



Symposium

Common Mechanisms of Learning in Motor and Cognitive Systems

Christos Constantinidis,¹  Alaa A. Ahmed,²  Joni D. Wallis,³ and Aaron P. Batista⁴

¹Department of Biomedical Engineering, Vanderbilt University, Nashville, Tennessee 37235, ²Department of Mechanical Engineering, University of Colorado Boulder, Boulder Colorado 80309, ³Department of Psychology, University of California Berkeley, Berkeley, California 94720, and

⁴Department of Bioengineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

Rapid progress in our understanding of the brain's learning mechanisms has been accomplished over the past decade, particularly with conceptual advances, including representing behavior as a dynamical system, large-scale neural population recordings, and new methods of analysis of neuronal populations. However, motor and cognitive systems have been traditionally studied with different methods and paradigms. Recently, some common principles, evident in both behavior and neural activity, that underlie these different types of learning have become to emerge. Here we review results from motor and cognitive learning, relying on different techniques and studying different systems to understand the mechanisms of learning. Movement is intertwined with cognitive operations, and its dynamics reflect cognitive variables. Training, in either motor or cognitive tasks, involves recruitment of previously unresponsive neurons and reorganization of neural activity in a low dimensional manifold. Mapping of new variables in neural activity can be very rapid, instantiating flexible learning of new tasks. Communication between areas is just as critical a part of learning as are patterns of activity within an area emerging with learning. Common principles across systems provide a map for future research.

Introduction

Learning to perform motor and cognitive tasks has been traditionally studied with different methods and paradigms. In recent years, some common principles that underlie different types of learning have become to emerge. It is thus becoming increasingly understood that movement is intertwined with cognitive operations, and its dynamics reflect cognitive variables (Korbisch et al., 2022). Training, in either motor or cognitive tasks, involves recruitment of previously unresponsive neurons, to perform a novel task (Meyers et al., 2012; Shenoy and Carmena, 2014). Engagement of neurons by a new task does not always imply increased firing rate; both increases and decreases of activity may be observed for acquisition of different task elements (Tang et al., 2022). What could be most critical for performance of either motor or cognitive tasks is reorganization of activity in a low dimensional manifold (Sadtlir et al., 2014) and transmission of some aspects of this information to other brain structure through a communication subspace (Shenoy and Carmena, 2014; Semedo et al., 2019). The difficulty of new task learning depends on whether the requisite pattern of activity maps well on existing low-dimensional manifold in neural activity (Losey et al., 2022).

Mapping of new variables in neural activity can be very rapid, instantiating flexible learning of new tasks (Knudsen and Wallis, 2021).

Understanding the principles underlying motor and cognitive learning has far-reaching implications. Brain–computer interfaces (BCIs) rely on learning to operate a computer system to restore movement in patients who have suffered injury at various levels of the nervous system. Cognitive rehabilitation depends on training regimens to ameliorate deficits in cognitive functions (Constantinidis and Klingberg, 2016). Optimizing training regimens and understanding the limitations of learning stand to benefit basic science and translational applications.

In this review, we bring together expertise from different fields to produce an integrative picture. We first review the evidence that movement is modulated by and reflects cognitive variables. Second, we review the principles of neural activity reorganization during motor learning. We then review activity reorganization in the PFC during learning of working memory tasks. We follow with a review of reversal learning, engaging the cortico-hippocampal system. We end with a summary of common principles and a roadmap for future research.

Movement is modulated by cognitive variables

The goal of any movement can be framed as the desire to place oneself in a more rewarding state. Which states are more rewarding or less rewarding is something that is continually learned over timescales that can range from seconds to years. Historically, the process of learning these values has been the focus of researchers in the cognitive domain, whereas the process controlling the movements that realize their acquisition

Received Aug. 7, 2023; revised Aug. 29, 2023; accepted Aug. 31, 2023.

This work was supported by National Eye Institute Award R01 EY017077 to C.C.; National Institute of Neurological Disorders and Stroke Award R01 NS096083 to A.A.A., and R01 NS129584 and R01 NS129098 to A.P.B.; and National Institute of Mental Health Awards R01 MH117763 and MH121448 to J.D.W.

The authors declare no competing financial interests.

Correspondence should be addressed to Christos Constantinidis at christos.constantinidis.1@vanderbilt.edu.

<https://doi.org/10.1523/JNEUROSCI.1505-23.2023>

Copyright © 2023 the authors

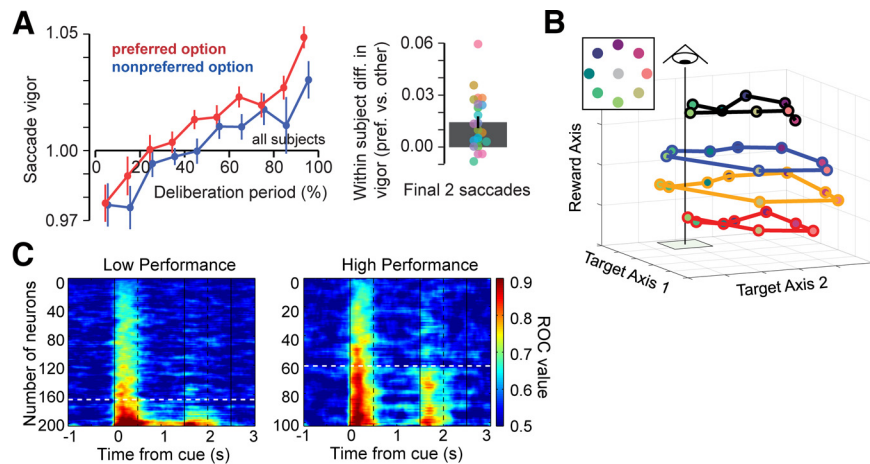


Figure 1. *A*, Saccade vigor (peak velocity normalized by amplitude) across subjects as they deliberated between two on-screen options. Vigor increased over the course of deliberation and was significantly greater toward the preferred option at the end of deliberation. Data from Korbisch et al. (2022). *B*, Neural population activity in M1 cortex corresponding to motor planning for different reach directions is pushed apart with increasing cued reward from small through large. For Jackpot rewards, the activity for different reach directions collapses back toward each other, diminishing their discriminability. We projected neural activity grouped by trial conditions defined by reward and direction and then averaged into a 3D space reflecting reward information (Reward Axis) and target information (Target Axis 1 and 2). The units (population neural activity, spikes/s) on the three axes are the same. Adjacent reach directions (dot color) are connected by a ring for each reward (line color). Data from Smoulder et al. (2023). *C*, Area under the receiver operating characteristic (ROC) curve for prefrontal neurons recorded during training, while the subject achieved low (left) or high-performance sessions (right). Dark red colors represent ability of neurons to discriminate between different displays that needed to be maintained in memory. Data from Tang et al. (2019).

has been the focus of researchers in the motor domain. Tucked away in their respective silos, cognitive and motor research has had little cause to imagine that the neural circuits that control cognitive learning and ultimately decision-making, would influence the neural circuits that control our movements. Yet recent research has highlighted that correlates of variables critical to learning and decision-making, such as reward, history of reward, and reward prediction error, are represented in movement dynamics. A potential reason for why this may be is that the brain is attempting to maximize a global currency that consists not only of the reward one hopes to acquire, but the energy cost of the movement required to do so (Shadmehr et al., 2016; Yoon et al., 2018). Here we will review recent findings pointing to a reflection of learning and decision variables in movement control and learning, and discuss them in the context of commonalities across motor and cognitive domains.

A cornerstone of decision-making is that humans and other animals prefer options with greater value. Findings on the neural basis of these decisions demonstrate that neural activity scales with two key decision variables: the linear value of individual options and their relative value (Knudsen and Wallis, 2021). Intriguingly, we and others show that both variables, linear and relative value, are reflected in the vigor of movement. Not only do animals prefer more valuable options, but they will move faster to acquire them. Monkeys saccade faster to stimuli promising greater reward, and greater expectation of reward (Takikawa et al., 2002). Humans also saccade faster to visual stimuli with greater implicit or explicit reward (Chen et al., 2020; Yoon et al., 2020) and reach faster to targets promising greater point rewards (Summerside et al., 2018). Thus, vigor scales with the linear value of an option. Surprisingly, movement vigor also appears to reflect the evolution of the decision-making process. As subjects deliberate between two options presented as visual stimuli, the time spent deliberating is a reflection of the relative value of the options. The greater the relative value, the quicker the decision and vice versa. When we record subjects' eye movements as they deliberate, we find that they harmoniously

align with the evolution of relative value. Saccade vigor increases over the course of deliberation and increases faster for saccades directed toward the option that is ultimately chosen (Korbisch et al., 2022). Furthermore, the greater the relative value of one option over the other, the greater the difference in the vigor of the saccade directed to each option. Together, vigor of movement appears to provide a real-time readout two key decision-making variables: both the linear value and relative value between options (Fig. 1A).

Another key variable that influences decision-making is the history of reward. Even when the options on immediate offer have identical value, the history of reward experienced up to that point will influence behavior. Humans and animals will linger and harvest longer following a history of poor reward, compared with a richer history of reward (Constantino and Daw, 2015). Remarkably, we observe a similar effect in movement vigor. Following a rich history of reward, subjects move faster between options and move slower following a poor history of rewards (Niv et al., 2007; Yoon et al., 2018; Sukumar et al., 2021).

Decisions are driven by value, but how are these values learned? Decades of research have shown that reward prediction error plays a key role in this process, supported by both computational models and a neurophysiological correlate in phasic dopamine activity (Schultz et al., 1997). Critically, dopamine is implicated in both the representation of value and the control of movement, and thus may provide a clue to understanding the interaction between the respective neural circuits. When learning the association between a cue and the probabilistic reward it predicts, DA activity scales with reward prediction error at both the time of cue presentation and reward presentation (Fiorillo et al., 2003). DA activity is greater in response to cues associated with greater expectation of reward. Upon reward presentation, DA activity is greater, the lower the expectation of reward. Conversely, if no reward is presented, DA activity decreases the greater the expectation of reward. If DA is indeed the bridge between value and vigor, then vigor should also track reward prediction error at both cue and reward presentation, and scale with reward prediction error magnitude and sign. Indeed, recent findings indicate

that saccade vigor immediately following a positive reward prediction error is greater than the vigor of a saccade following a negative reward prediction error (Sedaghat-Nejad et al., 2019). We also find that in a reaching task, where subjects were required to learn the probabilistic reward contingencies of individual targets and ultimately choose between them, reaching movements are faster following a positive reward prediction error compared with a negative prediction error. Intriguingly, reach vigor also tracked the learned value of the cue (Korbisich and Ahmed, 2022).

Cognitive variables such as reward and its counterpart, punishment, can surprisingly influence the process of motor learning itself. Until recently, it was thought that learning new movements and improving motor performance were driven by the need to reduce movement error. However, recent results clearly demonstrate an effect of reward, and even a dissociable effect of punishment. In a well-studied visuomotor adaptation task, where subjects must learn a novel transformation between the hand and the cursor they are trying to control, both reward and punishment can accelerate learning (Galea et al., 2015; Nikooyan and Ahmed, 2015). Reward can also lead to movements that are not only faster, but more accurate (Manohar et al., 2015; Codol et al., 2020). These cognitive variables can also influence the ability to recall and retain these motor memories (Abe et al., 2011; Galea et al., 2015). For example, reward can help retention of motor memories, while punishment leads to accelerated forgetting (Galea et al., 2015). Intriguingly, there is some evidence that individual subject performance in a high-level decision-making task can be predictive of their performance in a motor learning task (Chen et al., 2017).

Together, these findings point to a fascinating link between the control of movement and decision-making. Implicit control of the speed at which we move and the rate at which we learn are influenced by the same cognitive variables that guide our decisions. Explanations for how neural circuits for cognition and movement are harnessed to enable this interplay, and why this may be advantageous remain outstanding questions which we discuss in the following sections.

The view of cognition from motor cortex

The primary motor cortex (M1) provides the predominant pathway for the cortical control of movement, including command signals for the arm, hand, and fingers. Given this anatomic connectivity, we might expect M1 to mostly exhibit signals that are directly related to movement. Instead, recent research has revealed the presence of signals that appear unrelated to the details of movement, and that seem to inform movements without shaping them directly. These signals relate to higher-level “cognitive” functions, such as memory retention and the influence of context (set both internally and externally) on our sensory-motor repertoire. These include motor memories and motivation signals. Evidence for the existence of these two types of signals in M1 is presented below. Why these signals might be present in M1 in the first place remains an intriguing question.

Learning is a whole-brain process. Although we know much about the plasticity rules of learning at individual synapses, we still lack a cohesive understanding of how synaptic plasticity can lead to the storage of a specific memory. To begin to bridge that gap, it is useful to record the activity of populations of neurons, and to examine how they might reorganize their patterns of activity after a bout of learning. However, a challenge arises with this approach. During learning, many changes occur, and we

must somehow identify the ones that are directly responsible for retaining the newly learned concept or skill.

Several years ago, we (Sadtler et al., 2014) and others (reviewed in Shenoy and Carmena, 2014) came to recognize the value of a BCI approach to study learning. BCI systems are primarily envisioned as tools to restore motor function to individuals with paralysis, and as we sought to improve them in the laboratory, we came to recognize their utility for studying the function of the cerebral cortex, and of particular interest here, how it changes with learning. The brain stores what is learned in memory, sometimes only briefly (think of prism adaptation), and sometimes permanently (consider the quintessential skill of riding a bike), and we have recently begun to study the process of memory retention in our BCI context. A BCI has two key advantages for studying motor learning. First, we know precisely how neural activity causes behavior (in our case, the movement of an onscreen cursor). This is not currently knowable for arm movements. Second, we can examine whether neural activity is suitable for a behavior that is not currently being performed. This is also currently not possible in general.

Combining these advantages enabled us to make a novel discovery (Losey et al., 2022): After a bout of BCI learning, neural activity remains suitable for the newly learned behavior, even when it is not being performed. The experiments that demonstrate this begin with us presenting animals with an “intuitive” BCI decoder: that is, one that has been calibrated to work well without requiring any learning. After a few minutes, we switch to a new BCI decoder that the animal has never experienced before. This novel decoder changes the mapping from neural population activity to cursor movement; and thus, to restore proficient control of the BCI cursor, the animal must learn over time to generate new patterns of neural activity, suitable for the new BCI decoder. This learning process can take just a few hundred trials. Subsequently, we return the animal to the intuitive BCI decoder. Performance returns to normal, but neural activity remains some suitability for the newly learned BCI decoder. This is only possible because the dimensionality of M1’s neural population activity is considerably larger (~10D during our experiments) than the dimensionality of the BCI task (2D, i.e., horizontal and vertical on the computer screen). This means that a large redundant “null space” exists for any particular BCI mapping, which allows for the possibility that the ability to perform two tasks can jointly consolidate without interference.

Now, we consider the influence of motivation signals in M1’s activity. Many cortical areas are influenced by the anticipated outcome of a behavior, and an outcome of particular interest here is the payoff for a given action. For the monkeys in our study, this is the size of the fluid reward the animal will earn for successfully completing a task. When we increase the magnitude of a proffered reward, behavior tends to improve. But only up to a point. When rewards that are exceptionally larger (“jackpots”) are made available, monkeys behave much as humans do, and have a tendency to “choke under pressure,” failing at a difficult motor task more of than they would when ordinary-sized rewards are made available (Smoulder et al., 2021).

We recently found that there exists in M1’s activity a signal of the offered reward. Notably, this reward information is sequestered from the target-related information, and we presume, from the subspace that is “potent” for movement kinematics (Kaufman et al., 2014). Despite the orthogonality of the reward and target information in M1, there is nevertheless an interaction between reward information and movement information. Namely, when a jackpot is offered, neural activity attains a large

projection onto the reward axis, and target information reduces (Fig. 1B). To quantify this, we measured the extent to which neural activity can discriminate between two adjacent targets, and we found that this information is reduced in the presence of a potential jackpot reward. Furthermore, there are tight links between the extent to which neural information is influenced by the size of the potential reward and the fine details of the movement. Our findings comport nicely with those reported in the previous section, in which the vigor of human actions is closely related to the rewards they anticipate. This implies that the influence of motivation on movement might be mediated by M1, and perhaps these signals mix in many other brain areas as well. Together, this finding of a reduction in target-related neural information because of the anticipation of a jackpot reward offers a potential neural explanation for the frustrating phenomenon of choking under pressure.

Why might we find such “cognitive” signals of motor memories and motivation in motor cortex? As a partial answer to this conundrum, we note that the “null space” of M1’s population activity is vast, compared with its behaviorally relevant motor command signals. In other words, there are many more neurons in motor cortex than there are corticospinal neurons; and in turn, there are many more corticospinal neurons than there are muscles, or degrees of freedom in which the joints of the arm can rotate. This means that there is plenty of information capacity in M1 than can be accounted for by muscle activity or movement kinematics alone. Perhaps the brain derives no value in expending resources (whether those be computational, energetic, or neuron count) to eliminate those signals from M1, as long as they can be safely sequestered away from the motor-control signals. Or perhaps the motor cortex actually benefits from access to whole-brain information, using this information as needed to provide for the flexible, context-dependent control of our actions, and to drive learning.

Prefrontal cortical plasticity in cognitive learning

If motor cortex is the pinnacle of functional specialization of a cortical area, the PFC sits at the other extreme, being activated by all kinds of sensory stimuli, cognitive variables, and behavioral responses. In recent experiments, we and others studied how training to perform a task affects prefrontal neural activity. Learning a working memory task for the first time elicits plastic changes in persistent activity (Mendoza-Halliday and Martinez-Trujillo, 2017; Riley et al., 2018). More neurons become activated and generate a higher level of persistent activity (Meyer et al., 2007; Riley et al., 2017). Further increases in firing rate tended to accrue with cumulative training (Qi et al., 2011; Tang et al., 2019). Acquiring more complex tasks, such as requiring working memory for multiple stimuli, further recruits more neurons (Fig. 1C) that fire at a higher rate; however, decreases in baseline firing rate have also been described (Tang et al., 2019). In the human brain, BOLD activity decreases after training in complex tasks (Schneiders et al., 2011; Kuhn et al., 2013; Schweizer et al., 2013; Takeuchi et al., 2013), and this is often interpreted as improvements in efficiency (Constantinidis and Klingberg, 2016). Studies that track neural activity after training in the task being trained and in a control task that remained the same revealed that both increases and decreases in activity observed in the active task transferred to the passive task (Tang et al., 2022). Artificial neural networks provide a framework for understanding transfer learning: a network trained on one task produces changes in connection weights in the hidden layers of the network, which when probed with a

different task can generate some training-dependent output (Sinz et al., 2019), consistent with the findings in the motor cortex, reviewed in the preceding section. An important finding in our work was that, although prefrontal activity reflects training in specific tasks, a substantial fraction of stimulus-selective responses and persistent activity are generated automatically, in subjects not required or even trained to perform a task, particularly in posterior and dorsal subdivisions of the PFC (Riley et al., 2017).

Encoding of information in neuronal firing depends not only on the mean firing rate of neuronal responses, but also on how variable these responses are from trial to trial, and on whether firing rates of neurons are positively correlated with each other, which limits how much information can be stored in their collective discharges (Moreno-Bote et al., 2014). Consistent with this principle, we found that the effects of training affect not only mean firing rate but also the variability of persistent activity (Qi and Constantinidis, 2012b) and the correlation of firing rate between simultaneously recorded neurons (Qi and Constantinidis, 2012a). The Fano factor of spike counts, a measure of variability, generally decreases after practicing the task, with the greatest decreases observed in neurons that exhibit persistent activity, compared with neurons that do not. This decrease in trial-to-trial variability may be responsible for increasing the reliability of stimulus property representation after training. Similarly, the spike-count correlation of persistent firing rates between pairs of neurons (known as noise correlation) also decreases after training, which improves the information that can be decoded from simultaneously active neurons (Qi and Constantinidis, 2012a). As in the case of motor learning, the action of dopamine is thought to be critical for prefrontal plasticity, and the density of dopaminergic innervation, which differs between prefrontal subdivisions, is thought to constrain their capacity for plasticity, rendering ventral PFC more plastic (Li et al., 2020).

The persistent-activity model of spatial working memory posits that the appearance of a stimulus generates activity that is maintained during the delay period but may drift with time (Wimmer et al., 2014; Barbosa et al., 2019). In this context, working memory training is thought to rely on strengthening network connections between neurons that generate persistent activity, by virtue of recruiting more neurons during the delay period of the task; by achieving greater discharge rates during the delay period; and by realizing lower variability in firing rate from trial to trial (Meyer et al., 2011; Qi et al., 2011; Qi and Constantinidis, 2012b). Such changes in discharge patterns suggest enduring changes in the prefrontal circuitry after training, which would suggest that the excitability of prefrontal neurons and the ability to generate persistent activity are lastingly altered following training.

As has been recognized in the motor cortex, we found that task training does not alter only measures of single neuron firing rate, but time course and dynamics of population activity (Kobak et al., 2014; Tang et al., 2019). This realization has led to a new paradigm of learning by populations (Libby and Buschman, 2021). It is thus understood that subjects can learn to perform a working memory task to the extent that working memory can be maintained in a stable subspace, free of interference by distracting stimuli (Murray et al., 2017; Spaak et al., 2017; Parthasarathy et al., 2019). Training in a working memory task alters the dynamics of neuronal populations. New latent variables emerge or become more pronounced after training to perform a new task compared with the responses the same stimuli elicit in naive animals, when they view these passively (Kobak et al., 2014; Tang et al.

al., 2019). The consequence of this change is that a greater percentage of firing rate variance is accounted by “condition-independent” components not directly tied to the remembered stimulus location or identity, but presumably reflecting task rules and variables.

Cortico-hippocampal interactions during model-based learning

Neural activity related to a different type of learning, reinforcement learning, also reveals common neural principles. Reinforcement learning describes the process by which organisms learn to modify their behavior in response to rewards and punishments, which can be formalized using reinforcement learning algorithms (Sutton and Barto, 1998). These algorithms fall along a continuum between two extremes (Collins and Cockburn, 2020). At one end is model-free learning, which is associated with the acquisition of habits and skills. Model-free algorithms are computationally efficient, using trial-and-error learning to update the cached values of past actions and repeating the actions that predict higher rewards. On the other end of the continuum is model-based learning, which is associated with deliberative choice and planning. Model-based algorithms learn a model that describes the spatiotemporal structure of an environment, and then use this model to infer reward predictions and evaluate the best course of action. Although more computationally expensive, model-based learning is better at generalizing, performs better in complex and dynamic environments, and enables reasoning and inference.

Both model-free and model-based learning use prediction errors that are encoded by dopaminergic neurons (Fiorillo et al., 2003). However, a critical difference between the algorithms is the incorporation of an environmental model whose neural instantiation is much less clear. Damage to orbitofrontal cortex (OFC) typically produces deficits on tasks that require model-based inference more so than on those that rely on model-free cached values (McDannald et al., 2011; Gremel and Costa, 2013), but its exact contribution to the process of model-based learning has been a subject of considerable debate (Padoa-Schioppa and Schoenbaum, 2015; Wikenheiser and Schoenbaum, 2016; Hayden and Niv, 2021). Until recently, the dominant view of OFC function was that it was responsible for generating reward predictions that can be used to guide value-based decision-making (Bechara et al., 1994; Camille et al., 2011), consistent with a large literature showing that OFC neurons encode the value of choice options (Padoa-Schioppa and Assad, 2006; Kennerley et al., 2009). More recently, this view has been challenged by an alternate hypothesis, which argues that OFC encodes the hidden states that are used to construct a “cognitive map” (Niv, 2019). Although the concept of a cognitive map has existed since the middle of the 20th century (Tolman, 1948), it has often lacked a precise definition. Recent accounts define it as a state-transition graph, which specifies the spatiotemporal relationship of task states and the probability that one state can lead to another (Buzsáki and Tingley, 2018; Niv, 2019; Whittington et al., 2020). These states are often “hidden” in that they are not explicitly obvious from the environment but must be inferred. In this view, reward becomes just one type of hidden state that must be inferred from the environment, albeit one that is common to nearly every experimental task. Evidence to support this view comes from neuroimaging studies which have reported that OFC is the only cortical region to be activated when subjects use cognitive maps (Schuck et al., 2016).

The hippocampus (HPC), which bidirectionally connects to OFC (Barbas and Blatt, 1995; Carmichael and Price, 1995), has also long been associated with representing a cognitive map (O’Keefe and Nadel, 1978; Howard et al., 2014; Behrens et al., 2018). In the 1970s, researchers discovered that hippocampal neurons in rats encoded the animal’s location in space, consistent with the idea that animals formed maps of their environment (O’Keefe and Dostrovsky, 1971; Moser et al., 2008). HPC, like OFC, has also been associated with encoding more abstract relationships (Wikenheiser and Schoenbaum, 2016; Behrens et al., 2018). This raises a question as to the difference between HPC and OFC in terms of encoding the cognitive map. Our own research has shown that HPC and OFC encoding during performance of the same behavioral task is markedly different. We trained monkeys to choose between probabilistically rewarded pictures, in which the reward contingencies gradually changed across the course of a session. Closed-loop stimulation of either HPC or OFC, in which we specifically disrupted the theta oscillation, resulted in learning impairments (Knudsen and Wallis, 2020). This suggests that both structures are involved in learning the task contingencies, and that information may be communicated between the two brain structures via the theta oscillation. However, when we recorded the activity of single neurons, we found that the neural responses in the two areas were very different. Whereas the firing rate of OFC neurons correlated with the linear value of options on offer, HPC encoded the specific value of options relative to one another (Knudsen and Wallis, 2021).

These results suggest a division of labor between the two structures in terms of the implementation of model-based learning, whereby value-based decision-making is the preserve of OFC, and the representation of the cognitive map is primarily a hippocampal function (Knudsen and Wallis, 2022). However, there is likely a strong interaction between these two processes since calculating the value of an option frequently requires real-world knowledge of relationships. For example, when choosing a restaurant, I may favor Mexican food when I’m in California and Indian food when I’m in England, using my knowledge of the history of immigration to predict the likelihood of a good meal. Several studies have favored such an organization in humans. Neuroimaging results have shown that the BOLD response to paired cues becomes more similar in both HPC and OFC as a cognitive map is learned (Wang et al., 2020a), while stimulation of OFC in humans specifically impairs the ability to infer outcomes from cue associations (Wang et al., 2020b). However, brain activity measurements in humans lack the spatiotemporal resolution to address how the OFC and HPC interact mechanistically. Future experiments in nonhuman primates can test the model by training monkeys to perform value-guided decision-making tasks that benefit from understanding of task structure.

Commonalities and generalities

Together, some common themes and general principles emerge from these narratives. First, the value of the options we face influences processes, such as decision-making, the details of our actions, and how we learn. Second, learning is a neural-populations principle. Lasting changes in the correlation pattern among a population of neurons seem to encode the particulars of our memories, whether they are cognitive, motoric, or relational. Third, although not directly measured in any of our studies, we can infer the action of neuromodulators, chiefly dopamine, on the cortical representation of value, and the details of our movements. Fourth, communication between areas is as critical as patterns of activity within an area emerging with learning.

In the future, we can envision a rich interplay between the efforts to understand cognitive learning and motor learning. In many respects, findings from motor learning through the use of BCI technology provide a road map for studying the neural basis of cognitive learning. For example, a clear prediction of motor learning is that the cognitive tasks that are difficult to acquire are the ones whose performance requires the activation of patterns of neural population activity that are outside the low-dimensional manifold already present in the activity of the PFC and other areas. Findings from our studies of cognitive learning, such as in the interaction between brain areas, might similarly yield insight into how the motor system can flexibly adjust movements based on context. Together, convergence of ideas across tasks, across brain areas, and across species will enable a fuller picture of the principles of cognition in all its variegated forms to emerge.

References

- Abe M, Schambra H, Wassermann EM, Luckenbaugh D, Schweighofer N, Cohen LG (2011) Reward improves long-term retention of a motor memory through induction of offline memory gains. *Curr Biol* 21:557–562.
- Barbas H, Blatt GJ (1995) Topographically specific hippocampal projections target functionally distinct prefrontal areas in the rhesus monkey. *Hippocampus* 5:511–533.
- Barbosa J, Stein H, Martinez R, Galan A, Adam K, Li S, Valls-Sole J, Constantinidis C, Compte A (2019) Interplay between persistent activity and activity-silent dynamics in prefrontal cortex during working memory. *bioRxiv* 763938. <https://doi.org/10.1101/763938>.
- Bechara A, Damasio AR, Damasio H, Anderson SW (1994) Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 50:7–15.
- Behrens TE, Muller TH, Whittington JC, Mark S, Baram AB, Stachenfeld KL, Kurth-Nelson Z (2018) What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron* 100:490–509.
- Buzsáki G, Tingley D (2018) Space and time: the hippocampus as a sequence generator. *Trends Cogn Sci* 22:853–869.
- Camille N, Griffiths CA, Vo K, Fellows LK, Kable JW (2011) Ventromedial frontal lobe damage disrupts value maximization in humans. *J Neurosci* 31:7527–7532.
- Carmichael ST, Price JL (1995) Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J Comp Neurol* 363:615–641.
- Chen X, Mohr K, Galea JM (2017) Predicting explorative motor learning using decision-making and motor noise. *PLoS Comput Biol* 13:e1005503.
- Chen X, Zirnsak M, Vega GM, Moore T (2020) Frontal eye field neurons selectively signal the reward value of prior actions. *Prog Neurobiol* 195:101881.
- Codol O, Holland PJ, Manohar SG, Galea JM (2020) Reward-based improvements in motor control are driven by multiple error-reducing mechanisms. *J Neurosci* 40:3604–3620.
- Collins AG, Cockburn J (2020) Beyond dichotomies in reinforcement learning. *Nat Rev Neurosci* 21:576–586.
- Constantinidis C, Klingberg T (2016) The neuroscience of working memory capacity and training. *Nat Rev Neurosci* 17:438–449.
- Constantino SM, Daw ND (2015) Learning the opportunity cost of time in a patch-foraging task. *Cogn Affect Behav Neurosci* 15:837–853.
- Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299:1898–1902.
- Galea JM, Mallia E, Rothwell J, Diedrichsen J (2015) The dissociable effects of punishment and reward on motor learning. *Nat Neurosci* 18:597–602.
- Gremel CM, Costa RM (2013) Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat Commun* 4:2264.
- Hayden BY, Niv Y (2021) The case against economic values in the orbitofrontal cortex (or anywhere else in the brain). *Behav Neurosci* 135:192–201.
- Howard MW, MacDonald CJ, Tiganj Z, Shankar KH, Du Q, Hasselmo ME, Eichenbaum H (2014) A unified mathematical framework for coding time, space, and sequences in the hippocampal region. *J Neurosci* 34:4692–4707.
- Kaufman MT, Churchland MM, Ryu SI, Shenoy KV (2014) Cortical activity in the null space: permitting preparation without movement. *Nat Neurosci* 17:440–448.
- Kennerley SW, Dahmubed AF, Lara AH, Wallis JD (2009) Neurons in the frontal lobe encode the value of multiple decision variables. *J Cogn Neurosci* 21:1162–1178.
- Knudsen EB, Wallis JD (2020) Closed-loop theta stimulation in the orbitofrontal cortex prevents reward-based learning. *Neuron* 106:537–547.e534.
- Knudsen EB, Wallis JD (2021) Hippocampal neurons construct a map of an abstract value space. *Cell* 184:4640–4650.e4610.
- Knudsen EB, Wallis JD (2022) Taking stock of value in the orbitofrontal cortex. *Nat Rev Neurosci* 23:428–438.
- Kobak D, Brendel W, Constantinidis C, Feierstein CE, Kepecs A, Mainen ZF, Romo R, Qi XL, Uchida N, Machens CK (2014) Demixed principal component analysis of population activity in higher cortical areas reveals independent representation of task parameters. *arXiv* 1410.6031.
- Korbisch C, Ahmed AA (2022) Vigor of movement to probabilistic reward tracks reward prediction error. *Proceedings of Advances in Motor Learning and Motor Control*, San Diego, CA.
- Korbisch CC, Apuan DR, Shadmehr R, Ahmed AA (2022) Saccade vigor reflects the rise of decision variables during deliberation. *Curr Biol* 32:5374–5381.e5374.
- Kuhn S, Schmiedek F, Noack H, Wenger E, Bodammer NC, Lindenberger U, Lovden M (2013) The dynamics of change in striatal activity following updating training. *Hum Brain Mapp* 34:1530–1541.
- Li S, Zhou X, Constantinidis C, Qi XL (2020) Plasticity of persistent activity and its constraints. *Front Neural Circuits* 14:15.
- Libby A, Buschman TJ (2021) Rotational dynamics reduce interference between sensory and memory representations. *Nat Neurosci* 24:715–726.
- Losey DM, Henning JA, Oby ER, Golub MD, Sadtler PT, Quick KM, Ryu SI, Tyler-Kabara EC, Batista AP, Yu BM, Chase SM (2022) Learning alters neural activity to simultaneously support memory and action. *bioRxiv* 498856. <https://doi.org/10.1101/2022.07.05.498856>.
- Manohar SG, Chong TT, Apps MA, Batla A, Stamelou M, Jarman PR, Bhatia KP, Husain M (2015) Reward pays the cost of noise reduction in motor and cognitive control. *Curr Biol* 25:1707–1716.
- McDannald MA, Lucantonio F, Burke KA, Niv Y, Schoenbaum G (2011) Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *J Neurosci* 31:2700–2705.
- Mendoza-Halliday D, Martinez-Trujillo JC (2017) Neuronal population coding of perceived and memorized visual features in the lateral prefrontal cortex. *Nat Commun* 8:15471.
- Meyer T, Qi XL, Constantinidis C (2007) Persistent discharges in the prefrontal cortex of monkeys naive to working memory tasks. *Cereb Cortex* 17 Suppl 1:i70–i76.
- Meyer T, Qi XL, Stanford TR, Constantinidis C (2011) Stimulus selectivity in dorsal and ventral prefrontal cortex after training in working memory tasks. *J Neurosci* 31:6266–6276.
- Meyers EM, Qi XL, Constantinidis C (2012) Incorporation of new information into prefrontal cortical activity after learning working memory tasks. *Proc Natl Acad Sci USA* 109:4651–4656.
- Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A (2014) Information-limiting correlations. *Nat Neurosci* 17:1410–1417.
- Moser EI, Kropff E, Moser MB (2008) Place cells, grid cells, and the brain's spatial representation system. *Annu Rev Neurosci* 31:69–89.
- Murray JD, Bernacchia A, Roy NA, Constantinidis C, Romo R, Wang XJ (2017) Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. *Proc Natl Acad Sci USA* 114:394–399.
- Nikooyan AA, Ahmed AA (2015) Reward feedback accelerates motor learning. *J Neurophysiol* 113:633–646.
- Niv Y (2019) Learning task-state representations. *Nat Neurosci* 22:1544–1553.
- Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 191:507–520.

- O'Keefe J, Dostrovsky J (1971) The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. *Brain Res* 34:171–175.
- O'Keefe J, Nadel L (1978) *The hippocampus as a cognitive map*. Oxford: Oxford UP.
- Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. *Nature* 441:223–226.
- Padoa-Schioppa C, Schoenbaum G (2015) Dialogue on economic choice, learning theory, and neuronal representations. *Curr Opin Behav Sci* 5:16–23.
- Parthasarathy A, Tang C, Herikstad R, Cheong LF, Yen SC, Libedinsky C (2019) Time-invariant working memory representations in the presence of code-morphing in the lateral prefrontal cortex. *Nat Commun* 10:4995.
- Qi XL, Constantinidis C (2012a) Correlated discharges in the primate prefrontal cortex before and after working memory training. *Eur J Neurosci* 36:3538–3548.
- Qi XL, Constantinidis C (2012b) Variability of prefrontal neuronal discharges before and after training in a working memory task. *PLoS One* 7: e41053.
- Qi XL, Meyer T, Stanford TR, Constantinidis C (2011) Changes in prefrontal neuronal activity after learning to perform a spatial working memory task. *Cereb Cortex* 21:2722–2732.
- Riley MR, Qi XL, Constantinidis C (2017) Functional specialization of areas along the anterior-posterior axis of the primate prefrontal cortex. *Cereb Cortex* 27:3683–3697.
- Riley MR, Qi XL, Zhou X, Constantinidis C (2018) Anterior-posterior gradient of plasticity in primate prefrontal cortex. *Nat Commun* 9:3790.
- Sadtler PT, Quick KM, Golub MD, Chase SM, Ryu SI, Tyler-Kabara EC, Yu BM, Batista AP (2014) Neural constraints on learning. *Nature* 512:423–426.
- Schneiders JA, Opitz B, Krick CM, Mecklinger A (2011) Separating intramodal and across-modal training effects in visual working memory: an fMRI investigation. *Cereb Cortex* 21:2555–2564.
- Schuck NW, Cai MB, Wilson RC, Niv Y (2016) Human orbitofrontal cortex represents a cognitive map of state space. *Neuron* 91:1402–1412.
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Schweizer S, Grahn J, Hampshire A, Mobbs D, Dalgleish T (2013) Training the emotional brain: improving affective control through emotional working memory training. *J Neurosci* 33:5301–5311.
- Sedaghat-Nejad E, Herzfeld DJ, Shadmehr R (2019) Reward prediction error modulates saccade vigor. *J Neurosci* 39:5010–5017.
- Semedo JD, Zandvakili A, Machens CK, Yu BM, Kohn A (2019) Cortical areas interact through a communication subspace. *Neuron* 102:249–259. e244.
- Shadmehr R, Huang HJ, Ahmed AA (2016) A representation of effort in decision-making and motor control. *Curr Biol* 26:1929–1934.
- Shenoy KV, Carmena JM (2014) Combining decoder design and neural adaptation in brain-machine interfaces. *Neuron* 84:665–680.
- Sinz FH, Pitkow X, Reimer J, Bethge M, Tolias AS (2019) Engineering a less artificial intelligence. *Neuron* 103:967–979.
- Smoulder AL, Pavlovsky NP, Marino PJ, Degenhart AD, McClain NT, Batista AP, Chase SM (2021) Monkeys exhibit a paradoxical decrease in performance in high-stakes scenarios. *Proc Natl Acad Sci USA* 118: e2109643118.
- Smoulder AL, Marino PJ, Oby ER, Snyder SE, Miyata H, Pavlovsky NP, Bishop WE, Yu BM, Chase SM, Batista AP (2023) A neural basis of choking under pressure. *bioRxiv* 537007. <https://doi.org/10.1101/2023.04.16.537007>.
- Spaak E, Watanabe K, Funahashi S, Stokes MG (2017) Stable and dynamic coding for working memory in primate prefrontal cortex. *J Neurosci* 37:6503–6516.
- Sukumar S, Shadmehr R, Ahmed A (2021) Effects of reward history on decision-making and movement vigor. *bioRxiv* 453376. <https://doi.org/10.1101/2021.07.22.453376>.
- Summerside EM, Shadmehr R, Ahmed AA (2018) Vigor of reaching movements: reward discounts the cost of effort. *J Neurophysiol* 119:2347–2357.
- Sutton RS, Barto AG (1998) *Reinforcement learning: an introduction* (adaptive computation and machine learning). Cambridge, MA: Massachusetts Institute of Technology.
- Takeuchi H, Taki Y, Nouchi R, Hashizume H, Sekiguchi A, Kotozaki Y, Nakagawa S, Miyauchi CM, Sassa Y, Kawashima R (2013) Effects of working memory training on functional connectivity and cerebral blood flow during rest. *Cortex* 49:2106–2125.
- Takikawa Y, Kawagoe R, Hikosaka O (2002) Reward-dependent selectivity of anticipatory activity in monkey caudate neurons. *J Neurophysiol* 87:508–515.
- Tang H, Qi XL, Riley MR, Constantinidis C (2019) Working memory capacity is enhanced by distributed prefrontal activation and invariant temporal dynamics. *Proc Natl Acad Sci USA* 116:7095–7100.
- Tang H, Riley MR, Singh B, Qi XL, Blake DT, Constantinidis C (2022) Prefrontal cortical plasticity during learning of cognitive tasks. *Nat Commun* 13:90.
- Tolman EC (1948) Cognitive maps in rats and men. *Psychol Rev* 55:189–208.
- Wang F, Schoenbaum G, Kahnt T (2020a) Interactions between human orbitofrontal cortex and hippocampus support model-based inference. *PLoS Biol* 18:e3000578.
- Wang F, Howard JD, Voss JL, Schoenbaum G, Kahnt T (2020b) Targeted stimulation of an orbitofrontal network disrupts decisions based on inferred, not experienced outcomes. *J Neurosci* 40:8726–8733.
- Whittington JC, Muller TH, Mark S, Chen G, Barry C, Burgess N, Behrens TE (2020) The Tolman-Eichenbaum machine: unifying space and relational memory through generalization in the hippocampal formation. *Cell* 183:1249–1263. e1223.
- Wikenheiser AM, Schoenbaum G (2016) Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. *Nat Rev Neurosci* 17:513–523.
- Wimmer K, Nykamp DQ, Constantinidis C, Compte A (2014) Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nat Neurosci* 17:431–439.
- Yoon T, Geary RB, Ahmed AA, Shadmehr R (2018) Control of movement vigor and decision making during foraging. *Proc Natl Acad Sci USA* 115: E10476–E10485.
- Yoon T, Jaleel A, Ahmed AA, Shadmehr R (2020) Saccade vigor and the subjective economic value of visual stimuli. *J Neurophysiol* 123:2161–2172.