

Perspective

How learning unfolds in the brain: toward an optimization view

Jay A. Hennig,^{1,2,3,*} Emily R. Oby,^{2,4} Darby M. Losey,^{1,2,3} Aaron P. Batista,^{2,4,7} Byron M. Yu,^{1,2,5,6,7} and Steven M. Chase^{1,2,6,7}

¹Neuroscience Institute, Carnegie Mellon University, Pittsburgh, PA, USA

²Center for the Neural Basis of Cognition, Pittsburgh, PA, USA

³Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA, USA

⁴Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA, USA

⁵Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

⁶Department of Biomedical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

⁷These authors contributed equally

*Correspondence: jhennig@fas.harvard.edu

<https://doi.org/10.1016/j.neuron.2021.09.005>

SUMMARY

How do changes in the brain lead to learning? To answer this question, consider an artificial neural network (ANN), where learning proceeds by optimizing a given objective or cost function. This “optimization framework” may provide new insights into how the brain learns, as many idiosyncratic features of neural activity can be recapitulated by an ANN trained to perform the same task. Nevertheless, there are key features of how neural population activity changes throughout learning that cannot be readily explained in terms of optimization and are not typically features of ANNs. Here we detail three of these features: (1) the inflexibility of neural variability throughout learning, (2) the use of multiple learning processes even during simple tasks, and (3) the presence of large task-nonspecific activity changes. We propose that understanding the role of these features in the brain will be key to describing biological learning using an optimization framework.

INTRODUCTION

Learning is the process by which individuals accumulate knowledge and develop skillful behavior. Over the course of minutes, years, or even a lifetime of practice, we can learn a variety of skills including how to move and control our bodies, efficiently navigate our surroundings, and obtain resources. What principles underlie the brain’s ability to learn such a variety of different behaviors? Answers to this question have spanned different levels of description. At the microscopic level, studies have revealed how plasticity laws modify synapse strengths between neurons during learning (Feldman, 2009). At the macroscopic level, studies have revealed that behavioral changes during learning are guided by different types of feedback such as supervision (Brainard and Doupe, 2002), reward (Schultz et al., 1997), and sensory prediction errors (Shadmehr and Holcomb, 1997), and develop on different timescales (Newell and Rosenbloom, 1981; Boyden et al., 2004; Smith et al., 2006; Yang and Lisberger, 2010).

We view neural population activity as the link between synaptic plasticity and behavior. Changes in plasticity affect which neural population activity patterns a network can express, and the firing activity of neural populations in turn drives behavior. At the same time, coordinated patterns of spiking activity can drive changes in synapse strengths via plasticity rules (Feldman, 2012). These considerations make the neural population an ideal focal point for studying how the brain changes to improve behavior (Sohn et al., 2020). As detailed below, studying learning in neural populations has already begun to provide unique in-

sights into the brain’s learning process. While many studies have compared neural activity before versus after a learning experience, many key insights necessary for understanding how the brain learns are likely encoded in how neural population activity changes *throughout* learning. Here we place a special emphasis on these latter studies, highlighting the unique benefits of this approach.

How should we interpret the changes in neural population activity observed during learning? Taking inspiration from machine learning may provide a useful starting point. After all, while the brain is 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, even exceeding human levels of performance at some tasks (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; Barak et al., 2013; Cadieu et al., 2014; Sussillo et al., 2015; Rajan et al., 2016; Yamins and DiCarlo, 2016; Chaisangmongkon et al., 2017; Cueva and Wei, 2018; Wang et al., 2018b; Pospisil et al., 2018; Haesemeyer et al., 2019; Saxe et al., 2021; Bakhtiari et al., 2021). These results make the tantalizing proposal that much of the complexity and structure observed in neural population activity can be understood as the outcome of an optimization process, similar to the one used to train ANNs. While these correspondences between artificial and biological networks have been observed

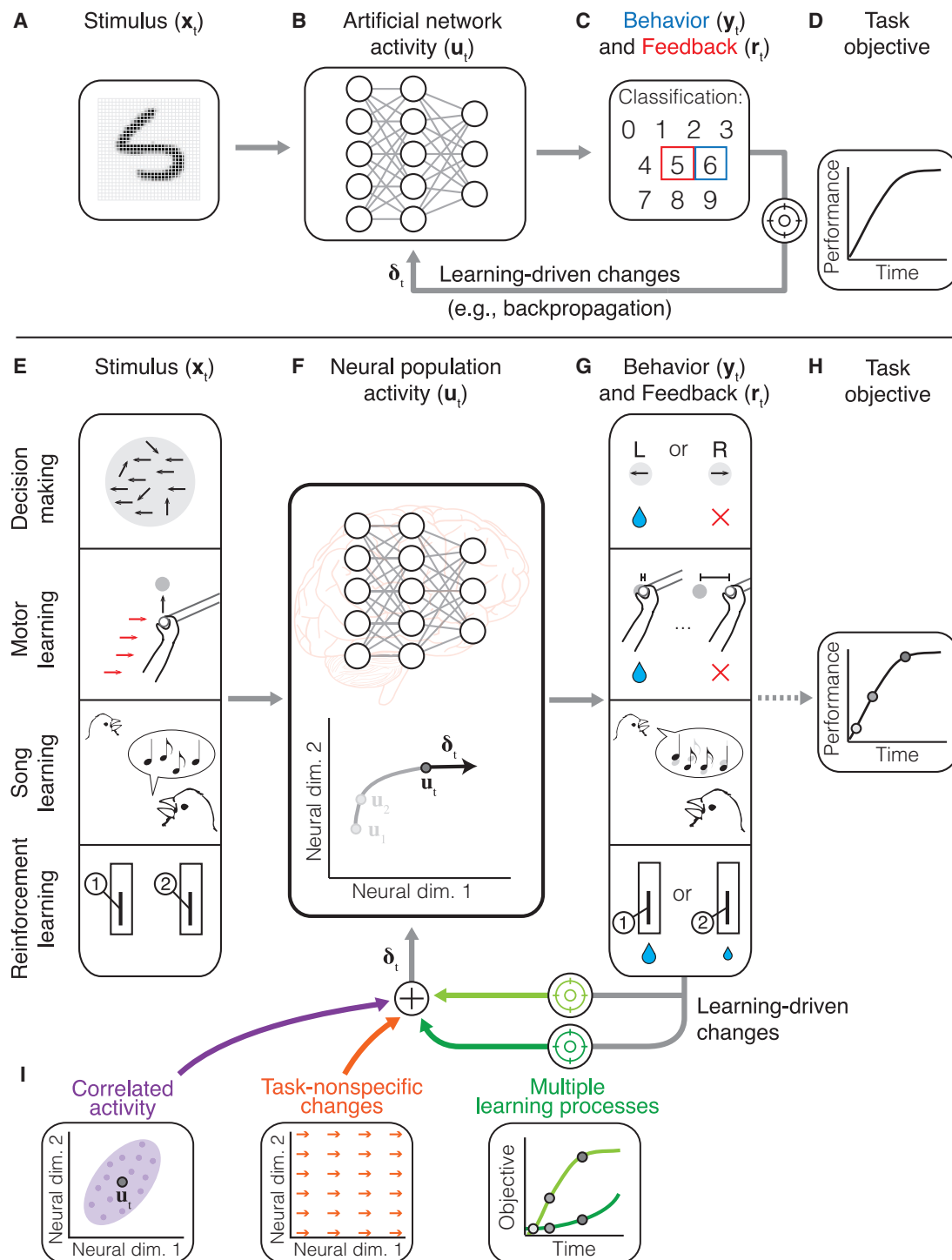


Figure 1. Learning in artificial and biological networks under an optimization framework

Figure 1: Learning in a neural and regular network under an optimization framework. (A–D) Schematic of learning in an artificial neural network (ANN) during a handwritten digit classification task. Given an image stimulus (A), the network’s activity (B) determines the network’s behavior (blue box, C) (i.e., its guess as to which digit is present in the image). The network is given feedback about the correct response (red box, C), which is used by a learning rule (e.g., backpropagation) to modify the network (δ_i) to improve the task objective (D) when the same image is encountered on subsequent trials.

(legend continued on next page)

only *after* each network has finished learning, more generally these results suggest that we may be able to understand the learning process in the brain similar to how we understand learning in artificial networks (Marblestone et al., 2016; Richards et al., 2019).

According to this point of view, which we will refer to as the “optimization framework” (Figure 1), learning is the outcome of optimizing an objective function (also called a “cost function”) using a 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 (Figures 1A–1D). In neuroscience, 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 and decision making (Engel et al., 2015; Yamins and DiCarlo, 2016), motor learning (Haith and Krakauer, 2013), song learning (Fiete et al., 2007), and reinforcement learning (Niv, 2009; Neftci and Averbeck, 2019) (Figures 1E–1H). While learning in the brain is certainly more complicated than in artificial networks, identifying the brain’s objective function and learning rule for a given behavior may be a promising approach for providing a normative account of changes in neural population activity during learning.

However, as we will explain here, studies of neural population activity during learning have identified a number of distinct features not typically present in the activity of ANNs (Figure 1I). These include (1) the inflexibility of neural variability throughout learning; (2) the existence of multiple learning processes, even during simple tasks; and (3) task-nonspecific changes in network activity, which can persist even when they negatively impact task performance. As we will discuss, these features refine our understanding of the extent to which learning in the brain can be described through the lens of optimization.

Below we will present these three features of learning in the brain, with an emphasis on results from studies using a brain-computer interface (BCI) learning paradigm. One key advantage of a BCI for studying learning is that the causal link between neural population activity and behavior can be defined precisely and controlled by the experimenter (Box 1 and Figure 2). We suggest ways in which each feature refines our understanding of learning in the brain as an optimization process, as well as how this feature may differ from learning in artificial agents. While we emphasize results from BCI learning, the features we discuss are relevant to our understanding of learning in the brain in general, spanning different brain areas and timescales. Overall, these observations highlight the importance of understanding not just the endpoint of learning, but also the path that neural activity takes to get to that endpoint. Taking these observations into account may lead to a better understanding of how the brain learns, while also inspiring new methods for learning in artificial networks.

1. Neural variability shapes learning, but is often inflexible

Can the brain control behavioral and neural variability?

Behavior is notoriously variable. Even star basketball players miss free throws, and professional musicians sometimes play the wrong note. But instead of being “the unintended consequence of a noisy nervous system,” behavioral variability may in fact be critical for learning, allowing us to fully explore reward landscapes and adapt to ever-changing environments (Dhawale et al., 2017). 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). The potential benefits of behavioral variability during learning can be appreciated through the lens of reinforcement learning (RL) (Sutton and Barto, 2018). In RL, an agent tries to maximize their cumulative reward in a particular task by finding a balance between exploring new actions (i.e., using behavioral variability) and exploiting successful ones. Evidence in a variety of species suggests that the brain may regulate its behavioral variability dynamically based on the needs of the task, such that behavioral variability is reduced when the stakes are high (Dhawale et al., 2017). For example, as a songbird ages, its song “crystallizes,” becoming more accurate and less variable (Konishi, 1985; Tumer and Brainard, 2007). Similarly, in primates, variability in both arm reaches and eye movements has been shown to decrease given larger reward expectation (Takikawa et al., 2002; Pekny et al., 2015). These results suggest that the brain regulates behavioral variability to facilitate learning.

To what extent can the brain control *neural* variability during learning? Of course, if variability is present at the level of behavior, it will also be present in the brain areas that drive that behavior. But surprisingly, a substantial amount of the variability present in neural population activity does not appear to be modified during learning. Like behavior, the spiking activity of neurons shows substantial variability. Some of this spiking variability, which we refer to as “population covariability,” is not independent to individual neurons but is correlated (i.e., shared) across a population of neurons. There are many potential sources of population covariability, including variability in shared inputs to the network, architectural considerations (e.g., clustered connectivity), neuromodulation, and others (Doiron et al., 2016). As a result, even in the context of performing a familiar task (i.e., when there is no learning), neural population activity exhibits covariability, both across different task conditions, as well as across trials with the same task condition. Multiple lines of evidence suggest that the structure of population covariability is somewhat inflexible—persisting over the course of days or weeks of practice—even when it limits performance. As we explain below, this inflexibility is an important consideration in the context of learning.

(E–H) Schematic of learning in the brain during a variety of different tasks. As in an ANN, the stimulus (\mathbf{x}_t , E) drives neural population activity (\mathbf{u}_t , F), which in turn drives the subject’s behavior (G), resulting in feedback (r_t) that can be used for learning. The goal is for the subject to improve their performance relative to the task objective (H).

(I) In contrast to ANN activity, changes in neural population activity during learning may be influenced by other factors beyond the task objective. These involve correlated activity between neurons (purple ellipse), task-nonspecific changes (orange vectors) (e.g., neural drift), and multiple learning processes with objectives (green curves) that may differ from the task objective.

Box 1. Studying biological learning using a brain-computer interface (BCI) paradigm

Systems neuroscientists usually do not fully know the causal relationship between neural activity and behavior. This makes it difficult to interpret how changes in neural activity during learning lead to improved behavior. In a brain-computer interface (BCI) paradigm, by contrast, the precise causal relationship between neural activity and behavior is determined by the *BCI mapping*, which is known and set by the experimenter (Jarosiewicz et al., 2008; Ganguly and Carmena, 2009; Suminski et al., 2010; Koralek et al., 2012; Hwang et al., 2013; Sadtler et al., 2014; Law et al., 2014; Stavisky et al., 2017; Athalye et al., 2018; Vyas et al., 2018; Mitani et al., 2018). This paradigm allows one to understand how changes in neural activity during learning lead to improved behavior (Golub et al., 2016; Orsborn and Pesaran, 2017). While BCI has certain limitations for studying the full learning system—such as the absence of natural sensory feedback (e.g., proprioception) and involvement of areas downstream of cortex (e.g., spinal cord)—BCI is known to engage similar cognitive processes and brain areas as arm reaching (Golub et al., 2016). Thus, BCI provides a powerful tool for studying how neural activity reorganizes in the brain during learning. In our BCI learning studies, monkeys are trained to modulate the activity of neurons in motor cortex to control the velocity of a computer cursor on a screen (Sadtler et al., 2014; Zhou et al., 2019; Oby et al., 2019) (Figure 2A). At each moment in time (e.g., a 45 ms time window), the cursor's velocity, $\mathbf{v} \in \mathbb{R}^2$, is determined by the simultaneous spiking activity of K neural units, $\mathbf{u} \in \mathbb{R}^K$. For simplicity, suppose we have a linear relationship, with $\mathbf{v} = M\mathbf{u}$, where $M \in \mathbb{R}^{2 \times K}$ is the BCI mapping. On each trial, the subject is rewarded for moving the cursor to acquire a cued visual target. Subjects first use an “intuitive” BCI mapping, chosen so that subjects can proficiently control the cursor. To induce learning, the experimenter changes the BCI mapping, requiring the subject to modify the neural activity they produce in order to restore proficient behavior. Knowing the BCI mapping allows us to link neural population activity to performance during learning. Consider each moment in time when the subject wants to move the cursor toward a cued target. Early in learning (Figure 2B), subjects generate cursor velocities (blue arrow) that do not point directly toward the target, leading to indirect cursor trajectories (gray line), and thus less frequent rewards. Later in learning (Figure 2C), subjects produce more direct cursor movements, resulting in a higher reward rate. To quantify the subject's moment-by-moment performance, consider the speed of the cursor's velocity in the direction of the target at some time t : $p_t = \mathbf{v}_t^T \mathbf{x}_t \in \mathbb{R}$, where $\mathbf{x}_t \in \mathbb{R}^2$ is a unit vector pointing from the cursor to the target. Because cursor velocity is a function of the neural activity, we can rewrite this quantity directly in terms of neural activity: $p_t = (M\mathbf{u}_t)^T \mathbf{x}_t$. This link between neural activity and performance lets us think concretely about how neural activity might evolve during learning according to different hypotheses. By considering neural activity and performance in a joint “neural performance” space (Figure 2D), we can ask how neural activity might change as a function of the feedback received on each trial. For example, suppose that during trial t , the subject generates the neural population activity pattern \mathbf{u}_t . The subject receives visual feedback about the performance, p_t , that resulted from generating \mathbf{u}_t . To improve their performance on subsequent trials, they must use this feedback to modify their neural activity in some direction in neural activity space, δ (Figure 2E). Subjects may learn to control a new BCI mapping in a process akin to supervised learning, reinforcement learning, or both. In any case, as long as δ moves the neural activity “uphill” along the performance axis, performance will improve on the next trial. Repeating this process for multiple trials will lead to changes in neural population activity that gradually lead to improvements in performance. Thus, the causal link between neural activity and performance provided by a BCI paradigm can reveal important clues about how the brain learns. As we explain in later sections, actual changes in neural population activity during BCI learning do not always align with what we might expect from this purely optimization-based perspective.

Constraints on neural variability limit performance

One line of evidence that the structure of population covariability can be inflexible comes from BCI learning studies (Sadtler et al., 2014). In these experiments, population activity recorded in primary motor cortex (M1) controlled the movement of a cursor on a screen according to a BCI mapping (Figure 2A). The utility of a BCI in this study was that it provided a direct causal test of whether subjects could learn to modify their neural population covariance structure. That is, by designing particular BCI mappings, the experimenters could challenge subjects to break that structure, and observe whether or not they were able to do so. Within a single day, subjects showed more performance improvements for the BCI mappings that were aligned with dimensions of population activity that reflected shared variability among neurons (Sadtler et al., 2014). Constraints on population covariability limited which neural activity patterns subjects could generate, which in turn impacted which BCI mappings could be easily learned (Golub et al., 2018). These constraints were evident not only across trials with different intended movement

directions, but also across trials with the same intended movement (Hennig et al., 2018). Being able to modify the population covariance structure during learning has other advantages beyond improving task performance, such as minimizing energy use. But in fact, the high-dimensional structure of population covariability changed remarkably little before and after learning, regardless of these considerations (Hennig et al., 2018). Overall, these results show that the structure of population covariability may not be readily modifiable, despite the apparent benefits of doing so.

Studies of perceptual learning provide another line of evidence that the structure of population covariability can be somewhat inflexible, even when it interferes with task performance. This relates to the existence of population covariability 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). When subjects perform a familiar task (i.e., when there is no learning pressure), neural population activity exhibits substantial trial-to-trial covariability across trials

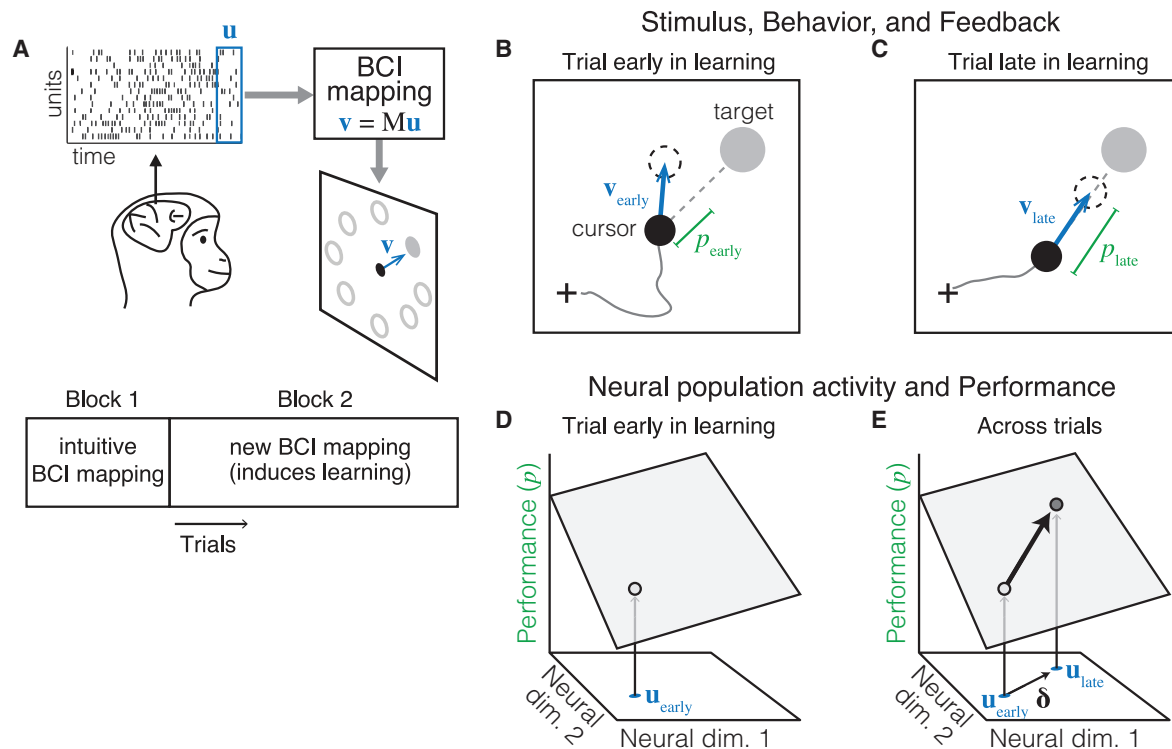


Figure 2. Studying learning using a brain-computer interface (BCI) paradigm

(A) Monkeys guide a computer cursor to one of eight visual targets by modulating the spiking activity of ~90 neurons recorded in primary motor cortex. The relationship between neural activity and the cursor's velocity is defined by the BCI mapping. When a new BCI mapping is introduced, subjects must learn to modify their neural activity to improve performance (i.e., guide the cursor to the target).

(B) Cursor position (black circle), target position (gray circle), and cursor velocity (v_{early} , blue arrow) at a single time step during a trial early in learning. Performance (p_{early} , green line) can be quantified as the cursor velocity projected onto the line connecting the cursor and the target (dotted gray line; not visible to monkey). Cursor positions at previous time steps (gray curve) and at the next time step (dashed circle) are shown for reference.

(C) Later in learning, monkeys are able to generate cursor velocities that move the cursor straighter toward the target, resulting in improved performance. Same conventions as (B).

(D) Visualizing neural activity and performance during the trial shown in (B). The relationship between neural activity and performance (given by the BCI mapping) is depicted as a gray plane. When the subject generates neural activity, u_{early} , they receive visual feedback about performance (e.g., the cursor's speed toward the target direction, p_{early}), which can be used to guide improvements to their neural activity on future trials.

(E) During learning, monkeys modify their average neural activity in a particular direction δ (black arrow). This change will lead to improved performance if the projection of the arrow onto the performance plane points "uphill." Same conventions as (D).

with the same task condition. When these correlated fluctuations in firing activity are aligned with stimulus-encoding dimensions of the population, they can potentially limit behavioral performance. While the magnitude of information-limiting correlations can be reduced with learning (Gu et al., 2011; Jeanne et al., 2013; Ni et al., 2018), these correlations persist even in over-trained tasks (Rumyantsev et al., 2020; Bartolo et al., 2020b).

The structure of neural variability can influence the path of learning

Even if the structure of neural covariability is largely fixed, it may nevertheless be a critical component of how neural population activity evolves over time during learning. To explain why this is the case, consider that the brain must use feedback from the task to estimate the direction, or gradient, in which it should modify its activity in order to improve future performance. In an ANN, this gradient can be calculated using the chain rule from calculus, in an algorithm called backpropagation. For the brain, however, there are a number of reasons that backpropagation may be biologically implausible (Bengio et al., 2015; Lillicrap et

al., 2020). If the brain has a different way of estimating gradients, perhaps the structure of neural covariability can provide some clues. For example, suppose the brain estimates the gradient by exploring the outcomes of different neural activity patterns (Figures 3A and 3B). (This is similar to the approach taken by many "black-box" learning rules, including weight perturbation, node perturbation, REINFORCE, and evolution strategies [Williams, 1992; Werfel et al., 2005; Wierstra et al., 2014; Salimans et al., 2017].) In this case, the gradient might be estimated as a weighted sum of the sampled neural population activity patterns, with weights given by their resulting performance (Figure 3B). From a statistical perspective, for this gradient estimate to be unbiased, it must be multiplied by the inverted covariance matrix that generated the sampled population activity patterns. Thus, if the neural circuit is unable to account for the covariance, the resulting gradient estimate will not point directly uphill (Figures 3C and 3D). This dependence will then influence how neural population activity evolves throughout learning (Figure 3E). A similar outcome can occur even if gradients are calculated directly (e.g.,

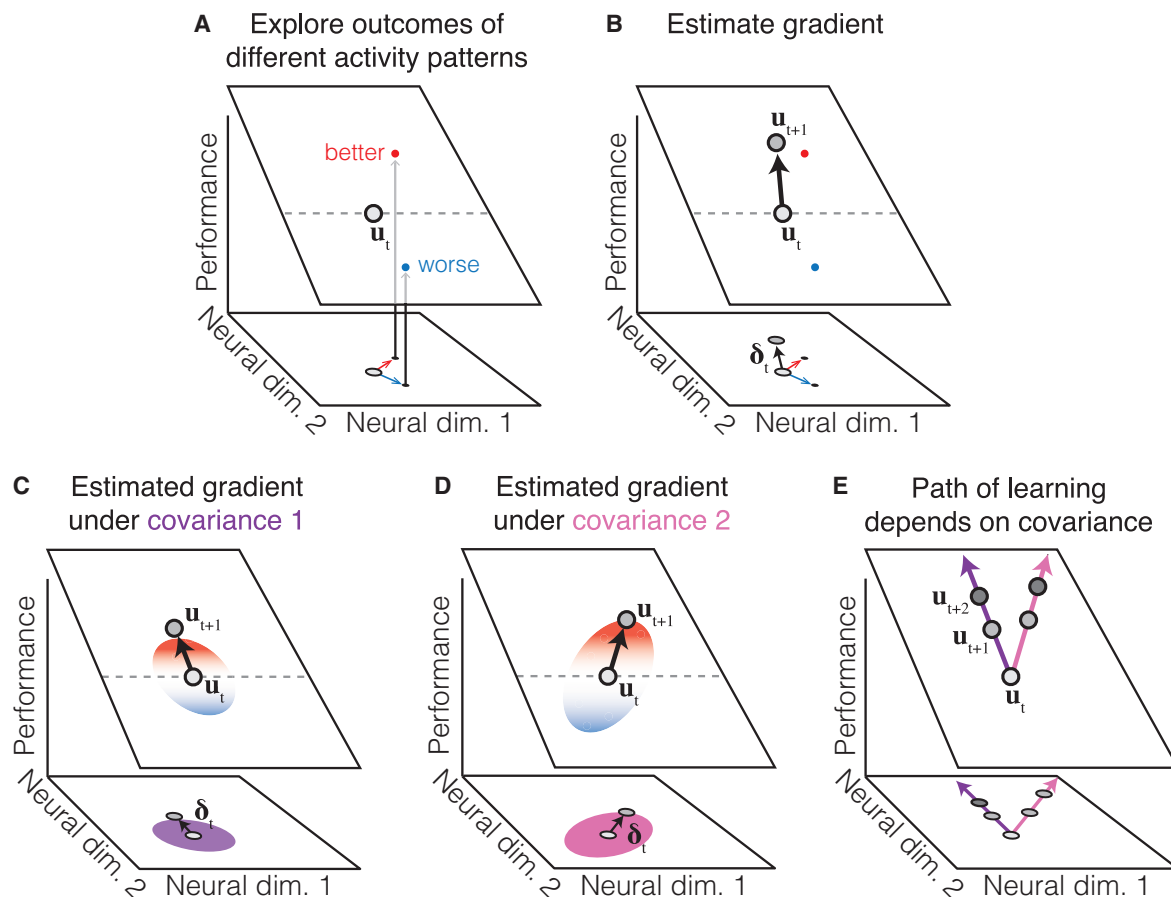


Figure 3. Neural population covariability influences the path of learning

(A) During learning, samples of population activity (black dots), drawn from an underlying covariance structure, may help the brain explore the relationship between neural activity patterns and performance (where the plane depicts this relationship, similar to Figures 2D and 2E). Relative to the average activity (\mathbf{u}_t), some activity patterns may result in improved performance (red dot), while other activity patterns may make performance worse (blue dot). Gray dashed line depicts points on the plane with the same performance as \mathbf{u}_t . Here we suppose that on trial t we sample activity patterns as $\mathbf{u}_t(j) = \mathbf{u}_t + \mathbf{e}_t(j)$, for $j = 1, \dots, J$, with $\mathbf{e}_t(j) \sim \mathcal{N}(\mathbf{0}, \Sigma)$. Let p_j be the performance associated with $\mathbf{u}_t(j)$.

(B) To improve performance on subsequent trials, the brain may use its past experience (e.g., the red and blue dots) to move its activity in a particular direction (black arrow, δ_t). In this example, δ_t weights each explored activity pattern by its performance: $\delta_t = \sum_j p_j \mathbf{e}_t(j)$ (see text). Thus, δ_t moves toward the red arrow and away from the blue arrow.

(C and D) Under the scheme illustrated in (B), differences in the sampling distribution of neural activity patterns (purple and pink ellipses depict the sampling covariance in C and D, respectively) can lead to different neural activity changes (black arrows, δ_t), even when the relationship between neural activity and performance (depicted by the plane) is identical. Red-blue color scale depicts the relative performance of different activity patterns within each covariance ellipse.

(E) Schematic of repeating the process in (C) and (D) for multiple trials. Changes in neural activity during learning may thus take a path that depends on the covariance structure.

in a manner akin to backpropagation). These considerations suggest that understanding how the structure of neural covariability is accounted for by the learning process may provide insights into the observed path of learning.

Differences in network variability between biological and artificial networks

Neural variability in the brain is typically quite different from that in ANNs. First, as detailed earlier, population covariability cannot necessarily be attenuated or restructured by the learning process, even if doing so would improve task performance. This stands in contrast to what one might expect of an ANN trained to perform a given task, where typically all the parameters of the network are adaptable for the purposes of optimization. More generally, identifying which aspects of biological networks

are unchanged during learning can provide clues for developing better network models of learning. For example, one might consider which aspects of an ANN could be optimized such that learning a new task leaves the population covariance structure unchanged. Second, consider the presence of substantial population covariability across trials with the same intended movement, as exhibited in the BCI learning task. This form of covariability would not naturally occur in a rate-based ANN (i.e., a network whose units have continuous-valued outputs), because an ANN's activity given a fixed input is deterministic. This suggests that if ANNs are to serve as models of how the brain changes during learning, we may need to consider classes of artificial networks that can capture the properties of population covariability described above, such as stochastic rate-based

ANNs (Hennequin et al., 2018; Echeveste et al., 2020) or spiking neural networks (van Vreeswijk and Sompolinsky, 1998; Savin and Deneve, 2014; Litwin-Kumar and Doiron, 2014).

The inflexibility of neural variability in the brain during learning can also be contrasted with the typical role of variability in methods for reinforcement learning (RL). For example, consider deep RL, where an agent approximates particular functions using an ANN, with the outputs of the ANN determining the agent's current best guess of the optimal behavior (Li, 2018). We can think of the activations of the ANN as akin to the activity of a population of neurons in the brain (Botvinick et al., 2020). As discussed earlier, RL agents discover optimal behavioral policies by effectively exploring their environment, where exploration manifests as behavioral variability. However, variability in an artificial agent's network activity during learning is likely distinct from the variability observed in the brain. For one, RL agents' behavioral variability is often added to the outputs after the fact, as in the classical "ε-greedy" strategy, where with probability ε the agent takes a purely random action, ignoring its current policy (Sutton and Barto, 2018). In this case, there is behavioral variability without any concurrent network variability, whereas in animals, behavioral variability is driven by neural variability. Second, even when RL agents do exhibit network variability, the amount of variability is typically chosen by the modeler based on the task and attenuated throughout training or when evaluating performance (Mnih et al., 2015; Salimans et al., 2017; Plappert et al., 2017). By contrast, as discussed above, population covariability in the brain may be largely inflexible and unattenuated, and likely only some of this variability reflects active exploration. Understanding how learning proceeds alongside, or even in spite of, substantial neural variability may inform the development of RL agents that can learn complex tasks without precisely tuning or attenuating the amount of network variability.

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

Learning in the brain involves multiple learning processes, even in simple tasks

A major hurdle to understanding learning in the brain is to identify the learning objectives underlying the observed neural and behavioral changes. In an artificial network, by contrast, this is not an issue because the modeler decides how feedback from the task is used to drive learning. For example, consider teaching an ANN and a child to identify which animals are shown in different images. For the ANN, the modeler can choose the objective function (e.g., cross-entropy) that the ANN tries to optimize during the training process. But when teaching a child, while the teacher may have a particular task objective in mind (e.g., naming the animals correctly), there is no guarantee that this is the objective driving changes in the child's brain during learning.

This issue of identifying the objective underlying learning is complicated by the fact that, even in simple tasks, behavioral changes during learning appear to reflect multiple learning processes (McDougle et al., 2016; Marblestone et al., 2016; Neftci and Averbeck, 2019; Morehead and de Xivry, 2021). Here we define a "learning process" as an optimization process with its own objective function, learning rule, and/or instantiation in neu-

ral circuitry. Because behavior is multi-faceted, there are multiple aspects of behavior that might be optimized to achieve a given task goal. Consider the seemingly simple task of moving one's hand toward an object. Behavioral studies of adaptation have identified a number of different learning processes used by the brain to solve this task (McDougle et al., 2016; Morehead and de Xivry, 2021). For example, some learning processes have been distinguished behaviorally based on the type of feedback that engages them. In the case of reaching toward an object, some learning processes use visual or proprioceptive feedback to minimize errors in motor execution (e.g., so that the intended movement direction of the hand matches its actual movement direction) (Tseng et al., 2007), in a manner akin to supervised learning. Other processes seem to be driven only by whether a trial was successful or not (Vaswani et al., 2015), akin to reinforcement learning. More generally, it has been proposed that the brain learns new motor skills via supervised learning, unsupervised learning, and reinforcement—perhaps even simultaneously (Doya, 2000; Izawa and Shadmehr, 2011). These findings and others illustrate that, even for the simple task of reaching one's hand toward an object, behavioral changes appear to be driven by a variety of different learning processes.

Disentangling the presence of multiple learning processes in neural population activity

As we discussed above, even a simple task such as reaching for an object may involve multiple learning processes. Although this may be beneficial from the brain's perspective, it poses a challenge for the neuroscientists seeking to identify the neural substrates of these learning processes. Neural signatures of multiple learning processes have been identified in a variety of different tasks, including motor learning (Zhou et al., 2019; Oby et al., 2019), perceptual learning (Poort et al., 2015), rule learning (Sigurdardottir and Sheinberg, 2015), and reinforcement learning (Cohen et al., 2015; Schultz, 2007, 2019). While these processes may originate in distinct brain areas, signatures of multiple learning processes can also coexist within the same population of neurons. When this happens, how can we, as neuroscientists, tease them apart?

One of the primary means of teasing apart the influence of distinct learning processes in the brain is to identify distinct timescales of behavioral or neural change (Schultz, 2007; Lak et al., 2016; Schultz, 2019; Zhou et al., 2019; Morehead and de Xivry, 2021). For example, after subjects are first introduced to a new task, one might observe rapid performance improvements due to learning the rules and basic strategies, while mastery may take weeks or even years (Newell and Rosenbloom, 1981). While observing multiple timescales of behavioral or neural change does not on its own guarantee the presence of distinct learning processes (Newell et al., 2009), it can serve as a clue that there may indeed be multiple learning processes at play. BCI learning paradigms similar to those mentioned in the previous section have begun to identify the neural signatures distinguishing different learning processes in motor cortex. 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 improving distinct aspects of task performance (directional errors and reward rate, respectively).

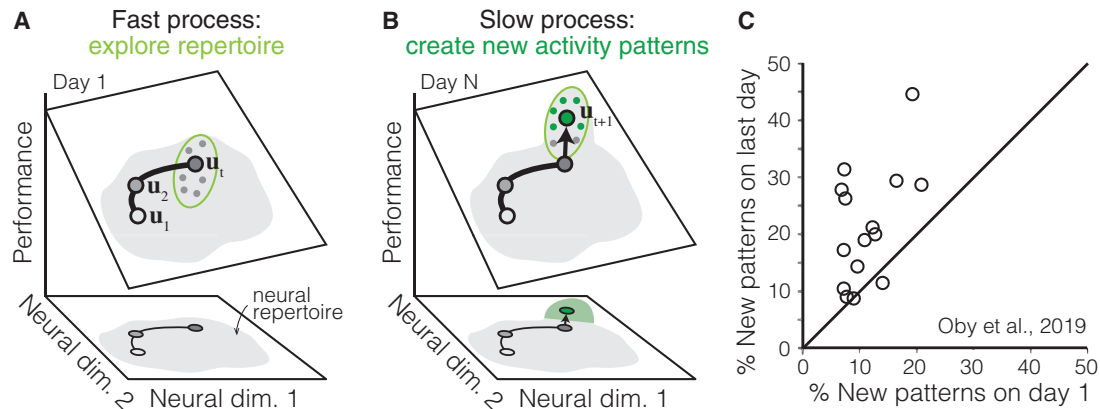


Figure 4. Fast and slow learning processes during BCI learning

(A) Within a day, the set of different neural activity patterns (“neural repertoire”) that subjects can generate may be largely fixed. Thus, the initial stages of learning may consist of choosing which of the different activity patterns within this neural repertoire will lead to maximum performance (Golub et al., 2018).

(B) Given multiple days of practice, new neural activity patterns (green dots) emerge, expanding the neural repertoire in ways that lead to improved behavior (Oby et al., 2019).

(C) Comparing the percentage of new neural activity patterns observed on the first versus last day of learning a new BCI mapping. Adapted from Oby et al. (2019).

Another study of long-term learning using a BCI paradigm found that, given weeks of practice, subjects learned a difficult BCI mapping by developing the capacity to generate new patterns of population activity—that is, patterns of spiking activity across the population that were not observed before the learning experience (Oby et al., 2019). In both of these studies, long-term learning appeared to be made possible by changes to the covariability in firing across neurons. This stands in contrast to the studies of short-term BCI learning discussed in the previous section (Sadtlir et al., 2014; Golub et al., 2018; Hennig et al., 2018; Zhou et al., 2019), where the structure of neural covariability was largely unchanged before and after learning.

Together, these results suggest the presence of at least two distinct neural processes: one fast, one slow (Figure 4). The fast process, observable within a day, may be explained by changes to the inputs to the network (e.g., changes in the intended movement direction) (Golub et al., 2018; Hennig et al., 2021), a change that can presumably be achieved more quickly as it requires less network reorganization (Sohn et al., 2020). By contrast, the slow process—resulting in changes to the correlations between neurons—may take more time to complete because it requires more synaptic modifications (Wärnberg and Kumar, 2019; Feulner and Clopath, 2021). The different timescales of these neural processes may explain why some tasks can be learned within a few hours while others take weeks or years of practice.

The importance of identifying the brain's prior assumptions about learning

One of the primary differences between learning in biological and artificial agents is that, for a living creature, the environment and reward contingencies change continually (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). Inductive biases are the prior assumptions made by a learner to predict the

outcomes of novel inputs. Given the variety of different timescales and contexts in which animals need to adapt their behavior, the brain may employ a suite of inductive biases, each reflected in a different learning process. For example, learning processes with fast timescales may enable rapid learning for the tasks an animal is most likely to encounter (Braun et al., 2009; Sinz et al., 2019; Sohn et al., 2020), while those with slower timescales may enable generalization across tasks (Braun et al., 2009; Wang et al., 2018a).

Thus, understanding learning in the brain may involve not only a description of the multitude of learning processes engaged by the task, but also a characterization of the brain's inductive biases. We provide two examples of cases where identifying the brain's inductive biases is a critical part of the puzzle for understanding changes in neural activity during learning. First, 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; de Rugy et al., 2012; Zhou et al., 2019; Gardner, 2019; Rosenberg et al., 2021; Mochol et al., 2021). While these heuristics may enable rapid learning in many tasks, other tasks—even those with simple rules—may be surprisingly difficult to learn (e.g., mirror reversal learning in primates, or decision tasks in mice) (Hadjiosif et al., 2021; Rosenberg et al., 2021). In these cases, it may be that the brain is making incorrect assumptions about how to solve the task. These incorrect assumptions may be reflected in terms of how the brain estimates the gradient of neural activity with respect to errors in the task (Hadjiosif et al., 2021). Second, as mentioned earlier, the objective specified by a task is not necessarily the objective used by the subject to learn that task. Differences in these objectives may influence how a task is learned (Gershman and Niv, 2013). ANN models can help test hypotheses regarding the relevant objective. For example, predicting human performance at a particular task using an ANN can require first training the ANN to perform a *different* task (Nicholson and Prinz, 2021), presumably because

the pre-training encourages the network to acquire the relevant inductive biases. Overall, these results suggest that the brain's inductive biases may be just as critical as the task objective itself when interpreting changes in population activity during learning.

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

"Task-nonspecific" changes in population activity can interact with the learning process

During learning, neural activity changes to improve behavior. One common approach for understanding this process is to characterize which aspects of neural activity change, or do not change, during or after a learning experience (Schoups et al., 2001; Li et al., 2001; Law and Gold, 2008; Gu et al., 2011; Uka et al., 2012; Jeanne et al., 2013; Peters et al., 2014; Yan et al., 2014; Makino and Komiyama, 2015; Golub et al., 2018; Perich et al., 2018; Singh et al., 2019). This approach can offer clues as to how different parts of the brain contribute to improvements in task performance. For example, following perceptual learning, changes in earlier visual areas may suggest improved representations of stimuli shown in the task (Schoups et al., 2001; Yan et al., 2014), while changes in downstream areas may suggest improved readout of those representations (Law and Gold, 2008; Uka et al., 2012).

But not all changes in neural activity during learning are directly interpretable in terms of the task objective. As a growing list of studies has observed, changes in neural activity during learning are not always driven by performance considerations (Rokni et al., 2007; Okun et al., 2012; Singh et al., 2019; Hennig et al., 2021). We refer to these changes, which appear to occur without regard to the details of the task at hand, as being "task-nonspecific." Operationally, we consider changes in population activity to be task nonspecific if they occur with or without any learning pressure, or if they occur despite making performance worse. For example, even when an animal is performing a familiar task, neural activity is not stable over time, but can show substantial "drift" (Druckmann and Chklovskii, 2012; Fraser and Schwartz, 2012; Liberti et al., 2016; Driscoll et al., 2017; Mau et al., 2020; Cowley et al., 2020; Schoonover et al., 2020). Task-nonspecific 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), or the influence of internal states such as changes in arousal or satiety (Musall et al., 2019a; Allen et al., 2019; Stringer et al., 2019; Cowley et al., 2020; Hennig et al., 2021). Regardless of their source, task-nonspecific changes in neural activity must be accounted for by the brain if they are not to interfere with downstream processing or behavior (Rokni et al., 2007; Clopath et al., 2017; Rule et al., 2020; Xia et al., 2021). However, as we will discuss next, task-nonspecific changes can and do impact behavior, making them an integral part of the learning process.

Changes in an animal's internal state, such as its arousal, satiety, or attention, are reflected in brainwide changes in neural activity. Because these signals are known to fluctuate over time, internal states are one potential source of task-nonspecific changes in neural population activity during learning. Importantly,

internal state changes can impact behavior even when doing so impairs performance. For example, professional athletes can "choke" under pressure, failing to succeed at practiced behaviors when the stakes are high (Baumeister, 1984; Gucciardi et al., 2010; Yu, 2015; Hsu et al., 2019). Similarly, sustained changes in satiety or fatigue represent task-nonspecific factors that have clear impacts on behavior, and thus performance. As highlighted by a recent BCI learning study (Hennig et al., 2021), task-nonspecific changes can be large and may even make up the bulk of changes in population activity during learning. In that study, population activity in primary 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 experimenter-controlled events such as pauses in the task or changes to the BCI mapping. During learning, these task-nonspecific changes in neural activity impacted the speed and the direction in which the cursor moved, which in turn impacted how neural activity evolved on subsequent trials. This result illustrates that task-nonspecific factors can make up a large part of changes in population activity during learning and that these changes can influence the path of learning via their impact on behavior.

Reverse-engineering learning in the brain in the presence of task-nonspecific changes

We suspect that accounting for task-nonspecific changes in neural activity will prove to be a critical ingredient in attempts to reverse-engineer the process by which the brain learns (e.g., Lim et al. (2015); Lak et al. (2016); Roy et al. (2018); Ashwood et al. (2020); Nayebi et al. (2020)). In particular, accounting for task-nonspecific changes in population activity may be necessary for the challenge of determining the brain's learning objective. This is true even if, rather than providing clues about how the brain learns, task-nonspecific changes are simply a nuisance factor we have to deal with when interpreting neural activity changes. For example, suppose the observed population activity changes in the opposite direction with respect to a hypothesized objective function (Richards et al., 2019). Rather than this implying that the hypothesized objective is wrong, this may instead indicate the presence of task-nonspecific changes in activity, which may drive neural activity independently of the proposed learning objective. In this case, changes in neural population activity during learning may be best thought of as the combination of task-specific (e.g., related to the learning objective) and task-nonspecific influences (Figure 5). This idea may even explain why performance sometimes fails to improve during learning: in these situations, the task-nonspecific influence on neural activity may be overpowering the task-specific influence.

Overall, a complete description of how neural activity changes during learning may require accounting for the influence of various task-nonspecific processes (e.g., arousal) the brain manages concurrently with learning. This may be a challenge given that many task-nonspecific processes are often not under experimental control, resulting in these processes having a variable and latent influence on population activity. One potential approach to this problem involves leveraging natural behavior. Animals often show rich, uncued behaviors such as fidgeting even during over-trained tasks (Musall et al., 2019a). These uncued movements are themselves modulated by an animal's

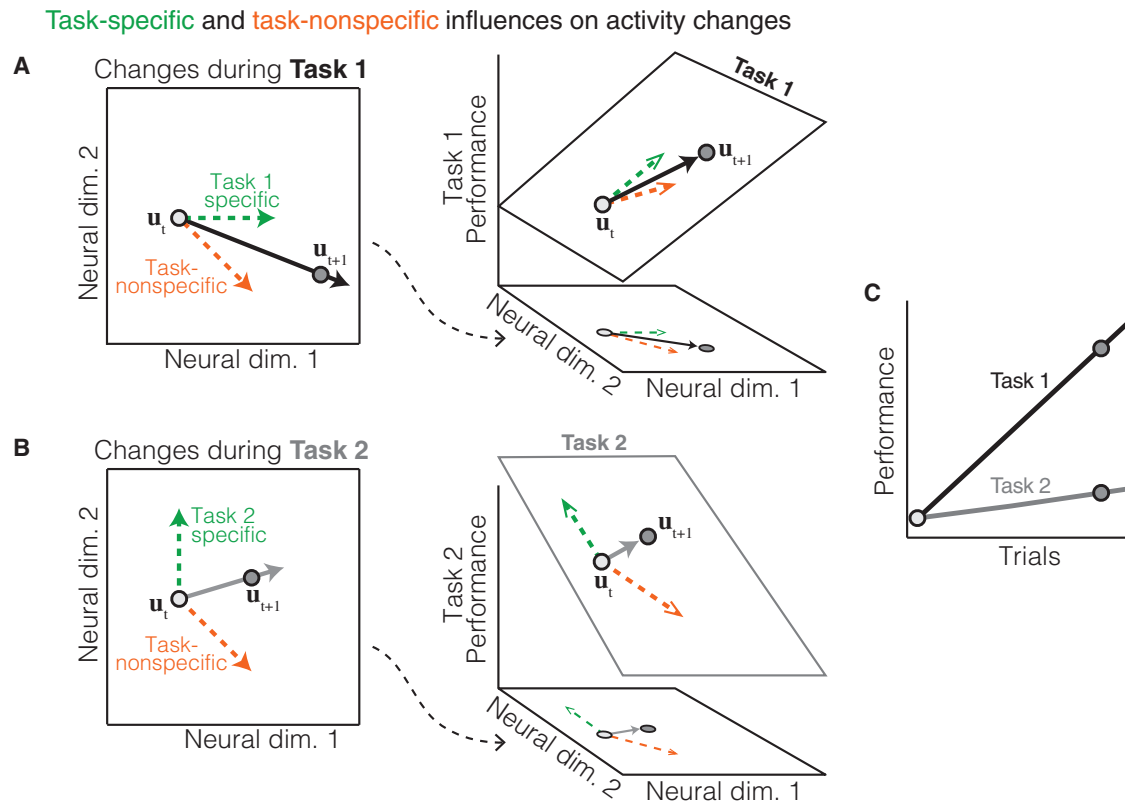


Figure 5. Task-specific and task-nonspecific changes in neural population activity during learning

(A and B) During learning, changes in neural population activity (black arrow) may be the sum of both task-specific (green arrow) and task-nonspecific (orange arrow) processes. In these schematics, the task-specific change in neural activity depends on the relationship between neural activity and performance (as in Figure 3). For example, in Task 1 (A), performance is a function only of neural dimension 1, while in Task 2 (B), performance is a function only of neural dimension 2. By contrast, the task-nonspecific change occurs in the same direction regardless of the task. The task-specific and nonspecific changes then combine to determine how neural population activity evolves with learning.

(C) Performance may improve more quickly with learning when the task-nonspecific process is more aligned with the task-specific process (as in Task 1) than when it is less aligned with the task-specific process (as in Task 2).

internal state such as its arousal or satiety (Musall et al., 2019a; Allen et al., 2019; Stringer et al., 2019). Thus, one might be able to use these behaviors to infer an animal's internal state, and its resulting influence on population activity (Musall et al., 2019b). Experimentally, this would involve monitoring more aspects of an animal's behavior that are not directly related to the task, such as recording pupil size during an arm movement task (Hennig et al., 2021) or movements during a decision-making task (Musall et al., 2019a). Thus, continuous whole-body behavioral monitoring, alongside statistical methods for relating behavior to animals' latent internal states, may provide a means of accounting for task-nonspecific changes in population activity during learning.

Potential benefits of task-nonspecific activity

Task-nonspecific changes in the brain during learning can be contrasted with network activity in ANNs, where task-nonspecific changes are typically small or absent. For example, in an ANN trained by gradient descent, changes in network activity are determined by the learning rule (e.g., given by the gradient of the task objective), such that all changes are, by construction, task specific. By contrast, as we saw above, a substantial pro-

portion of changes in neural population activity during learning may have little to do with the task at hand.

Why might the brain exhibit such large task-nonspecific changes? Answering this question may help us build better models of biological learning, while also inspiring new methods for artificial learning. When performing any given task, the brain must also manage a whole host of processes that are not directly related to task execution, such as arousal, attention, and memory. These processes may have their own objectives, which, while relevant to whatever process they subserve, are nevertheless irrelevant to executing the task itself, and therefore drive task-nonspecific changes in population activity. As one example of this, a recent study had subjects reach toward different targets in a perturbed environment, where some but not all targets required learning (Sun et al., 2020). The subjects' population activity in motor cortex exhibited a "uniform shift" across all targets, including those that did not require learning. In contrast to the task-nonspecific changes discussed earlier (Hennig et al., 2021), this uniform shift in population activity did not appear to influence motor output, indicating that task-nonspecific changes need not always impact performance. Because

the shift persisted even when the perturbation was removed, the authors propose the shift in activity may help maintain, or “index,” a motor memory of the learned experience. In other words, the task-nonspecific changes observed in [Sun et al. \(2020\)](#) may reflect the influence of a memory-related objective, perhaps one that ensures the newly acquired memory is maintained in population activity even when subjects go on to perform different tasks. This may serve as a mechanism for preventing catastrophic forgetting—that is, learning new tasks without overwriting previously learned tasks—a well-known issue for ANNs ([French, 1999](#)), and one of the key goals of methods for lifelong or continual learning ([Chen and Liu, 2018](#)).

Another potential benefit of task-nonspecific changes in the brain is that they may encourage robustness—by making the learner insensitive to large variance changes in network activity—or provide a means of escaping local optima or saddle points. In both cases, the brain may be willing to accept any negative impacts on performance caused by the task-nonspecific changes in the short-term, so as to achieve a more robust or optimal solution in the long run. Overall, understanding the role of task-nonspecific activity in the brain may inspire methods for developing flexible, robust artificial agents.

Challenges and future directions

Understanding synaptic plasticity from neural population activity

Learning in the brain is typically thought to be governed by changes in the synaptic strengths between neurons. Similarly, ANNs are trained to perform new tasks via modifications to the synaptic strengths between artificial neurons. Here we expand this view to consider learning in terms of changes in neural population activity. The differing viewpoints of studying learning in terms of plasticity versus neural population activity raises several potential questions for future study. First, what governs whether a given task is learned via plasticity? While learning may often require plasticity, this does not always have to be the case ([Mayo and Smith, 2017](#); [Wang et al., 2018a](#); [Sohn et al., 2020](#)). In particular, learning may sometimes unfold as changes in neural population activity without any corresponding changes in synaptic strengths. This might be possible if, for example, the underlying neural circuit always tries to counteract any changes in sensory feedback that indicate errors ([Sohn et al., 2020](#)). In other cases, improving performance may require plasticity. Perhaps the brain learns via plasticity only when the benefits of doing so offset the associated metabolic cost of adding or modifying synapses.

Second, which aspects of learning can be understood using only recordings of neural population activity? Because changes in synaptic strength influence downstream behavior via their influence on the spiking activity of postsynaptic neurons, recordings of neural population activity during learning may offer a sufficient view of how the brain modifies its activity to drive improved behavior. This would be beneficial from a technical standpoint given that measuring plasticity *in vivo* in single neurons is challenging, and methods for measuring plasticity across multiple neurons simultaneously are still in their infancy. By contrast, studies incorporating recordings of the simultaneous postsynaptic spiking activity of hundreds or even thousands of

neurons are increasingly common ([Ahrens et al., 2012](#); [Steinmetz et al., 2019](#); [Stringer et al., 2019](#); [Bartolo et al., 2020a](#); [Urai et al., 2021](#)).

Third, can leveraging studies of neural population activity during learning guide the search for new plasticity rules? For example, numerous studies have observed that population activity shows low-dimensional structure, and that this structure is highly conserved during learning on short timescales ([Chase et al., 2010](#); [Hwang et al., 2013](#); [Sadler et al., 2014](#)). Understanding how plasticity, which occurs at the level of single neurons, can maintain low-dimensional structure at the level of neural population activity is not easily explained by our current understanding of plasticity rules. Overall, studying neural population activity during learning offers a unique viewpoint of the neurobiological changes driving learning.

Understanding incomplete or suboptimal learning

One puzzling aspect of biological learning that may benefit from ideas in mathematical optimization and machine learning is the question of why biological learning is often incomplete. For example, performance at some tasks often converges to a sub-optimal level ([Vaswani et al., 2015](#); [Golub et al., 2018](#); [Langsdorf et al., 2021](#)), and sometimes learning fails to occur at all. If learning in the brain occurs via an optimization process, how are we to understand these failure modes? We have mentioned earlier two possibilities: mismatches between the brain’s inductive bias and task demands, and the presence of task-nonspecific changes in neural activity. Other possible explanations might come from machine learning, where the causes of incomplete or failed learning have been studied extensively. For example, the failure of an artificial network to learn can be due to premature convergence, the existence of local optima or “good enough” solutions, or other reasons. These ideas may provide hypotheses for understanding suboptimal learning in the brain. Another promising direction is to build probabilistic models of different learning hypotheses and use statistical inference to infer the (potentially suboptimal) principles underlying animal learning directly from data ([Linderman and Gershman, 2017](#); [Roy et al., 2018](#); [Ashwood et al., 2020, 2021](#)).

Using ANNs as artificial model organisms for understanding the brain

Using ANNs as artificial model organisms may be a key tool for understanding how the brain learns via optimization. Here we sketch out different approaches one might take. First, ANNs can serve as testbeds for developing new methods that infer the components of the optimization framework from recordings of neural activity. This is possible because we can analyze the activity of ANNs just like we analyze neural recordings. For example, recent work demonstrated how one might infer a network’s learning rule from neural population activity, regardless of the network’s architecture or learning objective ([Nayebi et al., 2020](#)).

Second, ANNs can shed light on how the components of the optimization framework interact. For example, recent work suggests it might not be possible to identify the brain’s objective function without also considering its architecture, as these two components interact when determining a network’s activity ([Bakhtiari et al., 2021](#)). This raises the question of whether building ANNs with architectures even more similar to the brain—for

example, networks incorporating distinct cell types (Doty et al., 2021), or even networks whose architectures match an identified connectome (Yan et al., 2017; Morales and Froese, 2019)—will be necessary for understanding how population activity changes with learning.

Finally, ANNs may help us build models of how learning proceeds throughout an animal's lifetime. For example, learning in a mature animal may resemble an ANN changing its synaptic weights in a network with a fixed architecture. By contrast, learning early in life, when a huge number of synapses are grown and then pruned, may more resemble an ANN where the architecture itself changes throughout learning (Millán et al., 2018; Elsken et al., 2019; Gaier and Ha, 2019). Developing a stronger parallel between learning in ANNs and brains has the potential to benefit both our understanding of the brain and methods for artificial learning.

Conclusion

To learn, the brain must discover the changes in neural activity that lead to improved behavior. How can we begin to understand the complex processes in the brain that drive these changes? The optimization framework for learning suggests that changes in neural activity during learning might be understood as the natural outcome of an objective function, learning rule, and network architecture. As we propose here, applying the optimization framework to biological networks requires us to focus on the key ways in which neural activity in the brain differs from network activity in typical artificial networks. In particular, we have presented three key observations about learning from studies of neural population activity that we believe need to be accounted for by models of learning in the brain: (1) the inflexibility of neural variability throughout learning, (2) the use of multiple learning processes even during simple tasks, and (3) the presence of large task-nonspecific activity changes. These challenges were readily apparent when considering neural population activity but would have been harder to detect from the vantage points either of synaptic weight changes or single-unit tuning properties.

The optimization framework is a promising starting point for understanding learning in the brain. But as we have seen, even in relatively simple tasks, changes in neural population activity during learning are not always easy to interpret as an optimization process. This difficulty may be even more salient in the context of understanding how more complex, naturalistic tasks are learned. Moving forward, accounting for the features of population activity described here into new computational models of learning in the brain and new experimental designs may be a useful next step for reverse engineering the process by which the brain learns.

ACKNOWLEDGMENTS

The authors would like to thank Patrick Mayo and Leila Wehbe for their helpful feedback on the manuscript. This work was supported by the Richard King Mellon Presidential Fellowship (J.A.H.), Carnegie Prize Fellowship in Mind and Brain Sciences (J.A.H.), NSF Graduate Research Fellowship (D.M.L.), NIH R01 HD071686 (A.P.B., B.M.Y., and S.M.C.), NSF NCS BCS1533672 (S.M.C., B.M.Y., and A.P.B.), NSF CAREER award IOS1553252 (S.M.C.), NIH CRCNS R01 NS105318 (B.M.Y. and A.P.B.), NSF NCS BCS1734916

(B.M.Y.), NIH CRCNS R01 MH118929 (B.M.Y.), NIH R01 EB026953 (B.M.Y.), and Simons Foundation 543065 (B.M.Y.).

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Ahrens, M.B., Li, J.M., Orger, M.B., Robson, D.N., Schier, A.F., Engert, F., and Portugues, R. (2012). Brain-wide neuronal dynamics during motor adaptation in zebrafish. *Nature* 485, 471–477.
- Allen, W.E., Chen, M.Z., Pichamoorthy, N., Tien, R.H., Pachitariu, M., Luo, L., and Deisseroth, K. (2019). Thirst regulates motivated behavior through modulation of brainwide neural population dynamics. *Science* 364, 253.
- Ashwood, Z.C., Roy, N.A., Bak, J.H., The International Brain Laboratory, and Pillow, J.W. (2021). Inferring learning rules from animal decision-making. In *Advances in Neural Information Processing Systems, Volume 33*, H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, eds. (Curran Associates, Inc.), pp. 3442–3453.
- Ashwood, Z.C., Roy, N.A., Stone, I.R., Churchland, A.K., Pouget, A., and Pillow, J.W. (2021). Mice alternate between discrete strategies during perceptual decision-making. *bioRxiv*. Published online September 12, 2021. <https://doi.org/10.1101/2020.10.19.346353>.
- Athalye, V.R., Santos, F.J., Carmona, J.M., and Costa, R.M. (2018). Evidence for a neural law of effect. *Science* 359, 1024–1029.
- Bakhtiari, S., Mineault, P., Lillicrap, T., Pack, C., and Richards, B. (2021). The functional specialization of visual cortex emerges from training parallel pathways with self-supervised predictive learning. *bioRxiv*. <https://doi.org/10.1101/2021.06.18.448989>.
- Barak, O., Sussillo, D., Romo, R., Tsodyks, M., and Abbott, L.F. (2013). From fixed points to chaos: three models of delayed discrimination. *Prog. Neurobiol.* 103, 214–222.
- Bartolo, R., Saunders, R.C., Mitz, A.R., and Averbeck, B.B. (2020a). Dimensionality, information and learning in prefrontal cortex. *PLoS Comput. Biol.* 16, e1007514.
- Bartolo, R., Saunders, R.C., Mitz, A.R., and Averbeck, B.B. (2020b). Information-limiting correlations in large neural populations. *J. Neurosci.* 40, 1668–1678.
- Baumeister, R.F. (1984). Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *J. Pers. Soc. Psychol.* 46, 610–620.
- Beck, J.M., Ma, W.J., Pitkow, X., Latham, P.E., and Pouget, A. (2012). Not noisy, just wrong: the role of suboptimal inference in behavioral variability. *Neuron* 74, 30–39.
- Bengio, Y., Lee, D.-H., Borscheim, J., Mesnard, T., and Lin, Z. (2015). Towards biologically plausible deep learning. *arXiv*, 1502.04156.
- Botvinick, M., Wang, J.X., Dabney, W., Miller, K.J., and Kurth-Nelson, Z. (2020). Deep reinforcement learning and its neuroscientific implications. *Neuron* 107, 603–616.
- Boyden, E.S., Katoh, A., and Raymond, J.L. (2004). Cerebellum-dependent learning: the role of multiple plasticity mechanisms. *Annu. Rev. Neurosci.* 27, 581–609.
- Brainard, M.S., and Doupe, A.J. (2002). What songbirds teach us about learning. *Nature* 417, 351–358.
- Braun, D.A., Aertsen, A., Wolpert, D.M., and Mehring, C. (2009). Motor task variation induces structural learning. *Curr. Biol.* 19, 352–357.
- Brown, N., and Sandholm, T. (2019). Superhuman AI for multiplayer poker. *Science* 365, 885–890.
- Cadiou, C.F., Hong, H., Yamins, D.L., Pinto, N., Ardila, D., Solomon, E.A., Majaj, N.J., and DiCarlo, J.J. (2014). Deep neural networks rival the representation

of primate IT cortex for core visual object recognition. *PLoS Comput. Biol.* **10**, e1003963.

Chaisangmongkon, W., Swaminathan, S.K., Freedman, D.J., and Wang, X.J. (2017). Computing by robust transience: how the fronto-parietal network performs sequential, category-based decisions. *Neuron* **93**, 1504–1517.e4.

Chase, S.M., Schwartz, A.B., and Kass, R.E. (2010). Latent inputs improve estimates of neural encoding in motor cortex. *J. Neurosci.* **30**, 13873–13882.

Chen, Z., and Liu, B. (2018). Lifelong machine learning. In *Synthesis Lectures on Artificial Intelligence and Machine Learning*, R. Brachman, F. Rossi, and P. Stone, eds. (Morgan & Claypool Publishers), pp. 1–145.

Clopath, C., Bonhoeffer, T., Hübener, M., and Rose, T. (2017). Variance and invariance of neuronal long-term representations. *Philos Trans R Soc Lond B Biol Sci* **372**, 20160161. <https://doi.org/10.1098/rstb.2016.0161>.

Cohen, J.Y., Amoroso, M.W., and Uchida, N. (2015). Serotonergic neurons signal reward and punishment on multiple timescales. *eLife* **4**, e06346.

Cowley, B.R., Snyder, A.C., Acar, K., Williamson, R.C., Yu, B.M., and Smith, M.A. (2020). Slow drift of neural activity as a signature of impulsivity in macaque visual and prefrontal cortex. *Neuron* **108**, 551–567.

Cueva, C.J., and Wei, X.-X. (2018). Emergence of grid-like representations by training recurrent neural networks to perform spatial localization. *arXiv*, 1803.07770 <https://arxiv.org/abs/1803.07770>.

de Rugy, A., Loeb, G.E., and Carroll, T.J. (2012). Muscle coordination is habitual rather than optimal. *J. Neurosci.* **32**, 7384–7391.

Dhawale, A.K., Smith, M.A., and Ölveczky, B.P. (2017). The role of variability in motor learning. *Annu. Rev. Neurosci.* **40**, 479–498.

Doiron, B., Litwin-Kumar, A., Rosenbaum, R., Ocker, G.K., and Josić, K. (2016). The mechanics of state-dependent neural correlations. *Nat. Neurosci.* **19**, 383–393.

Doty, B., Mihalas, S., Arhipov, A., and Piet, A. (2021). Heterogeneous “cell types” can improve performance of deep neural networks. *bioRxiv*. <https://doi.org/10.1101/2021.06.21.449346>.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Curr. Opin. Neurobiol.* **10**, 732–739.

Driscoll, L.N., Pettit, N.L., Minderer, M., Chetih, S.N., and Harvey, C.D. (2017). Dynamic reorganization of neuronal activity patterns in parietal cortex. *Cell* **170**, 986–999.e16.

Druckmann, S., and Chklovskii, D.B. (2012). Neuronal circuits underlying persistent representations despite time varying activity. *Curr. Biol.* **22**, 2095–2103.

Echeveste, R., Aitchison, L., Hennequin, G., and Lengyel, M. (2020). Cortical-like dynamics in recurrent circuits optimized for sampling-based probabilistic inference. *Nat. Neurosci.* **23**, 1138–1149.

Elsken, T., Metzen, J.H., and Hutter, F. (2019). Neural architecture search: A survey. *J. Mach. Learn. Res.* **20**, 1–21. <https://jmlr.org/papers/v20/18-598.html>.

Engel, T.A., Chaisangmongkon, W., Freedman, D.J., and Wang, X.J. (2015). Choice-correlated activity fluctuations underlie learning of neuronal category representation. *Nat. Commun.* **6**, 6454.

Feldman, D.E. (2009). Synaptic mechanisms for plasticity in neocortex. *Annu. Rev. Neurosci.* **32**, 33–55.

Feldman, D.E. (2012). The spike-timing dependence of plasticity. *Neuron* **75**, 556–571.

Feulner, B., and Clopath, C. (2021). Neural manifold under plasticity in a goal driven learning behaviour. *PLoS Comput. Biol.* **17**, e1008621.

Fiete, I.R., Fee, M.S., and Seung, H.S. (2007). Model of birdsong learning based on gradient estimation by dynamic perturbation of neural conductances. *J. Neurophysiol.* **98**, 2038–2057.

Fraser, G.W., and Schwartz, A.B. (2012). Recording from the same neurons chronically in motor cortex. *J. Neurophysiol.* **107**, 1970–1978.

French, R.M. (1999). Catastrophic forgetting in connectionist networks. *Trends Cogn. Sci.* **3**, 128–135.

Gaier, A., and Ha, D. (2019). Weight agnostic neural networks. *arXiv*, 1906.04358 <https://arxiv.org/abs/1906.04358>.

Ganguly, K., and Carmena, J.M. (2009). Emergence of a stable cortical map for neuroprosthetic control. *PLOS Biology* **7**, 1–13.

Gardner, J.L. (2019). Optimality and heuristics in perceptual neuroscience. *Nat. Neurosci.* **22**, 514–523.

Gershman, S.J., and Niv, Y. (2013). Perceptual estimation obeys Occam’s razor. *Front. Psychol.* **4**, 623.

Golub, M.D., Chase, S.M., Batista, A.P., and Yu, B.M. (2016). Brain-computer interfaces for dissecting cognitive processes underlying sensorimotor control. *Curr. Opin. Neurobiol.* **37**, 53–58.

Golub, M.D., Sadtler, P.T., Oby, E.R., Quick, K.M., Ryu, S.I., Tyler-Kabara, E.C., Batista, A.P., Chase, S.M., and Yu, B.M. (2018). Learning by neural reassociation. *Nat Neurosci* **21**, 1546–1726.

Gu, Y., Liu, S., Fetsch, C.R., Yang, Y., Fok, S., Sunkara, A., DeAngelis, G.C., and Angelaki, D.E. (2011). Perceptual learning reduces interneuronal correlations in macaque visual cortex. *Neuron* **71**, 750–761.

Gucciardi, D.F., Longbottom, J.L., Jackson, B., and Dimmock, J.A. (2010). Experienced golfers’ perspectives on choking under pressure. *J. Sport Exerc. Psychol.* **32**, 61–83.

Hadjiosif, A.M., Krakauer, J.W., and Haith, A.M. (2021). Did we get sensorimotor adaptation wrong? Implicit adaptation as direct policy updating rather than forward-model-based learning. *J Neurosci* **41**, 2747–2761.

Haesemeyer, M., Schier, A.F., and Engert, F. (2019). Convergent temperature representations in artificial and biological neural networks. *Neuron* **103**, 1123–1134.e6.

Haith, A.M., and Krakauer, J.W. (2013). Model-based and model-free mechanisms of human motor learning. In *Progress in motor control*, M.J. Richardson, M.A. Riley, and K. Shockley, eds. (Springer), pp. 1–21.

Hennequin, G., Ahmadian, Y., Rubin, D.B., Lengyel, M., and Miller, K.D. (2018). The dynamical regime of sensory cortex: stable dynamics around a single stimulus-tuned attractor account for patterns of noise variability. *Neuron* **98**, 846–860.e5.

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., Yu, B.M., and Chase, S.M. (2018). Constraints on neural redundancy. *eLife* **7**, e36774.

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 Yu, B.M. (2021). Learning is shaped by abrupt changes in neural engagement. *Nat Neurosci* **24**, 727–736.

Holtmaat, A., and Svoboda, K. (2009). Experience-dependent structural synaptic plasticity in the mammalian brain. *Nat. Rev. Neurosci.* **10**, 647–658.

Hsu, N.-W., Liu, K.-S., and Chang, S.-C. (2019). Choking under the pressure of competition: A complete statistical investigation of pressure kicks in the NFL, 2000–2017. *PLoS ONE* **14**, e0214096.

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

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

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

Jeanne, J.M., Sharpee, T.O., and Gentner, T.Q. (2013). Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron* **78**, 352–363.

Konishi, M. (1985). Birdsong: from behavior to neuron. *Annu. Rev. Neurosci.* **8**, 125–170.

- Koralek, A.C., Jin, X., Long, J.D., 2nd, Costa, R.M., and Carmena, J.M. (2012). Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. *Nature* **483**, 331–335.
- Lak, A., Stauffer, W.R., and Schultz, W. (2016). Dopamine neurons learn relative chosen value from probabilistic rewards. *eLife* **5**, e18044.
- Lake, B.M., Ullman, T.D., Tenenbaum, J.B., and Gershman, S.J. (2017). Building machines that learn and think like people. *Behav. Brain Sci.* **40**, e253.
- Langsdorf, L., Maresch, J., Hegele, M., McDougle, S.D., and Schween, R. (2021). Prolonged response time helps eliminate residual errors in visuomotor adaptation. *Psychon Bull Rev* **28**, 834–844.
- Law, A.J., Rivlis, G., and Schieber, M.H. (2014). Rapid acquisition of novel interface control by small ensembles of arbitrarily selected primary motor cortex neurons. *J Neurophysiol* **112**, 1528–1548.
- Law, C.T., and Gold, J.I. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat. Neurosci.* **11**, 505–513.
- Li, C.S., Padoa-Schioppa, C., and Bizzi, E. (2001). Neuronal correlates of motor performance and motor learning in the primary motor cortex of monkeys adapting to an external force field. *Neuron* **30**, 593–607.
- Li, Y. (2018). Deep reinforcement learning. *arXiv*, 1810.06339 <https://arxiv.org/abs/1810.06339>.
- Liberti, W.A., 3rd, Markowitz, J.E., Perkins, L.N., Liberti, D.C., Leman, D.P., Guitchounts, G., Velho, T., Kotton, D.N., Lois, C., and Gardner, T.J. (2016). Unstable neurons underlie a stable learned behavior. *Nat. Neurosci.* **19**, 1665–1671.
- Lillicrap, T.P., Santoro, A., Marris, L., Akerman, C.J., and Hinton, G. (2020). Backpropagation and the brain. *Nat Rev Neurosci* **21**, 335–346. <https://doi.org/10.1038/s41583-020-0277-3>.
- Lim, S., McKee, J.L., Woloszyn, L., Amit, Y., Freedman, D.J., Sheinberg, D.L., and Brunel, N. (2015). Inferring learning rules from distributions of firing rates in cortical neurons. *Nat. Neurosci.* **18**, 1804–1810.
- Linderman, S.W., and Gershman, S.J. (2017). Using computational theory to constrain statistical models of neural data. *Curr. Opin. Neurobiol.* **46**, 14–24.
- Litwin-Kumar, A., and Doiron, B. (2014). Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nat Commun* **5**, 5319.
- Makino, H., and Komiyama, T. (2015). Learning enhances the relative impact of top-down processing in the visual cortex. *Nat. Neurosci.* **18**, 1116–1122.
- Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84.
- Marblestone, A.H., Wayne, G., and Kording, K.P. (2016). Toward an integration of deep learning and neuroscience. *Front. Comput. Neurosci.* **10**, 94.
- Mau, W., Hasselmo, M.E., and Cai, D.J. (2020). The brain in motion: How ensemble fluidity drives memory-updating and flexibility. *eLife* **9**, e63550.
- Mayo, J.P., and Smith, M.A. (2017). Neuronal adaptation: tired neurons or wired networks? *Trends Neurosci.* **40**, 127–128.
- McDougle, S.D., Ivry, R.B., and Taylor, J.A. (2016). Taking aim at the cognitive side of learning in sensorimotor adaptation tasks. *Trends Cogn. Sci.* **20**, 535–544.
- Millán, A.P., Torres, J.J., Johnson, S., and Marro, J. (2018). Concurrence of form and function in developing networks and its role in synaptic pruning. *Nat. Commun.* **9**, 2236.
- Mitani, A., Dong, M., and Komiyama, T. (2018). Brain-computer interface with inhibitory neurons reveals subtype-specific strategies. *Curr. Biol.* **28**, 77–83.e4.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fiedelnd, A.K., Ostrovski, G., et al. (2015). Human-level control through deep reinforcement learning. *Nature* **518**, 529–533.
- Mochol, G., Kiani, R., and Moreno-Bote, R. (2021). Prefrontal cortex represents heuristics that shape choice bias and its integration into future behavior. *Curr. Biol.* **31**, 1234–1244.e6.
- Morales, A., and Froese, T. (2019). Self-optimization in a hopfield neural network based on the *C. elegans* connectome. In *Artificial Life Conference Proceedings*, pp. 448–453, MIT Press.
- Morehead, J.R., and de Xivry, J.-J.O. (2021). A synthesis of the many errors and learning processes of visuomotor adaptation. *bioRxiv*. <https://doi.org/10.1101/2021.03.14.435278>.
- Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., and Pouget, A. (2014). Information-limiting correlations. *Nat. Neurosci.* **17**, 1410–1417.
- Musall, S., Kaufman, M.T., Juavinett, A.L., Gluf, S., and Churchland, A.K. (2019a). Single-trial neural dynamics are dominated by richly varied movements. *Nat. Neurosci.* **22**, 1677–1686.
- Musall, S., Urai, A.E., Sussillo, D., and Churchland, A.K. (2019b). Harnessing behavioral diversity to understand neural computations for cognition. *Curr. Opin. Neurobiol.* **58**, 229–238.
- Nayebi, A., Srivastava, S., Ganguli, S., and Yamins, D.L.K. (2020). Identifying learning rules from neural network observables. *arXiv*, 2010.11765 <https://arxiv.org/abs/2010.11765>.
- Neftci, E.O., and Averbeck, B.B. (2019). Reinforcement learning in artificial and biological systems. *Nature Machine Intelligence* **1**, 133–143.
- Newell, A., and Rosenbloom, P.S. (1981). Mechanisms of skill acquisition and the law of practice. In *Cognitive skills and their acquisition*, First Edition (Psychology Press).
- Newell, K.M., Liu, Y.-T., and Mayer-Kress, G. (2009). Time Scales, Difficulty/Skill Duality, and the Dynamics of Motor Learning. *Adv Exp Med Biol* **629**, 457–476.
- Ni, A.M., Ruff, D.A., Alberts, J.J., Symmonds, J., and Cohen, M.R. (2018). Learning and attention reveal a general relationship between population activity and behavior. *Science* **359**, 463–465.
- Nicholson, D.A., and Prinz, A.A. (2021). Deep neural network models of object recognition exhibit human-like limitations when performing visual search tasks. *bioRxiv*. Published online January 12, 2021. <https://doi.org/10.1101/2020.10.26.354258>.
- Niv, Y. (2009). Reinforcement learning in the brain. *J. Math. Psychol.* **53**, 139–154.
- Oby, E.R., Golub, M.D., Hennig, J.A., Degenhart, A.D., Tyler-Kabara, E.C., Yu, B.M., Chase, S.M., and Batista, A.P. (2019). New neural activity patterns emerge with long-term learning. *Proc Natl Acad Sci U S A* **116**, 15210–15215.
- Okun, M., Yger, P., Marguet, S.L., Gerard-Mercier, F., Benucci, A., Katzner, S., Busse, L., Carandini, M., and Harris, K.D. (2012). Population Rate Dynamics and Multineuron Firing Patterns in Sensory Cortex. *J. Neurosci.* **32**, 17108–17119.
- Ölveczky, B.P., Andalman, A.S., and Fee, M.S. (2005). Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol.* **3**, e153.
- Orsborn, A.L., and Pesaran, B. (2017). Parsing learning in networks using brain-machine interfaces. *Curr. Opin. Neurobiol.* **46**, 76–83.
- Pekny, S.E., Izawa, J., and Shadmehr, R. (2015). Reward-dependent modulation of movement variability. *J. Neurosci.* **35**, 4015–4024.
- Perich, M.G., Gallego, J.A., and Miller, L.E. (2018). A neural population mechanism for rapid learning. *Neuron* **100**, 964–976.e7.
- Peters, A.J., Chen, S.X., and Komiyama, T. (2014). Emergence of reproducible spatiotemporal activity during motor learning. *Nature* **510**, 263–267.
- Plappert, M., Houthoofd, R., Dhariwal, P., Sidor, S., Chen, R.Y., Chen, X., Asfour, T., Abbeel, P., and Andrychowicz, M. (2017). Parameter space noise for exploration. *arXiv*, 1706.01905 <https://arxiv.org/abs/1706.01905>.
- Poort, J., Khan, A.G., Pachitariu, M., Nemri, A., Orsolic, I., Krupic, J., Bauza, M., Sahani, M., Keller, G.B., Mrsic-Flogel, T.D., and Hofer, S.B. (2015). Learning enhances sensory and multiple non-sensory representations in primary visual cortex. *Neuron* **86**, 1478–1490.
- Pospisil, D.A., Pasupathy, A., and Bair, W. (2018). ‘Artiphsiology’ reveals V4-like shape tuning in a deep network trained for image classification. *eLife* **7**, e38242.

Rajan, K., Harvey, C.D., and Tank, D.W. (2016). Recurrent network models of sequence generation and memory. *Neuron* 90, 128–142.

Richards, B.A., Lillicrap, T.P., Beaudoin, P., Bengio, Y., Bogacz, R., Christensen, A., Clopath, C., Costa, R.P., de Berker, A., Ganguli, S., et al. (2019). A deep learning framework for neuroscience. *Nat. Neurosci.* 22, 1761–1770.

Rokni, U., Richardson, A.G., Bizzi, E., and Seung, H.S. (2007). Motor learning with unstable neural representations. *Neuron* 54, 653–666.

Rosenberg, M., Zhang, T., Perona, P., and Meister, M. (2021). Mice in a labyrinth: Rapid learning, sudden insight, and efficient exploration. *bioRxiv*. <https://doi.org/10.1101/2021.01.14.426746>.

Roy, N.A., Bak, J.H., Akrami, A., Brody, C.D., and Pillow, J.W. (2018). Efficient inference for time-varying behavior during learning. *Adv. Neural Inf. Process. Syst.* 31, 5695–5705.

Rule, M.E., Loback, A.R., Raman, D.V., Driscoll, L.N., Harvey, C.D., and O’Leary, T. (2020). Stable task information from an unstable neural population. *eLife* 9, e51121.

Rumyantsev, O.I., Lecoq, J.A., Hernandez, O., Zhang, Y., Savall, J., Chrapkiewicz, R., Li, J., Zeng, H., Ganguli, S., and Schnitzer, M.J. (2020). Fundamental bounds on the fidelity of sensory cortical coding. *Nature* 580, 100–105.

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

Salimans, T., Ho, J., Chen, X., Sidor, S., and Sutskever, I. (2017). Evolution strategies as a scalable alternative to reinforcement learning. *arXiv*, 1703.03864 <https://arxiv.org/abs/1703.03864>.

Savin, C., and Deneve, S. (2014). Spatio-temporal representations of uncertainty in spiking neural networks, Volume 27 (NIPS), pp. 2024–2032.

Saxe, A., Nelli, S., and Summerfield, C. (2021). If deep learning is the answer, what is the question? *Nat. Rev. Neurosci.* 22, 55–67.

Schoonover, C.E., Ohashi, S.N., Axel, R., and Fink, A.J.P. (2020). Representational drift in primary olfactory cortex. *bioRxiv*. <https://doi.org/10.1101/2020.09.24.312132>.

Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature* 412, 549–553.

Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., Guez, A., Lockhart, E., Hassabis, D., Graepel, T., et al. (2020). Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature* 588, 604–609.

Schultz, W. (2007). Multiple dopamine functions at different time courses. *Annu. Rev. Neurosci.* 30, 259–288.

Schultz, W. (2019). Recent advances in understanding the role of phasic dopamine activity. *F1000Res*. 8, 8.

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

Shadmehr, R., and Holcomb, H.H. (1997). Neural correlates of motor memory consolidation. *Science* 277, 821–825.

Sigurdardottir, H.M., and Sheinberg, D.L. (2015). The effects of short-term and long-term learning on the responses of lateral intraparietal neurons to visually presented objects. *J. Cogn. Neurosci.* 27, 1360–1375.

Singh, A., Peyrache, A., and Humphries, M.D. (2019). Medial Prefrontal Cortex Population Activity Is Plastic Irrespective of Learning. *J. Neurosci.* 39, 3470–3483.

Sinz, F.H., Pitkow, X., Reimer, J., Bethge, M., and Tolias, A.S. (2019). Engineering a less artificial intelligence. *Neuron* 103, 967–979.

Smith, M.A., Ghazizadeh, A., and Shadmehr, R. (2006). Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biol.* 4, e179.

Sohn, H., Meirhaeghe, N., Rajalingham, R., and Jazayeri, M. (2020). A network perspective on sensorimotor learning. *Trends Neurosci.* 44, 170–181.

Stavisky, S.D., Kao, J.C., Ryu, S.I., and Shenoy, K.V. (2017). Motor cortical visuomotor feedback activity is initially isolated from downstream targets in output-null neural state space dimensions. *Neuron* 95, 195–208.e9.

Steinmetz, N.A., Zatka-Haas, P., Carandini, M., and Harris, K.D. (2019). Distributed coding of choice, action and engagement across the mouse brain. *Nature* 576, 266–273.

Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C.B., Carandini, M., and Harris, K.D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. *Science* 364, 255.

Suminski, A.J., Tkach, D.C., Fagg, A.H., and Hatsopoulos, N.G. (2010). Incorporating feedback from multiple sensory modalities enhances brain-machine interface control. *J. Neurosci.* 30, 16777–16787.

Sun, X., O’Shea, D.J., Golub, M.D., Trautmann, E.M., Vyas, S., Ryu, S.I., and Shenoy, K.V. (2020). Skill-specific changes in cortical preparatory activity during motor learning. *bioRxiv*. <https://doi.org/10.1101/2020.01.30.919894>.

Sussillo, D., Churchland, M.M., Kaufman, M.T., and Shenoy, K.V. (2015). A neural network that finds a naturalistic solution for the production of muscle activity. *Nat. Neurosci.* 18, 1025–1033.

Sutton, R.S., and Barto, A.G. (2018). Reinforcement learning: An introduction (MIT press).

Takikawa, Y., Kawagoe, R., Itoh, H., Nakahara, H., and Hikosaka, O. (2002). Modulation of saccadic eye movements by predicted reward outcome. *Exp. Brain Res.* 142, 284–291.

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

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

Uka, T., Sasaki, R., and Kumano, H. (2012). Change in choice-related response modulation in area MT during learning of a depth-discrimination task is consistent with task learning. *J. Neurosci.* 32, 13689–13700.

Urai, A.E., Doiron, B., Leifer, A.M., and Churchland, A.K. (2021). Large-scale neural recordings call for new insights to link brain and behavior. *arXiv*, 2103.14662 <https://arxiv.org/abs/2103.14662>.

Vaswani, P.A., Shmuelof, L., Haith, A.M., Delnicki, R.J., Huang, V.S., Mazzoni, P., Shadmehr, R., and Krakauer, J.W. (2015). Persistent residual errors in motor adaptation tasks: reversion to baseline and exploratory escape. *J. Neurosci.* 35, 6969–6977.

van Vreeswijk, C., and Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural Comput.* 10, 1321–1371.

Vyas, S., Even-Chen, N., Stavisky, S.D., Ryu, S.I., Nuyujukian, P., and Shenoy, K.V. (2018). Neural population dynamics underlying motor learning transfer. *Neuron* 97, 1177–1186.e3.

Wang, J.X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J.Z., Hassabis, D., and Botvinick, M. (2018a). Prefrontal cortex as a meta-reinforcement learning system. *Nat Neurosci* 21, 860–868.

Wang, J., Narain, D., Hosseini, E.A., and Jazayeri, M. (2018b). Flexible timing by temporal scaling of cortical responses. *Nat. Neurosci.* 21, 102–110.

Wärnberg, E., and Kumar, A. (2019). Perturbing low dimensional activity manifolds in spiking neuronal networks. *PLoS Comput. Biol.* 15, e1007074.

Werfel, J., Xie, X., and Seung, H.S. (2005). Learning curves for stochastic gradient descent in linear feedforward networks. *Neural Comput.* 17, 2699–2718.

Wierstra, D., Schaul, T., Glasmachers, T., Sun, Y., Peters, J., and Schmidhuber, J. (2014). Natural evolution strategies. *J. Mach. Learn. Res.* 15, 949–980.

Williams, R.J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Mach. Learn.* 8, 229–256.

Xia, J., Marks, T.D., Goard, M.J., and Wessel, R. (2021). Stable representation of a naturalistic movie emerges from episodic activity with gain variability. *Nat. Commun.* **12**, 5170.

Yamins, D.L., and DiCarlo, J.J. (2016). Using goal-driven deep learning models to understand sensory cortex. *Nat. Neurosci.* **19**, 356–365.

Yan, G., Vértés, P.E., Towilson, E.K., Chew, Y.L., Walker, D.S., Schafer, W.R., and Barabási, A.L. (2017). Network control principles predict neuron function in the *Caenorhabditis elegans* connectome. *Nature* **550**, 519–523.

Yan, Y., Rasch, M.J., Chen, M., Xiang, X., Huang, M., Wu, S., and Li, W. (2014). Perceptual training continuously refines neuronal population codes in primary visual cortex. *Nat. Neurosci.* **17**, 1380–1387.

Yang, Y., and Lisberger, S.G. (2010). Learning on multiple timescales in smooth pursuit eye movements. *J. Neurophysiol.* **104**, 2850–2862.

Yu, R. (2015). Choking under pressure: the neuropsychological mechanisms of incentive-induced performance decrements. *Front. Behav. Neurosci.* **9**, 19.

Zador, A.M. (2019). A critique of pure learning and what artificial neural networks can learn from animal brains. *Nat. Commun.* **10**, 3770.

Zhou, X., Tien, R.N., Ravikumar, S., and Chase, S.M. (2019). Distinct types of neural reorganization during long-term learning. *J. Neurophysiol.* **121**, 1329–1341.