across dimensions, such that producing particular activity in output-potent dimensions requires changing the distribution of activity in output-null dimensions. The histograms of output-null activity predicted by the Fixed Distribution hypothesis were a striking visual match to the recorded activity, and accurately predicted the dependence of these distributions on the cursor direction (Fig. 3.4E-F). Overall, these predictions differed from the observed output-null activity by only 13.4% ± 0.5% (mean ± SE) across sessions.

The Fixed Distribution hypothesis yielded a lower histogram error than all other hypotheses across sessions from three different animals (Fig. 3.5A). In total, the Fixed Distribution hypothesis had the lowest histogram error in 41 of 42 sessions. The histogram error metric does not explicitly capture the degree to which hypotheses predicted the mean output-null activity, or any correlations that exist across output-null dimensions. We therefore assessed how well the predictions captured the mean and covariance of observed data in all output-null dimensions jointly (see Methods). In agreement with our findings for histogram error, the mean (Fig. 3.5B) and covariance (Fig. 3.5C) of output-null activity was best predicted by the Fixed Distribution hypothesis, with an average mean error of 23.5 ± 1.4 spikes/s (mean ± SE) and an average covariance error of 1.4 ± 0.1 (mean ± SE in arbitrary units; see Methods). These error metrics offer further evidence that the Fixed Distribution hypothesis provides a good match to the output-null distribution, as measured by the agreement between the first and second moments of the two distributions. Because these error metrics rely on a limited number of trials, they should not be compared relative to zero error. We estimated the smallest histogram, mean, and covariance errors achievable by any hypothesis, given the limited number of samples available to estimate the true output-null distributions (see Methods, and gray regions in Fig. 3.5). The errors of Fixed Distribution were exceedingly close to the lowest achievable error given the number of samples available (see Methods). Next, we found that the

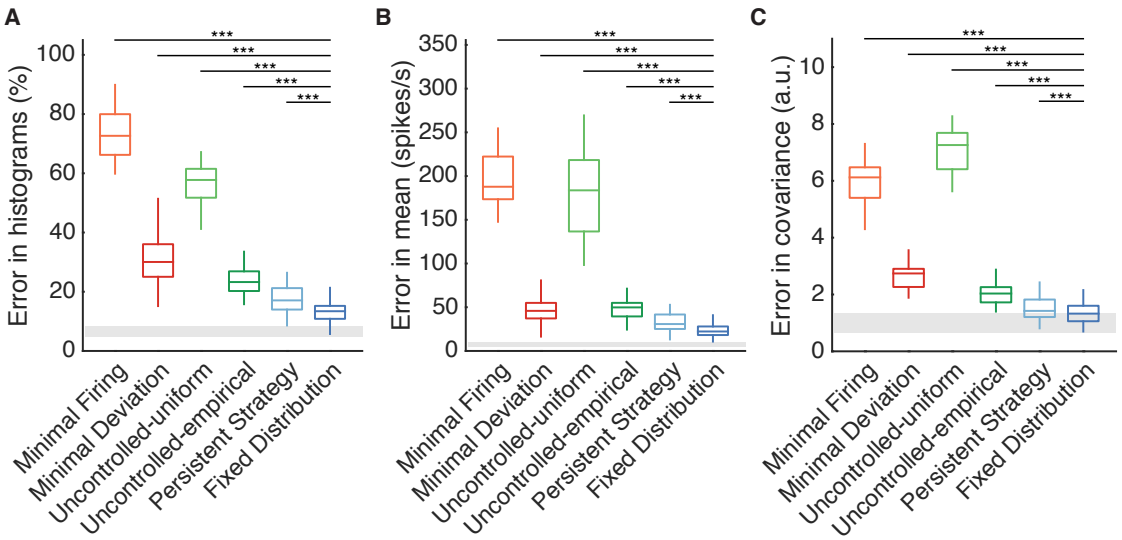


Figure 3.5. Fixed Distribution hypothesis best predicts output-null activity. Boxes depict the 25th, 50th, and 75th percentile of errors observed across sessions for all animals combined. Whiskers extend to cover approximately 99.3% of the data. Gray boxes depict the error floor across sessions (mean ± s.d.), estimated using half of the observed output-null activity to estimate the histogram, mean, and covariance of the other half (see Methods). Asterisks depict a significant difference between errors of Fixed Distribution and other hypotheses for a one-sided Wilcoxon signed rank test at the $\alpha = 0.001$ (***) level. (A) Error in predicted histograms of output-null activity. For each session, histogram error was averaged across all output-null dimensions and cursor directions. Average histogram error floor was 6.7% ± 1.9% (mean ± s.d., also shown in gray). (B) Error in predicted mean of output-null activity. For each session, mean error was averaged across all cursor directions, where the mean is an 8D vector of the average activity in each output-null dimension. Average mean error floor was 6.9 ± 2.5 spikes/s (mean ± s.d., also shown in gray). (C) Error in predicted covariance of output-null activity. For each session, covariance error was averaged across all cursor directions. Average covariance error floor was 1.0 ± 0.3 (mean ± s.d., also shown in gray).

Fixed Distribution hypothesis achieved the lowest prediction errors among all hypotheses when data for each monkey was considered individually (Supplemental Fig. A.2). We repeated our analyses to predict output-null activity produced during the first mapping using activity observed during the second mapping (Supplemental Fig. A.3). We also predicted output-null activity using the actual BCI mapping rather than the animal's internal model to define the output-null dimensions (Supplemental Fig. A.4). Both analyses yielded results similar to those in Fig. 3.5.

Predicting changes in neural variability when activity becomes output-null

So far we have shown that the Fixed Distribution hypothesis provides a better explanation for the structure of output-null activity than hypotheses incorporating constraints on firing rates or the minimal intervention principle. We next sought stronger evidence for the Fixed Distribution hypothesis by assessing our predictions in the particular dimensions of population activity where it is least likely to hold. Because cursor velocity is a two-dimensional quantity, all but two dimensions of population activity for each BCI mapping are output-null. Thus, given two different BCI mappings, most dimensions will be output-null under both mappings, and so most components of the population activity have no reason to change from one mapping to the other. Therefore, we assessed whether our results held in dimensions of population activity that were output-potent during the first mapping, but output-null during the second mapping (see Methods). These are the dimensions in which one would expect to see the most changes in the population activity between the first and second mappings.

Our hypotheses make distinct predictions about how the variance of activity should change if a dimension is output-potent under the first mapping and becomes output-null under the second mapping. For example, according to the Minimal Firing and Minimal Deviation hypotheses, the variance of activity will collapse in dimensions that are output-null because unnecessary spiking is undesirable. Thus, if a dimension becomes output-null, variance in this space should exhibit a marked decrease. On the other hand, the Uncontrolled hypotheses predict that, when conditioned on the cursor movement, variance will expand when the activity is output-null. This occurs because variability in this dimension will no longer affect cursor movement, and would therefore no longer need to be suppressed. Finally, the Fixed Distribution hypothesis posits that the same distributions of output-null activity will be observed regardless of whether a dimension was previously output-potent or output-null, and so this hypothesis predicts that there will be little to no change in the variance of activity in a particular dimension under the two mappings.

We asked whether the variance of population activity decreased, increased, or remained the same in dimensions that changed from being output-potent to output-null (Fig. 3.6A). Critically, we computed the variance of activity after first binning by the corresponding angle in the output-potent dimensions of the *second* mapping. This was done so that the neural activity in each bin would all result in similar

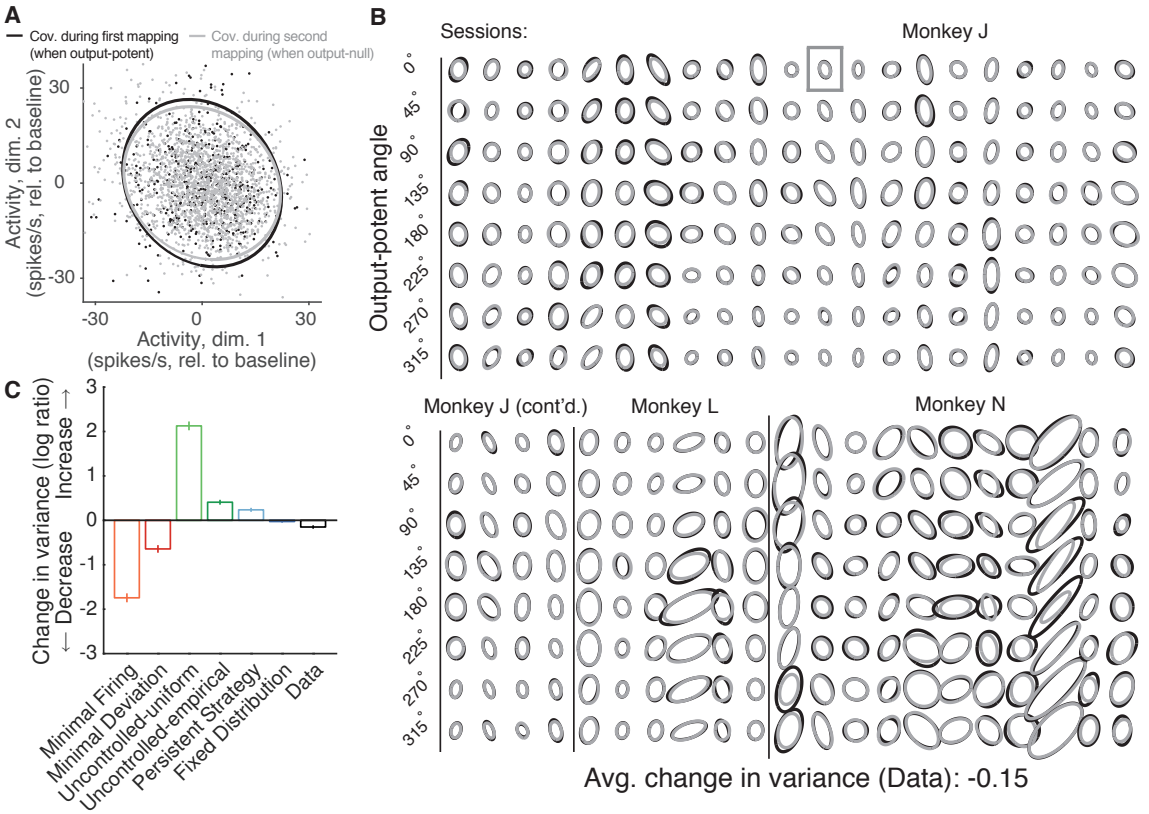


Figure 3.6. Variance of neural activity in dimensions that become output-null. (A) Observed activity from a representative session in the 2D subspace in which activity was output-potent under the first mapping and output-null under the second mapping. Activity recorded during use of the first mapping (black points) was output-potent while activity recorded during use of the second mapping (gray points) was output-null. The covariances during the first and second mapping (black and gray ellipses, respectively) are depicted as the 95% contours of a Gaussian density fit to the activity. Session J20120403, for all time steps when the activity would have moved the cursor to the right under the second mapping. (B) Covariance ellipses for all sessions and eight different cursor movement angles. Same conventions as in (A). Ellipses shown in (A) indicated by gray box. (C) Change in variance of neural activity in the same subspace as in (A), for the activity observed ("Data") and predicted by each hypothesis. Height of bars depicts the average change in variance across sessions (mean ± 2 SE).

cursor movements under the second mapping, and is identical to the procedure used previously to assess the errors of the hypotheses' predictions. Notably, binning in this way means that each bin may contain activity corresponding to different cursor movements under the *first* mapping, and so one might expect that in each bin the activity recorded under the first mapping would be more heterogeneous than the activity recorded under the second mapping.

We observed that the variance of population activity recorded under the first and second mappings was remarkably similar in the dimensions that changed from output-potent to output-null, even though these activity patterns usually corresponded to different cursor movements under the two mappings (Fig. 3.6B). Thus, the variance of activity did not change much when an output-potent dimension became output-null, in agreement with the predictions of the Fixed Distribution hypothesis. To quantify these observations, we computed the average change in variance in each session (see Methods). Across sessions, we found that the variance of observed activity showed a small but significant decrease when it became output-null (Fig. 3.6C, "Data") (t-test, $p < 0.001$). This is in contrast to the predictions of the Minimal Firing and Minimal Deviation hypotheses, which predicted much larger decreases.

The observed change in variance lies closest to the predictions of the Fixed Distribution hypothesis. In fact, we observed that the Fixed Distribution hypothesis also predicted a slight decrease in variance in dimensions that became output-null (Fig. 3.6C, "Fixed Distribution") (t-test, $p < 0.001$). This slight predicted change in variance occurs because the distributions of activity in the output-potent dimensions of the second mapping are different under the first and second mappings. Because the Fixed Distribution hypothesis predicts a fixed conditional distribution of output-null activity given the output-potent activity, slightly different sets of output-potent activity will result in a slightly different distribution of the corresponding output-null activity.

These analyses show that, contrary to the predictions of the minimal firing and uncontrolled hypotheses, the variance of population activity did not change dramatically in dimensions that were output-potent under the first mapping and output-null under the second mapping. We also assessed whether the reverse was true—if the variance of activity changed in dimensions that began as output-null and became output-potent. To measure this, we repeated the above analyses after predicting output-null activity produced during the first mapping using the activity observed under the second mapping (as in Supplemental Fig. A.3). We found that the activity showed little to no change in variance in these dimensions (t-test, $p > 0.5$), in agreement with the predictions of Fixed Distribution (Supplemental Fig. A.6).

Importantly, the agreement between the observed output-null activity and the predictions of the Fixed Distribution hypothesis in these analyses indicates that our ability to accurately predict the distribution of output-null activity is not merely a result of most activity being output-null under both mappings. Instead, the distribution of output-null activity remains consistent with the Fixed Distribution hypothesis even in the output-null dimensions that were previously output-potent.

In Fig. 3.6C, the observed output-null activity showed a larger decrease in variance than the predictions of the Fixed Distribution hypothesis, at least in the 2D subspace of output-null activity that was output-potent during the first mapping. This slight decrease in variance is in the direction of the predictions of Minimal Firing and Minimal Deviation. If this decrease in variance is to be explained by Minimal Firing or Minimal Deviation principles, we would expect that the observed *mean* output-null activity would also move in the direction of the predictions of Minimal Firing and Minimal Deviation, relative to what is predicted by Fixed Distribution. To see if this was the case, we first computed the distance of the observed mean output-null activity from the mean predicted by Minimal Deviation for each movement direction, and compared this to the distance of the mean output-null activity predicted by Fixed Distribution from the mean predictions of Minimal Deviation (Supplemental Fig. A.7A). We did not find evidence that the observed mean output-null activity was closer to the mean predicted by Minimal Deviation than was the mean predicted by Fixed Distribution (one-sided Wilcoxon signed rank test, $p > 0.5$; see Supplemental Fig. A.7B and Methods). Repeating the analysis with Minimal Firing instead of Minimal Deviation yielded similar results (one-sided Wilcoxon signed rank test, $p > 0.5$). Thus, while we observed a slight decrease in the *variance* of output-null activity in dimensions that changed from output-potent to output-null, we did not find any evidence that the *mean* output-null activity moved in the direction of the predictions of Minimal Firing or Minimal Deviation.

3.3 Discussion

Recent work has suggested that neural redundancy may be exploited for various computations (Kaufman et al., 2014; Elsayed et al., 2016; Druckmann and Chklovskii, 2012; Moreno-Bote et al., 2014; Murray et al., 2017; Driscoll et al., 2017). However, if the activity in output-null dimensions is constrained by the output-potent activity, then this may limit the ability of output-null activity to perform computations without affecting the readout. Here, we studied neural redundancy in the primary motor cortex using a BCI, where it is known exactly which population activity patterns are redundant, meaning they produce an identical cursor movement. We generated predictions of the distributions of output-null neural activity for subjects performing a BCI cursor control task, and compared them to the distributions observed in our experiments. We found that hypotheses inspired by minimal firing and minimal intervention principles, drawn from theories of muscle coordination, did not accurately predict the observed output-null activity. Instead, we found that the distribution of output-null activity was well predicted by the activity in the two output-potent dimensions. This coupling between the output-potent and output-null activity implies that, when output-potent activity is used to satisfy task demands, there are constraints on the extent to which neural circuits can use redundant activity to perform additional computations.

Our results indicate that the way in which neural redundancy is resolved is dif-

ferent from how muscle redundancy is resolved. There have been several prevalent proposals for how muscle redundancy is resolved, including minimal energy, optimal feedback control (OFC), and habitual control. Models incorporating minimal energy principles have helped to explain observed gait (McNeill, 2002) and arm reaches (Thoroughman and Shadmehr, 1999; Huang et al., 2012; Fagg et al., 2002; Farshchiansadegh et al., 2016). By analogy, it has been proposed that the brain may prefer an "economy of impulses" (Barlow, 1969; Softky and Kammen, 1991; Levy and Baxter, 1996), resolving neural redundancy by minimizing the production of action potentials. However, we found that minimal energy principles in terms of firing rates do not play a dominant role in the selection of output-null neural activity. Given that metabolic activity can decrease without corresponding changes in firing rates (Picard et al., 2013), the brain may implement minimal energy principles without influencing the way neural redundancy is resolved.

OFC posits that motor control signals are selected to minimize a cost function that depends on task requirements and other factors, such as effort or delayed reward. OFC models have been widely used to explain muscle activity during motor tasks (e.g., Todorov 2004; Scott 2004; Diedrichsen et al. 2010b). Our results for neural activity differ in two important respects from OFC predictions with standard cost functions involving task requirements and effort. First, those implementations of OFC predict that variability in task-irrelevant dimensions should be higher than variability in task-relevant dimensions, a concept often referred to as the "uncontrolled manifold" (Scholz and Schöner, 1999). We found that the variability of neural activity did not increase in dimensions that went from being task-relevant to task-irrelevant (Fig. 3.6C). Second, those implementations of OFC predict a "minimal intervention" strategy, whereby activity in task-relevant dimensions is corrected independently of activity in task-irrelevant dimensions (Todorov and Jordan, 2002; Valero-Cuevas et al., 2009; Diedrichsen et al., 2010b). Three of the hypotheses we tested incorporate this minimal intervention principle: Uncontrolled-uniform, Uncontrolled-empirical, and Persistent Strategy. None of these hypotheses predicted neural activity in task-irrelevant dimensions as accurately as did the Fixed Distribution hypothesis, which predicts that the distributions of task-relevant and task-irrelevant activity are yoked. Overall, our work does not rule out the possibility that OFC is appropriate for predicting neural activity. First, it may be possible to design a cost function such that OFC predictions are consistent with the findings presented here. Second, one could consider applying OFC with the control signal being the input to M1 (e.g., PMd activity), rather than the control signal being M1 activity (as we have done here) or muscle activity (where OFC has been traditionally applied). This could induce coupling between the output-potent and output-null dimensions of the M1 activity, and thereby yield predictions that are consistent with the findings presented here.

It has also been proposed that muscle recruitment is habitual rather than optimal, such that muscle recruitment under altered dynamics is a rescaled version of that under normal control (de Rugy et al., 2012). The results for habitual control are similar to what we found for neural activity, in that (1) we could predict activity from previously observed activity, and (2) we observed a tight coupling of the distributions

of task-relevant and task-irrelevant activity (in contrast to minimal intervention). However, the results for habitual control are different from our findings in that we found that subjects appear to use the same distribution of activity in each of two different BCI mappings, whereas different (overlapping) subsets of muscle activation patterns were used under different conditions in de Rugy et al. (2012).

Given how many dimensions of population activity there are (in this case, 10), it is somewhat surprising that conditioning on only the two output-potent dimensions could provide so much explanatory power for predicting the distribution in the remaining neural dimensions. This suggests that many of the dimensions of population activity are coupled, i.e., changing the activity along some dimensions may also lead to changes along other dimensions, even though those dimensions are mutually orthogonal. During arm movement control, output dimensionality and presumably the neural dimensionality are larger than in our BCI setup. We speculate that during arm movements, many of the null dimensions will remain coupled with the potent dimensions, thereby yielding results similar to what we found here. Future work could examine whether animals can be trained to uncouple dimensions, as well as the effects of larger output-potent dimensionality on redundancy, by repeating our analyses with a higher-dimensional effector, such as a multiple degree-of-freedom robotic limb (e.g., Wodlinger et al. 2014).

The results presented here are related to, and go beyond, those in Golub et al. (2018). Although the two studies analyzed data from the same experiments, they ask distinct questions. Golub et al. (2018) focused on explaining the changes in population activity underlying behavioral learning. By contrast, in the present work we seek to determine the constraints on activity in the task-irrelevant (i.e., output-null) dimensions. In other words, while Golub et al. (2018) focused on explaining the changes leading to *behavioral* learning, we focus here on the principles *other than behavior* that constrain population activity. As a result, all hypotheses we consider in the present work make predictions consistent with the observed amount of learning in the output-potent dimensions.

Golub et al. (2018) found that the amount of learning animals showed was consistent with a fixed neural repertoire of population activity patterns being reassociated to control the second BCI mapping. The repertoire of population activity refers to the set of population activity patterns that were observed, whereas here we focused on the *distribution*, which describes how often the animals produced different activity patterns. In other words, the finding of a fixed repertoire is a statement about the support of the distribution of population activity, whereas here we found that the distribution of population activity can be predicted in output-null dimensions, given the output-potent activity. Because many different distributions of neural activity can be constructed from a fixed repertoire, the present results represent a stronger constraint on population activity than that shown in Golub et al. (2018). Indeed, the majority of the hypotheses we tested were consistent with a fixed neural repertoire, and thus cannot be disambiguated based on our prior work. This is evidenced by the predicted distributions largely overlapping with the support of the actual data distributions (Figures 2-4). The two hypotheses that were not fully consistent

with a fixed repertoire are the Minimal Firing and Uncontrolled-uniform hypotheses. However, in the context of predicting the distribution of activity in redundant dimensions, these hypotheses represent interesting cases worth considering (i.e., where population activity either obeys minimal firing constraints, or that the output-null activity is fully unstructured, respectively), and so we included these hypotheses to cover these possibilities.

It is interesting to consider the relationship between arm movements and BCI cursor movements (Orsborn et al., 2014; Vyas et al., 2018). If the dimensions responsible for moving the arm overlap with both the output-potent and output-null dimensions of the BCI, this might explain the coupling we observe between the output-potent and output-null dimensions. However, in these experiments, the animal's arm was not moving during BCI control (see Extended Data Figure 5 in Sadtler et al. 2014). Thus, the activity we study here resides within the arm's output-null dimensions. This implies that in our recordings the arm's output-potent dimensions do not overlap with either the output-potent or the output-null dimensions of the BCI, and so arm movements (or the lack thereof) are unlikely to explain the coupling we observed between the output-potent and output-null dimensions of the BCI. Overall, being unaware of extra output-potent dimensions would likely make the predictions of the Fixed Distribution hypothesis worse, not better. The reason for this is as follows. The Fixed Distribution hypothesis predicts that the distribution of activity in output-null dimensions depends upon the corresponding output-potent activity. Under this hypothesis, the more we know of the output-potent activity, the better we can predict the output-null distribution. If there is an output-potent dimension that we have not accounted for in our analyses, accounting for this dimension would likely improve our predictions. The fact that we were able to accurately predict the output-null distributions (13% histogram error on average, with the lowest possible error being 7%) without knowing all the potent dimensions is then evidence that these extra potent dimensions, if they exist, would not provide substantial additional predictive power.

In this work, we define a set of population activity patterns as redundant if they all result in the same readout in downstream areas. This definition of redundancy comes from early work on motor control (Bernstein, 1967; Sporns and Edelman, 1993), where it was noted that different motor signals can result in the same movement kinematics. This is related to but distinct from the information-theoretic definition of redundancy (Schneidman et al., 2003; Latham and Nirenberg, 2005; Averbeck et al., 2006). In the information-theoretic case, redundancy describes the extent to which correlations among neurons limit decoding accuracy for different stimuli. This is distinct from the type of redundancy studied here, defined as the existence of multiple population activity patterns corresponding to the same readout. For example, by the information-theoretic definition, a system may have no redundancy (e.g., the population activity allows one to perfectly decode the encoded variable), but there may still be multiple population activity patterns that refer to this same encoded variable.

We found that the distribution of output-null activity could be well predicted

using activity recorded under a different BCI mapping. Two factors of our experimental design are particularly relevant when interpreting this result. First, we used a balanced center-out task design in which subjects made roughly equal numbers of movements in each direction. If we had, for example, required far more leftward than rightward movements, this would have altered the distribution of joint activity and skewed the estimates of output-null activity during the second mapping. Second, this study focused on short timescales, where we predicted output-null activity within one to two hours of subjects learning a new BCI mapping. On this timescale, the motor system must be able to rapidly learn a variety of different mappings between neural activity and behavior, and thus, a variety of different sets of redundant activity. An interesting avenue for further research would be to determine if the constraints we observe on neural redundancy remain over longer timescales. Given repeated practice with the same BCI mapping across days and weeks (Ganguly and Carmena, 2009), it is possible that there are different and perhaps fewer constraints on neural redundancy than what we found here.

We have tested six specific hypotheses about how neural redundancy is resolved. These hypotheses cover a spectrum of how strongly the activity in output-null dimensions is constrained, with the minimal firing hypotheses being the most constrained, the minimal intervention hypotheses being the least constrained, and the Fixed Distribution hypothesis lying in between. Although the hypotheses we tested are not exhaustive, the best hypothesis (Fixed Distribution) yielded predictions of the distributions of output-null activity whose marginal histograms differed from the data by only 13% on average (Fig. 3.4F), where we estimated the lowest error possible to be 7% on average. Further improvements to the prediction accuracy may be possible by incorporating additional constraints, such as dynamics (Shenoy et al., 2013). It should be stressed that our focus here was on predicting the *distribution* of output-null activity. Future work can assess whether output-null activity can be predicted on a time-step-by-time-step basis.

The central premise of the null space concept is that some aspects of neural activity are read out by downstream areas (output-potent) while other aspects are not (output-null) (Kaufman et al., 2014). This idea is related to the study of noise correlations, where it was recognized that activity fluctuations that lie outside of a stimulus encoding space (i.e., "stimulus-null") are not detrimental to the stimulus information encoded by the neurons (Averbeck et al., 2006; Moreno-Bote et al., 2014). Studies have also shown that structuring neural activity in an appropriate null space can allow for multiplexing of different types of information (Mante et al., 2013; Raposo et al., 2014), as well as stable behavior (Leonardo, 2005; Rokni et al., 2007; Ajemian et al., 2013) and stable working memory (Druckmann and Chklovskii, 2012; Murray et al., 2017) in the presence of time-varying neural activity. Additionally, the existence of output-null dimensions in the motor system may facilitate motor learning (Ranganathan et al., 2013; Moorman et al., 2017; Singh et al., 2016), or allow for motor preparation (Kaufman et al., 2014; Elsayed et al., 2016) or novel feedback processing (Stavisky et al., 2017) without causing overt movement. Our work suggests that there may be limits on the extent to which output-null activity

might be leveraged for neural computation. The coupling we observe between the distributions of output-null and output-potent activity suggests that output-null activity is not modified independently of output-potent activity. This coupling may cause activity fluctuations in a stimulus-null space to influence the downstream readout, or limit one's ability to plan the next movement without influencing the current movement. Moving forward, an important direction for understanding the computations performed by different brain areas is to find out which aspects of the neural activity are read out (Pagan et al., 2013; Kaufman et al., 2014) and to understand how the dependencies like those identified in this study impact the computations being performed.

4 Learning is shaped by abrupt changes in neural engagement

Internal states such as arousal, attention, and motivation modulate brain-wide neural activity, but how these processes interact with learning is not well understood. During learning, the brain modifies its neural activity to improve behavior. How do internal states affect this process? Using a brain-computer interface (BCI) learning paradigm in monkeys, we identified large, abrupt fluctuations in neural population activity in motor cortex indicative of arousal-like internal state changes, which we term "neural engagement." In a BCI, the causal relationship between neural activity and behavior is known, allowing us to understand how neural engagement impacted behavioral performance for different task goals. We observed stereotyped changes in neural engagement that occurred regardless of how they impacted performance. This allowed us to predict how quickly different task goals were learned. These results suggest that changes in internal states, even those seemingly unrelated to goal-seeking behavior, can systematically influence how behavior improves with learning.

Published as: Hennig, J.A., Oby, E.R., Golub, M.D., Bahureksa, L.A., Sadtler, P.T., Quick, K.M., Ryu, S.I., Tyler-Kabara, E.C., Batista, A.P., Chase, S.M., and Byron, M.Y. 2021. "Learning is shaped by abrupt changes in neural engagement." *Nature Neuroscience* (in press).

4.1 Introduction

As we move about the world, we experience fluctuations in internal states such as arousal, motivation, and engagement. Such fluctuations, which do not directly reflect sensory stimuli or intended movements, are governed by the modulation of neural activity throughout the brain (Aston-Jones and Cohen, 2005; McGinley et al., 2015; Allen et al., 2019; Stringer et al., 2019; Steinmetz et al., 2019). The manner in which these modulations relate to the ongoing computations performed by the cerebral cortex is not well understood. In predominantly sensory areas of cortex, changes in an animal's internal state are known to affect neural response magnitude, signal-to-noise ratio, timing, and variability (Mitchell et al., 2009; Cohen and Maunsell, 2009; McGinley et al., 2015; Vinck et al., 2015). Depending on how these changes align with respect to neural encoding of stimulus information or downstream readout,

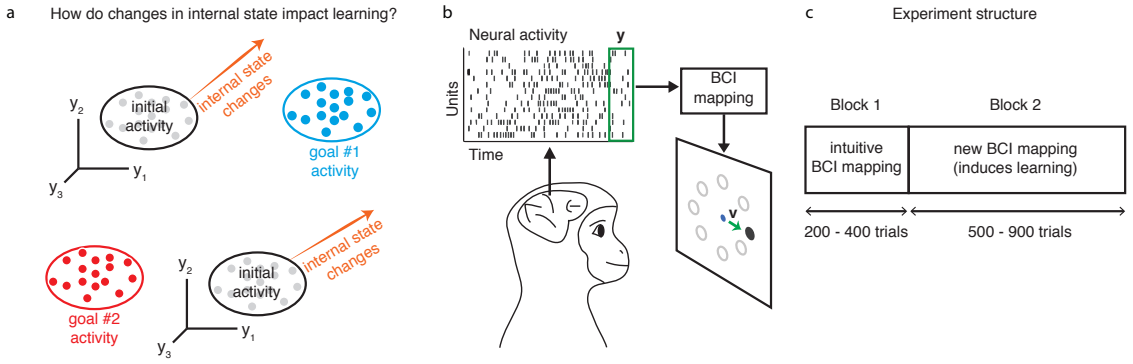


Figure 4.1. Studying how changes in neural activity during learning relate to changes in internal state. a. Here we ask whether changes in internal state relate to how neural population activity is modified during learning. Before learning, neural activity resides in some region ("initial activity") of population activity space, depicted here by the spiking activity of three neurons (y_1, y_2, y_3). During learning, the neural activity needs to migrate to a different region of population activity space to achieve a particular task goal ("goal #1 activity" and "goal #2 activity"). Changes in the animal's internal state can push the neural activity closer to (top orange arrow) or further from (bottom orange arrow) the region appropriate for achieving a given task goal. **b.** Monkeys performed an eight-target center-out task using a brain-computer interface (BCI). Neural activity was recorded using a multi-electrode array implanted in M1. Spike counts ($\mathbf{y}$) were taken in 45 ms bins (green box). The BCI mapping converted the neural activity into a cursor velocity ($\mathbf{v}$) at each 45 ms timestep, updating the position of a visual cursor on a screen. Monkeys were rewarded for successfully guiding the cursor to hit the visually instructed target. **c.** Each experiment consisted of two blocks of trials. In Block 1, a monkey completed 200-400 trials using an intuitive BCI mapping. In Block 2, the monkey completed 500-900 trials with a new BCI mapping he had not used before.

changes in an animal's internal state can impact perceptual processing and decision making (Averbeck et al., 2006; Moreno-Bote et al., 2014; Ruff and Cohen, 2019; Cowley et al., 2020). Changes in internal state are also known to impact motor control and behavior, as the speed and latency of both eye movements and arm reaches are known to be modulated by signals such as motivation, intrinsic value, and reward expectation (Sugrue et al., 2004; Mazzoni et al., 2007; Xu-Wilson et al., 2009; Leathers and Olson, 2012). These studies and others illustrate the importance of understanding the influence of internal states on sensory processing and behavior. What has been less well studied is the impact of internal state changes on learning (Fig. 4.1a). When we learn to perform a task, such as shooting a basketball, the firing activity of populations of neurons in the brain (Fig. 4.1a, gray clouds) is modified in a particular manner in order to drive improved behavior (Fig. 4.1a, blue and red clouds) (Li et al., 2001; Andalman and Fee, 2009; Ganguly and Carmena, 2009; Hwang et al., 2013; Jeanne et al., 2013; Law et al., 2014; Sadtler et al., 2014; Poort et al., 2015; Athalye et al., 2018; Golub et al., 2018; Vyas et al., 2018; Perich et al., 2018; Oby et al., 2019). We also know that while animals perform a task, neural activity undergoes internal state fluctuations that are not directly related to task performance (Fig. 4.1a, orange arrows) (Arieli et al., 1996; Churchland et al., 2010; Gu et al., 2011; Ecker et al., 2014; Lin et al., 2015; Stringer et al., 2019; Cowley et al., 2020). Depending on the task goals, changes in internal state have the potential to

make some learning-related neural changes easier to achieve (Fig. 4.1a, blue cloud), while other changes may be made more difficult (Fig. 4.1a, red cloud). When changes due to internal state are incongruous with learning, how do neural populations modify their activity to drive improved behavior? One possibility is that the internal state fluctuations that make learning more difficult might be suppressed. Alternatively, the impact of internal state fluctuations on learning may be unavoidable, such that some task goals are harder to achieve than others.

Answering this question is challenging because the causal relationship between neural activity and behavior is not known in general. This makes it difficult to understand which changes to neural activity would yield improved performance, as well as how fluctuations in internal state would either interfere or augment that performance. To address this difficulty we can leverage a brain-computer interface (BCI) (Shenoy and Carmena, 2014; Moxon and Foffani, 2015; Golub et al., 2016; Orsborn and Pesaran, 2017), where the causal relationship, or "mapping," between neural activity and behavior is known exactly and determined by the experimenter.

We trained three rhesus monkeys to modulate the activity of ∼90 units in primary motor cortex (M1) to move a computer cursor on a screen using a BCI (Sadtler et al., 2014). In previous work, we compared the neural population activity before versus after monkeys learned to use a new BCI mapping (Golub et al., 2018; Hennig et al., 2018). Here we study how neural activity changed throughout learning, and the degree to which these changes were influenced by fluctuations in the monkey's internal state.

We first identified the dimensions of the largest fluctuations in M1 population activity. Surprisingly, abrupt changes in population activity along these dimensions were triggered by changes in various aspects of the task, ranging from brief pauses in the task to perturbations of the BCI mapping. Furthermore, trial-to-trial changes in population activity along these dimensions were correlated with changes in the monkey's pupil size. These observations suggested that changes in population activity along these dimensions could be related to the monkey's arousal, engagement with the task, or motivation throughout the experiment. For this reason, we term these dimensions *neural engagement* axes.

To induce learning, we perturbed the mapping between neural activity and cursor movements, requiring monkeys to modify the neural activity they produced in order to restore proficient control of the cursor towards each target (Sadtler et al., 2014). This allowed us to study how changes in activity along neural engagement axes interacted with learning. We found that neural population activity did not take a direct path from the activity produced prior to learning to the activity produced at the end of learning. In particular, neural activity changed abruptly along the neural engagement axes at the start of learning. This change occurred regardless of the relationship between neural engagement axes and cursor movements, which led to an immediate improvement in behavioral performance for some targets and impaired performance for others. Following the abrupt change, neural activity retreated along neural engagement axes, which impacted performance differently for different targets. These findings enabled us to predict which targets would be learned more

quickly than others, based on how neural engagement interacted with the demands of the learning task. Our results suggest that changes in internal states, even those seemingly unrelated to goal-seeking behavior, can influence how behavior improves with learning.

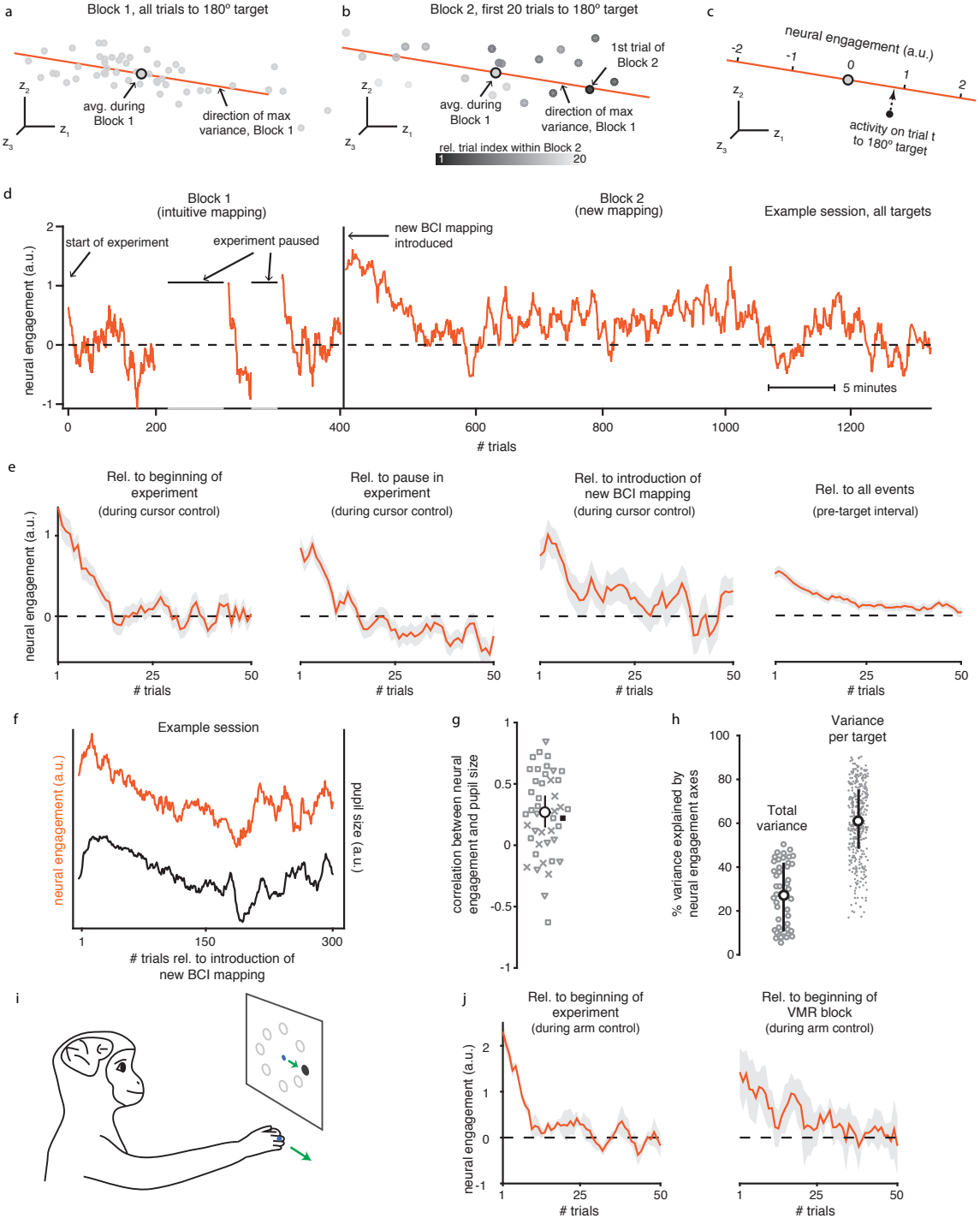
4.2 Results

To understand how changes in internal state might interact with learning (Fig. 4.1a), we studied three monkeys performing an eight-target center-out task using a brain-computer interface (BCI) (Fig. 4.1b; see Methods). On each trial, monkeys controlled a computer cursor by modulating neural activity recorded from primary motor cortex (M1). The relationship between the recorded neural activity and cursor velocity was specified by the BCI mapping. In each experimental session, monkeys used two different BCI mappings (Fig. 4.1c). During the first block of trials, monkeys used an "intuitive" BCI mapping, calibrated so as to provide the monkey with proficient control of the cursor. After monkeys performed the task for a few hundred trials using the intuitive mapping, we changed the mapping between neural activity and cursor movement to a new BCI mapping that the monkey had not used before. This new BCI mapping (a "within-manifold" perturbation (Sadtler et al., 2014)) was typically learned within one to two hours.

Prior to each experiment, we applied factor analysis to identify the top ten dimensions, or factors, capturing the most covariability of the neural population activity. The BCI mappings presented during each experiment were chosen such that the cursor velocity was determined by only these top ten factors. In order to ensure that our results captured changes in neural activity describing substantial covariance in the population, we analyzed neural activity only in these factors.

Internal state fluctuations in primary motor cortex

We first show that the neural population activity shifted abruptly in response to salient, experimenter-controlled events. We observed that, while monkeys used the intuitive mapping, the neural activity produced for a given target showed substantial trial-to-trial variability (Fig. 4.2a, gray dots). We found the direction of greatest variance of the neural activity for each target (Fig. 4.2a, orange line). Surprisingly, later in the session when the new BCI mapping was introduced, neural activity on the first trial to a given target showed an abrupt change from the neural activity produced during Block 1, with this change occurring almost directly along the axis identified earlier (Fig. 4.2b, compare "1st trial of Block 2" to "avg. during Block 1"). Interestingly, on subsequent trials, neural activity gradually retreated down this same axis (Fig. 4.2b, grayscale indicates trial index).



We next quantified how these trial-to-trial changes in neural activity progressed throughout the experiment (Fig. 4.2c). To do this, we identified the axis of greatest variability during Block 1 for each target separately (e.g., the orange axis in Fig. 4.2a-b), and projected the neural activity for each trial along the appropriate target-specific axis. So that we could compare these values across trials to different targets, we normalized the projected values for each target separately (see Methods). This

Figure 4.2. Neural activity increased abruptly along a neural engagement axis following experimental events. a. Neural activity in the top three factor dimensions of highest covariance (z_1, z_2, z_3) for trials to the same target from Block 1 of session J20120528. Each gray point is the average neural activity recorded within a single trial. Orange axis depicts the direction of maximum variance of all gray points. The axis was defined in the 10-dimensional factor space, although only the top three dimensions are depicted here. **b.** Same as **a**, but for the first 20 trials to the same target during Block 2 (color indicates trial index). Orange axis from **a** shown for reference. *Neural engagement* for each trial is the projection of neural activity onto the axis identified during Block 1 for trials to the same target. **c.** The value of neural engagement is given by the projection (dotted arrow) of the recorded neural activity (black circle) onto the neural engagement axis (orange) corresponding to the current target. Projections were normalized relative to the mean and standard deviation across trials to the same target during Block 1 (see Methods). **d.** Neural engagement over time from session J20120528, with annotations indicating timing of various events controlled by the experimenter. Position along horizontal axis indicates clock time (see legend), with trial indices marked for reference. Horizontal dashed line at zero indicates average neural engagement across all trials during Block 1. **e.** First three subpanels: Neural engagement averaged across sessions from all monkeys ($n = 46$ sessions) during cursor control relative to the start of the experiment, the longest pause during Block 1, and the start of Block 2. Last subpanel: Neural engagement during the interval of each trial before the monkey had seen the target (see Methods), averaged across all three experimental events. Shading indicates mean $\pm$ SE across sessions. **f.** Neural engagement during Block 2 from example session shown in **d**, alongside monkey's average pupil size during the same trials. **g.** Pearson's correlation between neural engagement and pupil size during Block 2 for each session, with sessions from monkeys J, L, and N indicated by squares, triangles, and crosses, respectively. Example session from **f** indicated as black square. White circle and black lines depict the bootstrapped median and 95% C.I. of the correlations across sessions, respectively. **h.** Percentage of shared covariance of neural population activity explained by neural engagement axes during Block 1, across trials to all targets ("Total variance"; $n = 46$ sessions), or across trials to a single target ("Variance per target"; $n = 368$ targets). White circle depicts median; error bar depicts median $\pm$ 25^{th} percentile of correlations across sessions/targets. **i.** In a different set of experiments, a monkey performed a center-out task by moving its hand to control the cursor's position (see Methods). **j.** Neural engagement averaged across sessions from hand control experiments ($n = 3$ sessions), relative to the beginning of the experiment (left), and relative to the introduction of a visuomotor rotation (right). Same conventions as **e**.

yielded a trial-by-trial measure we call *neural engagement*, for reasons we discuss below.

Neural engagement abruptly increased and gradually decreased following various experimental events, beyond just the introduction of the new BCI mapping (Fig. 4.2d). For example, neural engagement was initially elevated on the very first trials of the experiment, and then gradually decreased on later trials (Fig. 4.2d, "start of experiment"). Next, near the middle of Block 1, the experimenter would pause the experiment for a few minutes to choose the BCI mapping that would be introduced in the upcoming Block 2. Following these pauses (Fig. 4.2d, "experiment paused"), neural engagement increased, and then gradually subsided. Finally, a few minutes later when the experimenter seamlessly introduced the new BCI mapping (without pausing the experiment), neural engagement again abruptly increased and gradually subsided on subsequent trials (Fig. 4.2d, "new BCI mapping introduced"). The ensuing time course of neural engagement was similar following all three of these experimenter-controlled events, indicating that the changes in neural engagement during Block 2 were not due simply to the monkey trying to learn the new BCI mapping. These changes were not specific to the particular BCI mappings used during a given session, as we observed similar neural changes across multiple sessions from all three monkeys (Supplemental Fig. B.1). Nor could the abrupt increases in neural engagement be explained by hand movements, as monkeys showed little to no hand movements during these experiments, and no increase in hand speeds when the new BCI mapping was introduced (Supplemental Fig. B.2). Rather, these changes in neural activity appeared to reflect stereotyped changes in the monkey's internal state throughout the experiment, and could reflect changes in arousal (Vinck et al., 2015), engagement with the task (Steinmetz et al., 2019), or motivation (Mazzoni et al., 2007). While the specific source of these changes is as yet unknown (see Discussion), these changes have important consequences for learning, as we discuss below.

Two additional aspects of neural engagement were consistent with it reflecting variations in the monkey's internal state. First, when averaged across trials from all sessions, neural engagement showed a consistent time course following each experimental event (Fig. 4.2e). These changes in neural engagement appeared not only while the monkeys controlled the cursor (Fig. 4.2e, first three subpanels), but also during the beginning of each trial before the monkey had seen the visual target (Fig. 4.2e, last subpanel). Thus, neural engagement remained elevated even when the monkey was not actively performing the task, consistent with this signal reflecting a slowly-varying change in the monkey's internal state. Second, changes in an organism's internal state are typically correlated with changes in its pupil size (McGinley et al., 2015). On trials with elevated levels of neural engagement, the pupil was typically more dilated (Supplemental Fig. B.3), suggesting that neural engagement may be correlated with an arousal-like state. In agreement with this, we found that trial-to-trial fluctuations in neural engagement were often strikingly positively correlated with the monkey's pupil size (Fig. 4.2f). For the majority of sessions from all three monkeys, trial-to-trial changes in neural engagement and pupil size were positively correlated (Fig. 4.2g), with a median Pearson's correlation across sessions

of $\rho = 0.27$ (bootstrapped 95% C.I. $[0.15, 0.40]$, $n = 44$ sessions), similar to levels observed in other work(Cowley et al., 2020).

Changes in activity along the neural engagement axes accounted for a substantial amount of the covariance of the population activity. When considering population activity during Block 1 across trials to all eight targets—and thus also including the across-target variance in neural activity due to the monkey aiming towards different targets—changes in neural engagement explained ∼30% of the total trial-to-trial variance of the factor activity (Fig. 4.2h, "Total variance"). Within trials to the same target, changes along the neural engagement axis explained ∼60% of the trial-to-trial variance (Fig. 4.2h, "Variance per target"). These results indicate that the trial-to-trial changes in population activity along the neural engagement axes were substantial.

To assess whether similar changes in neural engagement were present during arm movements (as opposed to BCI control), we analyzed data from a fourth monkey performing an eight-target center-out task by controlling a computer cursor with his hand (Fig. 4.2i; see Methods). As with the BCI experiments, we identified a set of neural engagement axes in the population activity after applying factor analysis. We found that neural engagement was elevated both at the beginning of each experiment, and following the introduction of a visuomotor rotation (Fig. 4.2j), with a time course that was strikingly similar to that of BCI control (Fig. 4.2e). Taken together, we found that neural population activity in M1 during both BCI and hand control showed large, trial-to-trial variations with a consistent time course relative to experimental events. In the following, we focus on BCI control, where we know the causal relationship between neural activity and behavior. This enables us to directly assess how changes in neural engagement relate to behavior (i.e., cursor velocities).

Studying the impact of changes in neural engagement on behavior using a BCI paradigm

Having established the presence of large fluctuations in neural engagement in M1, we next wanted to understand how these fluctuations might interact with learning. Specifically, we asked whether the monkey's ability to learn to move the cursor in a given direction with the new BCI mapping could be understood in terms of the relationship between the neural engagement axes and the new mapping. First, we explain how a BCI paradigm allows us to quantify the interaction between neural engagement and behavior (i.e., cursor velocities). Consider a schematic of the neural activity produced by the monkey during Block 1 (Fig. 4.3a). For a given target, we can summarize the monkey's trial-averaged neural activity as a point in neural space ($\mathbf{z}$), where here we depict the neural activity in the three factor dimensions of highest variance. The cursor velocity under the intuitive BCI mapping ($\mathbf{v}$) is given by projecting the neural activity onto the intuitive BCI mapping ($\mathbf{v} = M_1\mathbf{z}$). During Block 1, the monkey's trial-averaged cursor velocities were near the target direction (Fig. 4.3b), indicating the monkey's ability to produce cursor movements

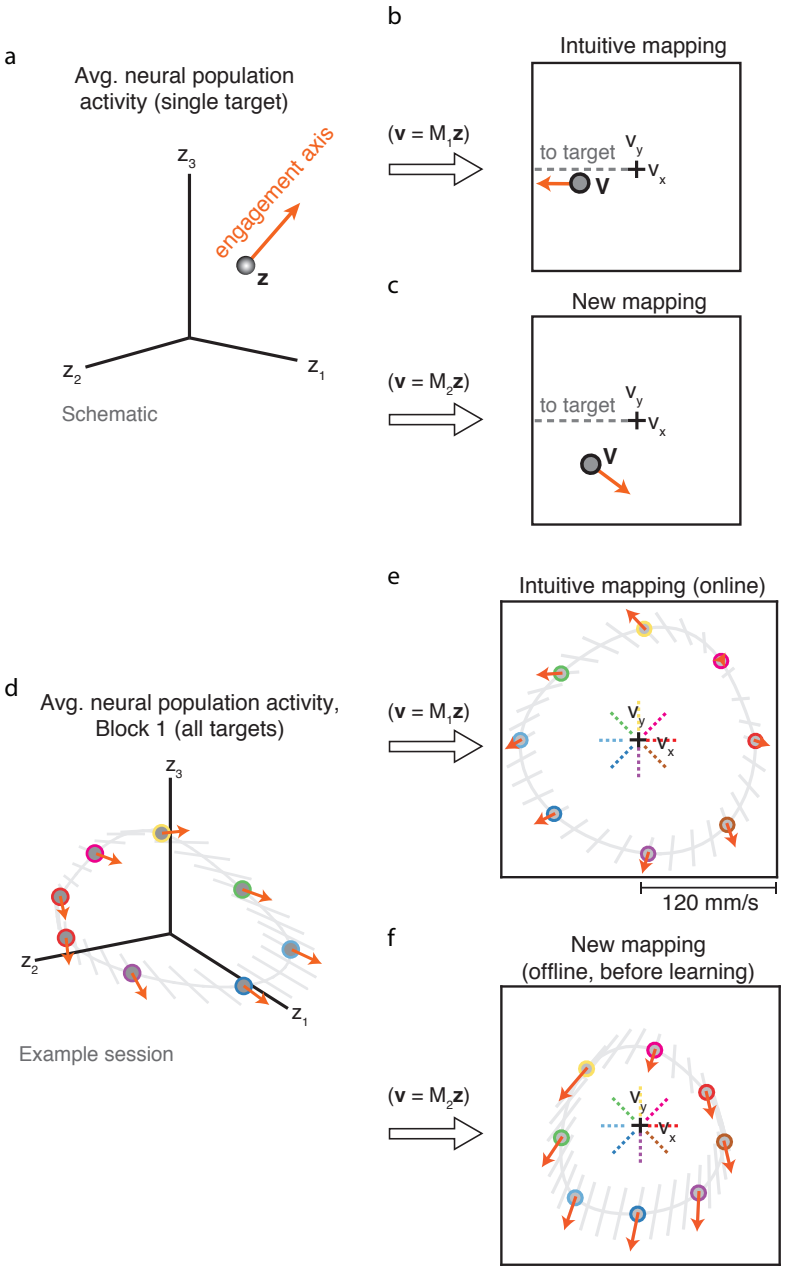


Figure 4.3. Understanding the impact of neural engagement on behavior during a BCI learning task. a. Left: Schematic of the average neural activity ($\mathbf{z}$) recorded across all trials to the same target during Block 1, along with the direction in which this activity is expected to move following an increase in neural engagement (orange arrow). Top right: Using the intuitive BCI mapping (M_1), we can inspect the intuitive cursor velocity ($\mathbf{v}$, gray circle) corresponding to $\mathbf{z}$, as well as how this velocity will change if neural engagement increases (orange arrow). In this case, increased neural engagement will result in faster cursor movements towards the target (gray dotted line). Zero velocity is indicated by the black cross. Bottom right: We can repeat the same procedure using the new BCI mapping (M_2) with the same neural activity $\mathbf{z}$ and neural engagement axis. **b.** For an example session, the trial-averaged neural activity (gray circles with colored outlines) and engagement axes (orange arrows) for all eight targets. Gray lines indicate interpolations between the neural engagement axes for each target. Dashed colored lines in the two right subpanels indicate the eight target directions. Extent of box is ± 120 mm/s.

that moved the cursor towards the target on average. We can also characterize the effect of an increase in neural engagement on cursor velocities by projecting the neural engagement axis (orange arrow in Fig. 4.3a) onto the intuitive BCI mapping (orange arrow in Fig. 4.3b). For this target, increasing neural engagement increases cursor speeds towards the target.

Next, consider the first trial of Block 2, when the monkey first encounters the new BCI mapping. If the monkey were to continue to produce the same average neural activity that he did during Block 1 (Fig. 4.3a), this would no longer result in cursor movements straight to the target (Fig. 4.3c). Thus, the monkey must learn how to modify the neural activity he produces in order to produce faster cursor speeds in the target direction. Importantly, the new BCI mapping also changes the manner in which neural engagement relates to cursor velocity. For this target, increasing neural engagement would move the cursor velocities even further from the target direction (orange arrow in Fig. 4.3c). In this manner, changes in neural engagement can interact with the monkey's attempts to move the cursor towards the target.

We can gain a more holistic picture of the interaction between neural engagement and cursor velocities by visualizing the neural activity produced for all eight targets together (Fig. 4.3d). We observed that, when visualized in factor space (Fig. 4.3d), the neural engagement axes identified for different targets often appeared quite similar. In fact, neural engagement axes were almost always consistent with the firing rates of all neural units changing in the same direction (Supplemental Fig. B.4). Increases in neural engagement corresponded to increased average firing rates in nearly all units, in a gain-like manner (Supplemental Fig. B.5). However, while the neural engagement axes for different targets were similar in terms of how they related to single unit firing rates, these axes also showed behaviorally relevant differences. For example, under the intuitive BCI mapping, increases in neural engagement typically led to faster speeds towards each target (Fig. 4.3e), which could not happen if the neural engagement axes were identical for all targets. A similar feature was also present during arm movements: By identifying the linear mapping of neural population activity most predictive of ensuing hand velocities, we found that increases in neural engagement typically predicted faster hand speeds towards each target (Supplemental Fig. B.6). Thus, the orientation of the neural engagement axis in population activity space depends on the monkey's intended movement direction.

We now focus on the velocities under the new BCI mapping (Fig. 4.3f), as this indicates the initial cursor velocities the monkey would expect to produce during Block 2, were he to continue producing the same activity he did during Block 1. As discussed above, neural engagement can have different effects on cursor velocities depending on the monkey's intended movement direction. For example, increased neural engagement may increase cursor speeds towards some targets (e.g., purple target in Fig. 4.3f) and decrease speeds towards other targets (e.g., pink target in Fig. 4.3b). Additionally, increased neural engagement can affect not just the speed but also the angular error of the velocity relative to the target direction (e.g., red and yellow targets in Fig. 4.3f). Overall, we observed that the new BCI mappings induced a variety of different relationships between neural engagement and cursor

velocity, both across sessions and within targets of the same session (Supplemental Fig. B.7). Thus, these experiments provided us with the means to assess how different relationships between neural engagement and cursor velocity related to how each target was learned.

Neural engagement increased initially regardless of its impact on performance

To study how changes in neural engagement might interact with learning, we first characterized the level of neural engagement on the very first trial to each target using the new BCI mapping. As shown earlier, monkeys' initial reaction to the introduction of the new mapping was, on average, to increase neural activity along the neural engagement axis (Fig. 4.2e, third subpanel). However, as we have also shown, there are a variety of ways in which the neural engagement axes affected velocities under the second mapping (Supplemental Fig. B.7). This raises the possibility that neural engagement might have increased by different amounts depending on whether increasing neural engagement was expected to increase (Fig. 4.4a) or decrease (Fig. 4.4b) the speed of the cursor towards the target under the new mapping.

We anticipated that neural engagement might increase more for targets where doing so would increase cursor speeds towards the target. To assess whether this was the case, for each target we used the trial-averaged activity from Block 1 to estimate the expected velocity under the new mapping (Fig. 4.4a-b, filled circles), as well as the expected impact on that velocity if neural engagement increased (Fig. 4.4a-b, orange axes). We then classified each target as belonging to one of two groups, based on whether an increase in neural engagement was expected to increase ("$T+$", Fig. 4.4a) or decrease ("$T-$", Fig. 4.4b) the speed of the cursor towards the target direction. We next assessed the levels of neural engagement on the first trial to each target in Block 2. Across targets from all sessions, the distribution of neural engagement on the first trial using the new mapping did not differ as a function of how performance for that target was impacted (Fig. 4.4c) ($p = 0.954$, two-sample Kolmogorov-Smirnov test, $n_1 = 220$ and $n_2 = 148$ targets). This indicates that initially, neural activity increased along the neural engagement axes even when doing so negatively impacted task performance. As a result, the initial increase in neural engagement made $T-$ targets more difficult than they would have been otherwise (relative to the average neural activity produced during Block 1), while $T+$ targets were made easier.

Differences in learning across targets can be explained by changes in neural engagement

We saw that changes in neural engagement on the first trials using the new BCI mapping occurred regardless of the impact on performance. We wondered whether, given repeated practice with the new mapping over subsequent trials, changes in neural

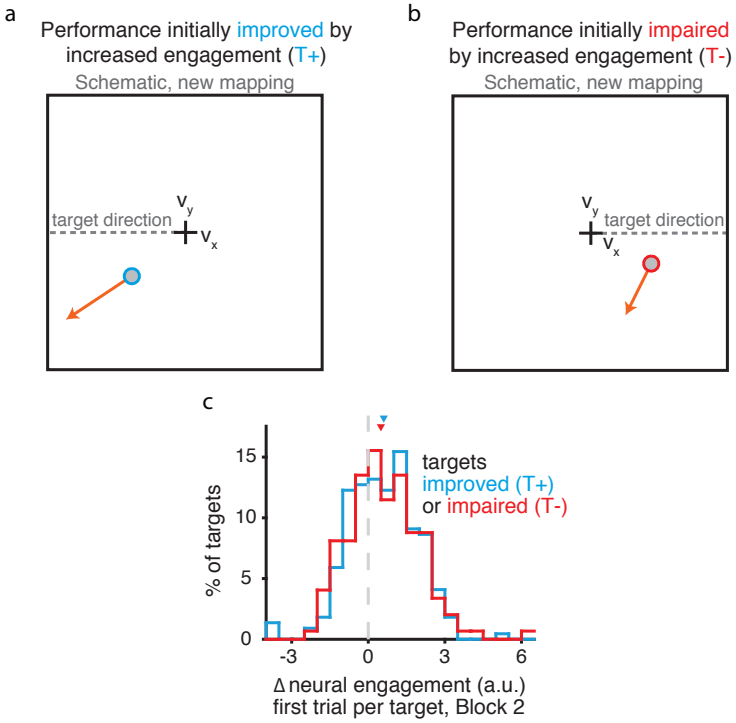


Figure 4.4. Neural engagement increased on the first trial of a learning task regardless of its impact on task performance. a-b. Schematics depicting how increased neural engagement can lead to either faster (**a**) or slower (**b**) cursor speeds towards the target direction under the new BCI mapping. Same conventions as Fig. 4.3c. **c.** Distribution of the increase in neural engagement on the first trial to each target during Block 2, as a function of whether performance under the new mapping was expected to be improved (blue) or impaired (red) by an increase in neural engagement (as in **a-b**). Triangles depict the median of each distribution.

engagement might interact with learning-driven changes for each type of target.

We visualized how cursor velocities under the second mapping changed throughout learning, as a function of whether the initial increase in neural engagement increased ($T+$) or decreased ($T-$) the speed of the cursor towards the target (Fig. 4.5a-b). For both types of targets, neural activity on the first trial jumped out abruptly along the neural engagement axis (Fig. 4.5a-b, white circles have moved along the orange arrows relative to the gray circles). Then, over tens of trials, velocities gradually aligned with the target direction, leading to increased speeds towards the target (Fig. 4.5a-b, projection of the blue and red traces increases along the target direction). Were these behaviorally beneficial changes to velocity driven by target-specific changes in neural engagement? We measured the levels of neural engagement for each target during Block 2 after accounting for any changes due to learning by neural reassociation (Golub et al., 2018) (see Methods). In agreement with what we observed earlier (Fig. 4.2e, third subpanel), we found that neural engagement gradually decreased throughout Block 2 (Fig. 4.5c). Importantly, this decrease in neural engagement was likely beneficial to $T-$ targets, the ones initially impaired by the increase in neural engagement. In fact, neural engagement decreased more for $T-$ targets than for $T+$ targets (Fig. 4.5c). These target-specific differences in neural engagement could not be explained by differences in the animal's arousal, as the average time course of pupil size did not differ between $T+$ and $T-$ targets (Fig. 4.5d). These results suggest that, as learning proceeded, changes along the neural engagement axis were driven by two components, one target-invariant (because neural engagement decreased throughout learning for both target types), and one target-specific (because neural engagement decreased by different amounts depending on the target type). This led us to ask whether these differential changes in neural engagement might explain how quickly the two types of targets were learned. To quantify the amount of learning for each target, we measured cursor speeds towards the target relative to the speeds monkeys would experience if they continued to use the neural activity they produced prior to the introduction of the new BCI mapping (Fig. 4.5e; see Methods). On the first trial of Block 2, the cursor speed towards the target increased for $T+$ targets and decreased for $T-$ targets (Fig. 4.5e, trial 1). This is in agreement with monkeys immediately increasing neural engagement at the start of Block 2, regardless of its impact on performance (Fig. 4.4c). As Block 2 continued, performance for both target types gradually improved (Fig. 4.5e, trials 1-75), indicating learning.

Interestingly, monkeys reached peak performance more quickly for $T+$ targets than for $T-$ targets ($p = 1.259 \times 10^{-4}$, two-sided Wilcoxon rank-sum test, $n_1 = 220$ and $n_2 = 148$ targets; Fig. 4.5f, Supplemental Fig. B.8). This was not due to a difference in learning rate, as the learning rates for the two target types were not statistically different ($p = 0.202$, two-sided Wilcoxon rank-sum test, $n_1 = 220$ and $n_2 = 148$ targets; see Methods). Additionally, performance levels at the end of Block 2 for the two target types were not statistically different ($p = 0.884$, two-sided Wilcoxon rank-sum test, $n_1 = 220$ and $n_2 = 148$ targets). Other groupings of targets agnostic to neural engagement did not predict differences in how quickly targets

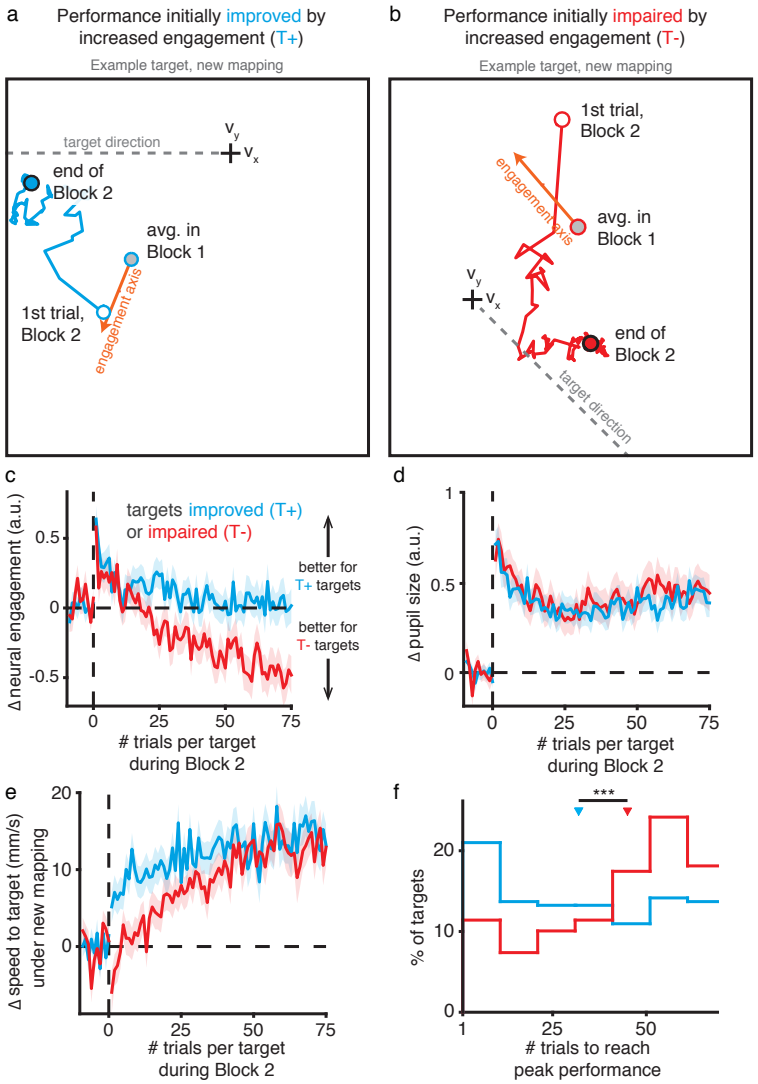


Figure 4.5. Relationship between neural engagement axes and task performance predicted which targets were learned more quickly. a. Average cursor velocities under the new mapping across trials during Block 2, for an example $T+$ target (180°, J20120528) where an increase in neural engagement initially improved performance relative to the average activity produced during Block 1 (gray circle). Same conventions as Fig. 4.4a. The blue line depicts how the trial-averaged velocity evolved throughout Block 2, starting with the first trial to that target (white circle) and ending with the average during the last trials (blue circle). Velocities gradually moved towards the target direction, both decreasing angular error and increasing the speed in the target direction, indicating learning. **b.** Same as **a**, but for an example $T-$ target (315°, J20120601) where an increase in neural engagement was initially expected to impair performance under the new mapping. **c.** Changes in neural engagement during Block 2, averaged across $T+$ ($n = 220$) and $T-$ ($n = 148$) targets (mean ± SE). Trial index is relative to the start of Block 2 for each target. **d.** Changes in pupil size during Block 2. Same conventions and sample sizes as **c**. **e.** Changes in cursor speed towards the target under the new mapping during Block 2, relative to the expected speed under the new mapping based on the neural activity produced during Block 1. Same conventions and sample sizes as **c**. **f.** Distribution of the number of trials at which each target attained its peak level of performance (see Methods), for all $T+$ and $T-$ targets. Triangles depict the median of each distribution; asterisks indicate that the medians were significantly different ($p = 1.259 \times 10^{-4}$, two-sided Wilcoxon rank-sum test, $n_1 = 220$ and $n_2 = 148$ targets).

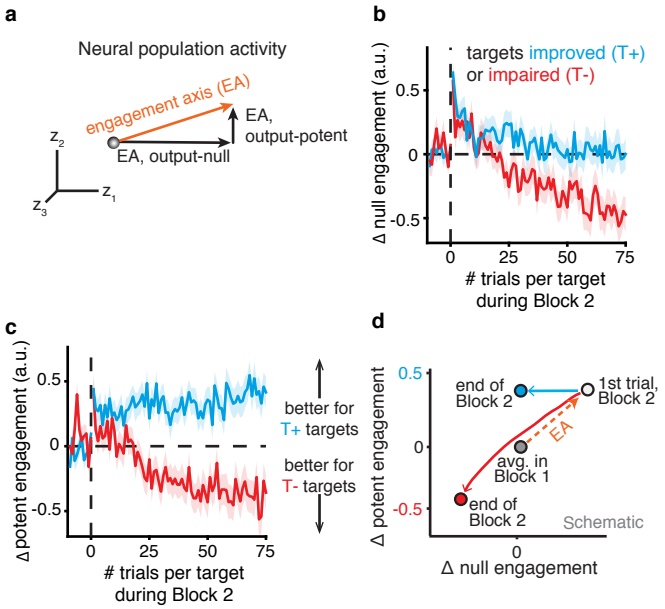


Figure 4.6. Neural engagement changed differently in output-potent versus output-null dimensions of the new BCI mapping. a. Schematic of decomposing a neural engagement axis (EA, orange arrow) into output-null and output-potent components. Given the new BCI mapping, this axis can be decomposed into output-null and output-potent axes, such that only changes in neural activity along the output-potent axis will affect cursor velocities under the new mapping. **b-c.** Changes in neural activity along the output-null (**b**) and output-potent (**c**) neural engagement axes during learning, averaged across $T+$ ($n = 220$) and $T-$ ($n = 148$) targets (mean ± SE). Same conventions as Fig. 4.5c. Changes in output-null neural engagement do not affect cursor movements, while changes in output-potent neural engagement do. **d.** Schematic summarizing how the evolution of neural activity during learning differed based on whether a neural engagement increase was predicted to initially improve ($T+$, blue) or impair ($T-$, red) performance to a given target. The average neural activity for each target type was similar on the first trial of Block 2 (white circle), relative to where activity was prior to the introduction of the new BCI mapping (gray circle), but gradually diverged during learning (blue and red arrows).

reached peak performance (Supplemental Fig. B.9). Overall, these results suggest that, although performance at the end of learning was similar for $T+$ and $T-$ targets, the initial increase in neural engagement gave performance for $T+$ targets a "head start," allowing monkeys to reach peak performance levels for these targets more quickly. This explanation is at apparent odds with the fact that neural engagement decreased throughout learning for both target types (Fig. 4.5c), which should have led to slower cursor speeds for the $T+$ targets. In the next section we explore how the initial performance improvements for $T+$ targets were maintained even as neural engagement decreased throughout learning.

Neural engagement changed differently in neural dimensions aligned with the new BCI mapping

Based on the relationship between neural engagement and cursor speed, one might expect that the gradual decrease in neural engagement (Fig. 4.5c) should have re-

sulted in gradually slower cursor speeds for $T+$ targets. But in Fig. 4.5e, one can see that performance for $T+$ targets actually *increased* over time. How is this possible? Crucially, our measurement of neural engagement does not account for which changes in neural engagement affect cursor movements, and which changes do not affect cursor movements. We therefore decomposed each neural engagement axis into two components (Fig. 4.6a; see Methods), where the first component was output-null to the new BCI mapping (i.e., changes in this direction would not impact cursor velocities under the new mapping), and the other component was output-potent (Kaufman et al., 2014; Stavisky et al., 2017; Hennig et al., 2018). This resulted in measures of output-null and output-potent neural engagement, which allowed us to look specifically at whether neural engagement changed differently depending on whether or not it impacted cursor movements. Changes along the output-null component of the neural engagement axis had no impact on cursor velocities (Fig. 4.6b), and followed the same pattern as the total neural engagement (Fig. 4.5c). By contrast, changes along the output-potent component of the neural engagement axis moved in the directions necessary to yield performance improvements for each target type (Fig. 4.6c). In particular, neural population activity for $T+$ targets remained elevated along the output-potent component of the neural engagement axis, where performance was initially improved by the increase in neural engagement (Fig. 4.6c, blue trace). This indicates that the net decrease in total neural engagement throughout learning was not entirely agnostic to task performance, as neural activity remained elevated specifically in the neural dimensions that were relevant to controlling the cursor.

Taken together, these results explain how monkeys reached peak performance more quickly for some targets than for others, based on the relationship between neural engagement and cursor movements (Fig. 4.6d). On the first trial of Block 2, neural activity increased along the neural engagement axis, regardless of its impact on performance (Fig. 4.6d, white circle). This led to immediate performance improvements for $T+$ targets and decrements for $T-$ targets (Fig. 4.5e, trial 1). As the trials continued, neural activity gradually decreased along the neural engagement axis for both target types (Fig. 4.6d, blue and red arrows). For $T-$ targets, this decrease in neural engagement was beneficial to performance, yielding progressively faster cursor speeds towards the target. For these targets, neural activity decreased similarly along the components of the neural engagement axis that were output-potent and output-null to cursor velocities under the new BCI mapping (Fig. 4.6d, red arrow). By contrast, for $T+$ targets, neural activity decreased along the output-null components of the neural engagement axis, but maintained the initial increase in the output-potent components (Fig. 4.6d, blue arrow). This allowed the immediate performance improvements from the increase in neural engagement on trial 1 to be maintained, even as total neural engagement decreased. As a result, monkeys reached peak performance more quickly for $T+$ targets than for $T-$ targets.

These results indicate that during learning, neural population activity did not change gradually from the activity observed before learning (Fig. 4.6d, "avg. in Block 1") to the activity at the end of learning (Fig. 4.6d, "end of Block 2"). Rather,

neural population activity underwent an abrupt change at the start of learning, improving performance for some targets and impairing performance for others. While the performance levels at the end of learning were similar for both target types (Fig. 4.5e), the path along which neural population activity changed to achieve this performance was quite different (Fig. 4.6d). These findings help to explain why some targets were learned more quickly than others.

4.3 Discussion

We have shown that large, trial-to-trial fluctuations in M1 population activity along neural engagement axes exhibit hallmarks of an arousal- or motivation-like process. On the first trials that monkeys used a new BCI mapping, neural activity increased abruptly along neural engagement axes, regardless of the effect on behavioral performance. On subsequent trials, neural activity retreated along neural engagement axes, impacting monkeys' performance using the new BCI mapping. The way that neural engagement axes related to behavior allowed us to predict which targets would be learned more quickly than others. Our findings indicate that 1) changes in neural activity during learning need not be a gradual transition between the activity produced prior to learning and the activity produced at the end of learning, and 2) changes in internal states seemingly unrelated to goal-seeking behavior can systematically influence how behavioral performance improves with learning.

In this study, we found that trial-to-trial changes in neural engagement were positively correlated with changes in the monkey's pupil size, a common psychophysical index for an animal's internal state (McGinley et al., 2015). The term "internal state" is used broadly, but typically refers to any neural signal that does not directly reflect, but may interact with, sensory encoding or behavior generation (McGinley et al., 2015). This includes internal states related to computation (e.g., internal models (Shadmehr and Holcomb, 1997), reward prediction (Schultz et al., 1997), working memory (Constantinidis and Klingberg, 2016)), as well as those reflective of more autonomic processes (e.g., arousal (Vinck et al., 2015), motivation (Mazzoni et al., 2007), task engagement (Steinmetz et al., 2019)). We have termed the internal state identified in the present work "neural engagement" because its stereotyped time course was suggestive of changes in the monkey's engagement with the task throughout the experiment (e.g., increases in neural engagement following pauses in the experiment and the introduction of a new BCI mapping). This is likely distinct from the concept of "task engagement" (Steinmetz et al., 2019), which refers to the difference between an animal actively performing a task versus passive sensory stimulation. Our results add to a growing line of work finding that there are large, systematic changes in M1 activity that are not related on a moment-by-moment basis to movement kinematics (Kaufman et al., 2016; Russo et al., 2018).

While our current study design does not allow us to identify the exact source of changes in neural engagement, here we briefly consider multiple possibilities and how they might explain (or fail to explain) the results in the present work. These

ideas are addressed in more detail in Supplemental Discussion. First, we observed that increased neural engagement predicted increased hand speed towards the target (Supplemental Fig. B.6), suggesting that neural engagement may simply reflect the monkey's intended movement speed. However, this seems unlikely given that neural engagement decreased over time even for $T+$ targets, which would decrease reward rate. Thus, neural engagement is likely not directly related to the monkey's intended movement speed. Second, neural engagement may reflect a default motor response such as muscle co-contraction. Co-contraction is thought to be a default strategy for reducing kinematic errors early in learning (Osu et al., 2002). However, we also observed an increase in neural engagement following pauses during Block 1 (Fig. 4.2e), when there were no unexpected kinematic errors to correct. Thus, if neural engagement does reflect a default motor response such as co-contraction, this response may be the manifestation in motor cortex of an uncertainty- or arousal-driven response, rather than a response to kinematic errors. Finally, neural engagement may reflect changes in animals' arousal, as indicated by its correlation with pupil diameter. We speculate that neural engagement may have a similar origin as the "neural drift" identified in V4 and prefrontal cortex (Cowley et al., 2020).

How might neural engagement impact learning? We found that neural engagement increased abruptly early in learning, and then gradually decreased on subsequent trials, regardless of its impact on behavior. We can imagine two mechanisms by which these stereotyped changes in neural engagement might impact the learning process. First, modifying one's future behavior to improve performance requires feedback. Thus, whenever changes in neural engagement directly impact behavior (i.e., BCI cursor movements), this will also impact the monkey's feedback about his performance, which will necessarily influence the learning process. This is likely to be the case during arm movements as well, as we found that changes in neural engagement were related to hand speeds during hand control experiments (Supplemental Fig. B.6). A second, not mutually exclusive possibility is that neural engagement may influence which neural activity patterns the monkey returns to on later trials via a reinforcement-like process (Athalye et al., 2018). For example, early in learning when levels of neural engagement were higher, animals produced neural activity patterns with an increased likelihood of reward for $T+$ (as compared to $T-$) targets. If neural activity patterns are more likely to be revisited when they lead to higher rewards, then the animals may revisit these same neural activity patterns on later trials for $T+$ targets. This could explain why $T+$ targets were learned more quickly than $T-$ targets.

Previous work has observed an increase in neural variability early in learning (Athalye et al., 2017). It was proposed that this increased variability may allow the brain to improve behavior by exploring new neural activity patterns. Our results confirm this observation, and may offer a new interpretation. We found that increased neural engagement corresponded with increased modulation depths and neural variability in most individual neural units (Supplemental Fig. B.5), consistent with findings of increased neural variability at the beginning of learning in previous studies (Athalye et al., 2017). If changes in neural activity at the beginning of learning

are indeed related to an exploratory drive, our results suggest that this process may not be a general exploratory process (i.e., an increase in variability without changing the mean), but rather a stereotyped exploration process (i.e., neural activity moves in a particular direction along neural engagement dimensions, which also results in more variability). Importantly, we found that neural engagement (and thus, neural variability) also increased even in the absence of learning (i.e., following pauses in Block 1). This raises the possibility that the changes in neural variability observed during learning in previous work may be driven by different internal states–not only by an exploratory drive, but also by internal states that reflect the animal's arousal or uncertainty about its environment. Future work is needed to disentangle the effects of these diverse sources of variability in neural population activity, all of which may impact behavior, and thus learning.

Overall, our results add to a growing body of work finding population-level signatures of internal state fluctuations (Ecker et al., 2014; Rabinowitz et al., 2015; Lin et al., 2015; Ni et al., 2018; Stringer et al., 2019; Allen et al., 2019; Cowley et al., 2020). Changes in internal state can influence on a moment-by-moment basis how we perceive a sensory stimulus (e.g., through attention), or how we execute a movement (e.g., through vigor). Building upon these previous studies, we found that changes in internal state can influence how behavior evolves during learning. Thus, internal state fluctuations can influence not only concurrent behavior, but also future behavior through their interaction with the learning process.

5 Other contributions

5.1 New neural activity patterns emerge with long-term learning

Consider a skill you would like to learn, like playing the piano. How do you progress from "Chopsticks" to Chopin? As you learn to do something new with your hands, does the brain also do something new? We found that monkeys learned new skilled behavior by generating new neural activity patterns. We used a brain–computer interface (BCI), which directly links neural activity to movement of a computer cursor, to encourage animals to generate new neural activity patterns. Over several days, the animals began to exhibit new patterns of neural activity that enabled them to control the BCI cursor. Thus, new neural activity patterns emerge with learning. We demonstrate that these new neural activity patterns cause the new behavior. This suggests that learning to play the piano and other skills might also involve the generation of new neural activity patterns.

Published as: Oby, E.R., Golub, M.D., Hennig, J.A., Degenhart, A.D., Tyler-Kabara, E.C., Byron, M.Y., Chase, S.M. and Batista, A.P., 2019. "New neural activity patterns emerge with long-term learning." *Proceedings of the National Academy of Sciences*, 116(30), pp.15210-15215.

5.2 Intracortical brain-machine interfaces

A brain–machine interface, or BMI, directly connects the brain to the external world, bypassing damaged biological pathways. It replaces the impaired parts of the nervous system with hardware and software that translate a user's internal motor commands into action. In this book chapter, we discuss the four basic components of an intracortical BMI: an intracortical neural recording, a decoding algorithm, an output device, and sensory feedback. An introduction to these concepts can be found in Chapter 2.

Published as: Oby, E.R., Hennig, J.A., Batista, A.P., Byron, M.Y. and Chase, S.M., 2020. "Intracortical brain–machine interfaces." In *Neural Engineering* (pp. 185-221). Springer, Cham.

6 Implications for learning

6.1 Introduction

Learning is an essential ability by which individuals accumulate knowledge and develop skillful behavior. Within a single lifetime, animals must learn how to move and control their bodies, efficiently navigate their surroundings, obtain resources, and communicate with others–all in a constantly changing environment. While these abilities improve with experience due to changes in the firing activity of populations of neurons throughout the brain, the exact ways in which this happens is poorly understood.

What principles underlie the brain's ability to learn? Answers to this question have spanned a range of different levels of description. At the microscopic level, studies have revealed how plasticity laws modify synapse strengths between neurons during learning (Rioult-Pedotti et al., 2000). At the macroscopic level, studies of organisms' behavior have revealed that behavioral changes during learning develop on a range of different timescales (Smith et al., 2006; Wang et al., 2018), and are guided by different types of feedback such as supervision (Ölveczky et al., 2005), reward (Schultz et al., 1997), and sensory prediction errors (Shadmehr and Holcomb, 1997).

In this paper, we focus on an intermediary between behavior and plasticity: the neural population. The firing activity of populations of neurons ultimately drives behavior while also being affected by changes in plasticity, making the neural population an ideal focus point for studying how the brain changes to improve behavior (Sohn et al., 2020). We argue here that studying learning in neural populations has already begun to provide unique insights into how the brain learns. Advances in recording technology provide an opportunity to understand how large numbers of neurons change their activity throughout a learning experience (Oby et al., 2019; Zhou et al., 2019; Bartolo et al., 2020a). While many studies have considered how single neurons or populations of neurons change their activity before and after a learning experience, studies of how neural population activity changes throughout a learning experience are relatively rare. Here we place a special emphasis on these latter studies, highlighting the unique benefits of this approach.

How might we interpret changes in neural population activity observed during learning? Theories and methods from statistics and machine learning may provide a useful starting point. After all, while the brain may be the most generally powerful

learner we are aware of (Lake et al., 2017; Sinz et al., 2019), artificial neural networks (ANNs) can also learn complex behaviors, sometimes even exceeding human levels of performance (Mnih et al., 2015; Brown and Sandholm, 2019; Schrittwieser et al., 2020). An increasing number of studies of artificial networks have revealed remarkable similarities between the network activity of trained ANNs and the activity of neurons in animals performing the same task (Mante et al., 2013; Cadieu et al., 2014; Sussillo et al., 2015; Engel et al., 2015; Yamins and DiCarlo, 2016; Saxe et al., 2020). These results make the tantalizing proposal that much of the complexity and structure we observe in neural population activity during trained behaviors may arise simply from the task itself. While a correspondence between artificial and biological networks has been observed only after each network has finished learning, these results suggest that we may be able to understand learning in the brain similar to how we understand learning in artificial networks (Marblestone et al., 2016; Richards et al., 2019).

From this point-of-view, which we refer to as the optimization framework (Fig. 6.1), learning is a process by which a system optimizes a particular objective or cost function with respect to some learning rule, subject to constraints. For example, one can train an artificial network to recognize handwritten digits using an objective function known as cross-entropy, and a learning rule known as backpropagation. The idea that the brain learns by modifying its activity to improve an objective function is not new, and exists in many subfields such as those studying perceptual learning (Yamins and DiCarlo, 2016), motor learning (Haith and Krakauer, 2013), and reinforcement learning (Neftci and Averbeck, 2019). While learning in the brain is certainly more complicated than in artificial networks (e.g., different brain areas may have different objective functions), trying to identify the brain's objective function and learning rule for a given behavior may nevertheless be a promising approach to understanding changes in neural population activity during learning. However, as we will argue here, studies of neural population activity during learning have identified a number of distinct features not typically present in the activity of ANNs. These include 1) the presence of substantial neural variability throughout learning; 2) the existence of multiple objective functions, often with distinct time scales; and 3) task-irrelevant changes in network activity, which can persist even when they negatively impact task performance. As we will discuss, these results refine our understanding of the extent to which learning in the brain can be described as an optimization process.

In what follows we will outline three key observations about learning identified from studies of neural population activity. These ideas emphasize the importance of understanding not just the endpoint of learning, but also the path that neural activity takes to get that endpoint. For each observation, we suggest ways in which this result may apply to learning in artificial agents, as well as how it refines a description of learning in the brain as an optimization process. We hope that these results will spur refinements to our understanding of learning in the brain, which in turn may inspire improved methods for learning in artificial networks.

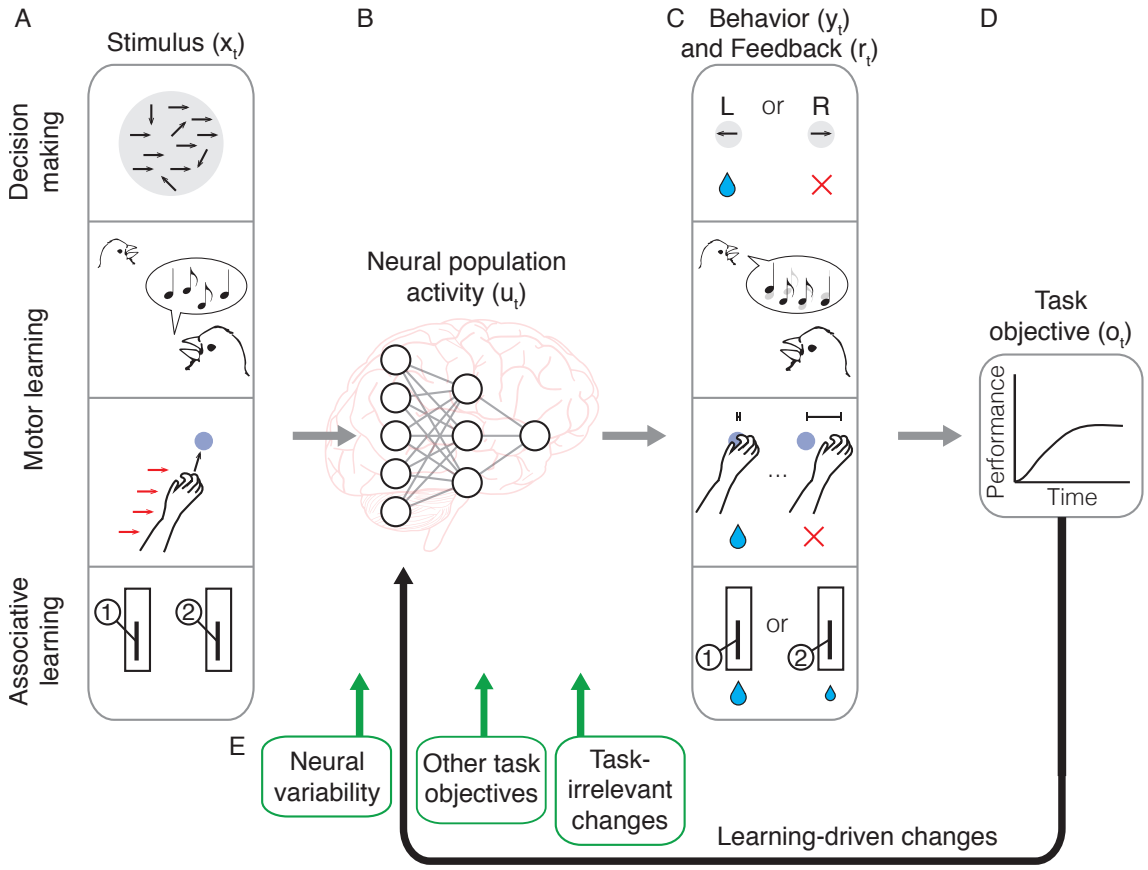


Figure 6.1. The optimization framework for understanding changes in neural population activity during learning. **A.** On each trial t, the provided stimulus, or sensory input, cues an animal to provide a particular behavior. **B.** This stimulus is provided to the network as input, resulting in a particular pattern of network activations, or population activity. **C.** The output of the network is the animal's chosen behavior given the stimulus, and the outcome of this behavior is available via feedback. **D.** Performance at the task on trial t can then be quantified as a scalar "task objective" or performance metric. Learning is defined as the improvement in this performance metric over time. **E.** At the end of the trial, the network will be updated (e.g., via changes in synapse strengths in the network) so as to try to improve performance on subsequent trials (black arrow, "Learning-driven changes"). As discussed in the main text, changes in the network may also be driven by other sources (green arrows).

6.2 Learning about learning in neural population activity

1. Neural "noise" during learning is important but inflexible

Given the same task and context, our behaviors are notoriously variable. Even star basketball players miss free throws, and professional musicians sometimes play the wrong note. Rather than being "the unintended consequence of a noisy nervous system" (Dhawale et al., 2017), behavioral variability may in fact be critical for learning, allowing us to fully explore reward landscapes and adapt to ever-changing environments. For example, juvenile songbirds learn the songs of their adult tutors by trial-and-error, effectively exploring different vocal strategies until they find the one that best matches their tutor's song (Ölveczky et al., 2005). As the bird ages, his song "crystallizes," becoming more accurate and less variable (Konishi, 1985; Tumer and Brainard, 2007). Evidence in a variety of species indicates that behavioral variability is reduced when the stakes are high, suggesting that the amount of variability can be dynamically regulated (Dhawale et al., 2017). When learning occurs by trial-and-error, the amount of task-relevant behavioral variability can predict how quickly different task goals will be learned (Wu et al., 2014). These results suggest that the brain may actively control behavioral variability to facilitate learning.

Of course, variability is present not only at the level of behavior, but also throughout the brain. In contrast to behavioral variability, the role of neural variability during learning is somewhat more complicated. On the one hand, any behavioral variability used for trial-and-error learning must be driven by variability in neural activity; and in birdsong, a long line of studies has established a theory of the relationship between neural variability, behavioral variability, and learning (Ölveczky et al., 2005; Kao et al., 2005; Ölveczky et al., 2011). On the other hand, a substantial amount of variability in neural population activity does not appear to be driven by considerations of improving performance or learning. Like behavior, the spiking activity of individual neurons shows substantial trial-to-trial variability even given repeated trials with the same context. Because this variability is not limited to single neurons, but is in fact shared across populations of neurons in the population, we refer to this as "population covariability." Importantly, there are converging lines of evidence suggesting that population covariability both limits performance and is relatively constrained, even over the course of days or weeks of practice. The first line of evidence relates to the existence of neural variability that interferes with the ability to decode information (e.g., about a stimulus) from neural activity, so-called "information-limiting" correlations (Moreno-Bote et al., 2014). Because these correlated fluctuations in firing activity are aligned with stimulus-encoding dimensions of the population, they cannot simply be averaged away by pooling over large numbers of neurons, potentially limiting behavioral performance. These correlations may be reduced by learning, given multiple days of training (Ni et al., 2018). Nevertheless, information-limiting correlations persist even in over-trained animals (Bartolo et al.,

2020b). While information-limiting correlations have been studied primarily in the context of perceptual learning, they have relevance to other types of learning as well: just as neural variability in sensory areas can interfere with accurate stimulus decoding, neural variability in motor areas can interfere with accurate movement readout. These findings highlight the fact that, depending on its structure, neural "noise" cannot always simply be averaged away or attenuated by continued experience.

The second line of evidence that population covariability can limit learning comes from learning studies using a brain-computer interface paradigm (BCI), in which the structure of population covariability predicted which tasks could be easily learned within a few hours. In these experiments, population activity recorded in primary motor cortex (M1) controlled the movements of a visual cursor on a screen according to a BCI "mapping." The utility of a BCI is that it provides a direct causal test of population covariance structure. That is, one can challenge animals to break that structure, and observe whether or not they are able to do so. Animals were rewarded for successfully navigating the cursor towards a cued visual target. To learn, animals used visual feedback about the cursor's position to modify the population activity they produced on subsequent trials. Within a single day, subjects showed more performance improvements for the BCI mappings that were aligned with dimensions of the population that reflected the bulk of the shared variance among neurons (Sadtler et al., 2014). Later work found that the structure of neural covariability both before and after learning a new BCI mapping was largely unchanged (Golub et al., 2018; Hennig et al., 2018), suggesting that constraints on population covariability throughout learning limited which types of BCI mappings could be learned within a day. These results indicate that population covariability may not always be easily modifiable, even when that variability interferes with task performance.

The role of variability in the brain during learning can be contrasted with the role of variability in modern methods for reinforcement learning (RL). Typically, RL agents must fully explore their environment in order to discover the optimal policy. One common approach to encouraging sufficient exploration, known as an "ϵ-greedy" strategy, is to add variability in the action selection process, such that the agent will occasionally select a suboptimal action with respect to its policy. In this case, the resulting variability is simple in structure (e.g., parametrized by a single scalar parameter), and exists only at the level of action selection (i.e., behavioral variability) rather than being internal to the agent (e.g., the added variability does not impact the policy itself). While more sophisticated approaches to adding variability to encourage exploration do exist, how that variability decreases over time is typically left up to the modeler, rather than being an invariant property of the network. By contrast, covariability in the brain is non-trivial (e.g., non-isotropic, and signal dependent), exists throughout the network, and is not easily modifiable on short timescales. While learning in the presence of substantial population covariability may potentially improve robustness or generalization, there is much we have yet to understand about how this covariability may limit, or even improve, the learning process.

How might population covariability influence how learning proceeds in the brain?

Here we consider two non-mutually exclusive possibilities. First, from the perspective of reinforcement learning, some aspects of population covariability may reflect exploratory changes in network activity that can be used to discover (or reinforce) the population activity patterns that lead to improved behavior. Because variability exists throughout the network as opposed to only at the level of behavioral output, this exploratory process may be akin to direct policy search methods such as evolutionary strategies or policy gradient, in which variability in population activity is used to explore new control policies. Second, as some have suggested, learning in the brain may proceed in a manner similar to gradient descent (Marblestone et al., 2016). For either possibility, the brain must use feedback from the task to estimate the direction, or gradient, in which population activity should be modified in order to improve future performance. Importantly, this gradient can depend on the structure of population covariability. For example, it may be easier to learn an accurate gradient when behavior depends on dimensions of the population with large, as opposed to small, variability (Feulner and Clopath, 2021). This may explain why some tasks can be learned in only a few hours while others cannot. Second, if some aspect of population covariability is external to the learner (and is in this sense "noise"), this will bias the resulting gradient, as well as the optimized solution (Todorov, 2005). For example, if the noise distribution is non-isotropic, population activity will not move along the steepest gradient direction, but will be skewed by the shape of the noise distribution. These ideas indicate how the structure of population covariability both limits and shapes how neural activity changes during learning.

Even given knowledge of the objective function and learning rule used by the brain to learn a given task, understanding the path taken by population activity during learning depends on characterizing the drivers of population covariability. For example, while some components of variability are driven by the learning process (e.g., due to a policy update or gradient step), other aspects may be exploratory variability accessible to the learner, while variability not driven by task considerations may be beyond the control of the learner and act effectively as "noise." The goal of characterizing covariability during learning is likely a moving target, as population covariability is typically signal-dependent (e.g., different covariance structure for different task goals), and nonstationary (e.g., due to changes in learning rate). Whether learning in the brain proceeds in a manner akin to gradient descent or reinforcement learning, understanding the structure of neural variability is critical for understanding how neural activity is modified to yield improved performance.

2. Multiple neural learning mechanisms are in play at all times

Learning in both artificial and biological systems is often equated with the optimization of a single, scalar objective function over time. From this perspective, one natural approach to understanding how learning proceeds in the brain is to see whether changes in population activity over time indicate signatures of optimizing that particular objective function. However, most real-world tasks are likely multi-

faceted, such that improving performance at a given task may require optimizing multiple different sub-objectives, each acting on different sets of information. For example, a child learning to communicate may use auditory feedback to minimize errors in motor execution (e.g., a child trying to say "dog" may actually say "dah"), as well as instruction from a parent or teacher to maximize accuracy (e.g., "No, that's not a *dog*, that's a cat."). Indeed, evidence from behavioral and electrophysiological studies indicates that, even in simple tasks, learning in the brain is driven by improvements in a multitude of objective functions, each acting with respect to distinct feedback signals, and at different time scales (Marblestone et al., 2016). Thus, even in simple tasks, changes in neural activity during learning are unlikely to be driven by a single task-related objective function.

Behavioral studies of motor learning provide perhaps the best example that even a seemingly simple behavior, such as moving one's hand towards a target, can simultaneously invoke a variety of different learning processes, each operating on a distinct aspect of feedback. For example, it has been proposed that the brain learns new motor skills via supervised learning, unsupervised learning, and reinforcement– either all at the same time (Doya, 2000), or switching between them dynamically (e.g., based on their reliability) (Izawa and Shadmehr, 2011). Distinct learning processes may also be present even within a given feedback modality. For example, errors in the extent of a reach may be learned separately from errors in direction (Krakauer et al., 2000), and errors in direction may themselves be composed of multiple processes, one fast and one slow (Smith et al., 2006). These errors, driven by visual feedback about the movement of the hand, are thought to drive learning of an internal model that predicts the outcome of motor commands (e.g., population activity in motor areas) on motor output (e.g., the resulting muscle force or arm kinematics). Learning processes driven by moment-by-moment error are thought to minimize error using a gradient-descent-like learning rule. This can be distinguished from other processes driven only by task success (e.g., the reward given at the end of the trial) that may resemble reinforcement learning (Diedrichsen et al., 2010a). These behavioral findings illustrate the variety of different learning processes that even an apparently simple task might invoke.

Evidence that distinct learning processes coexist even during simple tasks is not the exception but the rule. Neural correlates of multiple learning mechanisms have been identified in a variety of different systems, including systems for motor learning (Zhou et al., 2019; Oby et al., 2019), perceptual learning (Poort et al., 2015), rule learning, and reinforcement learning (Schultz, 2019). Because these learning processes must all ultimately act on a common downstream population in order to influence behavior (e.g., via the spinal cord), it may be inevitable that some of these learning processes interfere (Mazzoni and Krakauer, 2006). While sometimes these processes may proceed in distinct brain areas, multiple learning processes may also co-exist within the same population, either by acting on distinct time scales (Schultz, 2019), being multiplexed in time (Lak et al., 2016), or by acting on distinct modes of the neural population. Learning mechanisms may even proceed through entirely distinct substrates (e.g., via synaptic plasticity vs. population activity) (Sohn et al.,

2020).

The presence of different learning mechanisms in the brain, each with their own objective functions and time scales, impacts how quickly new behaviors can be learned. For example, after first being introduced to a new task–such as playing a new video game–learning the rules and rudimentary strategies may be possible within only a few minutes, while mastery takes weeks or even years. BCI learning paradigms similar to those mentioned in the previous sections have begun to identify the neural signatures distinguishing short vs. long-term learning. In a recent study (Zhou et al., 2019), population activity showed distinct changes following learning for a few hours versus learning for multiple weeks. These fast and slow changes to population activity appeared to be responsible for distinct aspects of the task–directional errors and reward rate, respectively. Another study of long-term learning using a BCI paradigm found that, given weeks of practice, animals learned a difficult BCI mapping by developing the capacity to generate new patterns of population activity that were not observed before the learning experience (Oby et al., 2019). In both of these studies, the neural mechanisms associated with long-term learning involved substantial changes to the correlations in firing across neural units. Dramatically changing the correlations between neurons may require altering their functional or even anatomical connectivity, a process that may be impossible on shorter timescales. Why might the brain limit how much learning can be expressed on shorter timescales? One of the primary differences between learning in biological and artificial agents is that, for a living creature, the environment and reward contingencies are constantly changing (Neftci and Averbeck, 2019). To deal with the complexities of learning in such a dynamic environment, rather than being a blank slate of optimization tools, the brain comes ready-prepared with a variety of ecologically-relevant inductive biases (Zador, 2019). These inductive biases are effectively "priors" about how to correct the errors in behavior that are most likely to be encountered, and facilitate rapid learning. As a consequence, more substantial (and energetically demanding) neural changes, such as changes to the brain's architecture and connectivity, may occur only when doing so is shown to be beneficial over longer timescales. The process of learning these inductive biases, also referred to as "meta-learning," may be the consequence of multiple, nested learning processes, such that the output of faster learning mechanisms "trains" the slower learning mechanisms (Pasupathy and Miller, 2005; Wang et al., 2018). This perspective highlights the fact that, in the short-term, rather than implementing a general-purpose optimization machine, the brain may employ a collection of heuristics and "good-enough" solutions sufficient for solving most problems (Beck et al., 2012; Zhou et al., 2019). When the brain's inductive bias is insufficient, even seemingly simple tasks can be impossible to learn (e.g., mirror reversal learning) (Hadjiosif et al., 2021). For more ecologically relevant problems, these heuristics may not only be sufficient, but they may also explain why humans can quickly learn some tasks that artificial agents struggle with (Sinz et al., 2019).

That the brain learns new tasks using multiple learning objectives strikes a contrast with reinforcement learning agents, which learn to optimize a single scalar

value, given by the reward function. Because maximizing this scalar reward is the primary goal of the agent, the modeler must carefully define the reward function so that an agent optimizing this reward will perform the behavior he wishes the agent to exhibit. While choosing the reward function for some tasks is obvious (e.g., winning a chess game), for other tasks the appropriate reward function may be more ambiguous. This problem, referred to as reward shaping or "reward engineering," is an important unsolved challenge in reinforcement learning. So how does the brain do it? In contrast to RL, the brain does not receive a scalar reward signal from the environment, but rather a multi-dimensional sensory feedback signal (Schultz, 2015). This rich feedback is available to the brain during learning, and as we've seen, multiple aspects of this feedback may then be used to train various learning mechanisms in parallel. As discussed above, the objectives of these learning mechanisms are likely a reflection of inductive biases regarding which objectives are likely to be most useful given statistics about the environment. Perhaps a focus on developing agents that can learn the optimal reward functions for a set of tasks, rather than these reward functions being defined by the modeler, could be a path forward for obviating the need for reward engineering (Singh et al., 2010, 2019b).

The coexistence of multiple learning processes in the brain poses a challenge to the optimization framework for understanding biological learning. For one, understanding learning in the brain would appear to require not only a description of the multitude of learning objectives, time scales, and neural mechanisms employed to learn any given task, but also a characterization of the organism's inductive biases. In particular, the brain may come with various heuristics for learning "out-of-the-box." For example, animals in decision-making tasks often exhibit substantial biases to choose one target over the other, regardless of the stimulus or task (Ashwood et al., 2020). Additionally, while the stimuli shown on subsequent trials may be independent in a statistical sense, this does not mean that the brain's learning mechanisms treat learning experiences on adjacent trials as independent (Mendonça et al., 2020). These results suggest that an understanding of how neural activity changes during learning may need to consider not just the presently learned task, but the history of tasks the animal has been previously exposed to.

3. Not all changes to neural activity during learning are driven by task performance

We tend to think of different brain regions as performing different functions. For example, visual cortex represents sensory stimuli, while motor cortex drives movements. Continuing this line of thought, sustained changes in a region's population activity during learning may reflect improvements in that region's function with respect to the task (Li et al., 2001). For example, following perceptual learning, changes in earlier visual areas may reflect improved representations of stimuli shown in the task (Schoups et al., 2001; Yan et al., 2014), while changes in downstream areas may reflect improved readout of those representations (Law and Gold, 2008; Uka et al., 2012). In this approach, correlating changes in neural activity with changes

in behavioral performance before and after learning provides a means for identifying which brain areas are recruited during learning.

However, even in the absence of learning pressure, neural activity is not stable, but shows substantial "drift" over time (Mau et al., 2020; Cowley et al., 2020). Thus, sustained changes in neural activity observed during learning are not solely driven by improving performance, and may also reflect task-irrelevant changes (Rokni et al., 2007; Okun et al., 2012; Singh et al., 2019a; Hennig et al., 2020). Task-irrelevant changes in neural activity may arise from a variety of sources, including synaptic turnover (Holtmaat and Svoboda, 2009), the accumulation of neural noise (Rokni et al., 2007), and the influence of task-irrelevant considerations such as changes in arousal (Cowley et al., 2020; Hennig et al., 2020). Regardless of their source, these neural changes have the potential to impact neural computation, and thus task performance. To prevent a task-irrelevant change from influencing behavior, the brain must account for the change so that it does not impact downstream processing or actions (Clopath et al., 2017; Rule et al., 2020). Indeed, some evidence suggests that the brain may do this by utilizing the redundant numbers of neurons encoding any given stimulus or action (Rokni et al., 2007; Xia et al., 2021).

Nevertheless, the impacts of "neural drift" and other task-irrelevant changes in neural activity are not always negated by downstream circuits, such that these changes impact behavior. For example, global changes in neural activity reflecting changes in an organism's arousal also have behavioral impacts. If you are chopping an onion and suddenly someone jumps out and scares you, the resulting change in your arousal will likely have an impact on your control of your knife. Similarly, sustained changes in arousal, due to fatigue, represent a task-irrelevant factor that has clear impacts on behavior, and thus performance.

While the impact of task-irrelevant changes in neural activity on concurrent behavior is clear, these changes also impact the learning process, influencing how neural activity changes over time. One recent example comes from a BCI learning study (Hennig et al., 2020). In this study, population activity in motor cortex varied not only with the direction in which subjects intended to move the cursor, but it also showed large and stereotyped changes following experimentally-controlled changes to the experiment such as pauses in the task, or changes to the BCI mapping. These task-irrelevant changes in neural activity impacted the speed and the direction in which the cursor moved, which in turn impacted how neural activity was updated on future trials. As a result, the manner in which these changes impacted cursor movements led to neural activity taking distinct paths in population space during learning, as well as predicted how quickly different task goals were learned. This study indicates that task-irrelevant changes in neural activity can nevertheless impact how learning proceeds via their impact on behavior.

Task-irrelevant changes in the brain during learning can be contrasted with the network activity in ANNs, where task-irrelevant changes are typically non-existent. For example, in an ANN trained by gradient descent, changes in network activity are determined by the learning rule, related to the gradient of the task objective, such that all changes are task-relevant. What relevance might the existence of task-

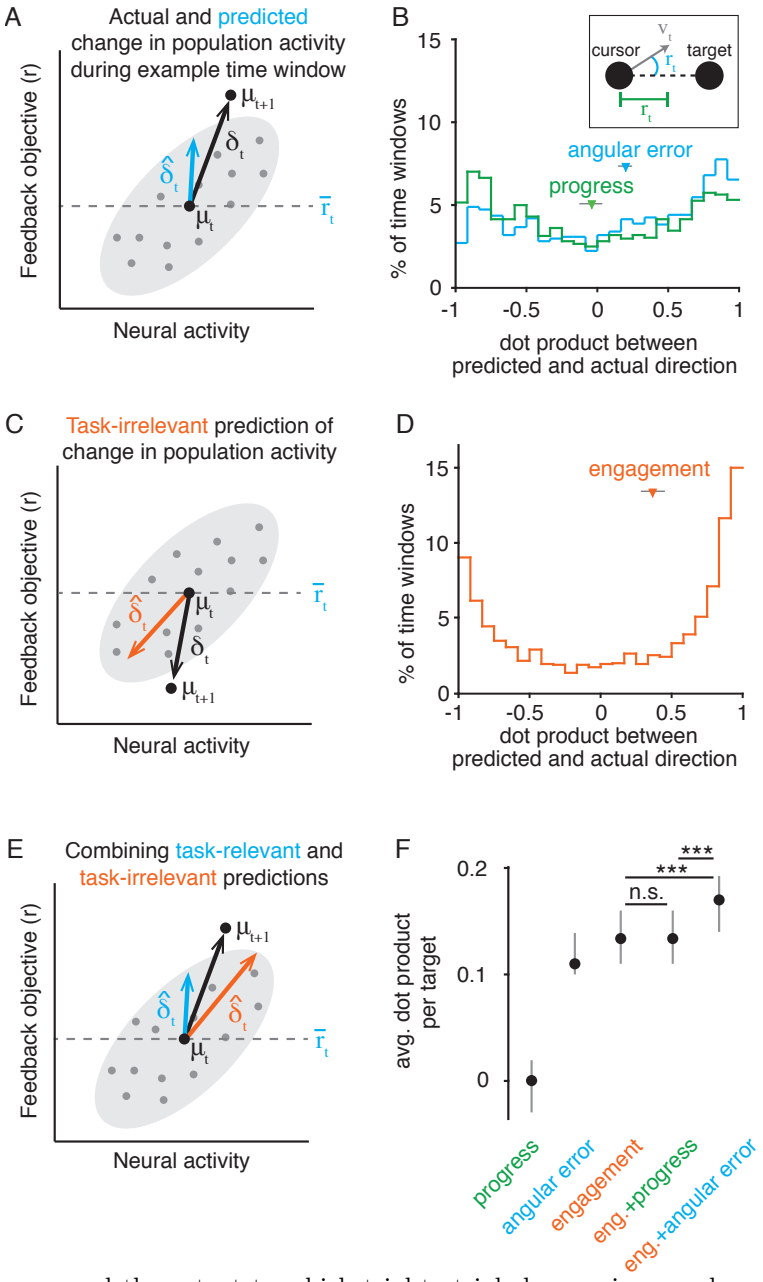


Figure 6.2. We assessed the extent to which trial-to-trial changes in neural population activity during learning in a BCI task (Sadtler et al., 2014) reflected the optimization of task-relevant feedback. **A.** During each block of trials (t), we compared the actual change in neural population activity ($\boldsymbol{\delta}_t$) during the next block of trials ($t+1$) to the prediction that population activity would change ($\widehat{\boldsymbol{\delta}}_t$) were it improving a hypothesized objective function (e.g., maximize r_t). **B.** Here we considered predictions made by two task-relevant objective functions (inset): the speed ("progress") and angular error of the cursor with respect to the target. Histogram depicts dot-product between predicted and actual directions across sessions. Neither hypothesized objective function reliably predicted the direction in which population activity changed on subsequent trials. **C.** According to a task-irrelevant hypothesis, population activity changes in a consistent direction regardless of the impact on task feedback. Same conventions as panel A. **D.** A task-irrelevant prediction based on decreasing "neural engagement" (Hennig et al., 2020) consistently predicted the direction population activity changed during learning. Same conventions as panel B. **E.** Combining task-relevant (as in panel A) and task-irrelevant (as in panel C) predictions. Same conventions as panel A. **F.** Changes in population activity during learning were best predicted by a combination of task-relevant and task-irrelevant influences. Black dots and error bars indicate mean ± SE across sessions.

irrelevant changes in the brain have for artificial learning? One possibility is that learning in the presence of task-irrelevant changes contributes to the robustness of the learner, allowing generalization to new contexts by encouraging insensitivity to large variance changes in neural activity. A second possibility is that studying how the brain manages both task-relevant and -irrelevant objectives has relevance to studies of multitask learning. For example, in addition to performing any given task, the brain is also responsible for maintaining all sorts of homeostatic processes (such as arousal, satiety, etc.) unrelated to the specifics of the task. Understanding how the brain distributes its available resources to maintain sufficient performance with respect to multiple objectives in parallel may be critical for developing artificial agents that can learn to perform a variety of different tasks.

The presence of substantial task-irrelevant changes in neural activity during learning may prove to be a critical ingredient in attempts to reverse-engineer the process by which the brain learns (Lak et al., 2016; Richards et al., 2019; Ashwood et al., 2020). Not accounting for task-irrelevant changes in population activity might yield incorrect conclusions about what the system's objective is. For example, one might observe population activity moving up a gradient rather than down (Richards et al., 2019); this doesn't mean your objective is wrong–it may instead be due to task-irrelevant changes in activity (Fig. 6.2). One important question is to what extent the different learning mechanisms in the brain are able to account for the influence of task-irrelevant changes in neural activity. For example, suppose that learning proceeds by exploration of new neural activity patterns, where the patterns that lead to improved reward are reinforced. Then it is important to understand whether the task-irrelevant components are accounted for by the learner, or whether they are treated as an additional exploratory pattern that may be reinforced. For example, performance during a learning session may begin as a monotonic curve indicating increased performance, but as time passes, mental and physical fatigue may set in, leading to performance decreasing back to baseline levels. If the learner is unaware of the influence of fatigue on neural activity, it may conclude that the neural activity generated at the end of the session was no better than its activity at the start of learning, given that performance was similar. Overall, understanding how neural activity changes during learning may require accounting for the influence of other processes being managed concurrently with task execution.

7 Future directions

In this chapter I will present some unpublished analyses and discussion relating to the structure of population activity during learning. These results, though tentative, represent my current understanding of these experiments, and might serve as a roadmap for future analyses or experiments.

First, the bulk of this chapter is Section 7.1, in which I will discuss how monkeys learn to use a particular type of new BCI mapping (a "within-manifold perturbation", or WMP; (Sadtler et al., 2014; Golub et al., 2018; Hennig et al., 2020)) given a few hours of practice. I will provide evidence that simply combining the results of neural reassociation and neural engagement is insufficient to describe the full range of results. I will then present a new, model-free hypothesis capturing many of the key features of how population activity changes during WMP learning. In addition, I will discuss how the covariance structure of population activity can influence the asymptote of learning, and I will also summarize what we currently know about how learning generalizes across adjacent targets.

Next, in Section 7.2 I will provide a simple firing rate model that explains differences in the "shapes" of population activity one might see when visualizing population activity in these experiments. In Section 7.3 I will show how the population activity on one trial can be used to decode the target identity on the *previous* trial, and I will provide a hypothesis for why this might be the case. In Section 7.4 I will summarize the variety of hypotheses for why learning WMPs is generally easier than learning OMPs. Finally, in Section 7.5 I will sketch out a probabilistic approach to inferring an animal's "objective function" during learning.

7.1 How are within-manifold perturbations (WMPs) learned?

Though we have gained many insights over the years about these BCI learning experiments, there is still much left to understand regarding how neural population activity changes throughout learning. In the following subsections I summarize my understanding on this topic.

Evidence for model-free learning

Introduction. To learn, the brain must use feedback to modify its neural activity in order so as to improve its future performance. Here we study how neural activity evolves during learning in a BCI experiment, where the relationship between neural activity and behavior is known. Say for a given target, we look at neural activity, $\mathbf{z}(t) \in \mathbb{R}^d$, from time $t = 1, \ldots, T$ during learning. (Suppose that t refers to a trial.) Let the BCI mapping from neural activity to cursor velocity be $g(\mathbf{z}) = \mathbf{v} \in \mathbb{R}^2$. What determines how neural activity evolves from time t to $t + 1$? We'll suppose that during learning, neural activity is changing in order to maximize some objective function, f, which is a function of the cursor velocity. In other words, the monkey's objective is to generate neural activity, $\mathbf{z}(t)$, that maximizes $r(t) = f(g(\mathbf{z}(t)))$.

Neural reassociation as "reaiming". From previous work (Golub et al., 2018), we know the repertoire of neural activity patterns produced at the beginning and end of learning is largely the same. The reassociation hypothesis, as implemented, actually makes an even stronger prediction: The relationship between the monkey's intended movement direction, $\theta \in [-\pi, \pi)$, and neural activity, $\mathbf{z} \in \mathbb{R}^d$, is fixed. Let h be a function that maps the intended movement direction θ into the high-d population activity associated with that movement direction, so we have $\mathbf{z} = h(\theta)$ (Fig. 7.1). Then the *Reaiming* model says that, after learning a new BCI mapping, the monkey has simply found the θ that maximizes $f(g(h(\theta)))$. (In previous work, f is cursor progress, but in general it could be anything.)

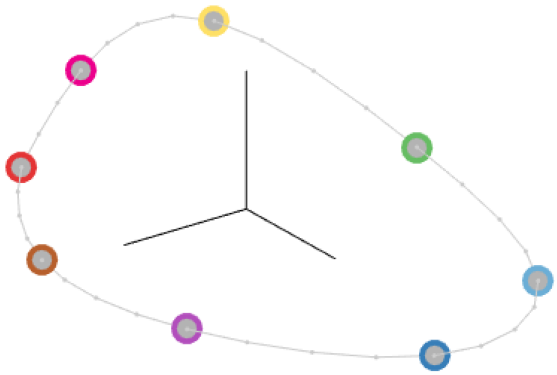


Figure 7.1. Average neural activity ($\mathbf{z} \in \mathbb{R}^d$) before learning during an example session, in the first three factor dimensions, for each intended movement direction ($\theta \in [-\pi, \pi)$), with activity for the eight target directions indicated by the colored circles. According to neural reassociation, neural activity throughout learning lives somewhere on the gray ring, because the relationship between θ and $\mathbf{z}$ is fixed throughout learning. In other words, the figure depicts $\mathbf{z} = h(\theta)$ for each value of θ, where h is some fixed function relating θ and $\mathbf{z}$.

We'd like to upgrade this hypothesis as a model that describes how neural activity changes from time $t = 1$ to $t = T$. For simplicity, let's suppose that the actual target direction is the 0° target, so that $\theta(t)$ can be thought of as how much reaiming the monkey is doing on trial t. To put the Reaiming hypothesis another way, we know that at the end of learning we have $\mathbf{z}(T) = h(\theta(T))$, where $\theta(T)$ is what's been

learned. So let's suppose that, during learning, the monkey is gradually reaiming, by updating $\theta(t)$. For simplicity let's suppose that our objective is to maximize $f(\mathbf{v}(t)) = -\frac{1}{2}e(t)^2$, where $e(t) \in [-\pi, \pi)$ is the angular error of the velocity $\mathbf{v}(t)$. Then our learning rule is:

$$\begin{aligned} \mathbf{z}(t) &= h(\theta(t)) + \boldsymbol{\epsilon}(t) \\ \theta(t+1) &= \theta(t) - \gamma_l e(t) \end{aligned} \tag{7.1}$$

where $\gamma_l \in \mathbb{R}$ is the learning rate, $\theta(1) = 0^\circ$ (since we're assuming 0° is the actual target direction), and $\boldsymbol{\epsilon}(t) \in \mathbb{R}^d$ is i.i.d. noise. This model essentially performs gradient descent on θ with respect to our objective of minimizing squared angular error. You can think of this model as describing how neural activity during learning moves along the gray ring in the figure above.

Reaiming + engagement. From our work on neural engagement (Hennig et al., 2020), we know that the model above is incomplete. This is because during learning we also have changes in neural engagement, $\boldsymbol{\delta}_{ne}(t) \in \mathbb{R}^d$, where $\|\boldsymbol{\delta}_{ne}(t)\|$ is large on $t = 1$, and exponentially decays towards zero on subsequent trials. Here we'll assume that changes in neural engagement during learning can be written as:

$$\boldsymbol{\delta}_{ne}(t+1) = \gamma_{ne}\|\boldsymbol{\delta}_{ne}(t)\|\mathbf{x}(t+1) \tag{7.2}$$

where $\gamma_{ne} < 1$ is the decay rate, and $\mathbf{x}(t) \in \mathbb{R}^d$ is the neural engagement axis at time t. Note that $\mathbf{x}(t)$ depends on $\theta(t)$, since we know the neural engagement axis varies with the intended movement direction.

So to improve on the *Reaiming* model, we have that neural activity during learning is described by learning-driven changes (characterized by reaiming), and learning-irrelevant changes (characterized by neural engagement):

$$\begin{aligned} \mathbf{z}(t) &= h(\theta(t)) + \boldsymbol{\delta}_{ne}(t) + \boldsymbol{\epsilon}(t) \\ \theta(t+1) &= \theta(t) - \gamma_l e(t) \\ \boldsymbol{\delta}_{ne}(t+1) &= \gamma_{ne}\|\boldsymbol{\delta}_{ne}(t)\|\mathbf{x}(t+1) \end{aligned} \tag{7.3}$$

***Reaiming + engagement* is insufficient.** To assess how well this learning model describes our data, we can see if it can reproduce our main results at a qualitative level. For each target in sessions from monkey J, I simulated the time course of neural activity during learning using the *Reaiming + engagement* model with $\gamma_l = 0.01$, $\gamma_{ne} = 0.95$, and $\|\boldsymbol{\delta}_{ne}(1)\| = 2$. $\mathrm{Var}[\boldsymbol{\epsilon}(t)]$ was 0.01 times the variance of factor activity during intuitive control.

Reaiming + engagement can reproduce the differences in performance we see between T+ and T- targets (Fig. 7.3, middle column; top row is the data, bottom row is the model). However, there are a few signs here that this model is an insufficient description of the data. First, this model *cannot* reproduce the differences in neural engagement between T+ and T- targets (left column). This is because, according to this model, changes in neural engagement during learning occur regardless of their impact on performance. Second, in the data we see more speed improvements than in

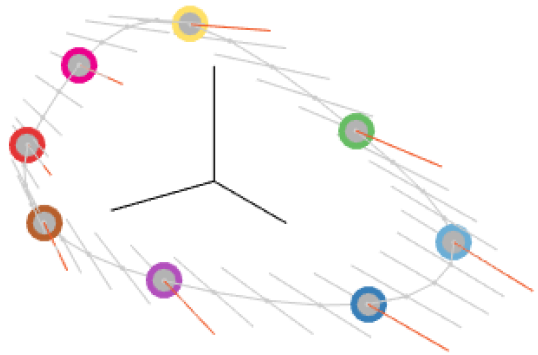


Figure 7.2. Same as previous figure, but with the neural engagement axes, $\mathbf{x}$, corresponding to each intended movement direction shown as orange vectors. According to the *Reaiming + engagement* model, neural activity during learning can move anywhere along the gray ring, while also decreasing down the orange axes.

the model (compare y-axes in middle column). Finally, in the data we have potent-engagement asymptoting to different values for T+ vs. T- targets, whereas under Reaiming it decays to zero in both cases (right column). All of these points suggest there are more (or different) learning-driven changes beyond simply reaiming.

Weight perturbation (model-free). According to the above model, all of the learning-driven changes in neural activity are due to changes in θ, the amount of reaiming. This greatly limits the space that neural activity can explore during learning, because our model parameter, θ, has lower dimensionality than our neural activity, $\mathbf{z}$. Alternatively, we could consider a model-free approach, where we can fully explore neural activity space, and maintain any changes that happen to improve our objective.

$$\begin{aligned} \mathbf{z}(t) &= h(\theta) + \boldsymbol{\delta}_l(t) + \boldsymbol{\delta}_{ne}(t) + \boldsymbol{\epsilon}(t) \\ \boldsymbol{\delta}_l(t+1) &= \gamma_d \boldsymbol{\delta}_l(t) + \gamma_l (r(t) - \overline{r}(t)) \boldsymbol{\epsilon}(t) \end{aligned} \tag{7.4}$$

where θ is our initial aiming direction (which we assume does not change), $\overline{r}(t)$ refers to a running mean of $r(t)$ (e.g., computed via exponential smoothing), and $\gamma_d > 0$ is a weight decay term, discussed below.

Above, our learning-driven activity is $\boldsymbol{\delta}_l(t+1)$. This term accumulates the noise, $\boldsymbol{\epsilon}(t)$, based on whether or not we just had above- or below-average evaluations of the objective value, $r(t)$. This is a REINFORCE learning rule Williams (1992) that approximates gradient ascent on neural activity to maximize our objective Jabri and Flower (1992); Cauwenberghs (1993), provided $\boldsymbol{\delta}_{ne}(t)$ is constant, and $\gamma_d = 1$. (Compare to *Reaiming*, where we optimize our model parameters, θ, instead of neural activity.) However, if our objective is unbounded (e.g., cursor progress) this model predicts that neural activity will continue to change without ever converging (i.e., eventually leaving the neural repertoire). To prevent this, we suppose that our learning-driven changes, $\boldsymbol{\delta}_l(t)$, have some decay, by setting $\gamma_d < 1$. This is also known as gradient descent with *weight decay* (with parameter $\lambda = \frac{1-\gamma_d}{\gamma_l}$), which is

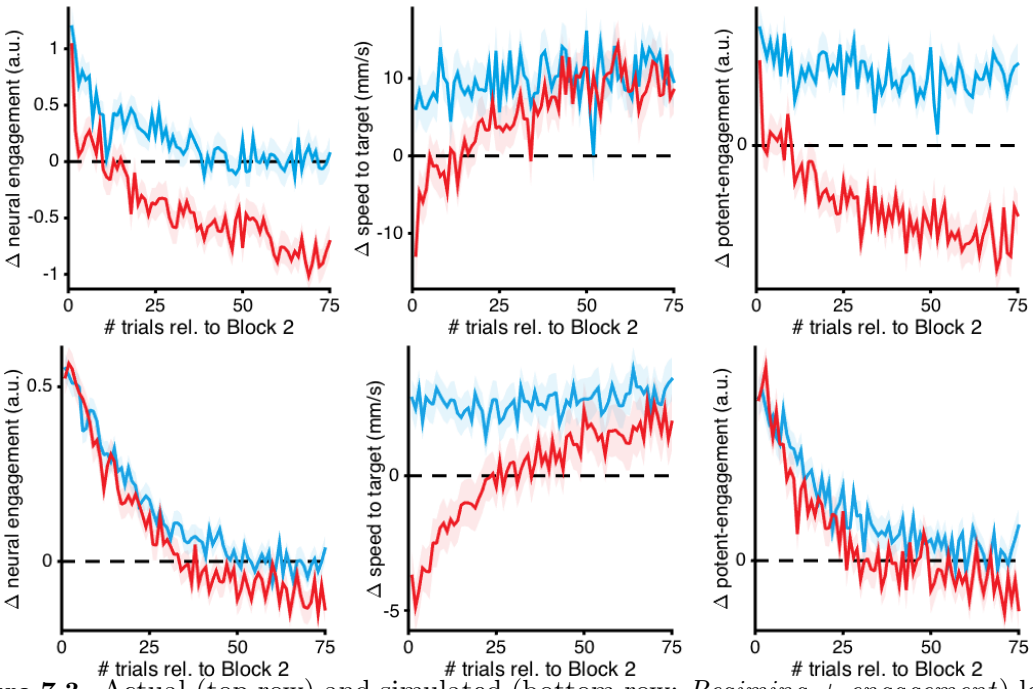


Figure 7.3. Actual (top row) and simulated (bottom row; *Reaiming + engagement*) levels of neural engagement and progress for T+ (blue) and T- (red) targets from all monkey J sessions. The *Reaiming + engagement* model was able to capture differences in performance (middle column) between T+ and T- targets during learning, but not differences in neural engagement (left column) or potent engagement (right column) between T+ and T- targets.

equivalent to saying that our objective is to minimize f plus a penalty on the squared ℓ_2 norm of δ_l.

Simulations for this Weight perturbation model used $f(\mathbf{v}) = \mathbf{v}^\top \theta$ (i.e., cursor progress, aka "speed to target" as shown below), $\gamma_l = 0.01$, $\gamma_d = 0.95$, $\gamma_{ne} = 0.95$, $\|\boldsymbol{\delta}_{ne}(1)\| = 2$. $\mathrm{Var}[\boldsymbol{\epsilon}(t)]$ was 0.01 times the variance of factor activity during intuitive control. In this case, we were able to fully capture all of the qualitative features of the data during learning (Fig. 7.4).

Results for this model were largely the same when using the squared angular error objective instead of progress. Results were also similar when changing the noise covariance of $\boldsymbol{\epsilon}$ (i.e., letting it be anisotropic, and target-specific). This suggests that we will need a more detailed analysis to distinguish between different forms of the *Weight perturbation* model.

Note on terminology. I had been calling this model "Node perturbation". If you have inputs x and a network output $h(x)$, node perturbation compares $f(h(x)+\epsilon)$ to $f(h(x))$, while weight perturbation compares $f(h(x+\epsilon))$ to $f(h(x))$. This distinction becomes important if we think about learning as happening in terms of some internal parameters, with the factor activity being the (potentially nonlinear) readout of these parameters (i.e., via h). I think really though, this model is more safely called simply an "evolutionary strategy."

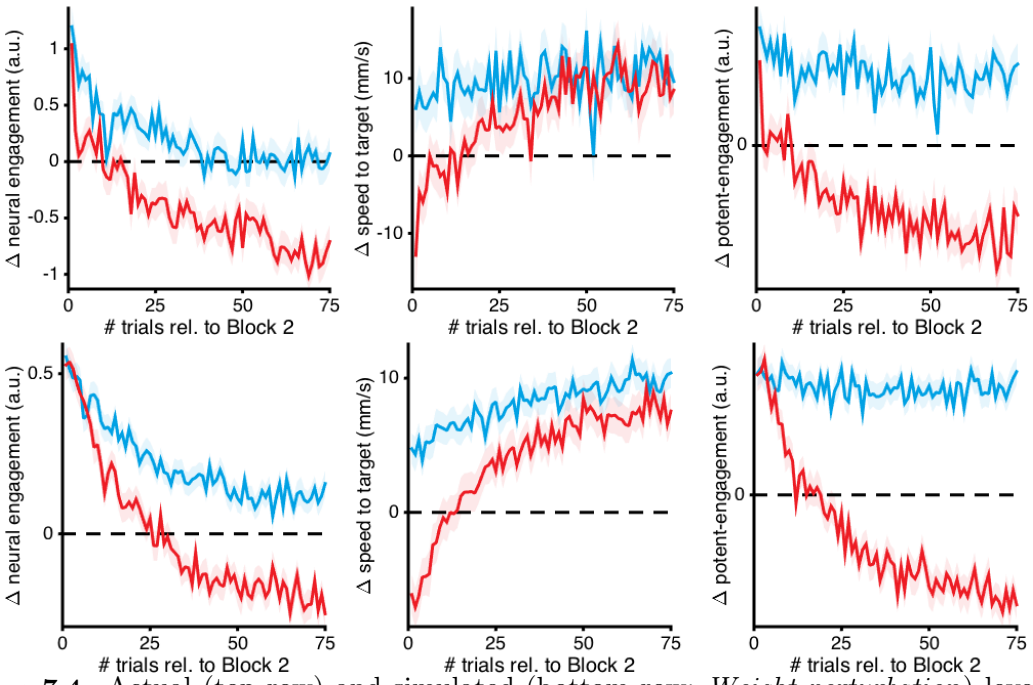


Figure 7.4. Actual (top row) and simulated (bottom row; *Weight perturbation*) levels of neural engagement and progress for T+ (blue) and T- (red) targets from all monkey J sessions. The *Weight perturbation* model was able to capture differences in performance (middle column) between T+ and T- targets during learning, as well as differences in neural engagement (left column) and potent engagement (right column).

Does noise influence the asymptote of learning?

Variability exists throughout the nervous system, both at the level of neural populations as well as behavior. During learning, we modify our neural activity based on feedback about our past performance. Thus, variability in neural activity can drive variability in our performance, resulting in variability in how we *modify* our future neural activity based on that feedback.

But does this variability really *matter*? In other words, can this variability not simply be averaged out over time? As I will show here, for any aspect of variability in neural population activity that is not accessible to the learning process, this variability can influence the asymptote of learning. In particular, we will show in two examples (one model-based, one model-free) how the asymptote of learning depends on the shape of the noise covariance of population activity.

Model-free gradients are impacted by noise. First, consider the *Weight perturbation* model discussed above. In this model (Eqn. 7.4), $\mathbf{z}(t)$ will eventually converge in expectation as long as $\gamma_d < 1$, and $\boldsymbol{\delta}_{ne}(t)$ converges for large enough t. The value of $\mathbf{z}(t)$ at convergence depends on the shape of the noise distribution. To see this, let $R(t) = r(t) - \overline{r}(t)$, $\Sigma = \mathrm{Cov}[\boldsymbol{\epsilon}(t)]$, and suppose $\mathbb{E}[\boldsymbol{\epsilon}(t)] = 0$. Then at

convergence (i.e., for large enough t) we will have:

$$\begin{aligned}
\mathbb{E}[\boldsymbol{\delta}_l(t+1)] &= \mathbb{E}[\gamma_d \boldsymbol{\delta}_l(t) + \gamma_l R(t)\boldsymbol{\epsilon}(t)] \\
\mathbb{E}[\boldsymbol{\delta}_l(t+1) - \gamma_d \boldsymbol{\delta}_l(t)] &= \mathbb{E}[\gamma_l R(t)\boldsymbol{\epsilon}(t)] \\
(1-\gamma_d)\mathbb{E}[\boldsymbol{\delta}_l(t)] &= \gamma_l \mathbb{E}[R(t)\boldsymbol{\epsilon}(t)] \\
\mathbb{E}[\boldsymbol{\delta}_l(t)] &= \frac{\gamma_l}{1-\gamma_d}\mathrm{Cov}[R(t), \boldsymbol{\epsilon}(t)] \\
&= \frac{\gamma_l}{1-\gamma_d}\mathrm{Cov}[r(t), \boldsymbol{\epsilon}(t)]
\end{aligned} \tag{7.5}$$

where in the last line we are assuming $\overline{r}(t)$ and $\boldsymbol{\epsilon}(t)$ are uncorrelated (because $\overline{r}(t)$ depends only on previous values of $\boldsymbol{\epsilon}(t)$, and $\boldsymbol{\epsilon}(t)$ is i.i.d.). Now, remember that $r(t) = f(g(\mathbf{z}(t)))$. We'll also use the fact that, in our BCI mapping, $g(\mathbf{z}) = B\mathbf{z}$ for some matrix $B \in \mathbb{R}^{2\times d}$. Let's write $\mathbf{z}(t) = \mathbb{E}[\mathbf{z}(t)] + \boldsymbol{\epsilon}(t)$, so that $r(t) = f(B\mathbb{E}[\mathbf{z}(t)] + B\boldsymbol{\epsilon}(t))$.

First let's suppose that f is affine, so we can write $f(\mathbf{v}) = \mathbf{c}^\top \mathbf{v} + d$, for some $\mathbf{c} \in \mathbb{R}^2$ and $d \in \mathbb{R}$. This gives us $r(t) = \mathbf{a}^\top \mathbb{E}[\mathbf{z}(t)] + \mathbf{a}^\top \boldsymbol{\epsilon}(t) + d$, where $\mathbf{a} = B^\top \mathbf{c}$. Thus, $\mathrm{Cov}[r(t), \boldsymbol{\epsilon}(t)] = \mathrm{Cov}[\mathbf{a}^\top \boldsymbol{\epsilon}(t), \boldsymbol{\epsilon}(t)] = \mathbf{a}^\top \Sigma$. Putting it all together, when f is affine, at convergence we have:

$$\begin{aligned}
\mathbb{E}[\mathbf{z}(t)] &= h(\theta) + \mathbb{E}[\boldsymbol{\delta}_l(t)] \\
&= h(\theta) + \frac{\gamma_l}{1-\gamma_d}\mathbf{a}^\top \Sigma
\end{aligned} \tag{7.6}$$

where $\mathbf{a} = B^\top \mathbf{c}$ is determined by the BCI mapping and the (affine) objective function, and Σ is the covariance of the noise distribution (i.e., the covariance of $\boldsymbol{\epsilon}(t)$). Geometrically (and empirically), it seems like $\boldsymbol{\delta}_l(t)$ converges to the point on a particular covariance ellipse (i.e., the set of points with equal probability under the p.d.f. of $\boldsymbol{\epsilon}(t)$) that maximizes the objective. I suspect this might also be true if f is not affine but has only one maxima (though this is just a hunch).

Now, what happens if learning is happening in terms of some internal parameters θ, where $\mathbf{z} = h(\theta)$, and h is nonlinear? Here we're imagining we still have a *Weight perturbation* learning rule, with high-dimensional learnable parameters (e.g., inputs to M1 coming from a different brain area). But suppose we have a nonlinear relationship between our learned parameters and the actual neural activity. For example:

$$\mathbf{z}(t) = h(\boldsymbol{\theta}_0 + \boldsymbol{\delta}_l(t) + \boldsymbol{\epsilon}(t)) \tag{7.7}$$

where $\boldsymbol{\theta}_0$ is the initial parameter value (e.g., from intuitive control), and h is some unknown nonlinear function. In this case, all of our convergence equations above are still correct, in that we have $\boldsymbol{\delta}_l(t) \to \frac{\gamma_l}{1-\gamma_d}\mathrm{Cov}[r(t), \boldsymbol{\epsilon}(t)]$. But it's now harder to go any further with this equation since the relationship between $\boldsymbol{\epsilon}(t)$ and $\mathbf{z}(t)$ (and therefore $r(t)$) involves going through h.

Model-based gradients are also impacted by noise. Now, suppose that rather than directly modifying their neural activity during learning, monkeys learn

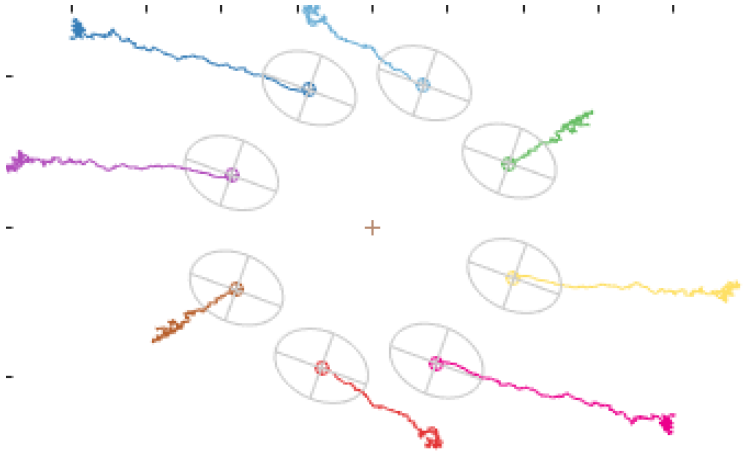


Figure 7.5. Example simulation of how neural activity ($\mathbf{z}_{1:T}$, colored lines) evolves during learning for eight different targets under the *Weight perturbation* model (Eqn. 7.4). The noise covariance, Σ, is shown as gray ellipses, and is the same for each target. The objective for each trace is to maximize its distance from the origin in a target-specific direction (e.g., the red trace wants to minimize its y-value), subject to a norm constraint. Because Σ is non-isotropic, the asymptotes across targets do not lie along a circle, but rather differ based on the alignment between the target direction and the orientation of Σ, as in Eqn. 7.6.

an internal forward model of the BCI mapping, which they then use to select the neural activity they expect will improve their performance. Let $\mathbf{z} \in \mathbb{R}^d$ be our neural activity. Suppose we receive feedback $v = \mathbf{b}^\top \mathbf{z} \in \mathbb{R}$, for some unknown $\mathbf{b} \in \mathbb{R}^d$. (For simplicity, we assume here that v is a scalar.) Our goal is to update our internal model estimate (IM), $\widehat{\mathbf{b}} \in \mathbb{R}$, to minimize our expected prediction error.

Suppose that at each time t, we sample $\mathbf{z}_t = \boldsymbol{\mu} + \boldsymbol{\epsilon}_t$, with $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \Sigma)$. Our model's prediction at each time t is $\widehat{v}_t = \widehat{\mathbf{b}}_t^\top \mathbf{z}_t$. Then our loss we want to minimize is $e_t = \mathbb{E}[\frac{1}{2}\|\widehat{v}_t - v_t\|_2^2]$, where the expectation is w.r.t. the distribution of $\mathbf{z}_t$. We learn by gradient descent:

$$\widehat{\mathbf{b}}_{t+1} = \widehat{\mathbf{b}}_t - \gamma \mathbf{z}_t (\widehat{v}_t - v_t)^\top \tag{7.8}$$

What I want to know is, what is our average gradient step? In other words, what is $\mathbb{E}[\mathbf{z}_t(\widehat{v}_t - v_t)^\top]$?

$$\begin{aligned} (\widehat{v}_t - v_t)^\top &= (\boldsymbol{\mu} + \boldsymbol{\epsilon}_t)^\top (\widehat{\mathbf{b}}_t - \mathbf{b}) \\ \rightarrow \mathbf{z}_t(\widehat{v}_t - v_t)^\top &= (\boldsymbol{\mu} + \boldsymbol{\epsilon}_t)(\boldsymbol{\mu} + \boldsymbol{\epsilon}_t)^\top (\widehat{\mathbf{b}}_t - \mathbf{b}) \\ \rightarrow \mathbb{E}[\mathbf{z}_t(\widehat{v}_t - v_t)^\top] &= \boldsymbol{\mu}\boldsymbol{\mu}^\top (\widehat{\mathbf{b}}_t - \mathbf{b}) + \Sigma(\widehat{\mathbf{b}}_t - \mathbf{b}) \end{aligned} \tag{7.9}$$

So what does this tell us? Well, suppose $\boldsymbol{\mu} = 0$, and $\Sigma = I$, so we have zero-mean, isotropic exploration noise. In that case, our expected gradient step is $\widehat{\mathbf{b}}_t - \mathbf{b}$, which is exactly what we want: With learning, our internal model of the BCI mapping will move, on average, directly towards the true BCI mapping. But what if Σ is something else? Then we will be updating our IM in the direction $\widehat{\mathbf{b}}_t - \mathbf{b}$ projected

along our search covariance Σ. In other words, our search covariance influences our expected gradient steps. (Note that if we have multiple target directions, such that the $\boldsymbol{\mu}$ updates will essentially cancel out across targets, this effectively leaves us in the same situation as if $\boldsymbol{\mu} = 0$.)

As our IM is updating, we will also want to modify our average neural activity, $\boldsymbol{\mu}$, according to some different objective. We'll assume for simplicity that our IM has already converged to $\mathbf{b}$. Now suppose that our inverse model's objective is to maximize the expected feedback, i.e., $\mathbb{E}[\widehat{\mathbf{b}}^\top \mathbf{z}] = \widehat{\mathbf{b}}^\top \boldsymbol{\mu}$. If we update $\boldsymbol{\mu}$ via gradient descent, then our updates will be in the direction $\widehat{\mathbf{b}}$. Because our inverse model's objective is unbounded, we will allow convergence by adding shrinkage, with parameter $\alpha > 0$, to our $\boldsymbol{\mu}_t$ update. So in total we have:

$$\boldsymbol{\mu}_{t+1} = \alpha \boldsymbol{\mu}_t + \rho \widehat{\mathbf{b}}_t \tag{7.10}$$

At convergence we have $\mathbb{E}[\boldsymbol{\mu}_{t+1}] = \mathbb{E}[\boldsymbol{\mu}_t] = \boldsymbol{\mu}$, and $\mathbb{E}[\mathbf{b}_{t+1}] = \mathbb{E}[\mathbf{b}_t] = \widehat{\mathbf{b}}$, for all $t > t_0$. This tells us the following:

$$\begin{aligned} \mathbb{E}[\boldsymbol{\mu}_{t+1}] &= \mathbb{E}[\alpha \boldsymbol{\mu}_t + \rho \widehat{\mathbf{b}}_t] \\ \rightarrow \boldsymbol{\mu} &= \alpha \boldsymbol{\mu} + \rho \widehat{\mathbf{b}} \\ \rightarrow \boldsymbol{\mu} &= \frac{\rho}{1-\alpha} \widehat{\mathbf{b}} \end{aligned} \tag{7.11}$$

Thus, at convergence, we can think of $\boldsymbol{\mu}$ as a stretched out estimate of $\widehat{\mathbf{b}}$ (i.e., we have $\boldsymbol{\mu}/\|\boldsymbol{\mu}\| = \widehat{\mathbf{b}}/\|\widehat{\mathbf{b}}\|$). Note that if we have $\boldsymbol{\mu} = \boldsymbol{\mu}_0 + \boldsymbol{\delta}$, where $\boldsymbol{\mu}_0$ is constant, $\boldsymbol{\delta}$ is the learned component, and shrinkage is applied to $\boldsymbol{\delta}$ instead of $\boldsymbol{\mu}$, we can instead think of $\boldsymbol{\delta}$ (i.e., $\boldsymbol{\mu} - \boldsymbol{\mu}_0$) as our stretched estimate of $\widehat{\mathbf{b}}$.

Finally, consider the case where we have $B \in \mathbb{R}^{2 \times d}$, so that our feedback $\mathbf{v} = B\mathbf{z}$ is 2D. For a given target direction $\mathbf{x} \in \mathbb{R}^2$ with $\|x\|_2^2 = 1$, suppose our inverse model's objective function is $\mathbf{x}B\mathbf{z}$. Then by the same logic above, at convergence we can think of $\boldsymbol{\delta}_x$ (i.e., our change in neural activity for direction $\mathbf{x}$ after learning) as our stretched estimate of $x\widehat{B}$. **Thus, whether learning be model-based or model-free the asymptote of neural population activity at the end of learning can depend on the noise covariance.**

Animals do not appear to learn a global internal model (IM)

In the WMP learning experiments, the BCI mapping specifies the relationship between neural activity and cursor movements. Previous work has suggested that animals may learn an internal estimate (IM) of this BCI mapping to account for sensory feedback delays (Golub et al., 2015). One possibility is that they also use this IM for learning. For example, to move straight toward the target, monkeys may generate whichever neural activity would do so according to their current internal estimate of the BCI mapping.

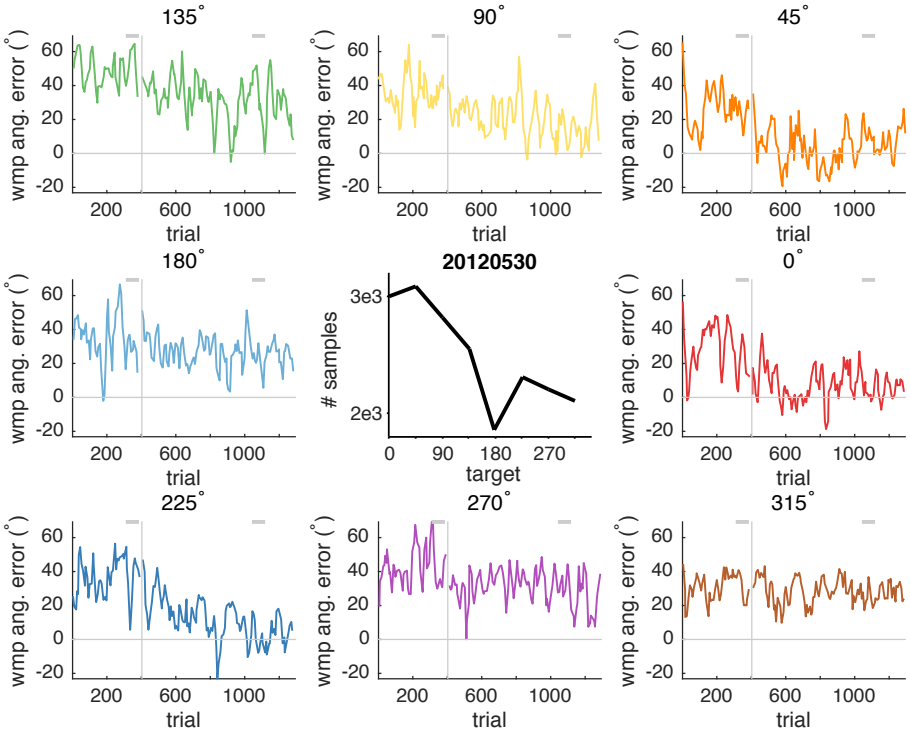


Figure 7.6. Changes in the cursor's (signed) angular error for all eight targets (colored lines) during an example session. Vertical gray lines indicate when the new BCI mapping was first introduced. Center panel indicates the total number of time steps ("samples") that the monkey received sensory feedback for each target throughout learning. Targets with more samples also exhibited larger improvements in the cursor's angular error during learning.

The true BCI mapping is "global" in the sense that it is the same for all targets. If the monkey's internal estimate is also global, this means they can use the error feedback from all targets combined to update their internal model estimate. One of the predictions of this hypothesis is therefore that the monkey's angular errors should improve with the same learning rate across targets, as any errors in their performance for a given target can be used to improve performance for all other targets, via the resulting improvement in the internal model estimate.

Instead, we frequently observe angular errors decreasing at different rates across targets. In the example session shown in Fig. 7.6, for example, angular errors rapidly decrease to zero for the 0° and 45° targets, whereas for the 270° and 315° targets, angular error barely improves at all. In fact, the targets that showed more improvement in angular errors are the targets where the monkey received more samples of error feedback (Fig. 7.6, center panel), because trials to these targets were longer. **This pattern of results is exactly what you would expect if monkeys are *not* learning a global internal model**, because in that case the errors for one target do not generalize to improvements for other targets.

Why would the monkey not learn a global internal model (IM)? This may be explainable by changes in neural engagement. A global IM assumes there to be a consistent relationship between neural activity and cursor movements. But if i)

neural engagement has target-specific effects on velocity, and ii) the animal does not have access to its level of neural engagement, then from the monkey's perspective there *is not* a global internal model.

Generalization across targets.

From the above results, we might conclude that learning occurs somewhat independently for different targets. However, from other studies of motor learning (Haith and Krakauer, 2013), we might expect that learning might generalize to *adjacent* targets (as opposed to all other targets).

If this is the case, and learning updates are shared across adjacent targets, what happens when the error feedback for these adjacent targets is inconsistent? For example, under a particular WMP, the 0° target might have an initial angular error of 45°, whereas the adjacent, 45° target might have an *opposite* angular error of −45°. If learning updates are shared between these "competing" targets, this might lead to interference, since the 0° target is "voting" to reaim clockwise, whereas the 45° target is voting to reaim counter-clockwise. By contrast, targets adjacent to targets with similar errors would likely be learned more quickly than the competing targets. I will refer to these targets as "supported."

To see if this was the case, I classified each target during WMP learning based on whether the target's adjacent targets had opposite initial WMP angular error signs ("competing," red) or with the same angular error sign ("supported," blue). We expected to see faster learning for supported targets than for competing targets, because for competing targets, experience from flanking targets interferes.

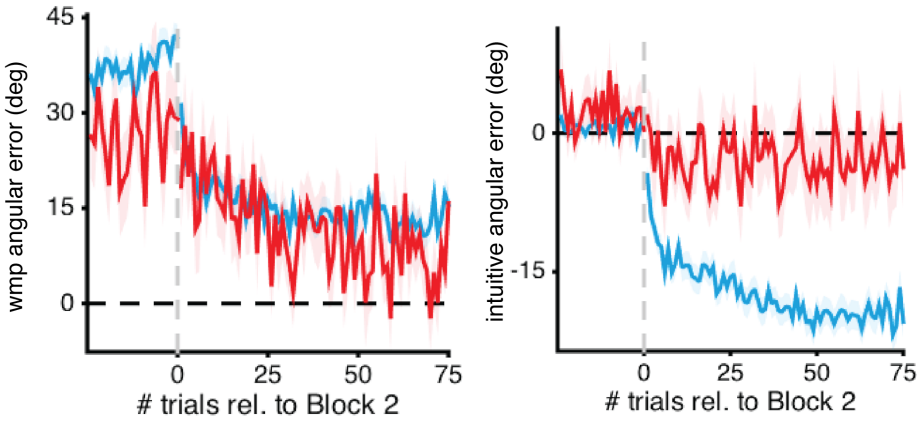


Figure 7.7. Changes in angular error during learning, split by targets whose neighbors had identical ("supported," blue) versus opposite ("competing," red) signs of initial WMP angular error.

In fact, what we found was that the WMP angular errors for both types of targets decreased at similar rates (Fig. 7.7, left). However, we only saw concomitant changes in intuitive angular error for the supported targets (Fig. 7.7, right). Thus, **competing and supported targets were learned differently, indicating that**

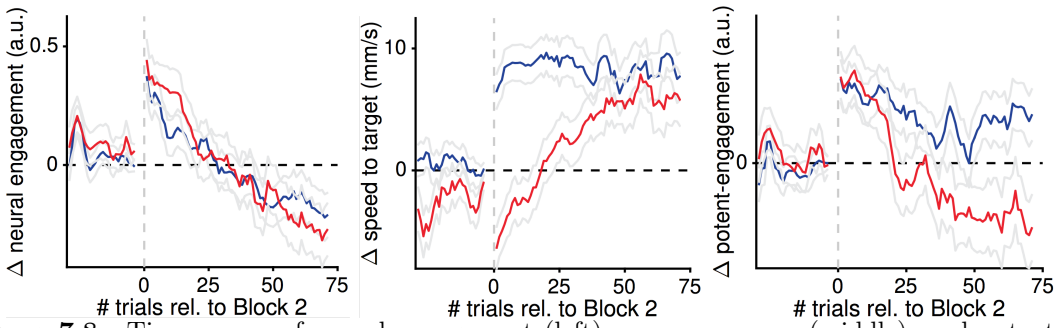


Figure 7.8. Time course of neural engagement (left), cursor progress (middle), and potent engagement (right), averaged across T+ (blue) and T- (red) targets, during the freeze period on each trial. Compare to Fig. 4.5c, Fig. 4.5e, and Fig. 4.6c, respectively, for the same quantities during the cursor control period.

how a target is learned is influenced by the error feedback from adjacent targets. In particular, these results suggest that we may be seeing signs of "reaiming" for supported targets (because we see angular errors under the intuitive and WMP mapping changing together), but something else entirely for competing targets.

Neural engagement has a distinct timecourse during movement preparation

In Chapter 4, we saw how changes in neural engagement during learning appeared to reflect both learning-driven and nonlearning-driven sources. For example, Fig. 4.5c shows how neural engagement decayed during learning for both T+ and T- targets, but to target-specific levels.

One interesting aspect of the data is that if you look at changes in neural engagement after the monkey has seen the target, but before he has begun controlling the cursor (i.e., during the "freeze period"), neural engagement appears to decay throughout learning (Fig. 7.8, left) to similar levels for both targets. At the same time, changes in both performance (Fig. 7.8, middle) and potent engagement during the freeze period are still target specific (Fig. 7.8, right). **This suggests that the target-specific changes we see in total neural engagement during the control period (Fig. 4.5c) may be due to *feedback*, whereas the target-specific changes in potent engagement we see are established during *preparation*.**

To assess whether this is the case, one may need to account for the fact that the freeze period is not a true "preparatory" period. Notably, different monkeys appear to have different strategies for what they do during the freeze period.

7.2 Interpreting geometries in neural population activity

Visualizing neural population activity often reveals interesting geometrical structure. For example, during BCI control, the average neural population activity produced within trials to different targets sometimes looks like a head massager (Fig. 7.9, top left), while for other sessions or animals it looks more like a gramophone (Fig. 7.9, top right). Can we understand the differences in these shapes in terms of differences in single neuron tuning (Fig. 7.9, bottom)?

I assume that for trials to target θ, each neuron's time-varying firing rate, $y(t, \theta)$, can be modeled as the sum of a target-specific (cosine tuned) component and a generic, target-irrelevant component. Specifically:

$$y(t,\theta) = max[0, y_{spont} + r_{evoked}(t,\theta)]$$
$$y_{evoked}(t,\theta) = g_{\tau_1}(t) r_{generic} + g_{\tau_2}(t) r_{stim} cos(\theta - \theta_{pd})$$

where $g_{\tau_1}(t)$ and $g_{\tau_2}(t)$ are gain terms that ramp up exponentially from 0 to 1 with timescales given by τ_1 and τ_2; $r_{generic}$ and r_{stim} are scalars controlling the strength of the target-irrelevant and tuned components, respectively; and θ_{pd} is the neuron's preferred direction.

I found that simply varying the gain and timescale of the target-specific and target-irrelevant components independently can result in a variety of different shapes in population space, including shapes that look like head massagers, gramophones, spiders, or snowflakes. This suggests that the different shapes one might observe when visualizing population activity can be understood in terms of the above model of single neuron firing. See `https://mobeets.github.io/psth-to-population/` for an interactive version of this model.

7.3 Decoding the target identity on previous trial using neural activity

We tend to think of primary motor cortex (M1) as reflecting only intended or upcoming movements. Surprisingly, in these BCI experiments, I have found strong evidence that M1 reflects information about *previous* movements. In particular, given the neural population activity on trial t, an LDA classifier can decode the target identity on trial $t-1$ (Fig. 7.10), even using activity roughly half a second after the previous trial has completed. This finding was incredibly robust, existing across nearly all sessions for all three monkeys.

(Note: While decoding accuracy was similar during the intuitive block and during the WMP block, the decoding planes were different—i.e., a decoder trained on the intuitive block would perform poorly on the WMP block.)

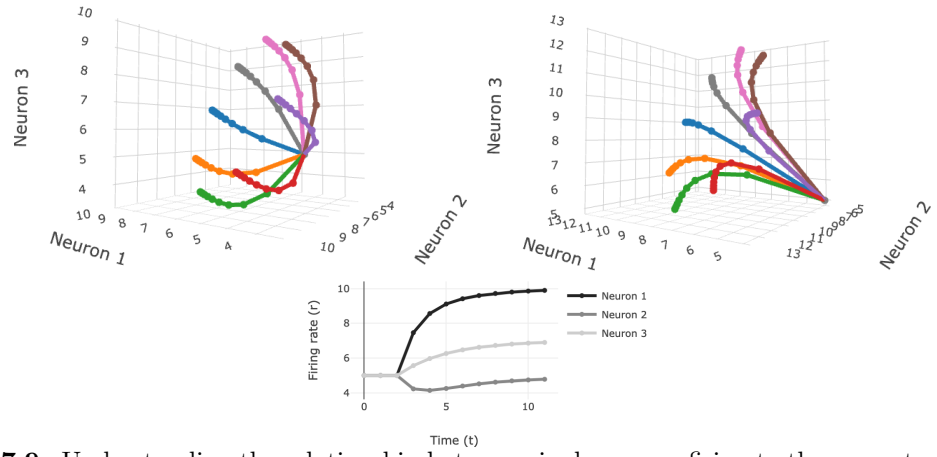


Figure 7.9. Understanding the relationship between single neuron firing to the geometry of neural population activity. When neural population activity looks more like a head massager (top left) than a gramophone (top right), can we understand these differences in terms of the PSTH of single neurons (bottom)?

The ability to decode the target on the previous trial even during cursor control on the *following* trial indicates that the previous target's identity can influence cursor velocities on the current trial (Fig. 7.11, top), depending on the alignment of the decoder. In fact, the influence of the previous trial's target on intuitive mapping velocities on the next trial seems to be consistent with slower velocities towards that previous target (Fig. 7.11, bottom). For example, if the previous target was 0°, on the next trial neural activity was consistent with slower velocities in the 0° direction. It's almost as if the prior is to be moving in the exact opposite direction. In fact, in these WMP experiments, the target shown on trial t was almost *never* the same as the target shown on trial $t-1$.). Thus, **I speculate that the ability to decode the target identity on the previous trial may reflect the fact that monkeys have a prior expectation that the following trial's target is more likely to be in the opposite direction.**

7.4 Explaining differences in single-day WMP and OMP learning

The difference between WMP and OMP learning is not fully understood. There are a few distinct explanations for why the intrinsic manifold was unchanged during single-day BCI learning, as well as why WMPs were learned more often than OMPs in a single day (Sadtler et al., 2014). I summarize these possible explanations below.

Feedforward synaptic changes. Wärnberg and Kumar (2019) found that synapse changes that respected the intrinsic manifold required smaller changes than those that did not respect the intrinsic manifold. This suggests that synaptic learning rules that respect the intrinsic manifold may requiring limiting the total magnitude of synaptic changes, e.g., with an ℓ_2 norm penalty.

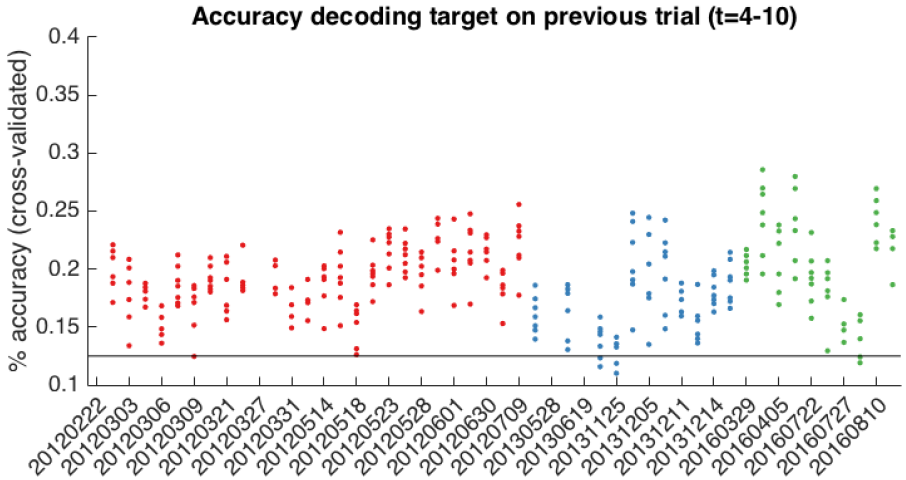


Figure 7.10. Percent of trials for each target (dots) where the target identity on trial $t-1$ was accurately decoded from the average neural activity on trial t (during time steps 4-10, approximately 400-850 ms after trial $t-1$ ended), using an LDA classifier trained on held-out trials for each session.

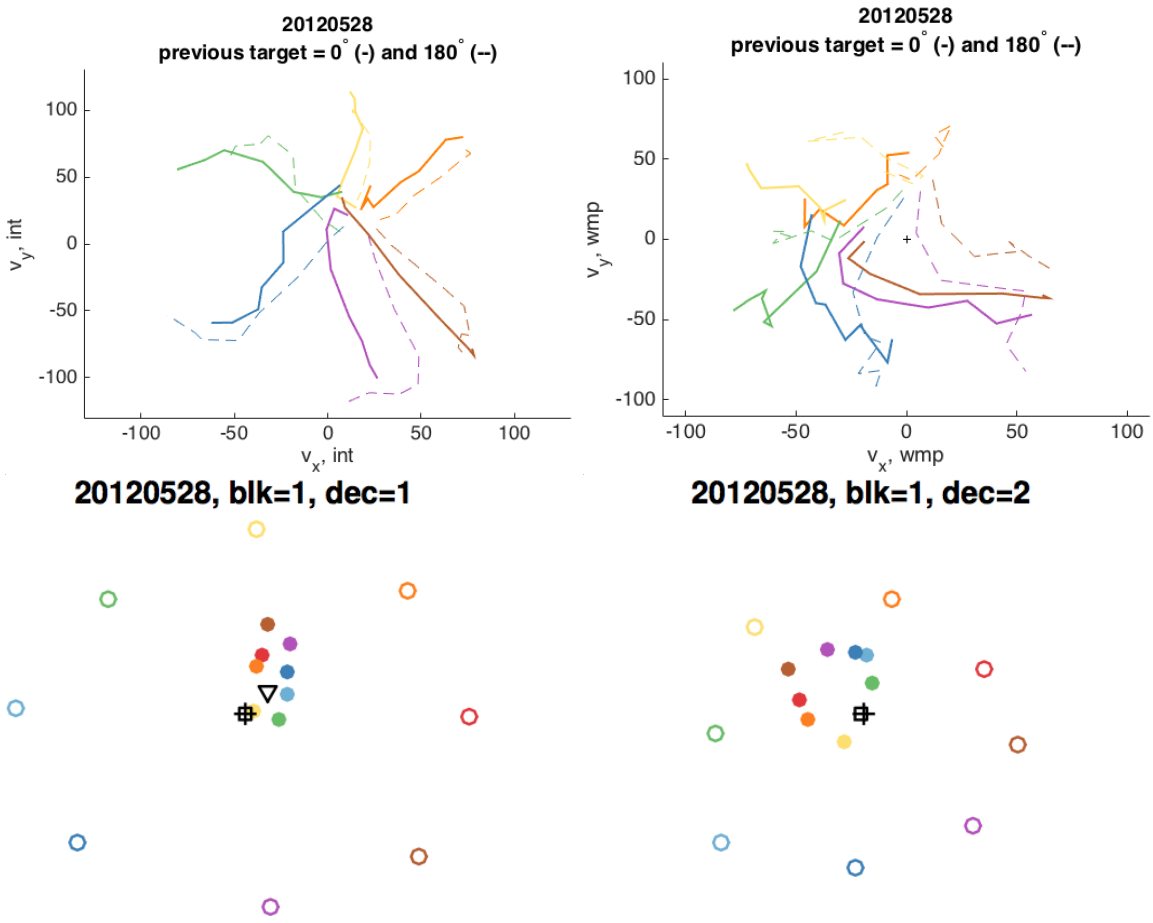


Figure 7.11. Impact of previous trial's target identity on cursor velocities during the following trial (L: intuitive mapping velocities; R: WMP mapping velocities) during an example session. Top: Trial-averaged cursor velocities to six targets conditioned on whether the previous trial's target was 0° (solid) versus 180° (dashed). Bottom: Average cursor velocities during the first two time steps of all trials, conditioned on the previous trial's target identity (colors).

Recurrent synaptic changes. Feulner and Clopath (2021) found they could explain differences in WMP and OMP learning given i) the presence of erroneous feedback signals, ii) sparse feedback signals (e.g., only some neurons get the feedback signal), or iii) if the number of plastic connections within the network is constrained. Also, the correct feedback weights could be learned for WMPs but not OMPs, suggesting another possible explanation. (Also note that they could not reproduce the findings of Wärnberg and Kumar (2019) as described above, though there were differences in their learning rules.)

Input changes. Perich et al. (2018) found evidence that the functional connectivity between PMd and M1 was unchanged during both force field perturbations and VMRs, and results suggested that adaptation in force field perturbations occurred by PMd changing its M1-null activity in order to provide new inputs to M1 (e.g., "setting the starting point" of the dynamical system). This suggests that synaptic plasticity may not be required/possible for either WMP or OMP learning. A 2020 Cosyne poster from Jorge Menendez and Peter Latham took the same perspective, simulating BCI learning where only the inputs to M1 could change with learning. They found that constraints on the dimensionality of the input signal to M1 (i.e., 2D reaiming signal) would preserve the intrinsic manifold, while also explaining why WMPs were better learned than OMPs.

Network structure. Williamson et al. (2016) showed that clustered networks (i.e., where excitatory neurons are more likely to synapse onto other excitatory neurons in their cluster than outside their cluster) yields an intrinsic manifold of population activity. From this it seems possible that the network structure itself (and not the learning rule) could constrain how M1 activity can change with learning. Note that this idea hasn't been directly applied to BCI learning in any way, so the connection is a little vague.

7.5 What is the brain's objective function?

Introduction

Learning a new behavior or skill requires coordinated changes in the neural population activity we use to perform the task. In particular, we must use feedback about our performance to modify the neural population activity we use to perform the task in the future. How does the brain do this? More specifically, what aspects of feedback does the brain utilize to drive changes to neural population activity?

In this section, I propose a simple, probabilistic method for inferring the objective function that is driving changes to neural population activity during learning. This method takes as input a time series of neural population activity, $\boldsymbol{z}_{1:T} \in \mathbb{R}^d$, alongside the corresponding feedback $\boldsymbol{v}_{1:T} \in \mathbb{R}^k$, recorded during learning. The method then attempts to use the changes in neural population activity that occur during learning to infer the objective function that is being optimized.

Importantly, this method assumes that sustained changes to neural population

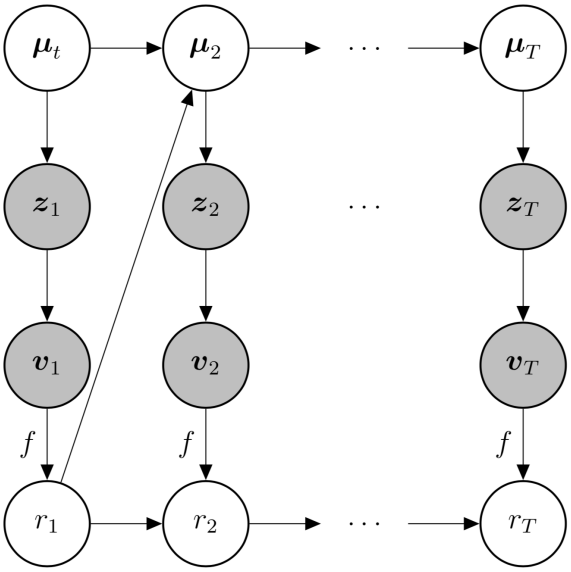


Figure 7.12. Assumed graphical model for inferring how reward (r_t), which is some unknown readout of task feedback ($\boldsymbol{v}_t$), influences changes in neural population activity ($\boldsymbol{\mu}_{t+1}$) during learning.

activity occur only in order to improve the task objective, f, which is some function of the feedback. While recent results indicate that feedback is not the only driver of changes in neural activity during learning (Hennig et al., 2020), it may be a useful starting point for future methods trying to infer the objective function driving learning in neural population activity.

Preliminaries

For each time t (e.g., a trial), let our observed neural activity be $\boldsymbol{z}_t \in \mathbb{R}^d$. Let the task feedback (e.g., cursor velocities) be $\boldsymbol{v}_t = B\boldsymbol{z}_t \in \mathbb{R}^k$, where $B \in \mathbb{R}^{k \times d}$ is known. Our assumption is that the recorded neural activity $\boldsymbol{z}_t$ is a noisy sample of some underlying latent state, $\boldsymbol{\mu}_t$, which evolves to optimize $r_t = f(\boldsymbol{v}_t) \in \mathbb{R}$, some scalar function of $\boldsymbol{v}_t$ (see Fig. 7.12).

To start, we must specify a generative model for our observations $\boldsymbol{z}_{1,..,T}$. Here, for concreteness, we will consider the following generative model:

$$\begin{aligned} \boldsymbol{z}_t &= \boldsymbol{\mu}_t + \boldsymbol{\epsilon}_t \\ \boldsymbol{\epsilon}_t &\sim \mathcal{N}(0, \Sigma) \\ \boldsymbol{\mu}_{t+1} &= \boldsymbol{\mu}_t - \lambda(r_t - \overline{r}_t)\boldsymbol{\epsilon}_t \end{aligned}$$

where $\lambda > 0$ is the learning rate, $r_t = f(\boldsymbol{v_t}) = f(B\boldsymbol{z}_t)$ is the "reward" (the objective value) at time t, and $\overline{r}_t = \alpha r_{t-1} + (1-\alpha)\overline{r}_{t-1}$ is the expected reward. In this model, the mean neural activity, $\boldsymbol{\mu}_t = \mathbb{E}[\boldsymbol{z}_t]$, is being updated in a model-free manner to optimize some function of the feedback. Specifically, $\boldsymbol{\mu}_t$ is changing in order to minimize $\mathbb{E}[f(\boldsymbol{v}_t)] \in \mathbb{R}$, where f is our objective function. The question is: Can we infer f?

Approach

To estimate f, which may be nonlinear, we will approximate it as a linear combination of basis functions:

$$f(\boldsymbol{v}) = \boldsymbol{a}^\top \boldsymbol{\Phi}(\boldsymbol{v})$$

where $\boldsymbol{\Phi}(\boldsymbol{v}) \in \mathbb{R}^K$ are our basis function responses to $\boldsymbol{v}$, e.g., radial basis functions $\Phi_i(\boldsymbol{v}) = \exp(-\|\boldsymbol{v} - \boldsymbol{c}_i\|^2/\sigma_i^2)$ for $i = 1, ..., K$.

We will treat α (used to define $\overline{r}_t$) and the parameters in $\boldsymbol{\Phi}$ as hyperparameters.

There are a few unknowns: $\boldsymbol{\mu}_{1,...,T}$, Σ, λ, and $\boldsymbol{a}$. (Note that $\boldsymbol{\epsilon}_t = \boldsymbol{z}_t - \boldsymbol{\mu}_t$.) We can't distinguish $\boldsymbol{a}$ and λ so we can only infer $\lambda\boldsymbol{a}$. We will use coordinate descent, where we (1) estimate Σ given the other parameters; (2) estimate $\boldsymbol{\mu}_{1,...,T}$ given the other parameters; (3) estimate $\lambda\boldsymbol{a}$. We repeat this process until convergence.

1. Estimating Σ

We know that $\boldsymbol{z}_t = \boldsymbol{\mu}_t + \boldsymbol{\epsilon}_t$, where $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \Sigma)$. In other words, $\Sigma = \mathbb{E}[(\boldsymbol{z} - \boldsymbol{\mu})(\boldsymbol{z} - \boldsymbol{\mu})^\top]$. Thus, we have $\widehat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{z}_t - \widehat{\boldsymbol{\mu}}_t)(\boldsymbol{z}_t - \widehat{\boldsymbol{\mu}}_t)^\top$. Initially, Σ could be estimated by assuming that $\boldsymbol{\mu}_t$ is the sample mean of $\boldsymbol{z}_{1,...,t}$.

2. Estimating $\boldsymbol{\mu}_{1,...,T}$

Estimating $\boldsymbol{\mu}_{1,...,T}$ is simply Kalman smoothing of our observations $\boldsymbol{z}_{1,...,T}$ under the following state-space model:

$$\begin{aligned}
\boldsymbol{\mu}_{t+1} &= \boldsymbol{\mu}_t - \lambda(r_t - \overline{r}_t)\boldsymbol{\epsilon}_t \\
&= A_t\boldsymbol{\mu}_t + B_t\boldsymbol{z}_t \\
\boldsymbol{z}_t &\sim \mathcal{N}(\boldsymbol{\mu}_t, \Sigma)
\end{aligned}$$

where $A_t = (1 + \lambda(r_t - \overline{r}_t))I$ is a time-varying dynamics matrix (known, given $\boldsymbol{a}$), $\boldsymbol{z}_t$ are our inputs to the system (known), and $B_t = -\lambda(r_t - \overline{r}_t)I$ is how the inputs impact the state.

3. Estimating $\lambda\boldsymbol{a}$

We first define $\boldsymbol{f}_t = \boldsymbol{\Phi}(B\boldsymbol{z}_t) \in \mathbb{R}^K$, which is observed. This lets us write $r_t = \boldsymbol{a}^\top\boldsymbol{\Phi}(B\boldsymbol{z}_t) = \boldsymbol{a}^\top\boldsymbol{f}_t$. Now let $\overline{\boldsymbol{f}}_t = \alpha\boldsymbol{f}_{t-1} + (1-\alpha)\overline{\boldsymbol{f}}_{t-1}$. Now note that $\boldsymbol{a}^\top\overline{\boldsymbol{f}}_t = \boldsymbol{a}^\top\alpha\boldsymbol{f}_{t-1} + \boldsymbol{a}^\top(1-\alpha)\overline{\boldsymbol{f}}_{t-1} = \overline{r}_t$. Putting this all together, we have $r_t - \overline{r}_t = \boldsymbol{a}^\top(\boldsymbol{f}_t - \overline{\boldsymbol{f}}_t)$. Now we can plug this into our learning rule and rearrange:

$$\begin{aligned}
\boldsymbol{\mu}_{t+1} &= \boldsymbol{\mu}_t - \lambda(r_t - \overline{r}_t)\boldsymbol{\epsilon}_t \\
\boldsymbol{\mu}_{t+1} &= \boldsymbol{\mu}_t - \lambda\boldsymbol{a}^\top(\boldsymbol{f}_t - \overline{\boldsymbol{f}}_t)\boldsymbol{\epsilon}_t \\
\boldsymbol{\mu}_{t+1} - \boldsymbol{\mu}_t &= -\lambda\boldsymbol{a}^\top(\boldsymbol{f}_t - \overline{\boldsymbol{f}}_t)\boldsymbol{\epsilon}_t
\end{aligned}$$

Let $X_t = -\boldsymbol{\epsilon}_t(\boldsymbol{f}_t - \overline{\boldsymbol{f}}_t)^\top \in \mathbb{R}^{d\times K}$, and $\boldsymbol{y}_t = \boldsymbol{\mu}_{t+1} - \boldsymbol{\mu}_t \in \mathbb{R}^d$. (For simplicity we will assume $\lambda = 1$, since we can't resolve any differences in λ.) We want to estimate $\boldsymbol{a}$ as follows:

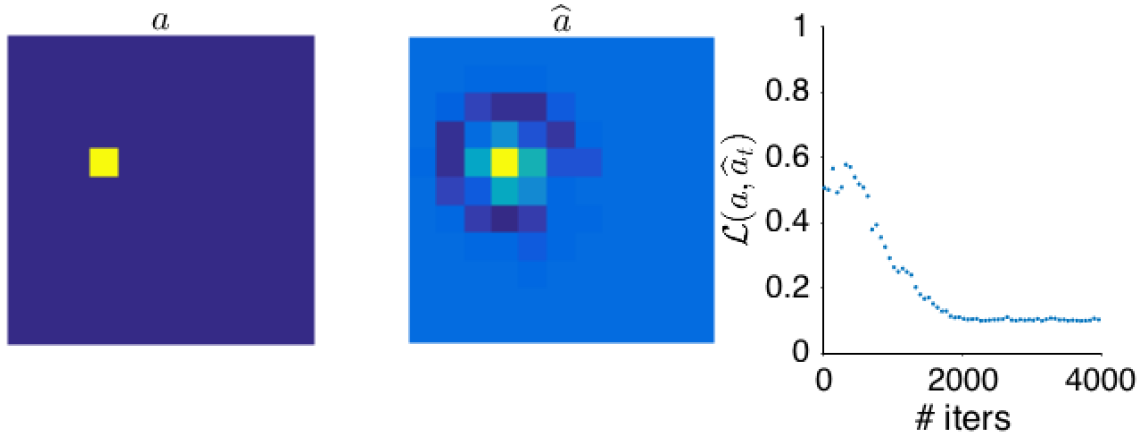
$$\begin{aligned}\widehat{\boldsymbol{a}} &= \arg\min_a \frac{1}{T}\sum_{t=1}^{T} \|\boldsymbol{y}_t - X_t\boldsymbol{a}\|_F^2 \\ &= (X^\top X)^{-1}(X^\top \boldsymbol{Y})\end{aligned}$$

where $X \in \mathbb{R}^{Td\times K}$ and $\boldsymbol{Y} \in \mathbb{R}^{Td}$ are constructed so that the i^{th} row of X corresponds to the i^{th} row of X_t, and the i^{th} row of $\boldsymbol{Y}$ corresponds to the i^{th} column of $\boldsymbol{y}_t$, for some t. In practice, we get better performance if we add a sparsity penalty on $\|\boldsymbol{a}\|_2$, so that $\widehat{\boldsymbol{a}} = (X^\top X + \rho I)^{-1}(X^\top \boldsymbol{Y})$ for some $\rho > 0$.

Results

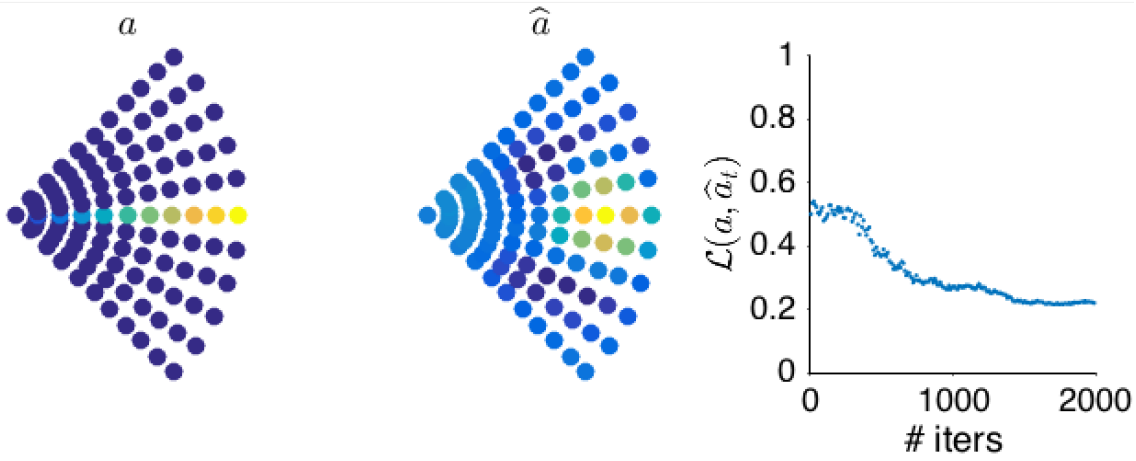
1. Target feedback.

In this simulation, you receive reward proportional to your distance from a target value of feedback, $\mathbf{v}^*$ (yellow). The true objective, a, is shown on the left. The estimated objective, $\boldsymbol{a}$ is shown in the middle. In the right panel, the error shown is based on the cosine angle between $\boldsymbol{a}$ and $\widehat{\boldsymbol{a}}_t$.



2. High speed with no angular error.

In this simulation, we suppose that your feedback is velocity, and that you receive reward for being near zero angular error, with higher rewards for faster speeds. Also note that here the RBFs are arranged using a polar coordinate grid.


a
$\widehat{a}$
$\mathcal{L}(a, \widehat{a}_i)$
1
0.8
0.6
0.4
0.2
0
0
1000
2000
iters

A Appendix to Chapter 2

A.1 Methods

Defining the mapping between neural activity and cursor movement

Experimental methods are described in detail in both Sadtler et al. (2014) and Golub et al. (2018). Briefly, we recorded from the proximal arm region of primary motor cortex (M1) in three male Rhesus macaques using implanted 96-channel microelectrode arrays (Blackrock Microsystems). All animal care and handling procedures conformed to the NIH Guidelines for the Care And Use of Laboratory Animals and were approved by the University of Pittsburgh's Institutional Animal Care and Use Committee. The population spiking activity in each non-overlapping 45 ms bin was computed as the number of threshold crossings on each channel. In each session, 85-94 neural units were recorded (25 sessions from monkey J, 6 sessions from monkey L, 11 sessions from monkey N). These sessions were analyzed previously in Golub et al. (2018). Data from monkeys J and L were first presented in Sadtler et al. (2014). The average firing rate of the neural units per session was 50 ± 8, 42 ± 4, and 55 ± 14 spikes/s (mean $\pm$ s.d.) for monkeys J, L, and N, respectively.

Each session began with a block of calibration trials. The calibration procedure for monkey J involved either passive observation of cursor movement, or closed-loop BCI cursor control using the previous day's BCI mapping. For monkeys L and N, we used a closed-loop calibration procedure that gradually stepped from passive observation to closed-loop control, as described in Sadtler et al. (2014). We then applied factor analysis (FA) to the spike counts recorded during these calibration trials to identify the 10D linear subspace (i.e., the "intrinsic manifold") that captured dominant patterns of co-modulation across neural units (Churchland et al., 2010; Harvey et al., 2012; Sadtler et al., 2014; Athalye et al., 2017). We then estimated the factor activity, $\mathbf{z}_t \in \mathbb{R}^{10\times 1}$, as the posterior expectation given the observed spike counts, $\mathbf{u}_t \in \mathbb{R}^{q\times 1}$, where q is the number of neural units:

$$\mathbf{z}_t = L^\top (LL^\top + \Psi)^{-1}(\mathbf{u}_t - \mathbf{d}) \tag{A.1}$$

Here, L, Ψ, and $\mathbf{d}$ are FA parameters estimated using the expectation-maximization algorithm, where Ψ is constrained to be a diagonal matrix. The factor activity, $\mathbf{z}_t$,

can be interpreted as a weighted combination of the activity of different neural units. We refer to $\mathbf{z}_t$ as a "population activity pattern."

We next orthonormalized $\mathbf{z}_t$ so that it had units of spike counts per time bin (Yu et al., 2009), using the following approach. In our FA model, L defines a mapping from low-dimensional factor space to the higher-dimensional neural space. Because the columns of L are not orthonormal, the factor activity does not have the same units (spikes counts per time bin) as the neural activity. However, we can fix this by finding an orthonormal basis for the columns of L (Yu et al., 2009). To do this, we apply the singular value decomposition, yielding $L = USV^\top$, where $U \in \mathbb{R}^{q \times 10}$ and $V \in \mathbb{R}^{10 \times 10}$ have orthonormal columns and $S \in \mathbb{R}^{10 \times 10}$ is diagonal. Then, we can write $L\mathbf{z}_t = U(SV^\top \mathbf{z}_t) = U\tilde{\mathbf{z}}_t$. Because U has orthonormal columns, $\tilde{\mathbf{z}}_t = SV^\top \mathbf{z}_t$ has the same units (spike counts per time bin) as $\mathbf{u}_t$. For notational simplicity, we refer to $\tilde{\mathbf{z}}_t$ as $\mathbf{z}_t$ throughout. The values in $\mathbf{z}_t$ appear larger than those expected for a single neuron because this value tends to grow with the total number of neural units.

Over the course of each experiment, animals used two different BCI mappings (see "Behavioral task" below). Each BCI mapping translated the resulting moment-by-moment factor activity ($\mathbf{z}_t$) into a 2D cursor velocity ($\mathbf{v}_t$) using a Kalman filter:

$$\mathbf{v}_t = A\mathbf{v}_{t-1} + B\mathbf{z}_t + \mathbf{c} \tag{A.2}$$

For the first BCI mapping, $A \in \mathbb{R}^{2 \times 2}$, $B \in \mathbb{R}^{2 \times 10}$, and $\mathbf{c} \in \mathbb{R}^{2 \times 1}$ were computed from the Kalman filter parameters, estimated using the calibration trials. For the second BCI mapping, we changed the relationship between population activity and cursor movement by randomly permuting the elements of $\mathbf{z}_t$ before applying Eqn. A.2. This permutation procedure can be formulated so that Eqn. A.2 still applies to the second BCI mapping, but with an updated definition of B (Sadtler et al., 2014).

Behavioral task

Each animal performed an 8-target center-out task by modulating its M1 activity to control the velocity of a computer cursor. Each session involved two different BCI mappings. The first mapping was chosen to be intuitive for the animal to use. The animal used this first mapping for 200-400 trials, after which the mapping was changed abruptly to a second BCI mapping. The second mapping was initially difficult for the animal to use, and the animal was given 400-600 trials to learn to use the second mapping. Both mappings were chosen to be within the animal's instrinic manifold, mappings that we found in previous work could be readily learned within one session (Sadtler et al., 2014).

At the beginning of each trial, a cursor appeared in the center of the workspace, followed by the appearance of one of eight possible peripheral targets (chosen pseudorandomly). For the first 300 ms of the trial, the velocity of the cursor was fixed at zero. After this, the velocity of the cursor was controlled by the animal through the BCI mapping. If the animal acquired the peripheral target with the cursor within

7.5 s, he received a water reward, and the next trial began 200 ms after target acquisition. Otherwise, the trial ended, and the animal was given a 1.5 s time-out before the start of the next trial.

Session and trial selection

The data analyzed in this study was part of a larger study involving learning two different types of BCI mapping changes: within-manifold perturbations (WMP) and outside-manifold perturbations (OMP) (Sadtler et al., 2014). We found that animals learned WMPs better than OMPs. Because we need animals to show stable cursor control under both mappings, we only analyzed WMP sessions in this study. Among the WMP sessions, we further selected those in which the animal learned stable control of the second mapping (42 selected and 12 discarded). This was important because performance with the second mapping was generally not as good as with the first mapping (Supplemental Fig. A.1), and we wanted to ensure that any potential results were not due to incomplete learning of the second mapping (see also "Internal model estimation" below). We further subselected from each session only those trials which exhibited stable behavioral performance, using a metric defined below. This was done to ensure that we were analyzing trials for which animals used a consistent strategy for selecting activity patterns.

We included sessions in which there existed a block of at least 100 consecutive trials that showed both substantial learning of the second mapping and consistent behavior. To identify trials showing substantial learning, we computed the running mean of the target acquisition time (on correct trials only), smoothed with a 100-trial boxcar shifted one trial at a time. The smoothed acquisition time for a trial corresponded to the average acquisition time within a 100-trial window centered on that trial. We then normalized these values so that 1 corresponded to the largest acquisition time in the first 50 trials using the second mapping, and 0 corresponded to the smallest acquisition time in the subsequent trials using the second mapping. We defined trials showing substantial learning as those with normalized acquisition times below 0.5. Next, to identify trials with consistent behavior, we computed the running variance of the target acquisition time. This was computed by taking the variance of the smoothed acquisition time above in a 100-trial boxcar, shifted one trial at a time. We then normalized these variances so that 1 corresponded to the largest variance in the first half of trials using the second mapping, and 0 corresponded to the smallest variance in any trial using the second mapping. We defined trials showing stable behavior as those with normalized variance below 0.5. We then identified blocks of consecutive trials that passed both of these criteria, joining blocks if they were separated by no more than 10 trials. We then selected the longest such block of at least 100 trials for our analyses. If no such block of trials was found, we excluded that session from our analyses. This procedure resulted in the 42 sessions across three monkeys that we included in our analyses.

We analyzed only successful trials. To avoid analyzing time steps with potentially idiosyncratic cursor control, we also ignored portions of the trial when the cursor

was closer than 50mm or more than 125mm away from the origin. We repeated our analyses without the latter exclusion and obtained quantitatively similar results.

Internal model estimation

When an animal uses a BCI mapping, its internal conception of the BCI mapping can differ from the actual BCI mapping, even during proficient control (Golub et al., 2015). As a result, the animal's conception of output-potent versus output-null dimensions can be different from those defined by the actual BCI mapping. To control for this possibility, we evaluated our predictions based on the animal's internal conception of the output-null dimensions, rather than the actual output-null dimensions of the BCI mapping. This is particularly important for the second mapping, but we also did this for the first mapping. We used a method (Internal Model Estimation, IME) that we developed previously for estimating the animal's internal model of the BCI mapping (Golub et al., 2015), with the exception that here we apply the model directly to the factor activity ($\mathbf{z}_t$) as opposed to the neural activity ($\mathbf{u}_t$), as was done in Golub et al. (2015).

The main idea of the IME framework is that the animal generates neural activity consistent with aiming straight to the target through an internal model of the BCI mapping. Due to natural visual feedback delay, the animal cannot exactly know the current cursor position, and thus aims from an internal estimate of the current cursor position. The internal estimate of the cursor position is a feedforward prediction based on previously issued neural activity and the most recently available visual feedback. Supplemental Fig. A.5A shows a single-trial BCI cursor trajectory (black), along with the animal's internal belief (red "whisker") about how cursor position (red dots) evolved from the cursor position known from the most recently available visual feedback. The final segments of the trajectories reflect the same neural activity, which produces the actual cursor velocity (black arrow) through the actual BCI mapping, or the animal's intended cursor velocity (red arrow) through the animal's internal model. The animal's velocity command viewed through the internal model points closer toward the target than the actual movement of the BCI cursor, corresponding to a smaller angular error. Across sessions, the animals' angular errors when using the second BCI mapping did not usually return to the original level of error that the animal achieved under the first mapping (Sadtler et al., 2014) (Supplemental Fig. A.5B). However, when viewed through the animals' internal models of the BCI mappings, angular errors during the second mapping were more similar to those observed during the first mapping (Supplemental Fig. A.5C). Thus, the internal model helps to control for possible incomplete learning of the second mapping.

We used IME to obtain the animal's internal model of the BCI mapping (in the form of $A, B, \mathbf{c}$ in Eqn. A.2), which yielded a corresponding set of cursor velocities ($\mathbf{v}_t$), cursor-target angles (θ_t), and bases for the output-potent and output-null dimensions of each mapping (see N and R below) that we used in our offline analyses. The results reported in the main text are based on these quantities obtained from IME. When we analyzed the data without using IME (i.e., using the actual

output-null dimensions of the BCI mapping), all of the results we report still held (Supplemental Fig. A.4).

Defining output-null activity

In Eqn. A.2, the matrix $B \in \mathbb{R}^{2\times 10}$ linearly projects a 10-dimensional input (factor activity) to a 2-dimensional output (cursor velocity). Thus, for any given cursor velocity ($\mathbf{v}_t$) there are multiple values of factor activity ($\mathbf{z}_t$) that would produce it. These multiple values of factor activity are all behaviorally equivalent, and we refer to their existence as "neural redundancy."

Mathematically, it is useful to consider the null space, $Nul(B)$, and the row space, $Row(B)$, of the matrix B. The critical property of $Nul(B)$ is that for any element $\mathbf{y} \in Nul(B) \subseteq \mathbb{R}^{10}$, we have $B\mathbf{x} = B(\mathbf{x} + \mathbf{y})$ for all $\mathbf{x} \in \mathbb{R}^{10}$. In other words, any change in activity within the null space of B has no effect on the cursor movement produced. On the other hand, to achieve a particular cursor velocity ($\mathbf{v}_t$), there is exactly one $\mathbf{x} \in Row(B)$ such that $B\mathbf{x} = \mathbf{v}_t$. Thus, the activity in the row space of B uniquely determines the cursor movement. To find a basis for $Row(B)$ and $Nul(B)$, we took a singular value decomposition of $B = USV^T$, where the diagonal elements of S were ordered so that only the first two values were nonzero. Then, we let $R \in \mathbb{R}^{10\times 2}$ be the first two columns of V, and $N \in \mathbb{R}^{10\times 8}$ be the remaining eight columns. The columns of N and R are mutually orthonormal and together form an orthonormal basis for the 10-dimensional space of factor activity. This allowed us to decompose the factor activity $\mathbf{z}_t$ at each time step into two orthogonal components: (1) activity in the row space of B that affects the cursor velocity, which we call the *output-potent* activity ($\mathbf{z}_t^r \in \mathbb{R}^2$); and (2) activity in the null space of B that does not affect the cursor movement, which we call the *output-null* activity ($\mathbf{z}_t^n \in \mathbb{R}^8$):

$$\begin{aligned} \mathbf{z}_t &= N\mathbf{z}_t^n + R\mathbf{z}_t^r \\ \text{where} \quad \mathbf{z}_t^n &:= N^\top \mathbf{z}_t, \quad \mathbf{z}_t^r := R^\top \mathbf{z}_t \end{aligned} \tag{A.3}$$

Note that all behaviorally equivalent activity will have the same output-potent activity ($\mathbf{z}_t^r$), but can differ in output-null dimensions. Thus, for time steps with similar cursor movements, the subject's choice of 8D output-null activity ($\mathbf{z}_t^n$) describes how the subject selected activity from a set of behaviorally equivalent options. Because the cursor velocity ($\mathbf{v}_t$) at each time step is a combination of output-potent activity and the cursor velocity at the previous time step (see Eqn. A.2), output-potent activity can be thought of as driving a change in the cursor velocity. Note that in the depictions of hypotheses in Fig. 3.1, Fig. 3.2, Fig. 3.3, and Fig. 3.4, we used $\mathbf{v}_t = B\mathbf{z}_t$ instead of Eqn. A.2 for clarity.

Predicting output-null activity

Our goal for each experiment was to predict the distribution of observed output-null activity during the second mapping across time steps corresponding to a given cursor movement direction (defined as the angle of $\mathbf{v}_t$ in Eqn. A.2). In the context of the center-out task, we assumed that cursor movements in the same direction but with different speeds were still behaviorally equivalent to the animal. This is supported by previous work that found substantially more direction-related information than speed-related information in both single-unit and population activity in M1 (Golub et al., 2014). For this reason we assessed the output-null distribution in bins of cursor movement direction rather than cursor velocity (i.e., direction × speed).

All hypotheses generated predictions of the distribution of output-null activity observed while animals used the second BCI mapping, unless otherwise noted. To generate predictions of the distributions of output-null activity, we made predictions of the output-null activity at each time step. This allowed us to ensure that our predictions were consistent with the cursor kinematics observed during the experiment. We then aggregated the predictions across all time steps during the experiment with a similar cursor movement direction. In all cases, the predicted output-null activity respected the intrinsic manifold (Sadtler et al., 2014), because the output-null activity lies in an 8-dimensional subspace of the 10-dimensional intrinsic manifold.

To generate a prediction of the output-null activity for a particular time step ($\mathbf{z}_t^n$), each hypothesis had access to three sources of information recorded during the experiments. First, all hypotheses used the observed output-potent activity ($\mathbf{z}_t^r$), in order to ensure that every prediction was physiologically plausible (see below). Second, all hypotheses except for the Minimal Firing hypothesis utilized factor activity recorded during use of the first BCI mapping to form their predictions of output-null activity. Finally, the Persistent Strategy hypothesis also utilized the current position of the cursor relative to the target, defined as the cursor-target angle (θ_t).

We ensured that all predictions of output-null activity ($\widehat{\mathbf{z}}_t^n$) corresponded to physiologically plausible neural activity ($\widehat{\mathbf{u}}_t$). By "physiologically plausible" we mean that the neural activity was non-negative, and no greater than the maximum number of spikes (per 45 ms time step) observed for that neural unit during trials using the first BCI mapping ($\mathbf{u}_{max}$). To enforce the constraint, we either incorporated the constraint $0 \leq \widehat{\mathbf{u}}_t \leq \mathbf{u}_{max}$ directly in the optimization problem (Minimal Firing hypothesis), or rejected predictions of neural activity that fell outside of the constraint (all other hypotheses). In the latter case, we combined the predicted output-null activity with the observed output-potent activity at that time step to form the predicted factor activity ($\widehat{\mathbf{z}}_t$). We then converted this value to neural activity using the FA generative model:

$$\widehat{\mathbf{u}}_t := L\widehat{\mathbf{z}}_t + \mathbf{d} \tag{A.4}$$

If this neural activity was not physiologically plausible, we attempted to generate a new prediction of $\widehat{\mathbf{z}}_t^n$ according to the hypothesis. This was possible because all

hypotheses incorporated some form of sampling to generate their predictions. If this procedure failed even after 100 attempts to generate a physiologically plausible prediction, we skipped making a prediction for that time step. This happened for less than 1% of all time steps.

Minimal firing hypotheses. According to the Minimal Firing hypothesis, generating spikes incurs a metabolic cost. Thus, the subject should select the population activity pattern that involves the fewest spikes among all patterns that generate the desired cursor movement. Predictions for this hypothesis were generated as follows. For each time step, we find the spiking activity closest to zero firing that produces the observed cursor velocity:

$$\begin{aligned} \widehat{\mathbf{u}}_t &:= \arg\min_u \|\mathbf{u}\|_2^2 \\ \text{subject to } \mathbf{v}_t &= A\mathbf{v}_{t-1} + Bf(\mathbf{u}) + \mathbf{c} \\ \text{and } \mathbf{0} &\leq \mathbf{u} \leq \mathbf{u}_{max} \end{aligned} \tag{A.5}$$

Above, $f(\mathbf{u})$ refers to the factor activity corresponding to $\mathbf{u}$, as in Eqn. A.1. Because $f(\mathbf{u})$ is a linear function of $\mathbf{u}$, the above minimization is a convex problem. $\mathbf{u}_{max}$ is the maximum activity level observed for each neuron, as described above. We solved for $\widehat{\mathbf{u}}_t$ at each time step t using Matlab's `quadprog`. All trends in results were the same if the L_2 norm in the optimization problem was changed to an L_1 norm. After solving the above minimization, we incorporated variability in spike generation by sampling from a Poisson: $\widehat{\mathbf{u}}'_t \sim \text{Poisson}(\widehat{\mathbf{u}}_t)$. We repeated this last step if necessary until $\widehat{\mathbf{u}}'_t$ was physiologically plausible. Finally, we converted the prediction to factor activity, so that the resulting prediction of $\widehat{\mathbf{z}}_t^n$ was $\widehat{\mathbf{z}}_t^n := N^\top f(\widehat{\mathbf{u}}'_t)$.

We chose to incorporate Poisson variability into the predictions of the Minimal Firing (above) and Minimal Deviation hypotheses (below), rather than the Gaussian noise assumed by our FA model. The observed spike counts are discrete, whereas adding Gaussian noise would make the spike counts predicted by these hypotheses continuous. For this reason, to ensure a fair comparison we used Poisson variability, which will ensure the predictions remain discrete even after adding variability.

For the Minimal Deviation hypothesis, we generalized the Minimal Firing hypothesis so that instead of predicting the spiking activity nearest zero spikes/s, we predicted factor activity closest to some unknown activity level $\eta \in \mathbb{R}^{10}$. Solving this problem in the 10-dimensional factor space for the optimal value of η yields lower prediction error than doing so in the q-dimensional neural space because we ultimately evaluate the hypotheses' predictions in factor space. After choosing η (see below), the predicted factor activity was obtained by solving the following optimization problem:

$$\begin{aligned} \widehat{\mathbf{z}}_t &:= \arg\min_z \|\mathbf{z} - \eta\|_2^2 \\ \text{subject to } \mathbf{v}_t &= A\mathbf{v}_{t-1} + B\mathbf{z} + \mathbf{c} \end{aligned} \tag{A.6}$$

The above problem is known as a "minimum norm" problem, and it turns out that the resulting solution's output-null activity, $\widehat{\mathbf{z}}_t^n$, is a constant, for all t:

$$\widehat{\mathbf{z}}_t^n = N^\top \eta \tag{A.7}$$

Because of the simple form of this solution, it was possible to choose the best value of η for each session by minimizing the resulting output-null prediction error across cursor directions (see "Error in mean" below). This value is:

$$\eta := N \left(\frac{1}{8} \sum_{i=1}^{8} \mu_i^n \right) \tag{A.8}$$

where μ_i^n is the average output-null activity in the i^{th} cursor direction bin, which we estimate using activity recorded during the first BCI mapping. This ensures that the data used to evaluate the predictions were not used to obtain η. Finally, we incorporated spiking variability just as we did for the Minimal Firing hypothesis. To do this, we first converted the above prediction ($\widehat{\mathbf{z}}_t$) to neural activity using the FA generative model ($\widehat{\mathbf{u}}_t := L\widehat{\mathbf{z}}_t + \mathbf{d}$). We then incorporated Poisson variability as described above, repeating the procedure until the resulting prediction was physiologically plausible, where the prediction of $\widehat{\mathbf{z}}_t^n$ was $\widehat{\mathbf{z}}_t^n := N^\top f(\widehat{\mathbf{u}}_t')$, with $\widehat{\mathbf{u}}_t' \sim \text{Poisson}(\widehat{\mathbf{u}}_t)$.

Uncontrolled-uniform hypothesis. According to the uncontrolled manifold concept (Scholz and Schöner, 1999), variability in output-null dimensions will be higher than that in output-potent dimensions. One explanation of this idea is the minimal intervention principle (Todorov and Jordan, 2002; Valero-Cuevas et al., 2009; Diedrichsen et al., 2010b), which states that the variability in output-potent dimensions is controlled independently of the output-null activity, with the output-null activity being unmodified. While this principle specifies that output-null activity is independent of output-potent activity, it does not specify what the distribution of output-null activity actually is. Thus, we considered two hypotheses about this distribution. First, we supposed that the output-null activity would be uniformly distributed within bounds determined by the physiological range of population activity. This hypothesis thus predicts that activity in output-null dimensions has maximal entropy within the physiological range. For each t, we sampled:

$$\widehat{\mathbf{z}}_t^{\mathbf{n}} \sim \text{Uniform}(\mathbf{z}_{min}^n,\ \mathbf{z}_{max}^n) \tag{A.9}$$

Above, $\mathbf{z}_{min}^n$ and $\mathbf{z}_{max}^n$ set the range on the minimum and maximum possible output-null activity. These bounds were set using population activity recorded during use of the first BCI mapping. We then resampled if necessary until our predictions generated physiologically plausible spiking activity when combined with the output-potent activity.

Note that in Fig. 3.2, Fig. 3.3, and Fig. 3.4 we applied PCA to the observed output-null activity to depict the three output-null dimensions with the most shared

variance in the observed activity. Because of this, our visualizations of the distributions predicted by the Uncontrolled-uniform hypothesis in Fig. 3.3 appear mound-shaped rather than uniform. To understand this, suppose we sample from a uniform distribution over a rectangle in 2D. If we rotate this rectangle slightly, and visualize the distribution of points along the x-axis, the distribution will be mound-shaped. Similarly, the Uncontrolled-uniform hypothesis samples from a uniform distribution in the 8-dimensional output-null space, where the bounds of the rectangle are determined by $\mathbf{z}^n_{min}$ and $\mathbf{z}^n_{max}$ above. Applying PCA rotates this activity, such that the density along the PC dimensions appears mound-shaped.

Uncontrolled-empirical hypothesis. Next, we considered a different hypothesis about the distribution of output-null activity under the minimal intervention principle. Rather than assuming output-null activity is uniformly distributed, we obtained an empirical distribution using population activity observed under the first mapping. To produce predictions of output-null activity during the second mapping, for each time step during the second mapping we sampled randomly from the population activity observed under the first mapping, and assessed the projection of that activity in the null space of the second mapping.

Concretely, let T_1 be the set of all time steps under the first mapping, and T_2 be the set of all time steps under the second mapping. Our prediction for each $t \in T_2$ is obtained by randomly sampling with replacement:

$$\widehat{\mathbf{z}}^{\mathbf{n}}_t \sim \{N^\top \mathbf{z}_i \mid \forall i \in T_1\} \tag{A.10}$$

In other words, at each time step using the second mapping, we randomly select factor activity observed during the first mapping ($\mathbf{z}_i$), and project it into the null space of the second mapping ($N^\top \mathbf{z}_i$). We then resampled if necessary until our predictions generated physiologically plausible spiking activity when combined with the output-potent activity.

Persistent Strategy hypothesis. An extension of the Uncontrolled-empirical hypothesis is motivated by the idea that the subject may select activity under one mapping by modifying the activity he used under the first mapping. For a given cursor-target angle, if the subject selects the same population activity under the second mapping as under the first mapping, that activity may not move the cursor towards the target under the second mapping. To correct the cursor movement, he modifies this activity according to the minimal intervention principle (Todorov and Jordan, 2002; Valero-Cuevas et al., 2009; Diedrichsen et al., 2010b), correcting activity only along output-potent dimensions of the current mapping. Concretely, for each $t \in T_2$, we sampled with replacement:

$$\widehat{\mathbf{z}}^{\mathbf{n}}_t \sim \{N^\top \mathbf{z}_i \mid \forall i \in T_1 \text{ such that } \theta_i \in \theta_t \pm 22.5^\circ\} \tag{A.11}$$

where θ_t is the cursor-to-target angle at time t. As before, we resampled if necessary until our predictions generated physiologically plausible spiking activity

when combined with the output-potent activity. This hypothesis is identical to the Uncontrolled-empirical hypothesis, except that at each time step we sampled only from time steps during the first mapping that had a similar cursor-target angle (i.e., within a 45° wedge around θ_t). We found no consistent improvements in varying the constraints on the cursor-target angle (i.e., using values other than 22.5° in Eqn. A.11), or when using the output-potent angle rather than the cursor-target angle.

Fixed Distribution hypothesis. According to the Fixed Distribution hypothesis, the activity in output-null dimensions is tightly coupled to the activity in output-potent dimensions, even under different BCI mappings when these dimensions are not necessarily still null and potent. This is in contrast to the three previous hypotheses (Uncontrolled-uniform, Uncontrolled-empirical, Persistent Strategy), which all incorporated a minimal intervention principle, whereby output-null activity can be modified independently of the output-potent activity, within the physiological limits on the firing rate of each unit.

Under the Fixed Distribution hypothesis, we predict that the distribution of output-null activity given the output-potent activity will be the same distribution as it was under the previous mapping. To implement this hypothesis, for each time step during the second mapping, we predict that the subject selects whichever activity pattern he produced under the previous mapping that would best match the current output-potent activity. Specifically, given the output-potent activity produced during the second mapping ($\mathbf{z}_t^r$), we found the time step during the first mapping ($i^* \in T_1$) where the factor activity $\mathbf{z}_{i^*}$ would have come closest to producing that output-potent activity using the second mapping. Our prediction for output-null activity was then the output-null component of $\mathbf{z}_{i^*}$ through the second mapping ($N^T \mathbf{z}_{i^*}$). Mathematically, for each $t \in T_2$ our prediction was:

$$\begin{aligned} \widehat{\mathbf{z}}_t^{\mathbf{n}} &:= N^\top \mathbf{z}_{i^*} \\ \text{where } i^* &= \arg\min_{i \in T_1} \|\mathbf{z}_t^r - R^\top \mathbf{z}_i\|_2^2 \end{aligned} \tag{A.12}$$

We observed that these predictions all satisfied physiological constraints, which suggests that the values of $\mathbf{z}_{i^*}$ selected at each time step each produced output-potent activity sufficiently close to $\mathbf{z}_t^r$.

This above implementation is also equivalent to the following: At each time step we identified the K previously observed population activity patterns that would produce output-potent activity closest to the current output-potent activity under the second mapping. We then selected one of these patterns at random, and used the output-null activity of that pattern as our prediction at that time step. In our above implementation, $K = 1$. We found that using other values of K (e.g., $K = 50$, $K = 200$) yielded similar results.

Evaluating predictions

For each session, we evaluated the predicted output-null distributions of the above hypotheses in terms of how well they matched the observed output-null distributions

for all time steps with similar cursor movements. To do this, we first grouped time steps by their corresponding cursor velocity into eight non-overlapping bins of cursor movement directions ($0° \pm 22.5°, 45° \pm 22.5°, ..., 315° \pm 22.5°$). We then evaluated the accuracy of the predictions for each cursor movement direction.

For consistency, all predictions were evaluated in terms of factor activity. The Minimal Firing hypothesis generated its predictions in terms of neural activity, and we converted these predictions to factor activity using Eqn. A.1.

Histogram overlap. We compared the predicted and observed distributions of output-null activity in each dimension in terms of the average overlap of their histograms. For each session, we selected a single bin size for all histograms using cross-validation (Rudemo, 1982). Then, for each cursor direction and output-null dimension, we computed the error between the observed ($\mathbf{y}$) and predicted ($\widehat{\mathbf{y}}$) histograms. Let $\mathbf{y}_i$ be the normalized frequency in the i^{th} bin, so that $\sum_{i=1}^{m} \mathbf{y}_i = 1$, and similarly for $\widehat{\mathbf{y}}$. Then the histogram error was computed as follows:

$$L(\mathbf{y}, \widehat{\mathbf{y}}) = \frac{1}{2} \sum_{i=1}^{m} |\mathbf{y}_i - \widehat{\mathbf{y}}_i| \tag{A.13}$$

Above, $\frac{1}{2}$ is included so that $L(\mathbf{y}, \widehat{\mathbf{y}}) = 1$ if the two histograms are completely non-overlapping. $L(\mathbf{y}, \widehat{\mathbf{y}}) = 0$ if the two histograms are identical. This error was then averaged across all cursor directions and output-null dimensions. We multiplied this value by 100 to yield the average histogram error percentages reported in the main text.

For the visualizations in Fig. 3.2, Fig. 3.3, and Fig. 3.4, we displayed the marginal histograms in the three output-null dimensions with highest variance in the observed output-null activity, as found by PCA. For all error calculations we considered all eight output-null dimensions without applying PCA.

Error in mean. We assessed how well our predictions matched the observed mean output-null activity for each cursor movement direction. For all time steps in the same cursor movement direction bin, let $\mu^{\mathbf{n}} \in \mathbb{R}^{8\times 1}$ be the vector of the mean observed output-null activity, and $\widehat{\mu}^{\mathbf{n}} \in \mathbb{R}^{8\times 1}$ the mean output-null activity predicted by a particular hypothesis. These are both vectors, and so we computed the distance between them using the ℓ_2 norm:

$$L(\mu^{\mathbf{n}}, \widehat{\mu}^{\mathbf{n}}) = \|\mu^{\mathbf{n}} - \widehat{\mu}^{\mathbf{n}}\|_2 \tag{A.14}$$

For each hypothesis, we computed the error in mean in each cursor movement direction bin, and took the average of these values as the error in mean for each session.

Error in covariance. We next assessed how well our predictions matched the observed covariance of output-null activity for each cursor movement direction. Let $C^n \in \mathbb{R}^{8\times 8}$ and $\widehat{C}^n \in \mathbb{R}^{8\times 8}$ be the covariance of the observed and predicted output-null activity, respectively. There are a variety of methods for comparing covariance

matrices, such as comparing their trace or determinant. We chose a metric invariant to affine transformations (e.g., scaling, translations, rotations) of the coordinate system (Dryden et al., 2009). Because the amount of variance in the recorded population activity might vary from session to session, this property of affine invariance helps ensure we can reasonably compare our covariance errors across sessions.

Let $\lambda_i(C^n, \widehat{C}^n)$ be the i^{th} generalized eigenvalue of C^n and $\widehat{C}^n$ (i.e., a value λ such that $\det(C^n - \lambda\widehat{C}^n) = 0$). Then following Lang (1999) and Förstner and Moonen (2003), we computed the distance between these two matrices as:

$$L(C^n, \widehat{C}^n) = \sqrt{\sum_i \log^2 \lambda_i(C^n, \widehat{C}^n)} \tag{A.15}$$

If $C^n = \widehat{C}^n$, then $L(C^n, \widehat{C}^n) = 0$. For each hypothesis, we computed the error in covariance in each cursor movement direction bin, and took the average of these values as the error in covariance for each session.

Error floor. To estimate the smallest errors achievable by any hypothesis (the "error floor"), given a limited number of samples to estimate the true output-null distributions, we performed the following analysis. For each session, we randomly split the data during the second mapping in half, and measured the histogram, mean, and covariance errors when using the output-null activity from one half to predict the distribution of the output-null activity during the other half. We repeated this process 100 times per session, and took the averages of the resulting errors as our estimates of the error floors for that session.

Activity that became output-null in the second mapping

We sought to assess whether the variance of population activity changed in dimensions that became output-null under the second mapping. To do this, we identified the subspace of activity that was output-potent under the first mapping, but output-null under the second mapping.

As before, let the columns of N be a basis for the null space of the second mapping. Now let the columns of R_1 be a basis for the row space of the first mapping. Then the space spanned by the columns of $(NN^\top)R_1 \in \mathbb{R}^{10\times 2}$ describes the activity that would move the cursor during the first mapping but would not move the cursor during the second mapping. Let $S \in \mathbb{R}^{10\times 2}$ be an orthonormal basis for $(NN^\top)R_1$, which we obtained by performing a singular value decomposition. Now let $Z \in \mathbb{R}^{10\times n}$ be a matrix of n factor activity patterns. To measure the amount of variance of Z in the subspace spanned by the columns of S, we computed $\text{Trace}(\text{Cov}(Z^\top S)) \in \mathbb{R}$.

To assess how the variance of activity changes when it becomes irrelevant to cursor control, we grouped the time steps based on the cursor movement angle under the *second* mapping, for activity recorded under both the first and second mappings. First conditioning on the movement angle under the second mapping is consistent with our earlier analyses, when comparing the predicted and observed output-null

distributions. To compute the cursor movement angle through the second mapping for activity recorded under the first mapping, we used the terms of Eqn. A.2 not involving the cursor velocity at the previous time step (i.e., we computed $\mathbf{v}_t = B\mathbf{z}_t + \mathbf{c}$). For consistency, we recomputed the cursor movement angle for activity recorded under the second mapping in the same way.

Let Z_1 and Z_2 be the factor activity in the same cursor movement angle bin recorded during the first and second mappings, respectively. We then computed the ratio of variance R as follows:

$$R = \log\left(\frac{\text{Trace}(\text{Cov}(Z_2^\top S))}{\text{Trace}(\text{Cov}(Z_1^\top S))}\right) \tag{A.16}$$

The sign of R specifies whether the variance of activity increased ($R > 0$) or decreased ($R < 0$) when that activity became irrelevant to cursor control under the second mapping. We took the average of this ratio across all cursor movement direction bins to compute a ratio for each session.

To compute this ratio for the predictions of our hypotheses, as in Fig. 3.6C, we substituted Z_2 with the predictions of our hypotheses, i.e., by combining their predicted output-null activity with the observed output-potent activity under the second mapping.

We also repeated the above analyses on our predictions of output-null activity produced during the first mapping using the activity observed under the second mapping, as shown in Supplemental Fig. A.3 and Supplemental Fig. A.6. This was done by swapping the roles of the first and second mappings in the above analysis description.

Distances of mean output-null activity from Minimal Firing and Minimal Deviation

For each cursor direction on each session, we computed the distance from the mean observed output-null activity to the mean predicted by the Minimal Deviation hypothesis, where the distance was computed as the ℓ_2 norm between the two 8D mean vectors. We then compared this distance to the distance between the mean predicted by Fixed Distribution and the mean predicted by Minimal Deviation (Supplemental Fig. A.7). If the latter distance was consistently smaller than the former, this would be evidence that the observed mean output-null activity had moved towards the predictions of Minimal Deviation, relative to what was predicted by Fixed Distribution. We did not find evidence that this was the case (one-sided Wilcoxon signed rank test, $p > 0.5$), suggesting that the mean observed output-null activity was not closer to Minimal Deviation than expected under Fixed Distribution. We repeated the same analysis using the mean predicted by Minimal Firing instead of Minimal Deviation, and reached the same results (one-sided Wilcoxon signed rank test, $p > 0.5$).

A.2 Supplemental Figures

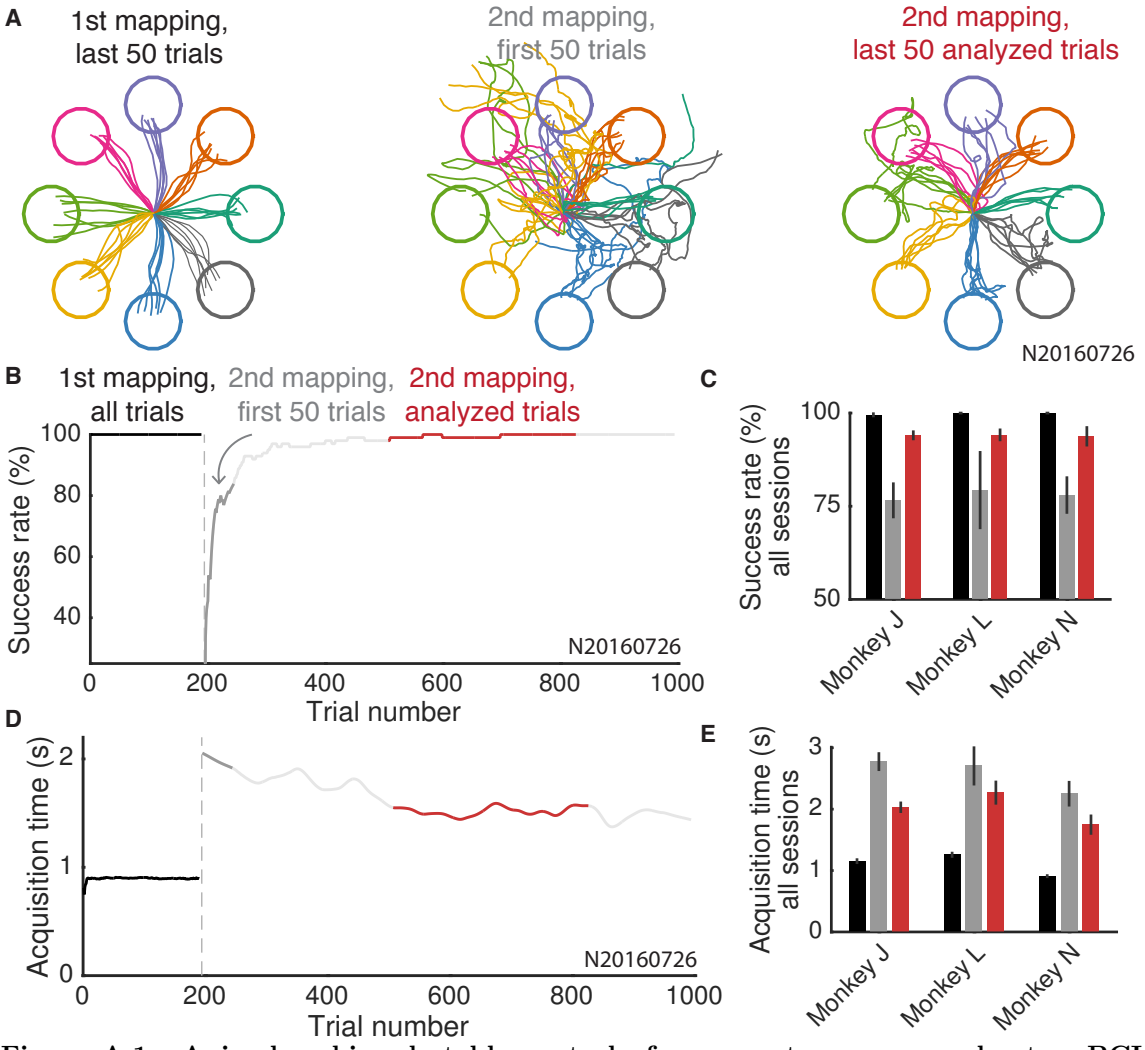


Figure A.1. Animals achieved stable control of a computer cursor under two BCI mappings in a 2D center-out task. (A) Cursor traces from 50 consecutive trials during three blocks in an example experiment (N2016726). Cursor traces are colored by the corresponding target on each trial. Cursor positions exceeding a cutoff distance from the workspace center are omitted from view. (B) Success rate during the same example experiment shown in (A), with colors indicating three blocks similar to those shown in (A): all trials under the first mapping (black), the first 50 trials under the second mapping (dark gray), and all analyzed trials under the second mapping (red). Success rates were smoothed using a 100-trial moving window. (C) Average success rate across all sessions for each of three animals, during the three blocks highlighted in (B). Error bars depict mean ± SE. (D) Same conventions as (B), for target acquisition time. Acquisition times were computed during successful trials only. For each session, for analysis we identified a consecutive block of at least 100 trials that showed both substantial learning of the second mapping and consistent behavior. Trials with substantial learning were identified by thresholding the smoothed acquisition times shown here (see Methods). Only correct trials within this block were analyzed. (E) Same conventions as (C), for acquisition time.

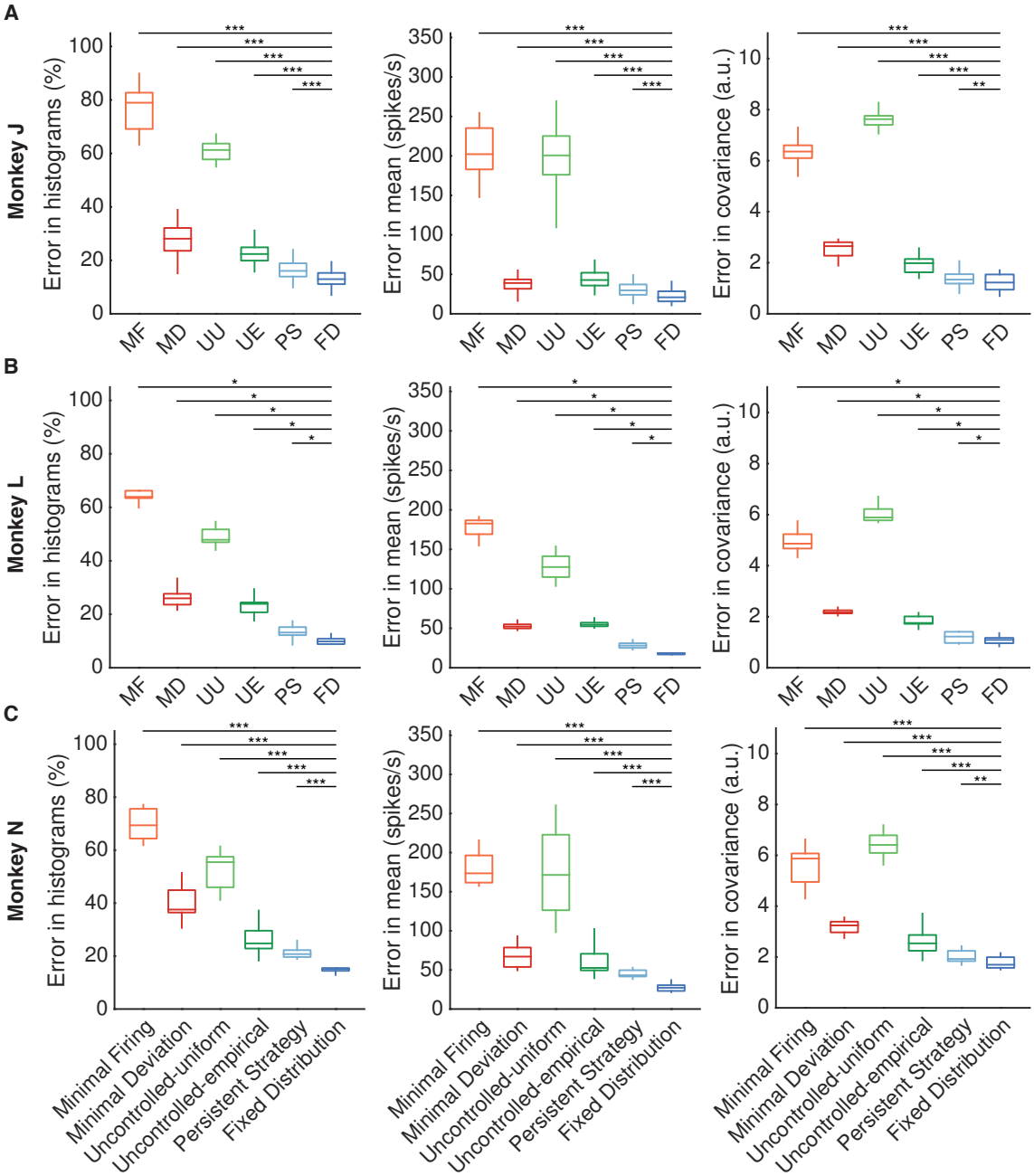


Figure A.2. Fixed Distribution hypothesis best predicts observed output-null activity for each animal. (A) Monkey J. (B) Monkey L. (C) Monkey N. Same conventions as Fig. 3.5. Asterisks denote a significance level of $\alpha = 0.05$ (*), 0.01 (**), and 0.001 (***).

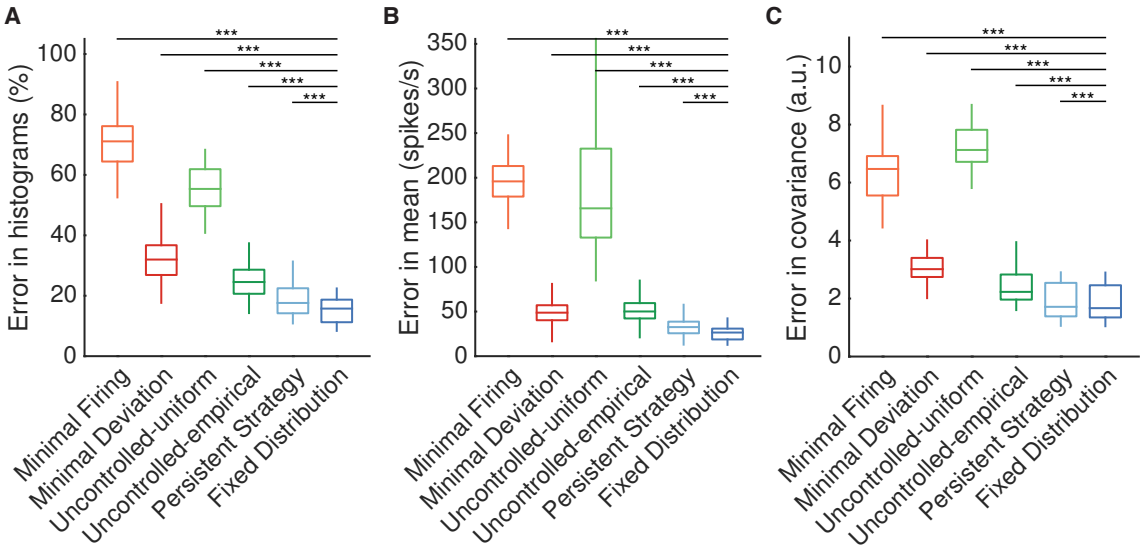


Figure A.3. Predicting output-null activity produced during the first mapping using activity observed during the second mapping. Throughout this work, we predict the output-null activity recorded while subjects used the second BCI mapping, and hypotheses can use activity recorded during use of the first BCI mapping to make their predictions. However, most of the hypotheses do not depend on the order in which the two mappings were presented to the subjects. Here we predict the output-null component of the activity recorded during use of the first BCI mapping, and hypotheses can use activity recorded under the second BCI mapping to make their predictions. Same conventions as Fig. 3.5.

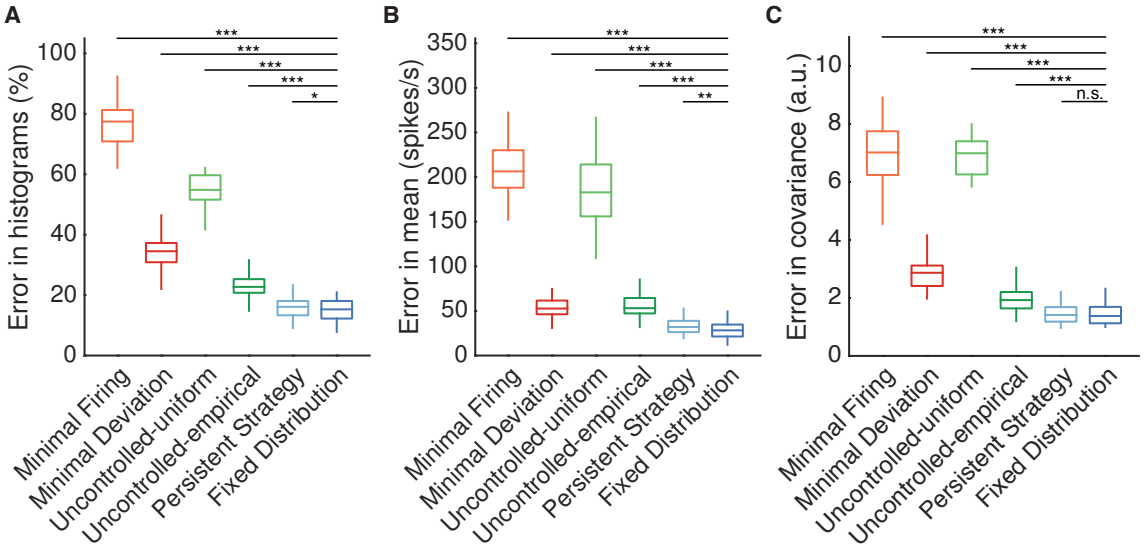


Figure A.4. Predicting output-null activity without using animal's internal model (IME) to define the output-null dimensions. Same conventions as Fig. 3.5. Asterisks denote a significance level of $\alpha = 0.05$ (*), 0.01 (**), and 0.001 (***). "n.s." is not significant.

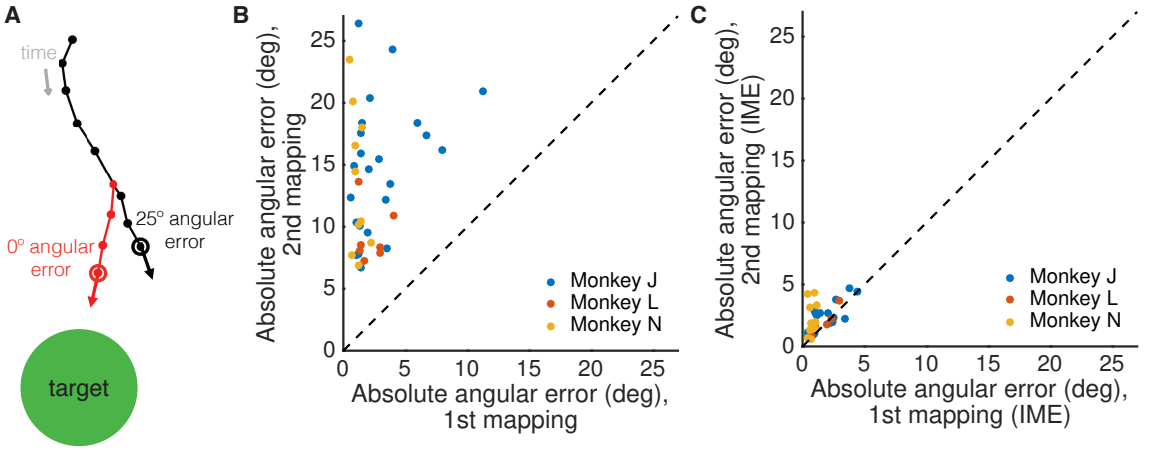


Figure A.5. Identifying the animal's internal model of the BCI mapping. (A) BCI cursor trajectory (black) for an example trial during the second mapping. At each time step, the subject takes its most recent visual feedback (where red whisker touches black trajectory) and propagates it forward in time (red whisker) using an internal model and neural activity produced in the recent past. This yields the subject's internal estimate of the current cursor position (red open circle), which is different from the actual BCI cursor position (black open circle). At this time step, the cursor movement according to the internal model (red arrow) points more directly to the target (green circle) than the actual BCI cursor movement (black arrow). Each dot indicates a 45 ms time step. (B) Average absolute angular cursor error (in degrees) for each session based on the actual BCI cursor movements (analogous to black arrow in (A)). Angular errors during stable control of the second mapping were larger than the angular errors during control of the first mapping. (C) Average absolute angular cursor error (in degrees) for each session based on IME (analogous to red arrow in (A)). When viewed through animals' internal estimates of the BCI, angular errors during control of the first and second mapping were similar.

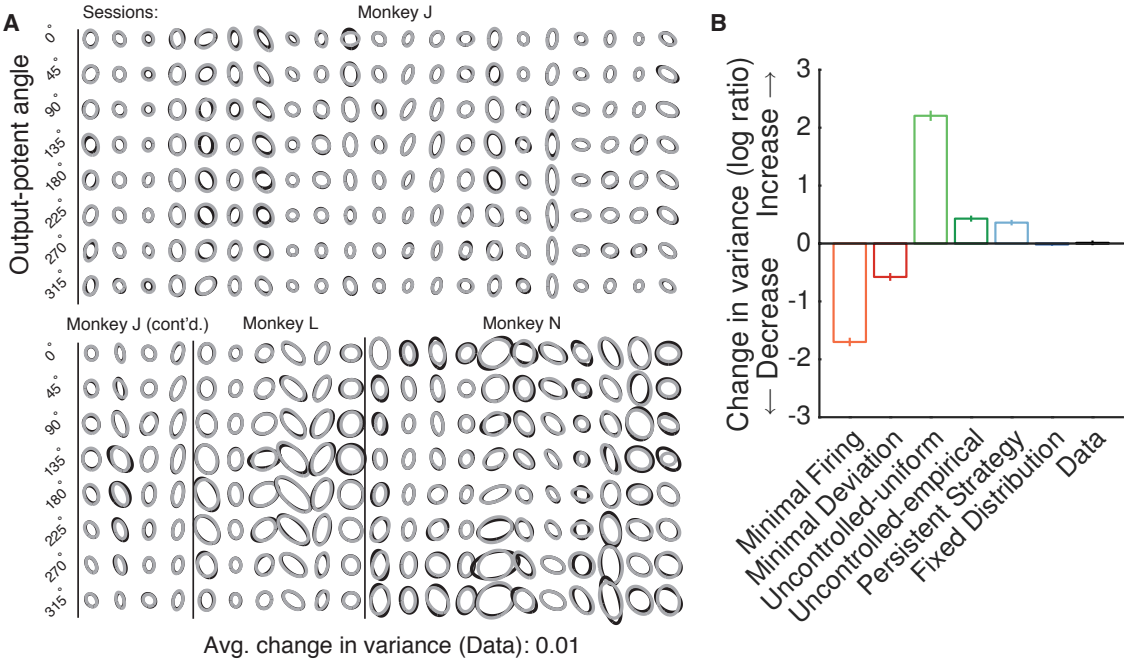


Figure A.6. Variance of neural activity in dimensions that became output-potent. We repeated the analyses shown in Fig. 3.6 on the predictions of output-null activity produced during the first mapping using activity observed during the second mapping (shown in Supplemental Fig. A.3). This analysis amounts to assessing the change in variance in dimensions that were output-null during the first mapping and output-potent during the second mapping. Same conventions as Fig. 3.6B-C.

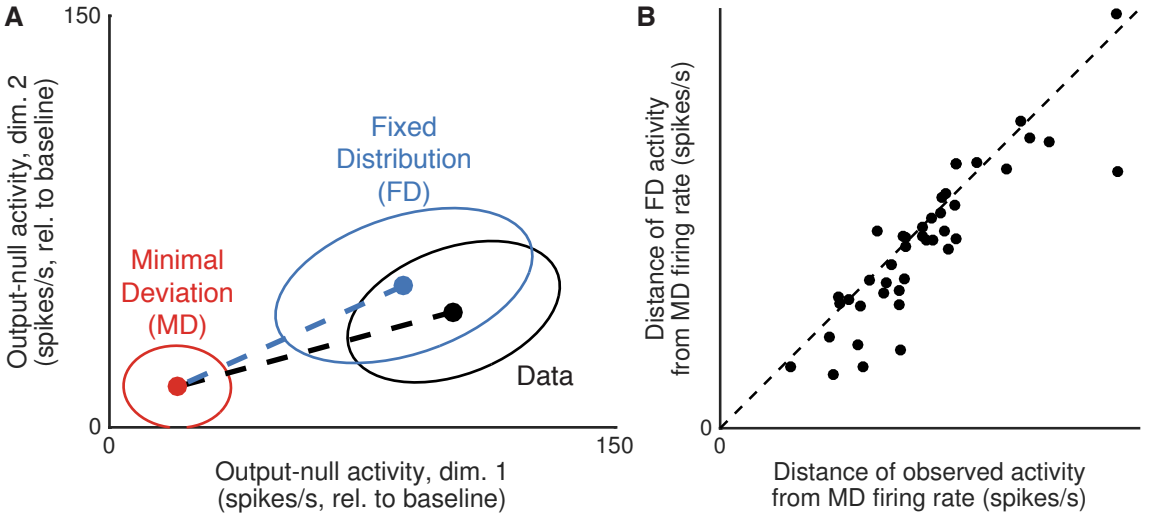


Figure A.7. Output-null activity was not closer to the mean predicted by Minimal Deviation than expected under Fixed Distribution. (A) Determining whether the observed mean output-null activity ("Data") was closer to the Minimal Deviation ("MD") prediction than expected under Fixed Distribution ("FD"). Dots and ellipses indicate the mean and covariance of output-null activity observed (black), predicted by Fixed Distribution (blue), and predicted by Minimal Deviation (red), for all time steps from session J20160714 corresponding to the same cursor movement, in the first two of eight output-null dimensions. Similar to Fig. 3.6, the observed covariance (black ellipse) is slightly smaller than that of Fixed Distribution (blue), suggesting the observed covariance is moving in the direction of the small covariance expected under Minimal Deviation (red). Do we see a similar trend in the mean activity, where the observed mean activity is closer to the Minimal Deviation mean than expected under Fixed Distribution? We can assess this by comparing the lengths of the dotted lines, which indicate the distances of the mean activity observed (black) and predicted by Fixed Distribution (blue) from the mean predicted by Minimal Deviation. (B) Distance of the observed and predicted output-null activity from the activity predicted by Minimal Deviation. Each dot indicates the average distance of the output-null activity observed (horizontal axis) and predicted by Fixed Distribution (vertical axis, 'FD') from the Minimal Deviation ('MD') in a session. For each cursor direction, the distance was computed as the ℓ_2 norm of the difference between the mean 8D output-null activity (predicted or observed) from the mean activity predicted by MD, similar to (A). Session distances were the average distances across all cursor directions. Most points lie below the diagonal, suggesting that the observed output-null activity was not closer to the MD predictions than expected under Fixed Distribution. Results were similar when comparing distances to the means predicted by Minimal Firing instead of Minimal Deviation.

Appendix

To understand how our results might change if we recorded from more neural units, we assessed the dimensionality and shared variance of population activity with a varying number of units (Williamson et al., 2016) (Fig. A.8). For each session, we fit factor analysis (FA) models (as defined in Eqn. A.1) to subsets of varying numbers of units and identified the number of factors needed to maximize the cross-validated data likelihood. This resulted in estimates of the model parameters L and Ψ. As in Williamson et al. (2016), dimensionality was defined as the number of eigenvectors of $LL^\top$ needed to explain 95% of the shared variance. Concretely, if the eigenvalues of $LL^\top$ are $\lambda_1, \lambda_2, ..., \lambda_D$, then d_{shared} is the smallest J such that $\left(\sum_{i=1}^{J} \lambda_i\right) / \left(\sum_{i=1}^{D} \lambda_i\right) \geq 0.95$. Note that the absolute dimensionality depends on the method (% shared variance threshold) and criterion (threshold = 95%) used for assessing dimensionality. This is the same method used in Williamson et al. (2016), but differs slightly from the method used in Sadtler et al. (2014). We found that the dimensionality of the population activity increased with the number of units (Fig. A.8A).

As in Williamson et al. (2016), we computed the percentage of each neural unit's activity variance that was shared with other recorded units (% shared variance). We calculated the average percent shared variance across neurons as follows:

$$\text{Percent shared variance for neuron } k = 100 \frac{L_k L_k^\top}{L_k L_k^\top + \Psi_k} \tag{A.17}$$

where L_k is the row of L corresponding to unit k. We found that the % shared variance initially increased with the number of units, then reached an asymptote, such that the % shared variance was similar with 30 and 85 units (Fig. A.8B).

The results in Fig. A.8A-B imply that the top ∼10 dimensions explain nearly all of the shared variance, and that additional dimensions identified by recording from more units explain only a small amount of additional shared variance. Thus, recording from more units beyond the ∼85 units that we recorded in these experiments is not likely to reveal additional dimensions with substantial shared variance.

We next measured the principal angles between modes identified using 30 units with those identified using 85 units (Fig. A.8C) (Björck and Golub, 1973). Modes were defined as the eigenvectors of the shared covariance matrices corresponding to units from the 30-unit set (i.e., the eigenvectors of $LL^\top$ where L includes only the rows corresponding to the same 30 units). To restrict the analysis to the number of modes used to estimate the intrinsic manifold, only the ten modes explaining the most shared variance were included in the principal angle calculations. The small principal angles between modes identified using 30 and 85 units indicate that the dominant modes remained largely unchanged when using more units, in agreement with Williamson et al. (2016). These modes define the intrinsic manifold, the space within which we perform all of our analyses in the current work. Thus, recording

from more units beyond the ∼85 units that we recorded in these experiments is not likely to substantially change the results reported in this work.

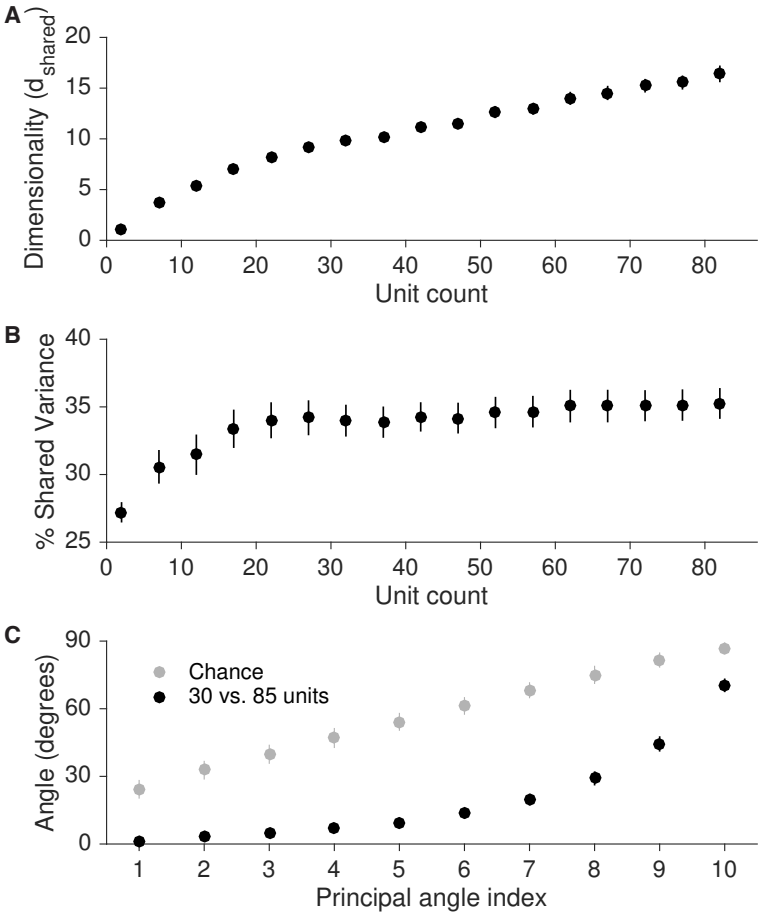


Figure A.8. Recording from more units is likely to reveal an intrinsic manifold similar to that identified in this study. (A) We assessed the dimensionality (d_{shared}) of population activity after applying factor analysis to varying numbers of units from each session. Dimensionality is defined as the number of factors needed to explain 95% of the shared variance. Dimensionality increased with the number of units. Error bars depict mean ± SE, across sessions. (B) We also computed the percentage of each neural unit's activity variance that was shared with other recorded units (% shared variance). The % shared variance is based on the same factor analysis models identified in (A). The % shared variance initially increased with the number of units, then reached an asymptote, such that the % shared variance was similar with 30 and 85 units. Error bars depict mean ± SE, across sessions. (C) We next measured the principal angles between the modes identified by factor analysis using 30 units with those identified using 85 units. Modes are defined as the eigenvectors of the shared covariance matrices corresponding to units from the 30-unit set. The small principal angles between modes identified using 30 and 85 units indicate that the dominant modes remained largely unchanged when using more units. Gray points represent principal angles between random 30-dimensional vectors. Error bars for black points depict mean ± SE, across sessions, while error bars for gray points depict mean ± s.d.

B Appendix to Chapter 3

B.1 Methods

Experimental details

Experimental methods are described in detail in previous work (Sadtler et al., 2014; Golub et al., 2018). Briefly, we recorded from the proximal arm region of primary motor cortex (M1) in three male rhesus macaques (Maccaca mulatta; age, monkey J: 7 years; monkey L: 8 years; monkey N: 7 years) using implanted 96 electrode arrays (Blackrock Microsystems). Data collection was performed using Labview (2012-2014) and Matlab (2011b-2015a). All animal care and handling procedures conformed to the NIH Guidelines for the Care and Use of Laboratory Animals and were approved by the University of Pittsburgh's Institutional Animal Care and Use Committee. We recorded from 85 to 94 neural units in each session. The activity of each neural unit is defined as the number of threshold crossings recorded by an electrode in non-overlapping 45 ms bins. The average firing rate of the neural units across sessions was 46 $\pm$ 7, 38 $\pm$ 8, and 56 $\pm$ 13 spikes/s (mean $\pm$ s.d.) for monkeys J, L, and N, respectively.

During each experimental session, a monkey performed an eight-target center-out task by modulating his recorded neural activity to control the velocity of a computer cursor on a screen. Each session involved two different BCI mappings. The first 'intuitive' mapping was chosen to provide the monkey with proficient control of the cursor. The animal used the intuitive mapping for 321 $\pm$ 96 trials (mean $\pm$ s.d.), after which the mapping was switched abruptly to a second, new BCI mapping that the monkey had never controlled before. This new mapping was chosen so as to be initially difficult for the monkey to use, and the monkey was given 698 $\pm$ 227 trials (mean $\pm$ s.d.) to learn the new mapping. Both BCI mappings were chosen so that they were controlled exclusively by the neural activity within the monkey's intrinsic manifold (defined below). During the BCI task, each animal's arms were loosely restrained. We monitored hand movements using an LED marker (PhaseSpace Inc.) on the hand contralateral to the recording array. During BCI control, animals showed little to no arm movements (Sadtler et al., 2014) (Supplemental Fig. B.2).

At the beginning of each trial, a cursor appeared in the center of the workspace, followed by the appearance of one of eight possible peripheral targets (chosen pseudorandomly among $\theta \in \{0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°\}$). For the first 300

ms of the trial, the velocity of the cursor was fixed at zero. After this, the velocity of the cursor was controlled by the animal through the BCI mapping. If the animal acquired the peripheral target with the cursor within 7.5 s, he received a water reward, and the next trial began 200 ms after target acquisition. Otherwise, the trial ended, and the animal was given a 1.5 s time-out before the start of the next trial.

During all experiments we monitored the monkey's pupil diameter (arbitrary units) using an infrared eye tracking system (EyeLink 1000; SR Research, Ottawa, Ontario). The eye tracker was first turned on while monkeys used the intuitive mapping, but this time varied from session to session. Pupil diameter was always measured while monkeys controlled the new BCI mapping.

Defining the BCI mappings

Each session began with the monkey performing a block of calibration trials, as described in previous work (Sadtler et al., 2014). Using these calibration trials, we z-scored the spike counts separately for each neural unit. We then applied factor analysis (FA) to the z-scored spike counts to identify the 10D linear subspace (i.e., the 'intrinsic manifold') that captured dominant patterns of co-modulation across neural units (Santhanam et al., 2009; Churchland et al., 2010; Harvey et al., 2012; Williamson et al., 2016; Athalye et al., 2017; Huang et al., 2019). We used 10 factors (i.e., dimensions), as this was the average dimensionality identified by cross-validation across experiments from monkeys J and L (Sadtler et al., 2014). In fact, in this paper we used exactly the same FA model and 10D latent factors that were found during the experiment to drive the BCI cursor (see below).

The factor activity, $\mathbf{z}_t \in \mathbb{R}^{10\times 1}$, was estimated as the posterior expectation given the z-scored spike counts, $\mathbf{y}_t \in \mathbb{R}^{q\times 1}$, where q is the number of neural units:

$$\mathbf{z}_t = L^\top (LL^\top + \Psi)^{-1}(\mathbf{y}_t - \mathbf{d}) \tag{B.1}$$

Above, L, Ψ and $\mathbf{d}$ are FA parameters estimated using the expectation-maximization algorithm, where L is termed the loading matrix, and Ψ is constrained to be a diagonal matrix. The factor activity, $\mathbf{z}_t$, can be interpreted as a weighted combination of the activity of different neural units. Prior to analysis, we orthonormalized $\mathbf{z}_t$ so that it had units of spike counts per time bin (Yu et al., 2009; Hennig et al., 2018). We refer to $\mathbf{z}_t$ as a "population activity pattern."

As discussed above, each experiment consisted of animals using two different BCI mappings. Each BCI mapping translated the resulting moment-by-moment factor activity ($\mathbf{z}_t$) into a 2D cursor velocity ($\mathbf{v}_t$) using a Kalman filter:

$$\mathbf{v}_t = A\mathbf{v}_{t-1} + M\mathbf{z}_t + \mathbf{c} \tag{B.2}$$

For the intuitive BCI mapping, $A \in \mathbb{R}^{2\times 2}$, $M = M_1 \in \mathbb{R}^{2\times 10}$, and $\mathbf{c} \in \mathbb{R}^{2\times 1}$ were computed from the Kalman filter parameters, estimated using the calibration

trials. For the second, new BCI mapping, we changed the relationship between population activity and cursor movement by randomly permuting the elements of $\mathbf{z}_t$ before applying Eqn. B.2. This permutation procedure can be formulated so that Eqn. B.2 still applies to the second BCI mapping, but for a new matrix $M_2 \in \mathbb{R}^{2\times 10}$ used in place of M_1 (Sadtler et al., 2014).

Hand control experiments

Data were collected from a fourth monkey (monkey G) for three sessions. During these experiments, the monkey performed an eight-target center-out task by moving his hand to control a computer cursor. An infrared marker was taped to the back of the monkey's hand and tracked optically using an Optotrak 3020 system. The marker position was used to update the position of the cursor in real-time on a stereoscopic computer monitor. During these experiments we recorded from the proximal arm region of primary motor cortex (M1) using an implanted 96 electrode array (Blackrock Microsystems).

Similar to the BCI control experiments, the targets shown on each trial were chosen pseudorandomly. At the beginning of each trial, a target (sphere; radius: 6 mm) was presented in the center of the reaching workspace. The animal was trained to move the cursor (sphere; radius: 6 mm) to this start target and hold for 0-100 ms. A peripheral target (sphere; radius: 6 mm) was presented at the end of this hold period. Water reward was delivered if the target was acquired within 1.5 s and the cursor was held on the target for a random hold period drawn uniformly from 150-550 ms. The next trial was initiated 200 ms after the trial ended, regardless of success or failure. The data analyzed includes 160 trials of baseline center-out trials, where the marker position was directly mapped to the cursor position, followed by 320 trials where a visuomotor rotation was applied to all reaches (40° CW, 40° CCW, and 30° CW for the three sessions, respectively).

To match the analysis procedure used in the BCI experiments, we took spike counts in non-overlapping 50 ms bins, and z-scored the spike counts using the mean and standard deviation of each neural unit during baseline reaches. We then applied FA to the z-scored spike counts recorded during all baseline reaches to identify a 12D linear subspace, where 12 was the number of dimensions that maximized the cross-validated log likelihood. We then orthonormalized the resulting 12D factor activity. All analyses of population activity considered only these top 12 factors.

Data analysis

Session and trial selection

The data analyzed in this study were part of a larger study involving learning two different types of BCI mapping changes: within-manifold perturbations (WMP), described above, and outside-manifold perturbations (OMP) (Sadtler et al., 2014). We found that animals learned WMPs better than OMPs, and so we only considered

WMP sessions in this study. We analyzed all sessions regardless of whether or not animals showed learning. For consistency, we excluded two sessions where the WMP was not the first perturbation shown during that experiment. In total, we analyzed 46 WMP sessions (monkey J: 25 sessions, monkey L: 10 sessions, monkey N: 11 sessions). No statistical method was used to pre-determine sample sizes but our sample sizes are similar to those reported in previous publications (e.g., (Oby et al., 2019; Golub et al., 2018; Hennig et al., 2018; Cowley et al., 2020)).

In the BCI experiments, spike counts were taken in non-overlapping 45 ms bins ("timesteps"), indexed here by $j = 1, ..., J$, where J is the number of timesteps in a given trial, and $j = 1$ is the timestep where the target first appeared. Each trial consisted of three intervals of interest: 1) the pre-target interval ($j \leq 2$, or 90 ms), during which the monkey had not yet perceived the target due to sensory processing delays; 2) the freeze interval ($j \leq 6$), during which the cursor was frozen in place at the center of the workspace; and 3) the cursor control interval ($j \geq 7$), where the cursor velocity was determined by Eqn. B.2. Unless otherwise noted, all analyses used data only during the cursor control interval.

We noted that when the cursor was near the target, or at the end of long trials, cursor movements were often idiosyncratic (e.g., reflecting small corrective movements), and so we discarded from our analyses any timesteps where the cursor was more than 65% of the way to the target, and any timesteps $j > 20$. To report trial-averaged quantities, we wanted to ensure that all neural activity within the same trial came from timesteps where the monkey attempted to push the cursor in the same direction. This was especially important given that we compared the time course of neural engagement during learning on a target-by-target basis (see Figs. 4.4, 4.5, and 4.6). We therefore analyzed only the timesteps where the angle between the cursor and target was within 22.5° of the target direction on that trial. Performing our analyses without this exclusion criterion did not change our results.

We analyzed both correct and incorrect trials in this study. We reasoned that sufficiently large increases in neural engagement (e.g., on the first trial using the new BCI mapping) may slow down the cursor's speed to the extent that the monkey is unable to obtain the target. Removing incorrect trials would then bias any analyses that compare levels of neural engagement between targets whose performance was improved versus impaired by neural engagement (see Figs. 4.4, 4.5, and 4.6). We did, however, remove trials where the monkey appeared to quit the task, by removing any sequence of more than 5 consecutive incorrect trials that occurred following at least one correct trial during Block 2. This occurred in only 11 of 46 sessions, and resulted in the removal of 0.71% of all trials. All results were qualitatively similar without this exclusion criterion.

For the hand control experiments, we analyzed data from the 15 timesteps of each trial immediately following the appearance of the target (which cued the monkey to begin moving his hand towards the target).

Quantifying behavioral performance

To quantify the monkey's moment-by-moment performance, we calculated the speed of the cursor in the target direction (i.e., "cursor progress" (Golub et al., 2018)). To do this, we first calculated the "single timestep" cursor velocity:

$$\mathbf{v}_j^{single-timestep} = M_2\mathbf{z}_j + \mathbf{c} \tag{B.3}$$

where $\mathbf{z}_j$ is the neural activity produced at timestep j, and M_2 and $\mathbf{c}$ are the new BCI mapping parameters (see Eqn. B.2). The cursor's speed towards the target, s_j, is then given by:

$$s_j = (\mathbf{v}_j^{single-timestep})^\top \mathbf{p}_j \tag{B.4}$$

where $\mathbf{p}_j$ is a unit vector pointing from the cursor position at timestep j to the target position. Assessing performance in this manner ensures that our measures of neural engagement and performance (as in Fig. 4.5) were both assessed using precisely the same neural activity, and at the same resolution (i.e., every timestep).

Changes in performance during Block 2 (Fig. 4.5e) were calculated as follows. Let $s_\theta(t)$ be the average of s_j for all timesteps j from the t^{th} trial to target θ. Let $s_\theta(0)$ be the expected speed to the target under the new BCI mapping before learning begins, given by projecting the monkey's trial-averaged neural activity to the same target during Block 1 into the new BCI mapping (i.e., using Eqn. B.3 and Eqn. B.4). For each t, Fig. 4.5e depicts $\Delta s_\theta(t) = s_\theta(t) - s_\theta(0)$ averaged across all targets in the same group (e.g., all $T+$ targets).

To find the trial at which performance for each target θ reached its peak (Fig. 4.5f), we first found the running mean of $\Delta s_\theta(t)$ in a sliding eight trial window. Let $\Delta\tilde{s}_\theta(t)$ be the resulting running mean. The trial at which performance for each target θ reached its peak was then $\arg\max_t \Delta\tilde{s}_\theta(t)$.

To test whether performance levels at the end of Block 2 differed between $T+$ and $T-$ targets, we used $\Delta\tilde{s}_\theta(\tilde{t}_\theta)$ as the performance level of target θ at the end of Block 2. To assess whether learning rates differed between $T+$ and $T-$ targets, for each θ we fit a saturating exponential to $s_\theta(t)$ with free parameter $\tau > 0$:

$$\widehat{s_\theta}(t) = s_\theta(1) + (\tilde{s}_\theta(\tilde{t}_\theta) - s_\theta(1))(1 - \exp(-(t-1)/\tau)) \tag{B.5}$$

where τ is the learning rate, governing how quickly $s_\theta(t)$ transitions from initial performance, $s_\theta(1)$ (unsmoothed because s changed more quickly early in learning), to performance at the end of Block 2, $\tilde{s}_\theta(\tilde{t}_\theta)$. For each target, τ was chosen so as to minimize the mean squared error between $\widehat{s_\theta}(t)$ and $s_\theta(t)$ for all t.

Identifying neural engagement axes

For each experimental session (for either BCI or hand control), we identified a set of *neural engagement axes*, capturing the dimension along which neural activity varied in the absence of learning pressure (i.e., while monkeys used the intuitive BCI mapping, or during baseline reaches, respectively) for each target $\theta \in \{0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°\}$. Let $\mathbf{a}_\theta \in \mathbb{R}^{10}$, with $\|\mathbf{a}_\theta\| = 1$ be the neural engagement axis for target θ. We defined $\mathbf{a}_\theta$ as the direction of greatest variance in the factor activity recorded during all trials to that target. Identifying this direction in the factor activity rather than in the spiking activity ensures that we focus on the shared covariance among neural units rather than variance that is independent to each unit.

We observed that the neural engagement axes involved the activity of nearly all neural units changing in the same direction (Supplemental Fig. B.4). We therefore choose the sign of $\mathbf{a}_\theta$ so that positive values of neural engagement corresponded to increases in the firing rate for the majority of units. This allowed us to average across values of neural engagement across targets and sessions, as presented in the main text.

Quantifying neural engagement

During Block 1, we defined the value of *neural engagement*, $e_j \in \mathbb{R}$, for each timestep j to target θ, as follows:

$$e_j = (\mathbf{z}_j - \bar{\mathbf{z}}_\theta)^\top \mathbf{a}_\theta \tag{B.6}$$

where $\bar{\mathbf{z}}_\theta$ is the mean neural activity produced to target θ during Block 1. The level of neural engagement on trial t was then defined as the average of e_j for all timesteps j from trial t.

To compute neural engagement during Block 2, we cannot simply use Eqn. B.6, because some of the changes in neural activity across trials will also be due to learning (e.g., by neural reassociation (Golub et al., 2018)). According to neural reassociation (Golub et al., 2018), to move the cursor in a particular direction $\theta \in [0, 2\pi)$ during Block 2, the monkey samples the neural population activity he used for movements in a potentially *different* direction $\theta' \in [0, 2\pi)$ during Block 1. Thus, to estimate neural engagement during Block 2 (as shown in Fig. 4.5 and Fig. 4.6), we used the following:

$$e_j = (\mathbf{z}_j - \bar{\mathbf{z}}_{\theta'_j})^\top \mathbf{a}_{\theta'_j} \tag{B.7}$$

where θ'_j is no longer necessarily equal to the target direction, θ. We estimated θ'_j from the neural activity, $\mathbf{z}_j$, which is reasonable provided that changes in neural activity

due to θ_j and e_j are not entirely overlapping. Specifically, θ'_j was defined as the direction that the cursor would have moved if $\mathbf{z}_j$ were produced under the intuitive mapping, as changes in neural engagement tended to have less effect on the cursor's movement direction using the intuitive mapping. This procedure allowed our estimate of θ'_j to vary as the monkey learned to control the new BCI mapping, thus accounting for any changes in neural activity due to neural reassociation. To compute $\bar{\mathbf{z}}_{\theta'_j}$ and $\mathbf{a}_{\theta'_j}$ for any continuous value of $\theta'_t \in [0, 2\pi)$, we used a cubic spline to interpolate between the values measured for each $\theta \in \{0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°\}$.

In the above procedure, the neural engagement axes corresponding to a given θ are assumed to be the same during both Block 1 and Block 2. We confirmed that the neural engagement axes estimated before learning (during Block 1) and after learning (at the end of Block 2) were similar (Supplemental Fig. B.10), indicating that the largest fluctuations in neural activity occurred along similar dimensions throughout the experiment.

To compare values of neural engagement across trials to different targets in Fig. 4.2 and Supplemental Fig. B.1, we z-scored the neural engagement for each target separately using the mean and standard deviation across all trials to each target during Block 1. In all other figures, neural engagement was z-scored using the mean of the last 10 trials to each target during Block 1. This was done so that the level of neural engagement on the first trial of Block 2 was relative to the average level at the end of Block 1 (e.g., see Fig. 4.4c and Fig. 4.5c).

Comparing neural engagement to pupil size. For each session, we estimated the trial-by-trial correlation between neural engagement and the monkey's pupil size (Fig. 4.2f-g). To do this, we first found the average pupil size and neural engagement during the control interval of each trial. Similar to previous work correlating population activity and pupil size (Cowley et al., 2020), we applied boxcar smoothing to the trial-averaged measurements of each quantity with a sliding window of 30 trials, and then computed the Pearson's correlation between the resulting time series. Pupil recordings began partway into Block 1 and continued throughout Block 2 (see *Experimental details* above). Thus, the analysis shown in Fig. 4.2g used trials throughout Block 2, for all sessions where Block 2 consisted of at least 200 trials (44 of 46 sessions). During the 13 sessions where pupil size was measured during Block 1 for at least 200 trials (all sessions were from monkey J), the median correlation between neural engagement and pupil size was $\rho = 0.67$ (bootstrapped 95% C.I. $[0.41, 0.79]$, $n = 13$ sessions).

Variance explained by changes in neural engagement. We sought to estimate the amount of variance in the neural population activity due to changes in neural engagement during Block 1 (Fig. 4.2h). To estimate the variance for trials to a given target, we first found the average neural activity $\mathbf{z}_t$ during each trial to that target, along with the corresponding neural engagement, e_t. The measure of the variance explained by changes in engagement for that target was then $\frac{\mathrm{Var}_t(e_t)}{\mathrm{Tr}\ \mathrm{Cov}_t(\mathbf{z}_t)}$. To compute the *total* amount of variance explained by changes in engagement, we computed the same metric above, but used the activity from all trials combined

rather than just the trials to a particular target.

Predicting the impact of neural engagement on performance under the new mapping. We labeled a target as $T+$ or $T-$ based on whether an initial increase in neural engagement would increase ($T+$) or decrease ($T-$) the cursor's speed towards the target under the new mapping (e.g., Figs. 4.4, 4.5, and 4.6). Specifically, let $\bar{\mathbf{z}}_\theta$ be the average neural activity recorded during Block 1 for target θ, and let $\mathbf{a}_\theta$ be the corresponding neural engagement axis. Then we labeled that target as $T+$ if an increase in neural engagement led to an increase in the cursor's speed to the target (see Eqn. B.4):

$$(M_2(\bar{\mathbf{z}}_\theta + \epsilon\mathbf{a}_\theta) + \mathbf{c})^\top \mathbf{p} > (M_2\bar{\mathbf{z}}_\theta + \mathbf{c})^\top \mathbf{p} \tag{B.8}$$

where M_2 and $\mathbf{c}$ are the parameters of the new BCI mapping, $\mathbf{p}$ is a unit vector corresponding to the target direction, and $\epsilon > 0$ is a small constant. Otherwise, the target was labeled as $T-$.

Identifying output-potent and output-null engagement axes. Given a neural engagement axis, $\mathbf{a} \in \mathbb{R}^{10}$, not all changes in neural activity along this axis will lead to changes in cursor velocity through the new BCI mapping, M_2. This is because the mapping between neural activity and cursor velocity, given by Eqn. B.2, is a linear mapping from 10D to 2D, implying that M_2 has a non-trivial null space, $Nul(M_2)$. To identify which components of $\mathbf{a}$ will result in changes in cursor velocity, we can find bases for the null space, $Nul(M_2)$, and the row (or potent) space, $Row(M_2)$ (Hennig et al., 2018). To do so, we took a singular value decomposition of $M_2 = USV^T$, with $U \in \mathbb{R}^{2\times2}$, $S \in \mathbb{R}^{2\times10}$, and $V \in \mathbb{R}^{10\times10}$, where the columns of S were ordered so that only the first two columns had non-zero elements. Then, we let $R \in \mathbb{R}^{10\times2}$ be the first two columns of V, and $N \in \mathbb{R}^{10\times8}$ be the remaining eight columns. The columns of N and R are mutually orthonormal and together form an orthonormal basis for the 10-dimensional space of factor activity. This allows us to rewrite the neural engagement axis for each target θ as the sum of a null-engagement axis, $\mathbf{a}_\theta^{null}$, and a potent-engagement axis, $\mathbf{a}_\theta^{potent}$:

$$\mathbf{a}_\theta = \mathbf{a}_\theta^{null} + \mathbf{a}_\theta^{potent} \tag{B.9}$$

$$\mathbf{a}_\theta^{null} = \mathbf{a}_\theta NN^\top \tag{B.10}$$

$$\mathbf{a}_\theta^{potent} = \mathbf{a}_\theta RR^\top \tag{B.11}$$

We then normalized $\mathbf{a}_\theta^{null}$ and $\mathbf{a}_\theta^{potent}$ to be unit vectors. We then used these axes in Eqn. B.7 to compute values of null and potent engagement (Fig. 4.6).

Statistics & Research Design

Data collection and analyses were not performed blind to the conditions of the experiments. Experiments were not grouped, and thus no group randomization procedures were performed. Statistical analyses were conducted in Matlab (2015a). To test

whether the distributions of neural engagement were different for $T+$ and $T-$ targets at the start of Block 2, we used a two-sample Kolmogorov-Smirnov test. To compare the medians of two distributions we used a two-sided Wilcoxon rank-sum test (unpaired) or a two-sided sign test (paired). Statistical tests were non-parametric and so did not assume normality. When depicting standard error (SE), data distributions were assumed to be normal but this was not formally tested. For additional information on statistics and research design, see the Life Sciences Reporting Summary.

Code availability

The code used in this study for performing analyses and generating figures can be found at `https://github.com/mobeets/neural-engagement`.

B.2 Supplemental Figures

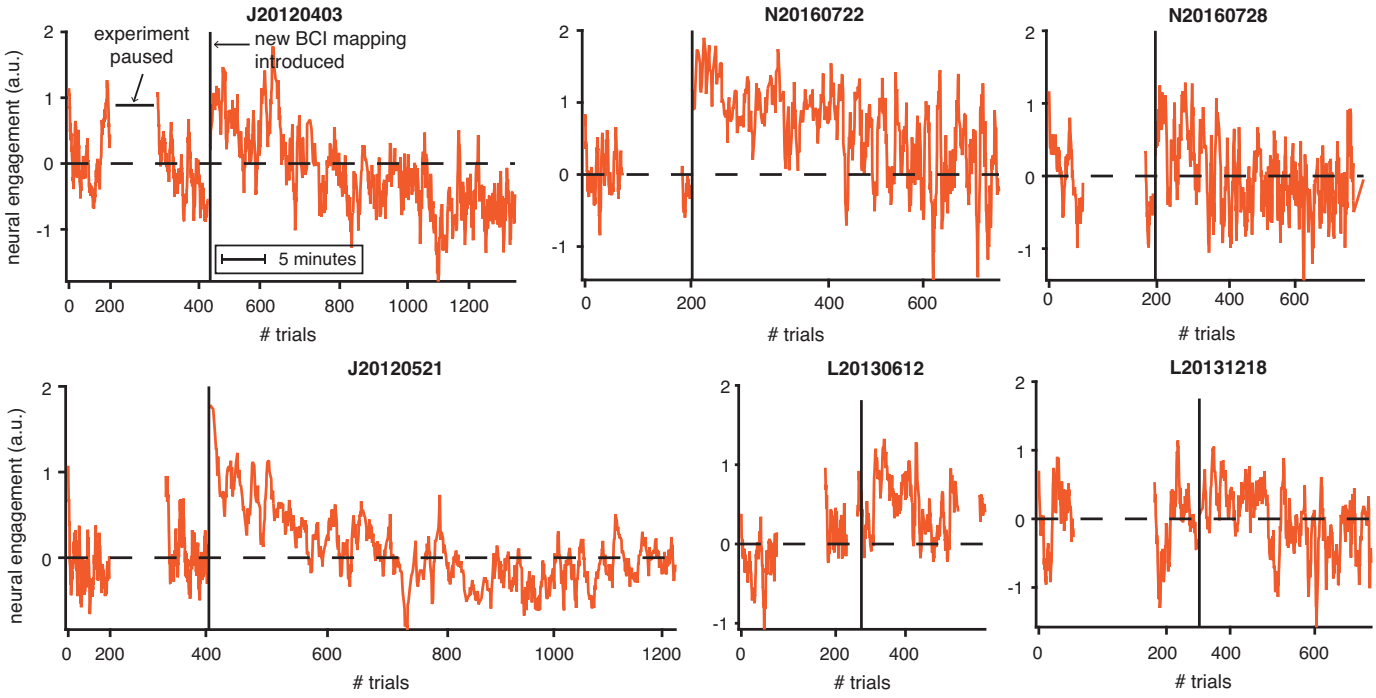


Figure B.1. Neural engagement showed stereotyped changes relative to experimental events, in multiple example sessions. Same conventions as Fig. 4.2d. Note that in contrast to other figures (e.g., Fig. 4.5c), here neural engagement is shown across trials to all eight instructed targets, where trials to different targets were interleaved. As a result, each time course shown here includes variability due to the target-specific differences in neural engagement during learning (e.g., see Fig. 4.5c). Position along the horizontal axis indicates clock time (see legend indicating "5 minutes"), so that pauses in the experiment are more visible. All sessions are plotted with the same time scale, and trial indices are marked for reference.

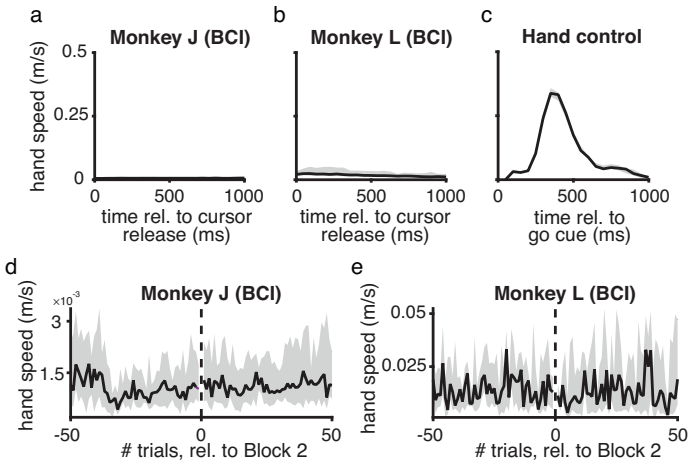


Figure B.2. Changes in neural engagement during BCI control could not be explained by hand movements. a-c. During the BCI experiments, we recorded the hand speed of two animals (monkey J, shown in panel **a**; and monkey L, shown in panel **b**), for the hand contralateral to the recording array (the other hand was restrained). Monkey N's hand speed was not recorded because his hand was restrained in a tube, and the reflection of the light on the tube made his hand difficult to track. We also recorded the hand speed of monkey G (shown in panel **c**), who performed a center-out arm reaching task (as shown in Fig. 4.2i-j). This allowed us to compare hand speeds across both types of experiments. We found that the arm movements during the BCI task (panels **a** and **b**) were substantially smaller than during the center-out arm reaching task. Black line indicates median across trials to all sessions, while shading indicates median ± 25th percentile (**a**, $n = 25$ sessions; **b**, $n = 10$ sessions; **c**, $n = 3$ sessions). **d-e.** Even if animals showed little to no arm movements (as shown in panels **a** and **b**), might it be the case that the increase in neural engagement at the start of Block 2 (Fig. 4.4c) can be explained by animals moving their hands more than they did on previous trials? We found no substantial increase in hand speed at the start of Block 2 for either monkey. Black line indicates median across sessions, while shading indicates median ± 25th percentile (**d**, $n = 25$ sessions; **e**, $n = 10$ sessions). Thus, the increase in neural engagement we observe at the start of Block 2 cannot be explained by animals suddenly moving their hands more than during Block 1.

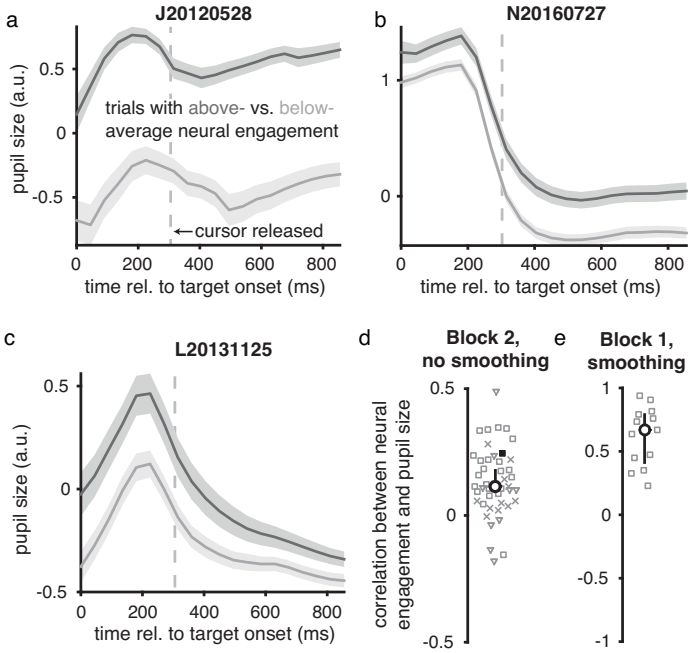


Figure B.3. Trials with elevated levels of neural engagement also showed increased pupil size. a-c. In Fig. 4.2g, we related neural engagement and pupil size by first averaging the pupil size across time points within a trial. To further explore this relationship, here we consider the time course of pupil size within a trial. Trial-averaged pupil sizes are shown for three example sessions after grouping trials separately based on whether neural engagement during the control interval of each trial during Block 2 was above- (dark gray) or below- (light gray) the median across trials during Block 2. Vertical dashed line indicates the time within each trial when the cursor was released (300 ms; see Methods), i.e., the beginning of the control interval. Shading indicates mean ± SE across trials (**a**, $n = 456$ trials; **b**, $n = 296$ trials; **c**, $n = 202$ trials). Within each example session, the time course of pupil size was similar for trials with above- versus below-average levels of neural engagement, but with a larger overall pupil size on trials with above-average neural engagement. **d.** Prior to computing the correlations between neural engagement and pupil size shown in Fig. 4.2g (and in the previous panel), we first smoothed the trial-by-trial time courses of pupil size and neural engagement with a 30-trial boxcar filter, similar to previous work correlating population activity and pupil size (Cowley et al., 2020). Here we show that neural engagement and pupil size were typically positively correlated even without smoothing. Without smoothing, the median Pearson's correlation across sessions was $\rho = 0.12$ (bootstrapped 95% C.I. $[0.09, 0.18]$, $n = 44$ sessions). Same conventions as Fig. 4.2g. **e.** Although the recording of pupil size began part way into Block 1 due to experimental constraints, we computed the trial-by-trial correlation between pupil size and neural engagement during Block 1 for the 13 sessions with a sufficient number of trials (all from monkey J). The median Pearson's correlation during these sessions was $\rho = 0.67$ (bootstrapped 95% C.I. $[0.41, 0.79]$, $n = 13$ sessions). Thus, a positive correlation between neural engagement and pupil size was also present before learning. Same conventions as Fig. 4.2g.

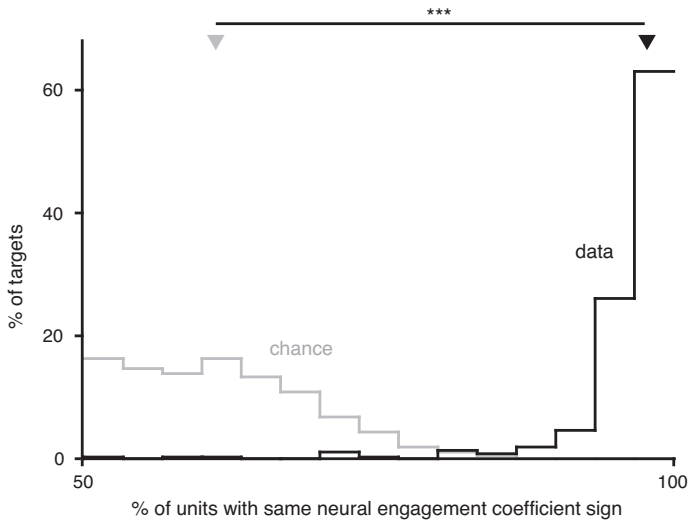


Figure B.4. Changes in neural engagement corresponded to nearly all neural units increasing or decreasing their activity together. We wanted to understand how changes in neural engagement were represented by the activity of individual units. For each target, a neural engagement axis was defined in 10-dimensional factor space. We used the $q \times 10$ loading matrix from factor analysis (see Methods) to define the neural engagement axis in the q-dimensional population activity space of the q recorded units. For example, if there were 90 units, the neural engagement axis would have 90 coefficients, describing how changes in neural engagement for a given target would be represented by the activity of each of the 90 units. For each target, we computed the percentage of units whose coefficients had the same sign (for whichever sign was in the majority, so that percentages could never be below 50%). Shown in black is the distribution of these percentages across the neural engagement axes for all targets across all sessions (bootstrapped 95% C.I. [97.6%, 97.7%], $n = 368$ axes (one per target)). This relationship means that an increase in neural engagement corresponds to an increase in the firing rate of most units (by an amount that is unit- and target-dependent). For reference, in gray, is the distribution after sampling random dimensions in factor space, and computing the corresponding effects on individual neural units (bootstrapped 95% C.I. [59.7%, 62.5%], $n = 368$ random axes). Triangles depict the medians of the "data" and "chance" distributions, which were significantly different ($p < 10^{-10}$, paired, two-sided sign test, $n = 368$ axes).

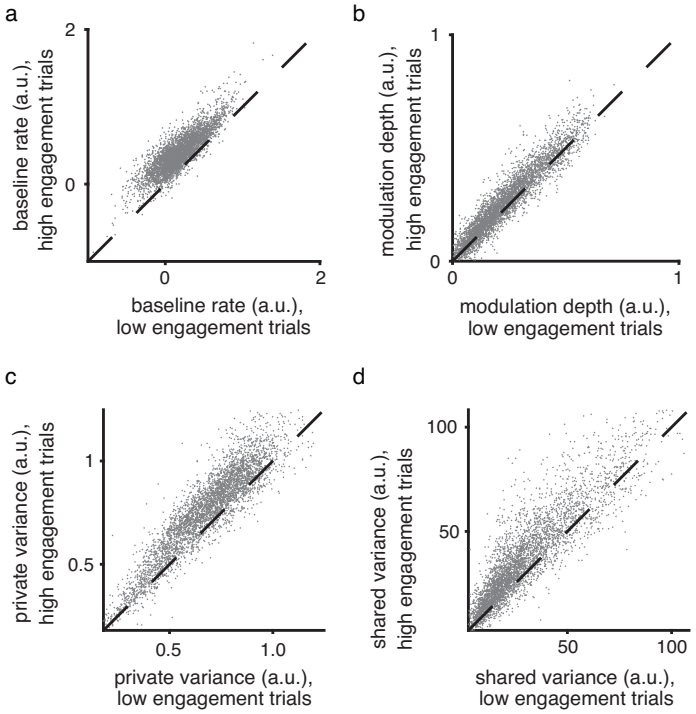


Figure B.5. Increased neural engagement corresponded with increased baseline firing rate, modulation depth, and spiking variance in single units. To understand the relationship between neural engagement and the firing properties of individual units, for each experiment we grouped trials during Block 1 based on whether they had above- versus below-average levels of neural engagement (similar to Supplemental Fig. B.3). **A-B.** For each individual unit from all sessions, we fit two cosine tuning models to each unit's z-scored spike counts: one model was fit to the average spike counts on trials with above-average levels of neural engagement ("high engagement trials"), while the other model fit to the spike counts on trials with below-average levels of neural engagement ("low engagement trials"). The cosine model was of the form $y = b + m\cos(\theta - \theta_{pref})$, where y is the unit's expected firing rate on a trial to target θ, b is the unit's baseline (average) firing rate, m is the unit's modulation depth, and θ_{pref} is the unit's preferred direction. We estimated b, m, and θ_{pref} using linear regression. Each dot corresponds to one unit. For most units, both the baseline firing rate (b; panel A) and the modulation depth (m; panel B) were higher on high engagement trials than on low engagement trials (in both cases: $p < 10^{-10}$, paired, two-sided sign test, $n = 4074$ units). **C-D.** For each session, we fit a factor analysis (FA) model to the z-scored spike counts of all units during low engagement trials, and then fit a separate FA model to the z-scored spike counts during high engagement trials. Each model had the same form as Eqn. B.1, resulting in parameter estimates of L and Ψ. The estimated private variance of unit i is given by Ψ_{ii}, while the shared variance is given by $(LL^\top)_{ii}$, where the ii subscript indicates the i^{th} diagonal element. Each dot corresponds to one unit. We found that both the private variance (panel C) and shared variance (panel D) of most units was higher on high engagement trials than on low engagement trials (in both cases: $p < 10^{-10}$, paired, two-sided sign test, $n = 4074$ units). This result is expected from Supplemental Fig. B.4 because the sum of a unit's shared and private variances equals its spike count variance. Because a unit's spike count variance tends to increase with its mean spike count, a higher firing rate will typically correspond with a higher shared and/or private variance.

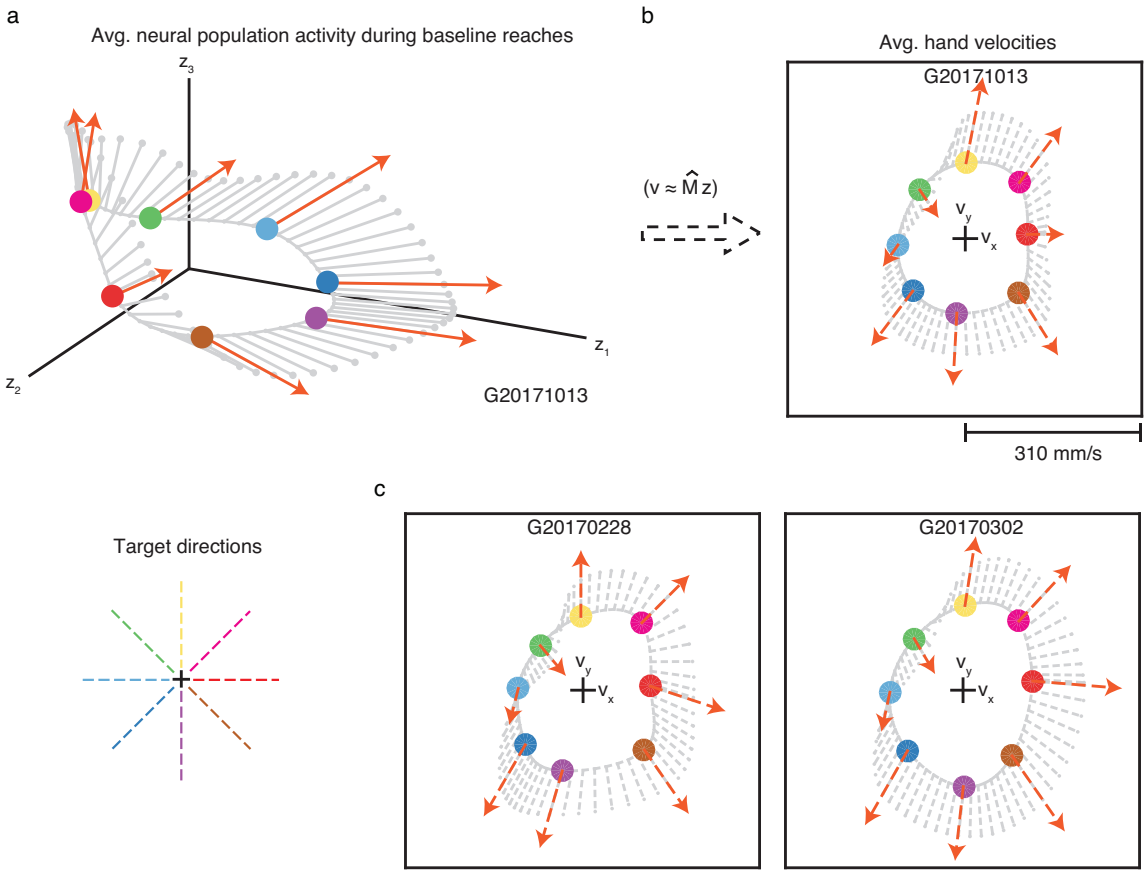


Figure B.6. Increased neural engagement during arm movements predicted faster hand speeds towards most targets. a. For the experiments involving arm movements (see Methods), we visualized the average neural population activity (circles) and neural engagement axes (orange arrows) during baseline reaches to each of eight targets. Same conventions as Fig. 4.3d. "Target directions" panel is a legend depicting the color corresponding to each target direction. **b.** We also visualized the monkey's average hand velocity during reaches to each target (circles), similar to Fig. 4.3e. Unlike during BCI control, we do not know the causal relationship between neural population activity and hand velocity. To understand how changes in neural engagement related to hand velocity, we used linear regression to predict the monkey's hand velocity during baseline reaches at each 50 ms timestep during the movement epoch of every trial, using the neural population activity recorded 100 ms prior. Cross-validated r^2 for the x- and y- components of hand velocity were 67% and 77%, respectively. The linear regression model ($\widehat{M}$) allowed us to estimate how increases in the neural engagement related to the monkey's average hand velocity towards each target (orange dashed arrows), and to intermediate target directions (gray dashed lines). In this session, an increase in neural engagement predicted an increase in the monkey's hand speed towards all but the 135° target. This suggests that differences in the neural engagement axes across targets may have behavioral relevance. **c.** We repeated the procedure from **b** during the other two arm movement sessions. Across sessions, increased neural engagement during arm movements predicted faster hand speeds towards most targets.

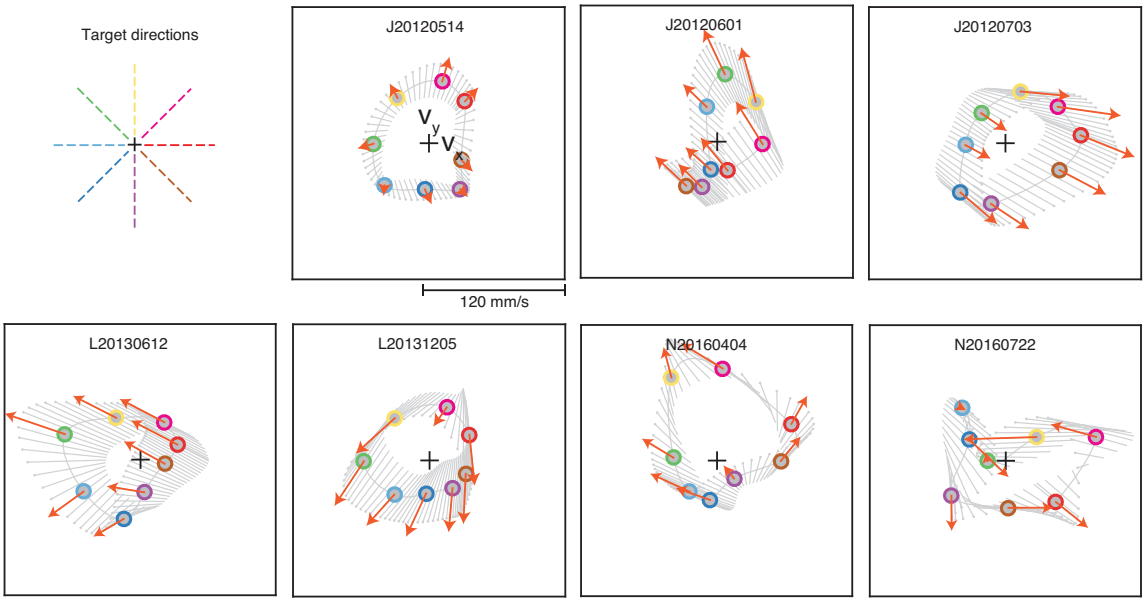


Figure B.7. New BCI mappings induced a variety of relationships between neural engagement and cursor velocity, across targets and sessions. Same conventions as Fig. 4.3f, for multiple example sessions (all with the same scale). "Target directions" panel is a legend depicting the color corresponding to each target direction.

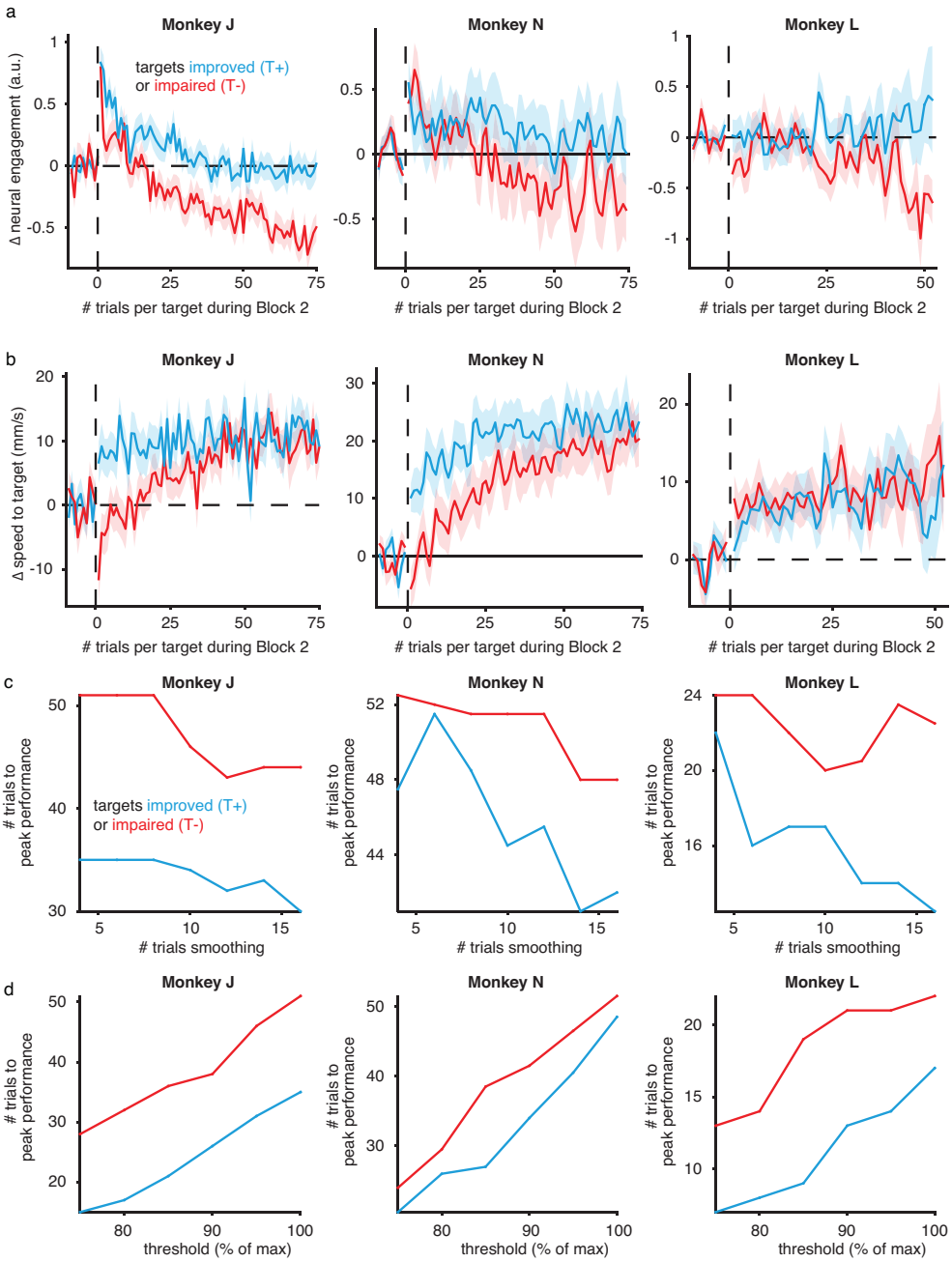


Figure B.8. Changes in neural engagement and performance per monkey. a-b. Changes in neural engagement (**a**) and performance (**b**) during Block 2, averaged across $T+$ and $T-$ targets for each monkey separately (J: $n_{T+} = 119, n_{T-} = 81$ targets; N: $n_{T+} = 50, n_{T-} = 38$; L: $n_{T+} = 51, n_{T-} = 29$). Same conventions as Fig. 4.5c (**a**) and Fig. 4.5e (**b**). **c-d.** Difference in learning speed between $T+$ and $T-$ targets is robust to amount of smoothing (**c**) and how the peak performance was determined (**d**). We found the number of trials at which performance for a given target reached x% of its maximum, after first smoothing the performance for each target with a k-trial boxcar filter (see Methods), where in Fig. 4.5f, $k = 8$ and $x = 100$. Here we sweep the amount of smoothing (k; panel **c**) while holding $x = 100$ constant, and sweep the threshold percentage (x; panel **d**) while holding $k = 8$ constant. Across all monkeys, the blue line was always below the red line, indicating that our result that $T+$ targets reached peak performance more quickly than $T-$ targets was robust to different parameter settings.

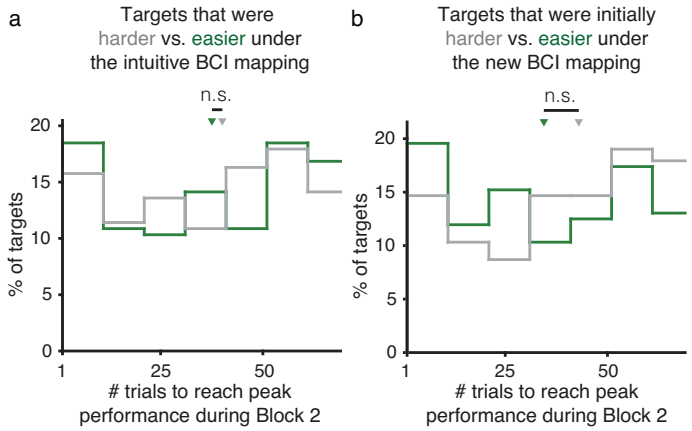


Figure B.9. Non-uniform task performance did not predict how quickly different targets reached peak performance. a-b. In BCI tasks, performance across targets is often non-uniform. Can differences in the animal's pre-learning performance across targets predict how quickly different targets were learned? We defined pre-learning performance in two ways: using the performance under the intuitive BCI mapping during Block 1 (panel **a**), and using the predicted initial performance under the new BCI mapping (panel **b**). For the latter quantity, we projected the trial-averaged neural activity for each target during Block 1 onto the new BCI mapping. For each definition of pre-learning performance, we divided all targets from each monkey into two groups based on whether pre-learning performance was above (green, "easier") or below (gray, "harder") the median performance level across all targets. We then found the number of trials needed for each group of targets to reach peak performance during Block 2, similar to Fig. 4.5f. The median number of trials needed to reach peak performance was not different for targets that were initially harder (gray triangle) versus easier (green triangle) during Block 1 using the intuitive BCI mapping ($p = 0.91$, two-sided Wilcoxon rank-sum test, $n_1 = 184$ and $n_2 = 184$ targets; panel **a**). Nor was there a difference in the median number of trials needed to reach peak performance for the targets that were predicted to be initially harder (gray triangle) versus easier (green triangle) under the new BCI mapping ($p = 0.06$, two-sided Wilcoxon rank-sum test, $n_1 = 184$ and $n_2 = 184$ targets; panel **b**).

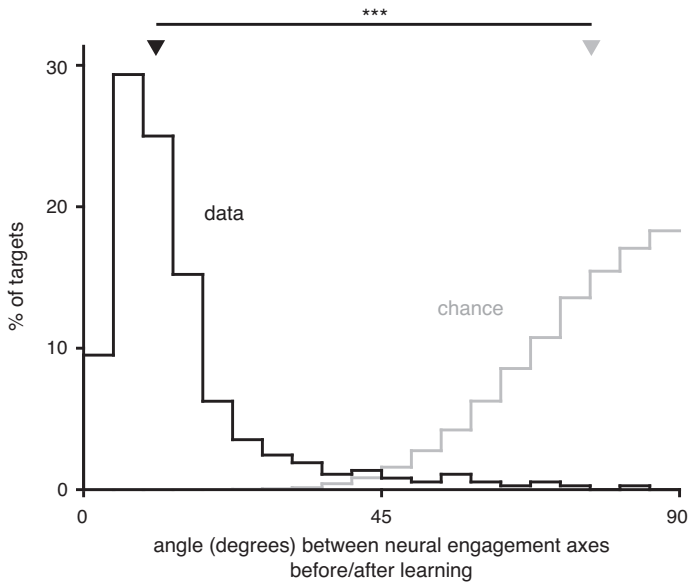


Figure B.10. Neural engagement axes were largely unchanged after learning. Distribution of the angle ("data", in black) between the neural engagement axis identified for each target during Block 1 ("before learning") vs. during the last 50 trials of Block 2 ("after learning"). To identify neural engagement axes during the last 50 trials of Block 2, we used the same procedure as used during Block 1 (i.e., the procedure used in the main text; see Methods), but applied to the last 50 trials of Block 2. "Chance" (in gray) indicates the distribution of the angle between random directions in ten-dimensional space. Triangles depict the medians of the "data" and "chance" distributions, which were significantly different ($p < 10^{-10}$, two-sided Wilcoxon rank-sum test, $n_1 = 368$ (data) and $n_2 = 50,000$ (chance) axes).

B.3 Supplemental Discussion

Here we discuss potential sources of neural engagement in more detail.

Is it intended speed? Neural engagement may be related to, but is likely distinct from, the monkey's intended movement speed. Neurons in M1 have long been known to reflect movement speed (Georgopoulos et al., 1986; Moran and Schwartz, 1999). We observed that during arm movements, increased neural engagement predicted increased hand speed towards the target (Supplemental Fig. B.6). This raises the possibility that neural engagement may simply reflect the monkey's intended movement speed. However, during BCI learning, we observed a gradual decrease in neural engagement during repeated trials to the same target, despite the fact that performance for many targets would have been improved by maintaining this increased neural engagement (Fig. 4.5c). Therefore, if neural engagement simply reflected intended movement speed, it would be necessary to explain why monkeys would intend to move slower when doing so would reduce their reward rate. One possible explanation might be that the monkey's intended movement speed is modulated by an internal state such as motivation or reward expectation. In fact, studies of "movement vigor," measured behaviorally as the reaction time and/or peak movement speed during eye or reaching movements (Mazzoni et al., 2007; Xu-Wilson et al., 2009; Dudman and Krakauer, 2016; Yttri and Dudman, 2018; Sedaghat-Nejad et al., 2019; Shadmehr et al., 2019), have found that the vigor (or speed) with which we execute a movement is not constant over time, but varies depending on context. Movement vigor is therefore thought to reflect a cost-benefit analysis, such that vigor increases when there is a higher subjective utility (e.g., expected reward) for doing so (Shadmehr et al., 2019). Consistent with this prediction, neural engagement was higher at the start of the experiment, and following pauses in the experiment (Fig. 4.2e); in both cases, the resumption of the experiment indicates to the monkey a higher expectation of reward, because completing trials resulted in a reward. However, we also saw an increase in neural engagement following the introduction of the new BCI mapping, a time when the monkey's reward expectation should be lower, given that the new BCI mapping will immediately decrease his reward rate. Thus, increases in neural engagement do not always reflect increased reward expectation, suggesting that neural engagement may not simply reflect movement vigor.

Is it a feedback response? Previous work has established that M1 population activity reflects sensory feedback following a perturbation, for both mechanical (Pruszynski et al., 2011, 2014; Omrani et al., 2014, 2016) and purely visual (Stavisky et al., 2017) perturbations. At first glance, these results may appear similar to our observation of an immediate increase in neural engagement following the introduction of a new BCI mapping. However, our results differ in two key ways. First, while we did find a fast increase in neural engagement (within a single trial), neural engagement then decreased gradually over subsequent trials (Fig. 4.2e). It is not known from these previous studies whether the magnitude of the sensory feedback signal should decay over subsequent trials, but we would expect it to decay more quickly than did neural engagement. Second, neural engagement followed a similar time course during

the portion of each trial before cursor feedback was available (Fig. 4.2e, last subpanel), indicating that this signal was not directly reflecting visual feedback. Thus, neural engagement does not simply reflect sensory feedback. Another line of work has observed that, in the presence of novel dynamics between intended and actual movement, subjects stiffen their limbs using muscle co-contraction (Osu et al., 2002; Heald et al., 2018). Co-contraction is thought to be a default strategy for reducing initial kinematic errors early in learning, as the amount of co-contraction gradually decreases as a new internal model of the environment is learned. The time course of co-contraction during learning might therefore resemble the time course of neural engagement that we observed during Block 2. Because we did not record EMG during our experiments, we cannot determine whether or not changes in neural engagement are related to co-contraction. We observed a similar time course of neural engagement following pauses during Block 1 of the experiment (Fig. 4.2e), when there is no reason for animals to co-contract, as there are no unexpected kinematics to correct. Thus, if neural engagement does reflect a default motor response such as co-contraction or a change in response gains, this response may be the manifestation in motor cortex of an uncertainty- or arousal-driven response (as discussed below), rather than a response to kinematic errors.

Is it arousal? Recent work identified a slowly varying correlate of internal state in the neural population activity of prefrontal cortex and visual area V4 while monkeys performed a perceptual decision-making task (Cowley et al., 2020). The authors present evidence that this "slow drift" in population activity reflected an arousal or impulsivity signal, which biased animals' decisions. The authors propose that this signal may arise from the release of a neuromodulator such as norepinephrine (NE), distributed by the locus coeruleus (LC) (Aston-Jones and Cohen, 2005; Joshi et al., 2016; McGinley et al., 2015). We speculate that the neural engagement signal identified in the present work may have a similar origin. This would also be consistent with recent work in rodents reporting brain-wide modulation associated with behavioral variables such as facial expression (Stringer et al., 2019) and licking (Stringer et al., 2019; Allen et al., 2019) that can indicate changes in arousal. What might be the role of an arousal signal, if any, in M1? It has been proposed that the LC signals uncertainty in the environment (Yu and Dayan, 2003; Sales et al., 2019), and that the release of NE modulates a trade-off between explorative-exploitative behaviors (Aston-Jones and Cohen, 2005). From this perspective, the increases in neural engagement that we observe following pauses in the experiment and at the start of learning may be due to the release of NE by the LC. If these changes indeed serve a function, such as indicating a change in the environment or driving exploration, our results suggest that this response is relatively coarse or stereotyped across task goals, because the increase in neural engagement persisted even when it was detrimental to behavior.

B.4 Finding neural engagement in your own data

The below procedure may be helpful for identifying neural engagement axes in your own data. I do not know how many of these steps are essential (e.g., the time bin size), but they are the ones I used, so they will likely be a helpful starting point.

Preprocessing

- Resample your data (if necessary) so that you have spike counts in ∼50 ms bins. Let $\mathbf{u} \in \mathbb{R}^q$ be the resulting vector of spike counts from q units in a single time bin.

- Identify an analysis window of interest for each trial, so that each trial contributes roughly the same amount of data (e.g., for 50 ms bins, using 250 ms starting from movement onset would give you 5 time steps per trial).

- Select a "baseline" block of trials in which there is no learning. (The idea is that for all trials in this block with the same task condition, we will be able to treat any fluctuations in neural activity as being unrelated to the task.)

- Normalize (z-score) the spike counts for each unit separately, using the baseline trials to find the mean and variance per unit. After z-scoring, each unit's spike counts during the baseline block should be zero mean and unit variance. Let each z-scored spike count vector be $\mathbf{y} \in \mathbb{R}^q$.

- Apply factor analysis (FA) to the z-scored spike counts from the baseline block, and use the resulting FA parameters (i.e., $L \in \mathbb{R}^{q \times d}$ and $\Psi \in \mathbb{R}^{d \times d}$, where d is the FA model's dimensionality) to convert each z-scored spike count ($\mathbf{y} \in \mathbb{R}^q$) to factor activity ($\mathbf{z} \in \mathbb{R}^d$). Make sure to orthonormalize the resulting factors. (If you're in a rush you can use PCA instead of FA.)

Estimating neural engagement axes

- For each task condition θ (e.g., reach direction), apply PCA to the factor activity from all baseline trials to get the eigenvector corresponding to the first principal component. This eigenvector, $\mathbf{a} \in \mathbb{R}^d$, is the *neural engagement axis* for this task condition.

- *Optional:* In the step above, even though we are applying PCA to trials with the same task condition, during trials with no learning, there may nevertheless be some task-related variability. For example, for arm reaches to the same target, there may be some trial-to-trial variability in reach direction due to the subject making small (but task relevant!) corrections to their reach. To get around this, you may want to identify a subspace in which this task-related neural variability is most salient, and then apply PCA to the null space of this subspace. This will help ensure that your resulting neural engagement axis is ignoring some of the task-related neural variability.

- The eigenvector resulting from PCA is unsigned, meaning there is no distinction between a "positive" versus "negative" direction. We will define positive neural engagement so that increasing factor activity in the direction $\mathbf{a}$ corresponds to most units increasing their spike count. To do this, for each engagement axis $\mathbf{a}$, ensure that we have $L\mathbf{a} > 0$ for as many units as possible, flipping the sign if necessary (i.e., setting $\mathbf{a} := -\mathbf{a}$).

BIBLIOGRAPHY

Robert Ajemian, Alessandro D'Ausilio, Helene Moorman, and Emilio Bizzi. A theory for how sensorimotor skills are learned and retained in noisy and nonstationary neural circuits. 110(52):E5078–E5087, 2013.

William E Allen, Michael Z Chen, Nandini Pichamoorthy, Rebecca H Tien, Marius Pachitariu, Liqun Luo, and Karl Deisseroth. Thirst regulates motivated behavior through modulation of brainwide neural population dynamics. 364(6437):253–253, 2019.

Aaron S Andalman and Michale S Fee. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. 106(30):12518–12523, 2009.

Amos Arieli, Alexander Sterkin, Amiram Grinvald, and AD Aertsen. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. 273(5283):1868–1871, 1996.

Zoe Ashwood, Nicholas A. Roy, Ji Hyun Bak, and Jonathan W Pillow. Inferring learning rules from animal decision-making. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 3442–3453. Curran Associates, Inc., 2020. URL `https://proceedings.neurips.cc/paper/2020/file/234b941e88b755b7a72a1c1dd5022f30-Paper.pdf`.

Gary Aston-Jones and Jonathan D. Cohen. An Integrative Theory of Locus Coeruleus-Norepinephrine Function: Adaptive Gain and Optimal Performance. 28(1):403–450, 2005. ISSN 0147-006X. doi: 10.1146/annurev.neuro.28.061604.135709.

Vivek R Athalye, Karunesh Ganguly, Rui M Costa, and Jose M Carmena. Emergence of coordinated neural dynamics underlies neuroprosthetic learning and skillful control. 93(4):955–970, 2017.

Vivek R Athalye, Fernando J Santos, Jose M Carmena, and Rui M Costa. Evidence for a neural law of effect. 359(6379):1024–1029, 2018.

Bruno B Averbeck, Peter E Latham, and Alexandre Pouget. Neural correlations, population coding and computation. 7(5):358–366, 2006.

HB Barlow. Trigger features, adaptation and economy of impulses. In *Information Processing in the Nervous System*, pages 209–230. New York: Springer Science & Business Media, 1969.

Ramon Bartolo, Richard C Saunders, Andrew R Mitz, and Bruno B Averbeck. Dimensionality, information and learning in prefrontal cortex. *PLoS computational biology*, 16(4):e1007514, 2020a.

Ramon Bartolo, Richard C Saunders, Andrew R Mitz, and Bruno B Averbeck. Information-limiting correlations in large neural populations. *Journal of Neuroscience*, 40(8):1668–1678, 2020b.

Jeffrey M Beck, Wei Ji Ma, Xaq Pitkow, Peter E Latham, and Alexandre Pouget. Not noisy, just wrong: the role of suboptimal inference in behavioral variability. *Neuron*, 74(1):30–39, 2012.

NA Bernstein. The problem of interrelation between coordination and localization. In *The Coordination and Regulation of Movements*, pages 15–59. New York: Pergamon, 1967.

Ake Björck and Gene H Golub. Numerical methods for computing angles between linear subspaces. 27(123):579–594, 1973.

Noam Brown and Tuomas Sandholm. Superhuman ai for multiplayer poker. *Science*, 365(6456):885–890, 2019.

Charles F Cadieu, Ha Hong, Daniel LK Yamins, Nicolas Pinto, Diego Ardila, Ethan A Solomon, Najib J Majaj, and James J DiCarlo. Deep neural networks rival the representation of primate it cortex for core visual object recognition. *PLoS Comput Biol*, 10(12):e1003963, 2014.

Jose M Carmena, Mikhail A Lebedev, Roy E Crist, Joseph E O'Doherty, David M Santucci, Dragan F Dimitrov, Parag G Patil, Craig S Henriquez, and Miguel AL Nicolelis. Learning to control a brain–machine interface for reaching and grasping by primates. 1(2):e42, 2003.

Gert Cauwenberghs. A fast stochastic error-descent algorithm for supervised learning and optimization. In *Advances in neural information processing systems*, pages 244–251, 1993.

Mark M Churchland, Byron M Yu, John P Cunningham, Leo P Sugrue, Marlene R Cohen, Greg S Corrado, William T Newsome, Andrew M Clark, Paymon Hosseini, Benjamin B Scott, et al. Stimulus onset quenches neural variability: a widespread cortical phenomenon. 13(3):369–378, 2010.

Claudia Clopath, Tobias Bonhoeffer, Mark Hübener, and Tobias Rose. Variance and invariance of neuronal long-term representations. *Philosophical Transactions of the*

Royal Society B: Biological Sciences, 372(1715):20160161, March 2017. doi: 10.1098/rstb.2016.0161. URL https://royalsocietypublishing.org/doi/full/10.1098/rstb.2016.0161.

Marlene R Cohen and John HR Maunsell. Attention improves performance primarily by reducing interneuronal correlations. 12(12):1594, 2009.

Christos Constantinidis and Torkel Klingberg. The neuroscience of working memory capacity and training. 17(7):438, 2016.

Benjamin R Cowley, Adam C Snyder, Katerina Acar, Ryan C Williamson, Byron M Yu, and Matthew A Smith. Slow drift of neural activity as a signature of impulsivity in macaque visual and prefrontal cortex. 108(3):551–567, 2020.

a. de Rugy, G. E. Loeb, and T. J. Carroll. Muscle Coordination Is Habitual Rather than Optimal. 32(21):7384–7391, 2012. ISSN 0270-6474. doi: 10.1523/JNEUROSCI.5792-11.2012.

Ashesh K Dhawale, Maurice A Smith, and Bence P Ölveczky. The role of variability in motor learning. *Annual review of neuroscience*, 40:479–498, 2017.

Jörn Diedrichsen, Olivier White, Darren Newman, and Níall Lally. Use-dependent and error-based learning of motor behaviors. *Journal of Neuroscience*, 30(15): 5159–5166, 2010a.

Jörn Diedrichsen, Reza Shadmehr, and Richard B Ivry. The coordination of movement: optimal feedback control and beyond. 14(1):31–39, 2010b.

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

Laura N Driscoll, Noah L Pettit, Matthias Minderer, Selmaan N Chettih, and Christopher D Harvey. Dynamic reorganization of neuronal activity patterns in parietal cortex. 170(5):986–999, 2017.

Shaul Druckmann and Dmitri B Chklovskii. Neuronal circuits underlying persistent representations despite time varying activity. 22(22):2095–2103, 2012.

Ian L Dryden, Alexey Koloydenko, and Diwei Zhou. Non-euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. 3:1102–1123, 2009.

Joshua T. Dudman and John W. Krakauer. The basal ganglia: From motor commands to the control of vigor. 37:158–166, 2016. ISSN 18736882. doi: 10.1016/j.conb.2016.02.005. URL http://dx.doi.org/10.1016/j.conb.2016.02.005.

Alexander S Ecker, Philipp Berens, R James Cotton, Manivannan Subramaniyan, George H Denfield, Cathryn R Cadwell, Stelios M Smirnakis, Matthias Bethge, and Andreas S Tolias. State dependence of noise correlations in macaque primary visual cortex. 82(1):235–248, 2014.

Gamaleldin F. Elsayed, Antonio H. Lara, Matthew T. Kaufman, Mark M. Churchland, and John P. Cunningham. Reorganization between preparatory and movement population responses in motor cortex. 7:13239, 2016. ISSN 2041-1723. doi: 10.1038/ncomms13239.

Tatiana A Engel, Warasinee Chaisangmongkon, David J Freedman, and Xiao-Jing Wang. Choice-correlated activity fluctuations underlie learning of neuronal category representation. *Nature communications*, 6(1):1–12, 2015.

G J C Ettema, G Styles, and V Kippers. The moment arms of 23 muscle segments of the upper limb with varying elbow and forearm positions: Implications for motor control. 17:201–220, 1998. ISSN 01679457. doi: 10.1016/S0167-9457(97)00030-4.

Edward V Evarts. Relation of pyramidal tract activity to force exerted during voluntary movement. *Journal of neurophysiology*, 31(1):14–27, 1968.

Andrew H Fagg, Ashvin Shah, and Andrew G Barto. A computational model of muscle recruitment for wrist movements. 88(6):3348–3358, 2002.

Ali Farshchiansadegh, Alejandro Melendez-Calderon, Rajiv Ranganathan, Todd D Murphey, and Ferdinando A Mussa-Ivaldi. Sensory agreement guides kinetic energy optimization of arm movements during object manipulation. 12(4):e1004861, 2016.

Bertram Feinstein, Bengt Lindegård, Eberhard Nyman, and Gunnar Wohlfart. Morphologic studies of motor units in normal human muscles. 23(2):127–142, 1955.

Barbara Feulner and Claudia Clopath. Neural manifold under plasticity in a goal driven learning behaviour. *PLoS computational biology*, 17(2):e1008621, 2021.

Wolfgang Förstner and Boudewijn Moonen. A metric for covariance matrices. In *Geodesy-The Challenge of the 3rd Millennium*, pages 299–309. Berlin, Heidelberg: Springer-Verlag, 2003.

Karunesh Ganguly and Jose M. Carmena. Emergence of a stable cortical map for neuroprosthetic control. 7(7):1–13, 07 2009. doi: 10.1371/journal.pbio.1000153.

Apostolos P Georgopoulos, John F Kalaska, Roberto Caminiti, and Joe T Massey. On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *Journal of Neuroscience*, 2(11):1527–1537, 1982.

Apostolos P Georgopoulos, Andrew B Schwartz, and Ronald E Kettner. Neuronal population coding of movement direction. 233(4771):1416–1419, 1986.

Vikash Gilja, Paul Nuyujukian, Cindy A Chestek, John P Cunningham, Byron M Yu, Joline M Fan, Mark M Churchland, Matthew T Kaufman, Jonathan C Kao, Stephen I Ryu, et al. A high-performance neural prosthesis enabled by control algorithm design. 15(12):1752–1757, 2012.

Matthew D Golub, Byron M Yu, Andrew B Schwartz, and Steven M Chase. Motor cortical control of movement speed with implications for brain-machine interface control. 112(2):411–429, 2014.

Matthew D. Golub, Byron M. Yu, and Steven M. Chase. Internal models for interpreting neural population activity during sensorimotor control. 4 (DECEMBER2015):1–28, 2015. ISSN 2050084X. doi: 10.7554/eLife.10015.

Matthew D Golub, Steven M Chase, Aaron P Batista, and Byron M Yu. Brain–computer interfaces for dissecting cognitive processes underlying sensorimotor control. 37:53–58, 2016.

Matthew D. Golub, Patrick T. Sadtler, Emily R. Oby, Kristin M. Quick, Stephen I. Ryu, Elizabeth C. Tyler-Kabara, Aaron P. Batista, Steven M. Chase, and Byron M. Yu. Learning by neural reassociation. 21:1546–1726, 2018. doi: 10.1038/s41593-018-0095-3.

Henry Gray. *Anatomy of the human body*. Philadelphia: Lea & Febiger, 1918.

Yong Gu, Sheng Liu, Christopher R Fetsch, Yun Yang, Sam Fok, Adhira Sunkara, Gregory C DeAngelis, and Dora E Angelaki. Perceptual learning reduces interneuronal correlations in macaque visual cortex. 71(4):750–761, 2011.

Alkis M. Hadjiosif, John W. Krakauer, and Adrian M. Haith. Did we get sensorimotor adaptation wrong? Implicit adaptation as direct policy updating rather than forward-model-based learning. *The Journal of Neuroscience*, pages JN–RM–2125–20, February 2021. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.2125-20.2021. URL `http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.2125-20.2021`.

Adrian M Haith and John W Krakauer. Model-based and model-free mechanisms of human motor learning. In *Progress in motor control*, pages 1–21. Springer, 2013.

Christopher D Harvey, Philip Coen, and David W Tank. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. 484(7392):62–68, 2012.

Markus Hauschild, Grant H Mulliken, Igor Fineman, Gerald E Loeb, and Richard A Andersen. Cognitive signals for brain-machine interfaces in posterior parietal cortex include continuous 3d trajectory commands. 109(42):17075–17080, 2012.

James B Heald, David W Franklin, and Daniel M Wolpert. Increasing muscle co-contraction speeds up internal model acquisition during dynamic motor learning. 8(1):1–11, 2018.

Jay A Hennig, Matthew D Golub, Peter J Lund, Patrick T Sadtler, Emily R Oby, Kristin M Quick, Stephen I Ryu, Elizabeth C Tyler-Kabara, Aaron P Batista, Byron M Yu, et al. Constraints on neural redundancy. 7:e36774, 2018.

Jay A Hennig, Emily R Oby, Matthew D Golub, Lindsay A Bahureksa, Patrick T Sadtler, Kristin M Quick, Stephen I Ryu, Elizabeth C Tyler-Kabara, Aaron P Batista, Steven M Chase, et al. Learning is shaped by abrupt changes in neural engagement. *bioRxiv*, 2020.

Leigh R Hochberg, Mijail D Serruya, Gerhard M Friehs, Jon A Mukand, Maryam Saleh, Abraham H Caplan, Almut Branner, David Chen, Richard D Penn, and John P Donoghue. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. 442(7099):164–171, 2006.

Anthony Holtmaat and Karel Svoboda. Experience-dependent structural synaptic plasticity in the mammalian brain. *Nature Reviews Neuroscience*, 10(9):647–658, 2009.

Chengcheng Huang, Douglas A Ruff, Ryan Pyle, Robert Rosenbaum, Marlene R Cohen, and Brent Doiron. Circuit models of low-dimensional shared variability in cortical networks. 101(2):337–348, 2019.

Helen J Huang, Rodger Kram, and Alaa A Ahmed. Reduction of metabolic cost during motor learning of arm reaching dynamics. 32(6):2182–2190, 2012. ISSN 1529-2401. doi: 10.1523/JNEUROSCI.4003-11.2012.

Eun Jung Hwang, Paul M Bailey, and Richard A Andersen. Volitional control of neural activity relies on the natural motor repertoire. 23(5):353–361, 2013.

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

Marwan Jabri and Barry Flower. Weight perturbation: An optimal architecture and learning technique for analog vlsi feedforward and recurrent multilayer networks. *IEEE Transactions on Neural Networks*, 3(1):154–157, 1992.

James M Jeanne, Tatyana O Sharpee, and Timothy Q Gentner. Associative learning enhances population coding by inverting interneuronal correlation patterns. 78(2): 352–363, 2013.

Siddhartha Joshi, Yin Li, Rishi M. Kalwani, and Joshua I. Gold. Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. 89(1):221–234, 2016. ISSN 10974199. doi: 10.1016/j.neuron.2015.11.028. URL `http://dx.doi.org/10.1016/j.neuron.2015.11.028`.

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

Matthew T Kaufman, Mark M Churchland, Stephen I Ryu, and Krishna V Shenoy. Cortical activity in the null space: permitting preparation without movement. 17 (3):440–8, 2014. ISSN 1546-1726. doi: 10.1038/nn.3643.

Matthew T Kaufman, Jeffrey S Seely, David Sussillo, Stephen I Ryu, Krishna V Shenoy, and Mark M Churchland. The largest response component in the motor cortex reflects movement timing but not movement type. 3(4), 2016.

Masakazu Konishi. Birdsong: from behavior to neuron. *Annual review of neuroscience*, 8(1):125–170, 1985.

John W Krakauer, Zachary M Pine, Maria-Felice Ghilardi, and Claude Ghez. Learning of visuomotor transformations for vectorial planning of reaching trajectories. *Journal of Neuroscience*, 20(23):8916–8924, 2000.

Armin Lak, William R Stauffer, and Wolfram Schultz. Dopamine neurons learn relative chosen value from probabilistic rewards. *Elife*, 5:e18044, 2016.

Brenden M Lake, Tomer D Ullman, Joshua B Tenenbaum, and Samuel J Gershman. Building machines that learn and think like people. *Behavioral and brain sciences*, 40, 2017.

Serge Lang. An example of seminegative curvature. In *Fundamentals of differential geometry*, pages 322–326. New York: Springer Science & Business Media, 1999.

Karl Spencer Lashley. Integrative functions of the cerebral cortex. 13(1):1–42, 1933.

Peter E Latham and Sheila Nirenberg. Synergy, redundancy, and independence in population codes, revisited. 25(21):5195–5206, 2005.

Simon B Laughlin. Energy as a constraint on the coding and processing of sensory information. 11(4):475–480, 2001.

Simon B Laughlin, Rob R de Ruyter van Steveninck, and John C Anderson. The metabolic cost of neural information. 1(1):36, 1998.

Andrew J Law, Gil Rivlis, and Marc H Schieber. Rapid acquisition of novel interface control by small ensembles of arbitrarily selected primary motor cortex neurons. (585):1528–1548, 2014. ISSN 1522-1598. doi: 10.1152/jn.00373.2013.

Chi-Tat Law and Joshua I Gold. Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nature neuroscience*, 11(4):505–513, 2008.

Marvin L Leathers and Carl R Olson. In monkeys making value-based decisions, lip neurons encode cue salience and not action value. 338(6103):132–135, 2012.

Roger N Lemon. Descending pathways in motor control. 31:195–218, 2008.

Anthony Leonardo. Degenerate coding in neural systems. 191(11):995–1010, 2005.

W B Levy and R A Baxter. Energy efficient neural codes. 8(3):531–543, 1996. ISSN 0899-7667. doi: 10.1162/neco.1996.8.3.531.

Chiang-Shan Ray Li, Camillo Padoa-Schioppa, and Emilio Bizzi. Neuronal correlates of motor performance and motor learning in the primary motor cortex of monkeys adapting to an external force field. 30(2):593–607, 2001.

I-Chun Lin, Michael Okun, Matteo Carandini, and Kenneth D Harris. The nature of shared cortical variability. 87(3):644–656, 2015.

Valerio Mante, David Sussillo, Krishna V Shenoy, and William T Newsome. Context-dependent computation by recurrent dynamics in prefrontal cortex. 503(7474): 78–84, 2013.

Adam H Marblestone, Greg Wayne, and Konrad P Kording. Toward an integration of deep learning and neuroscience. *Frontiers in computational neuroscience*, 10: 94, 2016.

William Mau, Michael E Hasselmo, and Denise J Cai. The brain in motion: How ensemble fluidity drives memory-updating and flexibility. *eLife*, 9:e63550, December 2020. ISSN 2050-084X. doi: 10.7554/eLife.63550. URL `https://doi.org/10.7554/eLife.63550`.

Pietro Mazzoni and John W Krakauer. An implicit plan overrides an explicit strategy during visuomotor adaptation. *Journal of neuroscience*, 26(14):3642–3645, 2006.

Pietro Mazzoni, Anna Hristova, and John W. Krakauer. Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. 27(27):7105–7116, 2007. ISSN 02706474. doi: 10.1523/JNEUROSCI.0264-07.2007.

Matthew J. McGinley, Martin Vinck, Jacob Reimer, Renata Batista-Brito, Edward Zagha, Cathryn R. Cadwell, Andreas S. Tolias, Jessica A. Cardin, and David A. McCormick. Waking State: Rapid Variations Modulate Neural and Behavioral Responses. 87(6):1143–1161, 2015. ISSN 10974199. doi: 10.1016/j.neuron.2015.09.012. URL `http://dx.doi.org/10.1016/j.neuron.2015.09.012`.

Alexander R. McNeill. Energetics and optimization of human walking and running: The 2000 raymond pearl memorial lecture. 14(5):641–648, 2002. ISSN 1520-6300. doi: 10.1002/ajhb.10067.

André G Mendonça, Jan Drugowitsch, M Inês Vicente, Eric EJ DeWitt, Alexandre Pouget, and Zachary F Mainen. The impact of learning on perceptual decisions and its implication for speed-accuracy tradeoffs. *Nature communications*, 11(1): 1–15, 2020.

Jude F Mitchell, Kristy A Sundberg, and John H Reynolds. Spatial attention decorrelates intrinsic activity fluctuations in macaque area v4. 63(6):879–888, 2009.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *nature*, 518(7540):529–533, 2015.

Helene G. Moorman, Suraj Gowda, and Jose M Carmena. Control of redundant kinematic degrees of freedom in a closed-loop brain-machine interface. 25(6):750–760, 2017. ISSN 1534-4320. doi: 10.1109/TNSRE.2016.2593696.

Daniel W Moran and Andrew B Schwartz. Motor cortical representation of speed and direction during reaching. 82:2676–2692, 1999.

Rubén Moreno-Bote, Jeffrey Beck, Ingmar Kanitscheider, Xaq Pitkow, Peter Latham, and Alexandre Pouget. Information-limiting correlations. 17(10):1410–1417, 2014.

Karen A Moxon and Guglielmo Foffani. Brain-machine interfaces beyond neuroprosthetics. 86(1):55–67, 2015.

John D. Murray, Alberto Bernacchia, Nicholas A. Roy, Christos Constantinidis, Ranulfo Romo, and Xiao-Jing Wang. Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. 114(2):394–399, 2017. doi: 10.1073/pnas.1619449114.

Emre O. Neftci and Bruno B. Averbeck. Reinforcement learning in artificial and biological systems. *Nature Machine Intelligence*, 1(3):133–143, March 2019. ISSN 2522-5839. doi: 10.1038/s42256-019-0025-4. URL `http://www.nature.com/articles/s42256-019-0025-4`.

Amy M Ni, Douglas A Ruff, Joshua J Alberts, Jen Symmonds, and Marlene R Cohen. Learning and attention reveal a general relationship between population activity and behavior. 359(6374):463–465, 2018.

Emily R Oby, Matthew D Golub, Jay A Hennig, Alan D Degenhart, Elizabeth C Tyler-Kabara, Yu Byron M, Steven M Chase, and Aaron P Batista. New neural activity patterns emerge with long-term learning. 116(30):15210–15215, 2019.

Michael Okun, Pierre Yger, Stephan L. Marguet, Florian Gerard-Mercier, Andrea Benucci, Steffen Katzner, Laura Busse, Matteo Carandini, and Kenneth D. Harris. Population Rate Dynamics and Multineuron Firing Patterns in Sensory Cortex. *Journal of Neuroscience*, 32(48):17108–17119, November 2012. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1831-12.2012. URL `https://www.jneurosci.org/content/32/48/17108`.

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

Bence P Ölveczky, Timothy M Otchy, Jesse H Goldberg, Dmitriy Aronov, and Michale S Fee. Changes in the neural control of a complex motor sequence during learning. *Journal of neurophysiology*, 106(1):386–397, 2011.

Mohsen Omrani, J Andrew Pruszynski, Chantelle D Murnaghan, and Stephen H Scott. Perturbation-evoked responses in primary motor cortex are modulated by behavioral context. 112(11):2985–3000, 2014.

Mohsen Omrani, Chantelle D Murnaghan, J Andrew Pruszynski, and Stephen H Scott. Distributed task-specific processing of somatosensory feedback for voluntary motor control. 5:e13141, 2016.

Amy L Orsborn and Bijan Pesaran. Parsing learning in networks using brain–machine interfaces. 46:76–83, 2017.

Amy L Orsborn, Helene G Moorman, Simon A Overduin, Maryam M Shanechi, Dragan F Dimitrov, and Jose M Carmena. Closed-loop decoder adaptation shapes neural plasticity for skillful neuroprosthetic control. 82(6):1380–1393, 2014.

Rieko Osu, David W Franklin, Hiroko Kato, Hiroaki Gomi, Kazuhisa Domen, Toshinori Yoshioka, and Mitsuo Kawato. Short-and long-term changes in joint co-contraction associated with motor learning as revealed from surface emg. 88(2): 991–1004, 2002.

Marino Pagan, Luke S Urban, Margot P Wohl, and Nicole C Rust. Signals in inferotemporal and perirhinal cortex suggest an untangling of visual target information. 16(8):1132–1139, 2013.

Anitha Pasupathy and Earl K Miller. Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature*, 433(7028):873–876, 2005.

Matthew G Perich, Juan A Gallego, and Lee E Miller. A neural population mechanism for rapid learning. 100(4):964–976, 2018.

Nathalie Picard, Yoshiya Matsuzaka, and Peter L Strick. Extended practice of a motor skill is associated with reduced metabolic activity in M1. 16(9):1340–1347, 2013. ISSN 1546-1726. doi: 10.1038/nn.3477.

Jasper Poort, Adil G Khan, Marius Pachitariu, Abdellatif Nemri, Ivana Orsolic, Julija Krupic, Marius Bauza, Maneesh Sahani, Georg B Keller, Thomas D Mrsic-Flogel, et al. Learning enhances sensory and multiple non-sensory representations in primary visual cortex. 86(6):1478–1490, 2015.

J Andrew Pruszynski, Isaac Kurtzer, Joseph Y Nashed, Mohsen Omrani, Brenda Brouwer, and Stephen H Scott. Primary motor cortex underlies multi-joint integration for fast feedback control. 478(7369):387–390, 2011.

J Andrew Pruszynski, Mohsen Omrani, and Stephen H Scott. Goal-dependent modulation of fast feedback responses in primary motor cortex. 34(13):4608–4617, 2014.

Neil C Rabinowitz, Robbe L Goris, Marlene Cohen, and Eero P Simoncelli. Attention stabilizes the shared gain of v4 populations. 4:e08998, 2015.

Rajiv Ranganathan, Adenike Adewuyi, and Ferdinando A Mussa-Ivaldi. Learning to be lazy: exploiting redundancy in a novel task to minimize movement-related effort. 33(7):2754–2760, 2013.

David Raposo, Matthew T Kaufman, and Anne K Churchland. A category-free neural population supports evolving demands during decision-making. 17(12): 1784–1792, 2014.

Blake A Richards, Timothy P Lillicrap, Philippe Beaudoin, Yoshua Bengio, Rafal Bogacz, Amelia Christensen, Claudia Clopath, Rui Ponte Costa, Archy de Berker, Surya Ganguli, et al. A deep learning framework for neuroscience. *Nature neuroscience*, 22(11):1761–1770, 2019.

Mengia-S Rioult-Pedotti, Daniel Friedman, and John P Donoghue. Learning-induced ltp in neocortex. *science*, 290(5491):533–536, 2000.

Uri Rokni, Andrew G Richardson, Emilio Bizzi, and H Sebastian Seung. Motor learning with unstable neural representations. 54(4):653–666, 2007. ISSN 0896-6273. doi: 10.1016/j.neuron.2007.04.030.

Mats Rudemo. Empirical choice of histograms and kernel density estimators. 9: 65–78, 1982.

Douglas A Ruff and Marlene R Cohen. Simultaneous multi-area recordings suggest that attention improves performance by reshaping stimulus representations. pages 1–8, 2019.

Michael E Rule, Adrianna R Loback, Dhruva V Raman, Laura N Driscoll, Christopher D Harvey, and Timothy O'Leary. Stable task information from an unstable neural population. *eLife*, 9:e51121, July 2020. ISSN 2050-084X. doi: 10.7554/eLife.51121. URL `https://elifesciences.org/articles/51121`.

Abigail A Russo, Sean R Bittner, Sean M Perkins, Jeffrey S Seely, Brian M London, Antonio H Lara, Andrew Miri, Najja J Marshall, Adam Kohn, Thomas M Jessell, et al. Motor cortex embeds muscle-like commands in an untangled population response. 97(4):953–966, 2018.

Patrick T. Sadtler, Kristin M. Quick, Matthew D. Golub, Steven M. Chase, Stephen I. Ryu, Elizabeth C. Tyler-Kabara, Byron M. Yu, and Aaron P. Batista. Neural constraints on learning. 512(7515):423–426, 2014. ISSN 0028-0836. doi: 10.1038/nature13665.

Anna C Sales, Karl J Friston, Matthew W Jones, Anthony E Pickering, and Rosalyn J Moran. Locus coeruleus tracking of prediction errors optimises cognitive flexibility: An active inference model. 15(1):e1006267, 2019.

Gopal Santhanam, Byron M Yu, Vikash Gilja, Stephen I Ryu, Afsheen Afshar, Maneesh Sahani, and Krishna V Shenoy. Factor-analysis methods for higher-performance neural prostheses. 102(2):1315–1330, 2009.

Andrew Saxe, Stephanie Nelli, and Christopher Summerfield. If deep learning is the answer, what is the question? *Nature Reviews Neuroscience*, pages 1–13, 2020.

Elad Schneidman, William Bialek, and Michael J Berry. Synergy, redundancy, and independence in population codes. 23(37):11539–11553, 2003.

John P. Scholz and Gregor Schöner. The uncontrolled manifold concept: Identifying control variables for a functional task. 126(3):289–306, 1999. ISSN 00144819. doi: 10.1007/s002210050738.

Aniek Schoups, Rufin Vogels, Ning Qian, and Guy Orban. Practising orientation identification improves orientation coding in v1 neurons. *Nature*, 412(6846):549–553, 2001.

Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, et al. Mastering atari, go, chess and shogi by planning with a learned model. *Nature*, 588(7839):604–609, 2020.

Wolfram Schultz. Neuronal Reward and Decision Signals: From Theories to Data. *Physiological Reviews*, 95(3):853–951, July 2015. ISSN 0031-9333, 1522-1210. doi: 10.1152/physrev.00023.2014. URL `https://www.physiology.org/doi/10.1152/physrev.00023.2014`.

Wolfram Schultz. Recent advances in understanding the role of phasic dopamine activity. *F1000Research*, 8, 2019.

Wolfram Schultz, Peter Dayan, and P Read Montague. A neural substrate of prediction and reward. 275(5306):1593–1599, 1997.

Stephen H Scott. Optimal feedback control and the neural basis of volitional motor control. 5(7):532–546, 2004.

Ehsan Sedaghat-Nejad, David J. Herzfeld, and Reza Shadmehr. Reward prediction error modulates saccade vigor. 39(25):5010–5017, 2019. ISSN 15292401. doi: 10.1523/JNEUROSCI.0432-19.2019.

Reza Shadmehr and Henry H Holcomb. Neural correlates of motor memory consolidation. 277(5327):821–825, 1997.

Reza Shadmehr, Thomas R. Reppert, Erik M. Summerside, Tehrim Yoon, and Alaa A. Ahmed. Movement Vigor as a Reflection of Subjective Economic Utility. 42(5):323–336, 2019. ISSN 1878108X. doi: 10.1016/j.tins.2019.02.003. URL https://doi.org/10.1016/j.tins.2019.02.003.

Krishna V Shenoy and Jose M Carmena. Combining decoder design and neural adaptation in brain-machine interfaces. 84(4):665–680, 2014.

Krishna V Shenoy, Maneesh Sahani, and Mark M Churchland. Cortical control of arm movements: a dynamical systems perspective. 36:337–359, 2013.

Abhinav Singh, Adrien Peyrache, and Mark D. Humphries. Medial Prefrontal Cortex Population Activity Is Plastic Irrespective of Learning. *Journal of Neuroscience*, 39(18):3470–3483, May 2019a. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1370-17.2019. URL https://www.jneurosci.org/content/39/18/3470.

Avi Singh, Larry Yang, Chelsea Finn, and Sergey Levine. End-To-End Robotic Reinforcement Learning without Reward Engineering. In *Robotics: Science and Systems XV*. Robotics: Science and Systems Foundation, June 2019b. ISBN 9780992374754. doi: 10.15607/RSS.2019.XV.073. URL http://www.roboticsproceedings.org/rss15/p73.pdf.

Puneet Singh, Sumitash Jana, Ashitava Ghosal, and Aditya Murthy. Exploration of joint redundancy but not task space variability facilitates supervised motor learning. 113(50):14414–14419, 2016. doi: 10.1073/pnas.1613383113.

Satinder Singh, Richard L Lewis, Andrew G Barto, and Jonathan Sorg. Intrinsically motivated reinforcement learning: An evolutionary perspective. *IEEE Transactions on Autonomous Mental Development*, 2(2):70–82, 2010.

Fabian H Sinz, Xaq Pitkow, Jacob Reimer, Matthias Bethge, and Andreas S Tolias. Engineering a less artificial intelligence. *Neuron*, 103(6):967–979, 2019.

Maurice A Smith, Ali Ghazizadeh, and Reza Shadmehr. Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biol*, 4 (6):e179, 2006.

William R Softky and Daniel M Kammen. Correlations in high dimensional or asymmetric data sets: Hebbian neuronal processing. 4(3):337–347, 1991.

Hansem Sohn, Nicolas Meirhaeghe, Rishi Rajalingham, and Mehrdad Jazayeri. A network perspective on sensorimotor learning. *Trends in Neurosciences*, 2020.

Olaf Sporns and Gerald M. Edelman. Solving Bernstein's problem: a proposal for the development of coordinated movement by selection. 64(4):960–981, 1993. ISSN 0009-3920. doi: 10.1111/j.1467-8624.1993.tb04182.x.

Sergey D Stavisky, Jonathan C Kao, Stephen I Ryu, and Krishna V Shenoy. Motor cortical visuomotor feedback activity is initially isolated from downstream targets in output-null neural state space dimensions. 95(1):195–208, 2017.

Nicholas A Steinmetz, Peter Zatka-Haas, Matteo Carandini, and Kenneth D Harris. Distributed coding of choice, action and engagement across the mouse brain. 576 (7786):266–273, 2019.

Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Charu Bai Reddy, Matteo Carandini, and Kenneth D. Harris. Spontaneous behaviors drive multidimensional, brainwide activity. 364(6437), 2019. ISSN 10959203. doi: 10.1126/science.aav7893.

Leo P Sugrue, Greg S Corrado, and William T Newsome. Matching behavior and the representation of value in the parietal cortex. 304(5678):1782–1787, 2004.

David Sussillo, Mark M Churchland, Matthew T Kaufman, and Krishna V Shenoy. A neural network that finds a naturalistic solution for the production of muscle activity. *Nature neuroscience*, 18(7):1025–1033, 2015.

Dawn M Taylor, Stephen I Helms Tillery, and Andrew B Schwartz. Direct cortical control of 3d neuroprosthetic devices. 296(5574):1829–1832, 2002.

Kurt A Thoroughman and Reza Shadmehr. Electromyographic Correlates of Learning an Internal Model of Reaching Movements. 19(19):8573–8588, 1999. ISSN 1529-2401.

Emanuel Todorov. Optimality principles in sensorimotor control. 7(9):907–15, 2004. ISSN 1097-6256. doi: 10.1038/nn1309.

Emanuel Todorov. Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. *Neural computation*, 17(5): 1084–1108, 2005.

Emanuel Todorov and Michael I Jordan. Optimal feedback control as a theory of motor coordination. 5(11):1226–1235, 2002. ISSN 1097-6256. doi: 10.1038/nn963.

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

Takanori Uka, Ryo Sasaki, and Hironori Kumano. Change in choice-related response modulation in area mt during learning of a depth-discrimination task is consistent with task learning. *Journal of Neuroscience*, 32(40):13689–13700, 2012.

F. J. Valero-Cuevas, M. Venkadesan, and E. Todorov. Structured Variability of Muscle Activations Supports the Minimal Intervention Principle of Motor Control. 102(1):59–68, 2009. ISSN 0022-3077. doi: 10.1152/jn.90324.2008.

Martin Vinck, Renata Batista-Brito, Ulf Knoblich, and Jessica A. Cardin. Arousal and Locomotion Make Distinct Contributions to Cortical Activity Patterns and Visual Encoding. 86(3):740–754, 2015. ISSN 10974199. doi: 10.1016/j.neuron.2015.03.028. URL http://dx.doi.org/10.1016/j.neuron.2015.03.028.

Saurabh Vyas, Nir Even-Chen, Sergey D Stavisky, Stephen I Ryu, Paul Nuyujukian, and Krishna V Shenoy. Neural population dynamics underlying motor learning transfer. 97(5):1177–1186, 2018.

Jane X. Wang, Zeb Kurth-Nelson, Dharshan Kumaran, Dhruva Tirumala, Hubert Soyer, Joel Z. Leibo, Demis Hassabis, and Matthew Botvinick. Prefrontal cortex as a meta-reinforcement learning system. *Nature Neuroscience*, 21(6):860–868, June 2018. ISSN 1097-6256, 1546-1726. doi: 10.1038/s41593-018-0147-8. URL http://www.nature.com/articles/s41593-018-0147-8.

Emil Wärnberg and Arvind Kumar. Perturbing low dimensional activity manifolds in spiking neuronal networks. *PLOS computational biology*, 15(5):e1007074, 2019.

Ronald J Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8(3-4):229–256, 1992.

Ryan C Williamson, Benjamin R Cowley, Ashok Litwin-Kumar, Brent Doiron, Adam Kohn, Matthew A Smith, and Byron M Yu. Scaling properties of dimensionality reduction for neural populations and network models. 12(12):e1005141, 2016.

B Wodlinger, JE Downey, EC Tyler-Kabara, AB Schwartz, ML Boninger, and JL Collinger. Ten-dimensional anthropomorphic arm control in a human brain-machine interface: difficulties, solutions, and limitations. 12(1):016011, 2014.

Howard G Wu, Yohsuke R Miyamoto, Luis Nicolas Gonzalez Castro, Bence P Ölveczky, and Maurice A Smith. Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nature Neuroscience*, 17 (2):312–321, February 2014. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.3616. URL http://www.nature.com/articles/nn.3616.

Ji Xia, Tyler Marks, Michael Goard, and Ralf Wessel. Stable representation of a naturalistic movie emerges from episodic activity with gain variability. 2021. doi: 10.21203/rs.3.rs-126977/v1. URL https://doi.org/10.21203/rs.3.rs-126977/v1.

Minnan Xu-Wilson, David S Zee, and Reza Shadmehr. The intrinsic value of visual information affects saccade velocities. 196(4):475–481, 2009.

Daniel LK Yamins and James J DiCarlo. Using goal-driven deep learning models to understand sensory cortex. *Nature neuroscience*, 19(3):356–365, 2016.

Yin Yan, Malte J Rasch, Minggui Chen, Xiaoping Xiang, Min Huang, Si Wu, and Wu Li. Perceptual training continuously refines neuronal population codes in primary visual cortex. *Nature neuroscience*, 17(10):1380–1387, 2014.

Eric Allen Yttri and Joshua Tate Dudman. A proposed circuit computation in basal ganglia: History-dependent gain. 33(5):704–716, 2018.

Angela Yu and Peter Dayan. Expected and unexpected uncertainty: Ach and ne in the neocortex. In *Advances in neural information processing systems*, pages 173–180, 2003.

Byron M. Yu, John P. Cunningham, Gopal Santhanam, Stephen I. Ryu, Krishna V. Shenoy, and Maneesh Sahani. Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. 102(1):614–635, 2009. ISSN 0022-3077. doi: 10.1152/jn.90941.2008.

Anthony M. Zador. A critique of pure learning and what artificial neural networks can learn from animal brains. *Nature Communications*, 10(1):3770, December 2019. ISSN 2041-1723. doi: 10.1038/s41467-019-11786-6. URL `http://www.nature.com/articles/s41467-019-11786-6`.

Xiao Zhou, Rex N. Tien, Sadhana Ravikumar, and Steven M. Chase. Distinct types of neural reorganization during long-term learning. *Journal of Neurophysiology*, 121(4):1329–1341, April 2019. ISSN 0022-3077, 1522-1598. doi: 10.1152/jn.00466.2018. URL `https://www.physiology.org/doi/10.1152/jn.00466.2018`.