Journal: **NETWORK NEUROSCIENCE**

PERSPECTIVES

Thalamocortical contribution to flexible learning in neural systems

Mien Brabeeba Wang[1,2], **and Michael M. Halassa**[1]

[1]Department of Brain and Cognitive Science, Massachusetts Institute of Technology, Cambridge MA, USA

[2]Computer Science & Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA

Keywords: [Meta-learning; Credit assignment; Continual learning; Thalamocortical interactions; Basal ganglia; Thalamus]

ABSTRACT

Animal brains evolved to optimize behavior in dynamic environments, flexibly selecting actions that maximize future rewards in different contexts. A large body of experimental work indicates that such optimization changes the wiring of neural circuits, appropriately mapping environmental input onto behavioral outputs. A major unsolved scientific question is how optimal wiring adjustments, which must target the connections responsible for rewards, can be accomplished when the relation between sensory inputs, action taken, environmental context with rewards is ambiguous. The credit assignment problem can be categorized into context-independent *structural credit assignment* and context-dependent *continual learning*. In this perspective, we survey prior approaches to these two problems and advance the notion that the brain's specialized neural architectures provide efficient solutions. Within this framework, the thalamus with its cortical and basal ganglia interactions serves as a systems-level solution to credit assignment. Specifically, we propose that thalamocortical interaction is the locus of meta-learning where the thalamus provides cortical control functions that parametrize the cortical activity association space. By selecting among these control functions, the basal ganglia hierarchically guide thalamocortical plasticity across two timescales to enable meta-learning. The faster timescale establishes contextual associations to enable behavioral flexibility while the slower one enables generalization to new contexts.

AUTHOR SUMMARY

Deep learning has shown great promise over the last decades allowing artificial neural networks to solve difficult tasks. The key to success is the optimization process by which task errors are translated to connectivity patterns. A major unsolved question is how the brain optimally adjusts the wiring of neural circuits to minimize task error analogously. In our perspective, we advance the notion that the brain's specialized architecture is part of the solution and spell out a path towards its theoretical, computational and experimental testing. Specifically, we propose the interaction between the cortex, thalamus and basal ganglia induces plasticity in two timescales to enable flexible behaviors. The faster timescale establishes contextual associations to enable behavioral flexibility while the slower one enables generalization to new contexts.

INTRODUCTION

Learning to flexibly choose appropriate actions in uncertain environments is a hallmark of intelligence Miller and Cohen (2001); Niv (2009); Thorndike (2017). When animals explore unfamiliar environments, they tend to reinforce actions that lead to unexpected rewards. A common notion in contemporary neuroscience is that such behavioral reinforcement emerges from changes in synaptic connectivity, where synapses that contribute to the unexpected reward are strengthened Abbott and Nelson (2000); Bliss and Lomo (1973); Dayan and Abbott (2005); Hebb (2002); Whittington and Bogacz (2019). A prominent model for connecting synaptic to behavioral reinforcement is dopaminergic innervation of basal ganglia (BG), where dopamine (DA) carries the **reward prediction error** (RPE) signals to guide synaptic learning Bamford, Wightman, and Sulzer (2018); Bayer and Glimcher (2005); Montague, Dayan, and Sejnowski (1996); Schultz, Dayan, and Montague (1997). This circuit motif is thought to implement a basic form of the reinforcement learning (RL) algorithm Houk, Davis, and Beiser (1994); Morris, Nevet, Arkadir, Vaadia, and Bergman (2006); Roesch, Calu, and Schoenbaum (2007); Suri and Schultz (1999); R. Sutton and Barto (2018); R. S. Sutton and Barto (1990); Wickens and Kotter

(1994), which has had much success in explaining simple Pavlovian and instrumental conditioning Ikemoto and Panksepp (1999); Niv (2009); R. Sutton and Barto (2018); R. S. Sutton and Barto (1990). However, it is unclear how this circuit can reinforce the appropriate connections in complex natural environments where animals need to dynamically map sensory inputs to different action in a context-dependent way. If one naively credits all synapses with the RPE signals, the learning will be highly inefficient since different cues, contexts and actions contribute to the RPE signals differently. To properly credit the cues, context and actions that lead to unexpected reward is a challenging problem, known as the **credit assignment** problem Lillicrap, Santoro, Marris, Akerman, and Hinton (2020); Minsky (1961); Rumelhart, Hinton, and Williams (1986); Whittington and Bogacz (2019).

One can roughly categorize the credit assignment into context-independent *structural credit assignment* and context-dependent ***continual learning***. In structural credit assignment, animals may make decisions in a multi-cue environment and should be able to credit those cues that contribute to the rewarding outcome. Similarly, if actions are being chosen based on internal decision variables, then the underlying activity states must also be reinforced. In such cases, neurons that are selective to external cues or internal latent variables need to adjust their downstream connectivity based on its contribution of their downstream targets to the RPE. This is a challenging computation to implement because, for upstream neurons, the RPE will be dependent on downstream neurons that are several connections away. For example, a sensory neuron needs to know the action chosen in the motor cortex to selectively credit the sensory synapses that contribute to the action. In *continual learning*, animals not only need to appropriately credit the sensory cues and actions that lead to the reward but also need to credit the sensorimotor combination in the right context to retain the behaviors learned from different contexts and even to generalize to novel contexts. Therefore, animals can continually learn and generalize across different contexts while retaining behaviors in familiar contexts. For example, when one is in the United States, one learns to first look left before crossing the street, whereas, in the United Kingdom, one learns to look right instead. However, after spending time in the UK, someone from the US should not unlearn the behavior of looking left first when they return home because their brain ought to properly assign the credit to a different context. Furthermore, once one learns how to cross the street in the US, it is much easier to learn how to cross the street in the UK because the brain flexibly generalize behaviors across contexts.

In this perspective, we will first go over common approaches from machine learning to tackle these two credit assignment problems. In doing so, we highlight the challenge in their efficient implementation within biological neural circuits. We also highlight some recent proposals that advance the notion of specialized neural hardware that approximate more general solutions for credit assignment Fiete and Seung (2006); Ketz, Morkonda, and O'Reilly (2013); Kornfeld et al. (2020); Kusmierz, Isomura, and Toyoizumi (2017); Lillicrap, Cownden, Tweed, and Akerman (2016); Liu, Smith, Mihalas, Shea-Brown, and Sümbül (2020); O'Reilly (1996); O'Reilly, Russin, Zolfaghar, and Rohrlich (2021); Richards and Lillicrap (2019); Roelfsema and Holtmaat (2018); Roelfsema and van Ooyen (2005); Sacramento, Ponte Costa, Bengio, and Senn (2018); Schiess, Urbanczik, and Senn (2016); Zenke and Ganguli (2018). Along these lines, we propose an efficient systems-level solution involving the thalamus and its interaction with the cortex and BG for these two credit assignment problems.

COMMON MACHINE LEARNING APPROACHES TO CREDIT ASSIGNMENT

One solution to structural credit assignment in machine learning is **backpropagation** Rumelhart et al. (1986). Backpropagation recursively computes the vector-valued error signal for synapses based on their contribution to the error signal. There is much empirical success of backpropagation in surpassing human performance in supervised learning such as image recognition He, Zhang, Ren, and Sun (2016); Krizhevsky, Sutskever, and Hinton (2012) and reinforcement learning such as playing the game of Go and Atari Mnih et al. (2015); Schrittwieser et al. (2020); Silver et al. (2016, 2017). Additionally, comparing artificial networks trained with backpropagation to neural responses from the ventral visual stream of non-human primates shows comparable internal representations Cadieu et al. (2014); Yamins et al. (2014). Despite its empirical success in superhuman level performance and matching the internal representation of actual brains, backpropagation may not be straightforward to implement in biological neural circuits as we explain below.

In its most basic form, backpropagation requires symmetric connections between neurons (forward and backward connections). Mathematically, we can write down the backpropagation in Equation 1:

$$\delta W_i \propto \frac{\partial E}{\partial W_i} = e_i f(a_{i-1})^\top \tag{1}$$

where

$$e_i = W_{i+1}^{\top} e_{i+1} \circ f'(a_i),$$

E is the total error, e_i is the vector error at layer i, W_i is the synaptic weight connecting layer $i-1$ to layer i and f is the nonlinearity. Intuitively, this is saying that the change of synaptic weight W_i is computed by a Hebbian learning rule between backpropagation error e_i and activity from last layer $f(a_{i-1})$ while the backpropagation error is computed by backpropagating the error in the next layer through symmetric feedback weights $W_{i+1}^{\top}$. Importantly, in this algorithm, error signals do not alter the activity of neurons in the preceding layers and instead operate independently from the feedforward activity. However, such arrangement is not observed in the brain; symmetric connections across neurons are not a universal feature of circuit organization, and biological neurons may encode both feedforward inputs and errors through changes in spike output (changes in activity) Crick (1989); Richards and Lillicrap (2019). Therefore, it is hard to imagine how the basic form of backpropagation (symmetry and error/activity separation) is physically implemented in the brain.

Furthermore, while an animal can continually learn to behave across different contexts, artificial neural networks trained by backpropagation struggle to learn and remember different tasks in different contexts: a problem known as **catastrophic forgetting** French (1999); Kemker, McClure, Abitino, Hayes, and Kanan (2018); Kumaran, Hassabis, and McClelland (2016); McCloskey and Cohen (1989); Parisi, Kemker, Part, Kanan, and Wermter (2019). Specifically, this problem occurs when the tasks are trained sequentially because the weights optimized for former tasks will be modified to fit the later tasks. One of the common solutions is to interleave the tasks from different contexts to jointly optimize performance across contexts by using an episodic memory system and replay mechanism Kumaran et al. (2016); McClelland, McNaughton, and O'Reilly (1995). This approach has received empirical success in artificial neural networks including learning to play many Atari game Mnih et al. (2015); Schrittwieser et al. (2020). However, since one needs to store past training data in memory to replay during learning, this approach demands a high computational overhead and can be is inefficient as the number of the contexts increases. On the other hand, humans and animals acquire diverse sensorimotor skills in different contexts throughout their life span: a feat that cannot be solely explained by memory replay M. M. Murray, Lewkowicz, Amedi, and Wallace (2016); Parisi et al. (2019); Power and Schlaggar

(2017); Zenke, Gerstner, and Ganguli (2017). Therefore, biological neural circuits are likely to employ other solutions to continual learning in addition to memory replay.

Therefore, to solve these two credit assignment problems in the brain, one needs to seek different solutions. One of the pitfalls of backpropagation is that it is a general algorithm that works on any architecture. However, actual brains are collections of specialized hardware put together in a specialized way. It can be conceived that through clever coordination between different cell types and different circuits, the brains can solve the credit assignment problem by leveraging its specialized architectures. Along this line of ideas, many investigators have proposed cellular Fiete and Seung (2006); Kornfeld et al. (2020); Kusmierz et al. (2017); Liu et al. (2020); Richards and Lillicrap (2019); Sacramento et al. (2018); Schiess et al. (2016) and circuit level mechanisms Lillicrap et al. (2016); O'Reilly (1996); Roelfsema and Holtmaat (2018); Roelfsema and van Ooyen (2005) to assign credit appropriately. In this perspective, we would like to advance the notion that the specialized hardware arrangement also happens at the system level and propose that the thalamus and its interaction with basal ganglia (BG) and the cortex serve as a system-level solution for these three types of credit assignment.

A PROPOSAL: THALAMOCORTICAL-BASAL GANGLIA INTERACTIONS ENABLE META-LEARNING TO SOLVE CREDIT ASSIGNMENT.

To motivate the notion of thalamocortical-basal ganglia interactions being a potential solution for credit assignment, we will start with a brief introduction. The cortex, thalamus and basal ganglia are the three major components of the mammalian forebrain – the part of the brain to which high level cognitive capacities are attributed to Alexander, DeLong, and Strick (1986); Badre, Kayser, and D'Esposito (2010); Cox and Witten (2019); Makino, Hwang, Hedrick, and Komiyama (2016); Miller (2000); Miller and Cohen (2001); Niv (2009); Seo, Lee, and Averbeck (2012); Wolff and Vann (2019). Each of these components has its specialized internal architectures; the cortex is dominated by excitatory neurons with extensive lateral connectivity profiles Fuster (1997); Rakic (2009); Singer, Sejnowski, and Rakic (2019), the thalamus is grossly divided into different nuclei harboring mostly excitatory neurons devoid of lateral connections Harris et al. (2019); Jones (1985); Sherman and Guillery (2005), and the basal ganglia are a series of inhibitory structures driven by excitatory inputs from the cortex and thalamus Gerfen and Bolam (2010); Lanciego, Luquin, and Obeso (2012); Nambu (2011) (Figure 1). A popular view within system

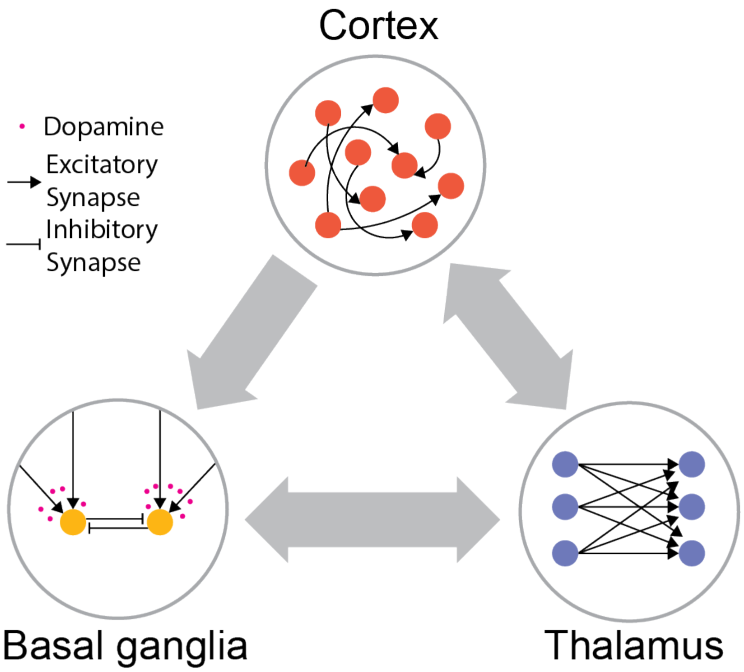

Figure 1. Distinct architectures of cortex, thalamus and basal ganglia Cortex is largely composed of excitatory neurons with extensive recurrent connectivity. Thalamus consists of mostly excitatory neurons without lateral connections. Basal ganglia consist of mostly inhibitory neurons driven by cortical and thalamic inputs, and the corticostriatal plasticity is modulated by dopamine.

neuroscience stipulates that BG and the cortex underwent different learning paradigms where BG is involved in reinforcement learning while the cortex is involved in unsupervised learning Doya (1999, 2000). Specifically, the input structure of the basal ganglia known as the striatum is thought to be where reward gated plasticity takes place to implement reinforcement learning Bamford et al. (2018); Cox and Witten (2019); Hikosaka, Kim, Yasuda, and Yamamoto (2014); Kornfeld et al. (2020); Niv (2009); Perrin and Venance (2019). One such evidence is the high temporal precision of DA activity in the striatum. To accurately attribute the action that leads to positive RPE, DA is released into the relevant corticostriatal synapses. However, DA needs to disappear quickly to prevent the next stimulus-response combination from being reinforced. In the striatum, this elimination process is carried out by dopamine active transporter (DAT) to maintain a high temporal resolution of DA activity on a timescale of around 100ms-1s to support reinforcement learning Cass and Gerhardt (1995); Ciliax et al. (1995); Garris and Wightman (1994). In contrast, although the cortex also has dopaminergic innervation, cortical DAT expression is low and therefore DA levels may change at a timescale that is too slow to support reinforcement learning Cass and Gerhardt (1995); Garris and Wightman (1994); Lapish, Kroener, Durstewitz, Lavin, and Seamans (2007); Seamans and Robbins (2010) but instead support other processes related to learning Badre et al. (2010); Miller and Cohen (2001). In fact, ample evidence indicates that cortical structures undergo Hebbian-like long term potentiation (LTP) and long term depression (LTD) Cooke and Bear (2010); Feldman (2009); Kirkwood, Rioult, and Bear (1996). However, despite the unsupervised nature of these processes, cortical representations are task-relevant and include appropriate sensorimotor mappings that lead to rewards Allen et al. (2017); Donahue and Lee (2015); Enel, Wallis, and Rich (2020); Jacobs and Moghaddam (2020); Petersen (2019); Tsutsui, Hosokawa, Yamada, and Iijima (2016). How could this arise from an unsupervised process? One possible explanation is that basal ganglia activate the appropriate cortical neurons during behaviors and the cortical network collectively consolidates high reward sensorimotor mappings via Hebbian-like learning Andalman and Fee (2009); Ashby, Ennis, and Spiering (2007); Hélie, Ell, and Ashby (2015); Tesileanu, Olveczky, and Balasubramanian (2017); Warren, Tumer, Charlesworth, and Brainard (2011). Previous computational accounts of this process have emphasized a consolidation function for the cortex in this process, which naively would beg the question of why duplicate a process that seems to function well in the basal ganglia and perhaps include a lot of details of the associated experience?

The answer to this question is the core of our proposal. We propose that the learning process is not a duplication, but instead that the reinforcement process in the basal ganglia selects thalamic control functions that subsequently activate cortical associations to allow flexible mappings across different contexts (Figure 2).

To understand this proposition, we need to take a closer look at the involvement of these distinct network elements in task learning. Learning in basal ganglia happens in corticostriatal synapses where the basic form of reinforcement learning is implemented. Specifically, the coactivation of sensory and motor cortical inputs generates eligibility traces in corticostriatal synapses that get captured by the presence or absence of DA Fee and Goldberg (2011); Fiete, Fee, and Seung (2007); Kornfeld et al. (2020). This RL algorithm is fast at acquiring simple associations but slow at generalization to other behaviors. On the other hand, the cortical plasticity operates in a much slower timescale but seems to allow flexible behaviors and fast generalization Kim, Johnson, Cilles, and Gold (2011); Mante, Sussillo, Shenoy, and Newsome (2013); Miller (2000); Miller and Cohen (2001). How does the cortex exhibit

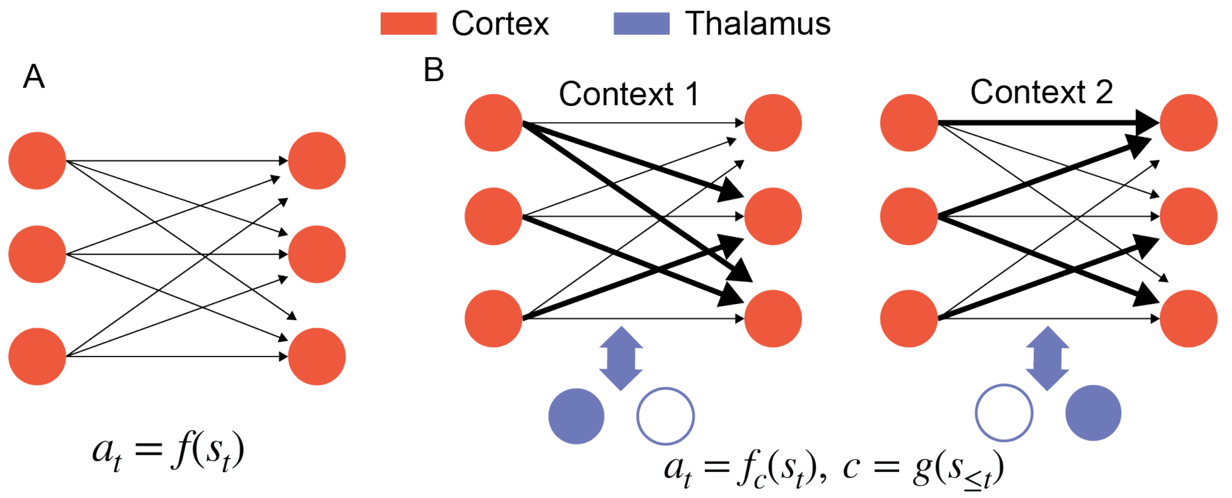

Figure 2. Two views of learning in the cortex A. One possible view is that the Hebbian cortical plasticity consolidates the sensorimotor mapping from BG to learn a stimulus-action mapping $a_t = f(s_t)$. B. We propose that thalamocortical systems perform meta-learning by consolidating the teaching signals from BG to learn a context-dependent mapping $a_t = f_c(s_t)$ where the context c is computed by past stimulus history and represented by different thalamic activities.

slow synaptic plasticity and flexible behaviors at the same time? An explanatory framework is **meta-learning** Botvinick et al. (2019); Wang et al. (2018), where the flexibility arises from network dynamics and the generalization emerges from slow synaptic plasticity across different contexts. In other words, synaptic plasticity stores a higher-order association between contexts and sensorimotor associations while the network dynamics switches between different sensorimotor associations based on this higher order association. However, properly arbitrating between synaptic plasticity and network dynamics to store such higher order association is a nontrivial task Sohn, Meirhaeghe, Rajalingham, and Jazayeri (2021). We propose that the thalamocortical system learns these dynamics, where the thalamus provides control nodes that parametrize the cortical activity association space. Basal ganglia inputs to the thalamus learn to select between these different control nodes directly implementing the interface between weight adjustment and dynamical controls. Our proposal rests on the following three specific points.

First, building on a line of the literature that shows diverse thalamocortical interaction in sensory, cognitive and motor cortex, we propose that thalamic output may be described as control functions over cortical computations. These control functions can be purely in the sensory domain like attentional filtering, in the cognitive domain like manipulating working memory or in the motor domain like preparation for movement Bolkan et al. (2017); W. Guo, Clause, Barth-Maron, and Polley (2017); Z. V. Guo et al. (2017); Mukherjee et al. (2020); Rikhye, Gilra, and Halassa (2018); Saalmann and Kastner (2015); Schmitt et al. (2017); Tanaka (2007); Wimmer et al. (2015); Zhou, Schafer, and Desimone (2016). These functions directly relate thalamic activity patterns to different cortical dynamical regimes and thus offer a way to establish higher order association between context and sensorimotor mapping within the thalamocortical pathways. Second, based on previous studies on direct and indirect BG pathways that influence most cortical regions Hunnicutt et al. (2016); Jiang and Kim (2018); Nakajima, Schmitt, and Halassa (2019); Peters, Fabre, Steinmetz, Harris, and Carandini (2021), we propose that BG hierarchically selects these thalamic control functions to influence activities of the cortex toward rewarding behavioral outcomes. Lastly, we propose that thalamocortical structure consolidate the selection of BG through a two timescales Hebbian learning process to enable meta-learning. Specifically, the faster corticothalamic plasticity learns the higher order association that enables flexible contextual switching with different thalamic patterns Marton, Seifikar, Luongo, Lee, and Sohal (2018); Rikhye et al. (2018) while the slower cortical plasticity learns the shared representations that allow generalization to new behaviors. Below, we will go over the supporting literature that leads us to this proposal.

MORE GENERAL ROLES OF THALAMOCORTICAL INTERACTION AND BASAL GANGLIA

Classical literature has emphasized the role of the thalamus in transmitting sensory inputs to the cortex. This is because some of the better studied thalamic pathways are those connected to sensors on one end and primary cortical areas on another Hubel and Wiesel (1961); Lien and Scanziani (2018); Reinagel, Godwin, Sherman, and Koch (1999); Sherman and Spear (1982); Usrey, Alonso, and Reid (2000). From that perspective, thalamic neurons being devoid of lateral connection transmit their inputs (e.g. from the retina in the case of the lateral geniculate nucleus (LGN)) to the primary sensory cortex (V1 in this same example case) and the input transformation (center-surround to oriented edges) occurs within the cortex Hoffmann, Stone, and Sherman (1972); Hubel and Wiesel (1962); Lien and Scanziani (2018); Usrey et al. (2000). In many cases, these formulations of thalamic “relay” have generalized to how motor and cognitive thalamocortical interactions may be operating. However, in contrast to the classical relay view of the thalamus, more recent studies have shown diverse thalamic functions in sensory, cognitive and motor processing Bolkan et al. (2017); W. Guo et al. (2017); Z. V. Guo et al. (2017); Rikhye et al. (2018); Saalmann and Kastner (2015); Schmitt et al. (2017); Tanaka (2007); Wimmer et al. (2015); Zhou et al. (2016). For example in mice, sensory thalamocortical transmission can be adjusted based on PFC-dependent, top-down biasing signals transmitted through non-classical basal ganglia pathways involving the thalamic reticular nucleus (TRN) Nakajima et al. (2019); Phillips, Kambi, and Saalmann (2016); Wimmer et al. (2015). Interestingly, these task-relevant PFC signals themselves require long range interactions with the associative mediodorsal (MD) thalamus to be initiated, maintained and flexibly switched Rikhye et al. (2018); Schmitt et al. (2017); Wimmer et al. (2015). One can also observe nontrivial control functions in the motor thalamus. Motor preparatory activities in the anterior motor cortex (ALM) show persistent activities that predicted future actions. Interestingly, the motor thalamus also shows similar preparatory activities that predict future actions and by optogenetically manipulate the motor thalamus activities, the persistent activities in ALM quickly diminished Z. V. Guo et al. (2017). Recently, Mukherjee, Lam, Wimmer, and Halassa (2021) discovers two cell types within MD thalamus differentially modulates the cortical evidence accumulation dynamics depending on if the evidences are conflicting or sparse to boost signal-to-noise ratio in decision making. Based on the above studies, we propose that the thalamus provides a set of control functions to the cortex. Specifically, cortical computations may be flexibly switched to different dynamical modes by activating a particular thalamic output that corresponds to that mode.

On the other hand, the selective role of BG in motor and cognitive control also has dominated the literature because thalamocortical-basal ganglia interaction is the most well studied in frontal systems Cox and Witten (2019); Makino et al. (2016); McNab and Klingberg (2008); Monchi, Petrides, Strafella, Worsley, and Doyon (2006); Seo et al. (2012). However, classical and contemporary studies have recognized that all cortical areas, including primary sensory areas project to the striatum Hunnicutt et al. (2016); Jiang and Kim (2018); Peters et al. (2021). Similarly, the basal ganglia can project to the more sensory parts of the thalamus through lesser-studied pathways to influence the sensory cortex Hunnicutt et al. (2016); Nakajima et al. (2019); Peters et al. (2021). Specifically, a non-classical BG pathway projects to TRN which in turn modulates the activities of LGN to influence sensory thalamocortical transmission Nakajima et al. (2019). On the other hand, it has also been argued that BG are involved in gating working memory McNab and Klingberg (2008); Voytek and Knight (2010). This shows that BG has a much more general role than classical action and action strategy selection. Therefore, combining with our proposals on thalamic control functions, we propose that BG hierarchically selects different thalamic control functions to influence all cortical areas in different contexts through reinforcement learning.

Furthermore, there are series of the work that indicates the role of BG to guide plasticity in thalamocortical structures Andalman and Fee (2009); Fiete et al. (2007); Hélie et al. (2015); Mehaffey and Doupe (2015); Tesileanu et al. (2017). In particular, there is evidence that BG is critical for the initial learning and less involved in the automatic behaviors once the behaviors are learned across different species. In zebra finches, the lesion of BG in adult zebra finch has little effects on song production, but the lesion of BG in juvenile zebra finch prevents the bird from learning the song Fee and Goldberg (2011); Scharff and Nottebohm (1991); Sohrabji, Nordeen, and Nordeen (1990). Similar patterns can be observed in people with Parkinson's disease. Parkinson's patients who have a reduction of DA and striatal defects have troubles in solving procedural learning tasks but can produce automatic behaviors normally Asmus, Huber, Gasser, and Schöls (2008); Soliveri, Brown, Jahanshahi, Caraceni, and Marsden (1997); Thomas-Ollivier et al. (1999). This behavioral evidence suggests that thalamocortical structures consolidate the learning from BG as the behaviors become more automatic. Furthermore, on the synaptic level, a songbird learning circuit also demonstrates this cortical consolidation motif Mehaffey and Doupe (2015); Tesileanu et al. (2017). In a zebra finch, the premotor nucleus HVC (a proper name) projects to the motor nucleus robust nucleus of the arcopallium (RA) to produce the song. On the other hand, RA also receives BG nucleus Area X mediated inputs from the lateral nucleus of the medial nidopallium (LMAN). The latter pathway is believed to be a locus of reinforcement learning in the songbird circuit. By burst stimulating both input pathways in different time lags, one can discover that HVC-RA and LMAN-RA underwent opposite plasticity Mehaffey and Doupe (2015). This suggests that the learning is gradually transferred from LMAN-RA to HVC-RA pathway Fee and Goldberg (2011); Mehaffey and Doupe (2015); Tesileanu et al. (2017). This indicates a general role of BG as the trainer for cortical plasticity.

THE THALAMOCORTICAL STRUCTURE CONSOLIDATES THE BG SELECTIONS ON THALAMIC CONTROL FUNCTIONS IN DIFFERENT TIMESCALES TO ENABLE META-LEARNING.

In this section, in addition to BG's role as the trainer for cortical plasticity, we further propose that BG is the trainer in two different timescales for thalamocortical structures to enable meta-learning. The faster timescale trainer trains the corticothalamic connections to select the appropriate thalamic control functions in different contexts while the slower timescale trainer trains the cortical connections to form a task-relevant and generalizable representation.

From the songbird example, we see how thalamocortical structures can consolidate simple associations learned through the basal ganglia. To enable meta-learning, we propose that this general network consolidation motif operates over two different timescales within thalamocortical-basal ganglia interactions (Figure 3). First, combining the idea of thalamic outputs as control functions over cortical network activity patterns and the basal ganglia selecting such functions, we frame learning in basal ganglia as a process that connects contextual associations (higher order) with the appropriate dynamical control that maximizes reward at the sensorimotor level (lower order). Under this framing, corticothalamic plasticity consolidates the higher order association within a fast timescale. This allows flexible switching between different thalamic control functions in different contexts. On the other hand, the cortical plasticity consolidates the sensorimotor association over a slow timescale to allow shared representation that can generalize across different contexts. As the thalamocortical structures learn the higher order association, the behaviors become less BG-dependent and the network is able to switch between different thalamic control functions to induce different sensorimotor mappings in different contexts. By having two learning timescales, animals can conceivably both adapt quickly in changing environments with fast learning of corticothalamic connections while maintaining the important information across the environment in the cortical connections. One should note that this separation of timescales is independent from different timescales across cortex Gao, van den Brink, Pfeffer, and Voytek (2020); J. D. Murray et al. (2014). While different timescales across cortex allows animals to process information differentially, the separation of corticothalmic and cortical plasticity allows the thalamocortcal system to learn the higher contextual association to modulate cortical dynamics flexibly.

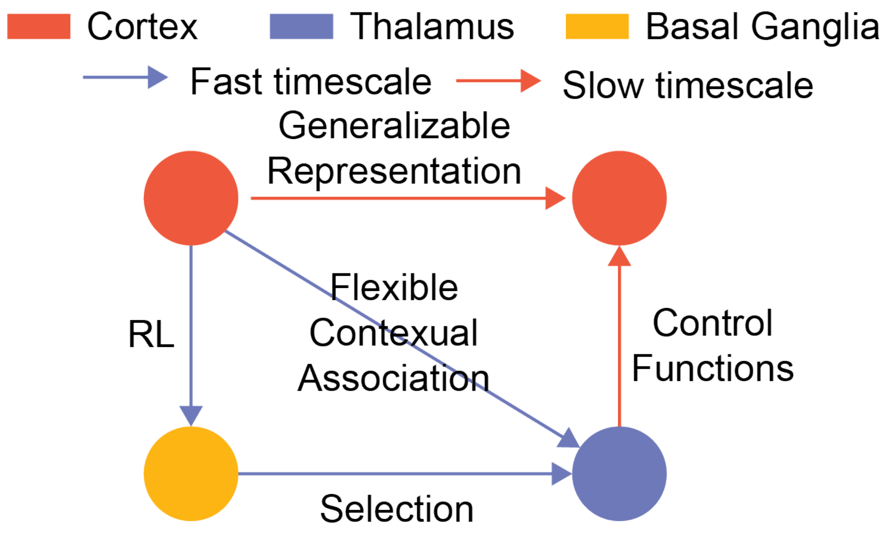

Figure 3. Two timescales learning in thalamocortical structures We propose that one can learn the thalamocortical structure to enable meta-learning by applying the general network motif in two different timescales. First, one can learn the corticothalamic connections by applying the motif on the blue loop with a faster timescale. This allows the network to consolidate flexible switching behaviors. Second, one can learn the cortical connections by applying the motif on the orange loop in a slower timescale. This allows cortical neurons to develop a task-relevant shared representation that can generalize across contexts.

Some anatomical observations support this idea. The thalamostriatal neurons has a more modulatory role to the cortical dynamics in a diffusive projection while thalamocortical neurons has a more driver role to the cortical dynamic in a topographically restricted dense projection Sherman and Guillery (2005). This indicates that thalamostriatal neurons might serve as the role of control functions in the faster consolidation loop with the feedback to striatum to conduct credit assignment. On the other hand, thalamocortical neurons might be more involved in the slower consolidation loop with the feedback to striatum coming from the cortex to train the common cortical representation across contexts.

In summary, this two timescales network consolidation scheme provides a general way for BG to guide plasticity in the thalamocortical architecture to enable meta-learning and thus solves structural credit assignment as a special case. Along these lines, experimental evidence supports the notion that when faced with multi-sensory inputs, the BG can selectively disinhibit a modality-specific subnetwork of the thalamic reticular nucleus (TRN) to filter out the sensory inputs that are not relevant to the behavior outcomes and thus solve the structural credit assignment problem.

In the discussion above, we discuss our proposal under a general formulation of thalamic control functions. In the next section, we will specify other thalamic control functions suggested by recent studies and observe how they can solve continual learning under this framework as well.

THE THALAMUS SELECTIVELY AMPLIFIES FUNCTIONAL CORTICAL CONNECTIVITY AS A SOLUTION TO CONTINUAL LEARNING AND CATASTROPHIC FORGETTING.

One of the pitfalls of the artificial neural network is catastrophic forgetting. If one trains an artificial neural network on a sequence of tasks, the performance on the older task will quickly deteriorate as the network learns the new task French (1999); Kemker et al. (2018); Kumaran et al. (2016); McCloskey and Cohen (1989); Parisi et al. (2019). On the other hand, the brain can achieve *continual learning*, the ability to learn different tasks in different contexts without catastrophic forgetting and even generalize the performance to novel context Lewkowicz (2014); M. M. Murray et al. (2016); Power and Schlaggar (2017); Zenke, Gerstner, and Ganguli (2017). There are three main approaches in machine learning in dealing with catastrophic forgetting. First, one can use the regularization method to mostly update the weights that are less important to the prior tasks Fernando et al. (2017); Jung, Ju, Jung, and Kim (2018); Kirkpatrick et al. (2017); Li and Hoiem (2018); Maltoni and Lomonaco (2019); Zenke, Poole, and Ganguli (2017). This idea is inspired by experimental and theoretical studies on how synaptic information is selectively protected in the brain Benna and Fusi (2016); Cichon and Gan (2015); Fusi, Drew, and Abbott (2005); Hayashi-Takagi et al. (2015); Yang, Pan, and Gan (2009). However, it is unclear how to biologically compute the importance of each synapse to prior tasks nor how to do global regularization locally. Second, one can also use a dynamic architecture in which the network expands the architecture by allocating a subnetwork to train with the new information while preserving old information Cortes, Gonzalvo, Kuznetsov, Mohri, and Yang (2017); Draelos et al. (2017); Rusu et al. (2016); Xiao, Zhang, Yang, Peng, and Zhang (2014). However, this type of method is not scalable since the number of neurons needs to scale linearly with the number of the task. Lastly, one can use a memory buffer to replay past tasks to avoid catastrophic forgetting by interleaving the experience of the past tasks with the experience of the present task Kemker and Kanan (2018); Kumaran et al. (2016); McClelland et al. (1995); Shin, Lee, Kim, and Kim (2017). However, this type of method cannot be the sole solution as the memory buffer needs to scale linearly with the number of the tasks and potentially the number of the trials.

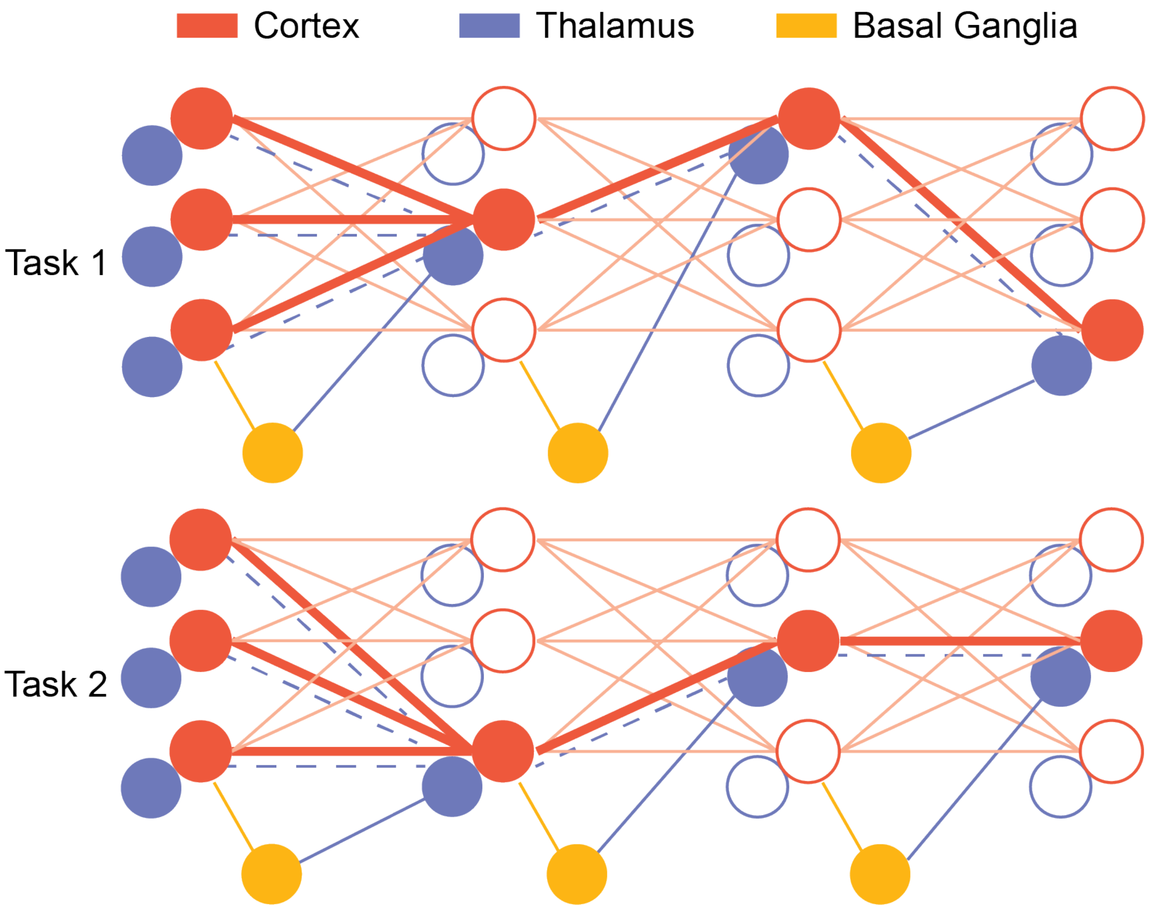

Figure 4. A thalamocortical architecture with interaction with BG for continual learning During task execution, BG selects thalamic neurons that amplify the relevant cortical subnetwork. This protects other parts of the network that are important for another context from being overwritten. When the other task comes, BG selects other thalamic neurons and since the synapses are protected from the last task, animals can freely switch from different tasks without forgetting the previous tasks. Furthermore, as the corticothalamic synapses learn how to select the right thalamic neurons in a different context (blue dash line), task execution can become less BG dependent.

We propose that the thalamus provides another way to solve continual learning and catastrophic forgetting via selectively amplifying parts of the cortical connections in different contexts (Figure 4). Specifically, we propose that a population of thalamic neurons topographically amplify the connectivity of cortical subnetworks as their control functions. During a behavioral task, BG selects subsets of the thalamus which selectively amplify the connectivity of cortical subnetworks. Because of the reinforcement learning in BG, the subnetwork that is the most relevant to the current task will be more preferentially activated and updated. By selecting only the relevant subnetwork to activate in one context, the thalamus protects other subnetworks which can have useful information in another context from being overwritten. The corticothalamic structures can then consolidate these BG-guided flexible switching behaviors via our proposed network motif and the switching becomes less BG-dependent. Furthermore, our proposed solution has implications on generalization as well. Different tasks can have principles in common that can be transferred. For example, although the rules of chess and Go are very different, players in both games all need to predict what the other players are going to do and counterattack based on the prediction. Since BG selects the subnetwork at each hierarchy that is most relevant to the current tasks, in addition to selecting different subnetworks to prevent catastrophic forgetting, BG can also select subnetworks that are beneficial to both tasks as well to achieve generalization. Therefore, the cortex can develop a modular hierarchical representation of the world that can be easily generalized.

The idea of protecting relevant information from the past tasks to be overwritten has been applied before computationally and has decent success in combating catastrophic forgetting in deep learning Kirkpatrick et al. (2017). Experimentally, we also have found thalamic neurons selectively amplify the cortical connectivity to solve the continual learning problem. In a task where the mice need to switch between different sets of task cues that guided the attention to the visual or auditory target, the performance of the mice does not deteriorate much after switching to the original context which is an indication of continual learning Rikhye et al. (2018). Electrophysiological recording of PFC and mediodorsal thalamic nucleus (MD) neurons, we discovered that PFC neurons preferentially code for the rule of the attention while MD neurons preferentially code for the contexts of different sets of the cues. Thalamic neurons that encode the task-relevant context translate this neural representation into the amplification of cortical activity patterns associated with that context (despite the fact that cortical neurons themselves only encode the context implicitly). These experimental observations are consistent with our proposed solution: by incorporating the thalamic population that can selectively amplify connectivity of cortical subnetworks, the thalamus and its interaction with cortex and BG solve the continual learning problem and prevent catastrophic forgetting.

CONCLUSION

In summary, in contrast to the traditional relay view of the thalamus, we propose that thalamocortical interaction is the locus of meta-learning where the thalamus provides cortical control functions, such as sensory filtering, working memory gating or motor preparation, that parametrize the cortical activity association space. Furthermore, we propose a two timescale learning consolidation framework where BG hierarchically selects these thalamic control functions to enable meta-learning, solving the credit assignment problem. The faster plasticity learns contextual associations to enable rapid behavioral flexibility while the slower plasticity establishes cortical representation that generalizes. By considering the recent observation of the thalamus selectively amplifying functional cortical connectivity, the thalamocortical-basal ganglia network is able to flexibly learn context-dependent association without catastrophic forgetting while generalizing to the new contexts. This modular account of the thalamocortical interaction may seem to be in contrast with the recent proposed dynamical perspectives Barack and Krakauer (2021) on thalamocortical interaction in which the thalamus shapes and constrains the cortical attractor landscapes Shine (2021). We would like to argue that both the modular and dynamical perspectives are compatible with our proposal. The crux of the perspectives is that the thalamus provides control functions that parametrize cortical dynamics and these control functions can be of modular nature or of dynamical nature depending on their specific input-output connectivity. Flexible behaviors can be induced either by selecting the control functions that amplify the appropriate cortical subnetworks or those that adjust the cortical dynamics to the appropriate regimes.

REFERENCES

Abbott, L. F., & Nelson, S. B. (2000). Synaptic plasticity: taming the beast. *Nat Neurosci*, *3 Suppl*, 1178–1183.

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

Allen, W. E., Kauvar, I. V., Chen, M. Z., Richman, E. B., Yang, S. J., Chan, K., … Deisseroth, K. (2017). Global Representations of Goal-Directed Behavior in Distinct Cell Types of Mouse Neocortex. *Neuron*, *94*(4), 891–907.

Andalman, A. S., & Fee, M. S. (2009). A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc Natl Acad Sci U S A*, *106*(30), 12518–12523.

Ashby, F. G., Ennis, J. M., & Spiering, B. J. (2007). A neurobiological theory of automaticity in perceptual categorization. *Psychol Rev*, *114*(3), 632–656.

Asmus, F., Huber, H., Gasser, T., & Schöls, L. (2008). Kick and rush: paradoxical kinesia in Parkinson disease. *Neurology*, *71*(9), 695.

Badre, D., Kayser, A. S., & D'Esposito, M. (2010). Frontal cortex and the discovery of abstract action rules. *Neuron*, *66*(2), 315–326.

Bamford, N. S., Wightman, R. M., & Sulzer, D. (2018). Dopamine's Effects on Corticostriatal Synapses during Reward-Based Behaviors. *Neuron*, *97*(3), 494–510.

Barack, D. L., & Krakauer, J. W. (2021). Two views on the cognitive brain. *Nature Reviews Neuroscience*, *22*(6), 359–371. Retrieved from `https://doi.org/10.1038/s41583-021-00448-6` doi: 10.1038/s41583-021-00448-6

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*(1), 129–141.

Benna, M. K., & Fusi, S. (2016). Computational principles of synaptic memory consolidation. *Nat Neurosci*, *19*(12), 1697–1706.

Bliss, T. V., & Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J Physiol*, *232*(2), 331–356.

Bolkan, S. S., Stujenske, J. M., Parnaudeau, S., Spellman, T. J., Rauffenbart, C., Abbas, A. I., … Kellendonk, C. (2017). Thalamic projections sustain prefrontal activity during working memory maintenance. *Nat Neurosci*, *20*(7), 987–996.

Botvinick, M., Ritter, S., Wang, J. X., Kurth-Nelson, Z., Blundell, C., & Hassabis, D. (2019). Reinforcement Learning, Fast and Slow. *Trends Cogn Sci*, *23*(5), 408–422.

Cadieu, C. F., Hong, H., Yamins, D. L. K., Pinto, N., Ardila, D., Solomon, E. A., … DiCarlo, J. J. (2014). Deep neural networks rival the representation of primate it cortex for core visual object recognition. *PLOS Computational Biology*, *10*(12), 1-18. Retrieved from `https://doi.org/10.1371/journal.pcbi.1003963` doi: 10.1371/journal.pcbi.1003963

Cass, W. A., & Gerhardt, G. A. (1995). In vivo assessment of dopamine uptake in rat medial prefrontal cortex: comparison with dorsal striatum and nucleus accumbens. *J Neurochem*, *65*(1), 201–207.

Cichon, J., & Gan, W. B. (2015). Branch-specific dendritic Ca(2+) spikes cause persistent synaptic plasticity. *Nature*, *520*(7546), 180–185.

Ciliax, B. J., Heilman, C., Demchyshyn, L. L., Pristupa, Z. B., Ince, E., Hersch, S. M., … Levey, A. I. (1995). The dopamine transporter: immunochemical characterization and localization in brain. *J Neurosci*, *15*(3 Pt 1), 1714–1723.

Cooke, S. F., & Bear, M. F. (2010). Visual experience induces long-term potentiation in the primary visual cortex. *J Neurosci*, *30*(48), 16304–16313.

Cortes, C., Gonzalvo, X., Kuznetsov, V., Mohri, M., & Yang, S. (2017). AdaNet: Adaptive structural learning of artificial neural networks. In *Proceedings of the 34th international conference on machine learning* (Vol. 70, pp. 874–883). Retrieved from `http://proceedings.mlr.press/v70/cortes17a.html`

Cox, J., & Witten, I. B. (2019). Striatal circuits for reward learning and decision-making. *Nat Rev Neurosci*, *20*(8), 482–494.

Crick, F. (1989). The recent excitement about neural networks. *Nature*, *337*(6203), 129–132.

Dayan, P., & Abbott, L. F. (2005). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. The MIT Press.

Donahue, C. H., & Lee, D. (2015). Dynamic routing of task-relevant signals for decision making in dorsolateral prefrontal cortex. *Nat Neurosci*, *18*(2), 295–301.

Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Netw*, *12*(7-8), 961–974.

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

Draelos, T. J., Miner, N. E., Lamb, C. C., Cox, J. A., Vineyard, C. M., Carlson, K. D., … Aimone, J. B. (2017). Neurogenesis deep learning: Extending deep networks to accommodate new classes. In *2017 international joint conference on neural networks (ijcnn)* (p. 526-533). doi: 10.1109/IJCNN.2017.7965898

Enel, P., Wallis, J. D., & Rich, E. L. (2020). Stable and dynamic representations of value in the prefrontal cortex. *Elife*, *9*.

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

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

Fernando, C., Banarse, D., Blundell, C., Zwols, Y., Ha, D., Rusu, A. A., … Wierstra, D. (2017). Pathnet: Evolution channels gradient descent in super neural networks. *CoRR*, *abs/1701.08734*. Retrieved from http://arxiv.org/abs/1701.08734

Fiete, I. R., Fee, M. S., & Seung, H. S. (2007). Model of birdsong learning based on gradient estimation by dynamic perturbation of neural conductances. *J Neurophysiol*, *98*(4), 2038–2057.

Fiete, I. R., & Seung, H. S. (2006). Gradient learning in spiking neural networks by dynamic perturbation of conductances. *Phys. Rev. Lett.*, *97*, 048104. Retrieved from https://link.aps.org/doi/10.1103/PhysRevLett.97.048104 doi: 10.1103/PhysRevLett.97.048104

French, R. M. (1999). Catastrophic forgetting in connectionist networks. *Trends Cogn Sci*, *3*(4), 128–135.

Fusi, S., Drew, P. J., & Abbott, L. F. (2005). Cascade models of synaptically stored memories. *Neuron*, *45*(4), 599–611.

Fuster, J. (1997). *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe*. Lippincott-Raven. Retrieved from https://books.google.com/books?id=YupqAAAAMAAJ

Gao, R., van den Brink, R. L., Pfeffer, T., & Voytek, B. (2020). Neuronal timescales are functionally dynamic and shaped by cortical microarchitecture. *Elife*, *9*.

Garris, P. A., & Wightman, R. M. (1994). Different kinetics govern dopaminergic transmission in the amygdala, prefrontal cortex, and striatum: an in vivo voltammetric study. *J Neurosci*, *14*(1), 442–450.

Gerfen, C., & Bolam, J. (2010). The neuroanatomical organization of the basal ganglia. *Handbook of Behavioral Neuroscience*, *20*, 3-28. doi: 10.1016/B978-0-12-374767-9.00001-9

Guo, W., Clause, A. R., Barth-Maron, A., & Polley, D. B. (2017). A Corticothalamic Circuit for Dynamic Switching between Feature Detection and Discrimination. *Neuron*, *95*(1), 180–194.

Guo, Z. V., Inagaki, H. K., Daie, K., Druckmann, S., Gerfen, C. R., & Svoboda, K. (2017). Maintenance of persistent activity in a frontal thalamocortical loop. *Nature*, *545*(7653), 181–186. Retrieved from `https://doi.org/10.1038/nature22324` doi: 10.1038/nature22324

Harris, J. A., Mihalas, S., Hirokawa, K. E., Whitesell, J. D., Choi, H., Bernard, A., … Zeng, H. (2019). Hierarchical organization of cortical and thalamic connectivity. *Nature*, *575*(7781), 195–202.

Hayashi-Takagi, A., Yagishita, S., Nakamura, M., Shirai, F., Wu, Y. I., Loshbaugh, A. L., … Kasai, H. (2015). Labelling and optical erasure of synaptic memory traces in the motor cortex. *Nature*, *525*(7569), 333–338.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *2016 ieee conference on computer vision and pattern recognition (cvpr)* (p. 770-778). doi: 10.1109/CVPR.2016.90

Hebb, D. (2002). *The organization of behavior: A neuropsychological theory*. Taylor & Francis. Retrieved from `https://books.google.com/books?id=gUtwMochAI8C`

Hikosaka, O., Kim, H. F., Yasuda, M., & Yamamoto, S. (2014). Basal ganglia circuits for reward value-guided behavior. *Annu Rev Neurosci*, *37*, 289–306.

Hoffmann, K. P., Stone, J., & Sherman, S. M. (1972). Relay of receptive-field properties in dorsal lateral geniculate nucleus of the cat. *J Neurophysiol*, *35*(4), 518–531.

Houk, J. C., Davis, J. L., & Beiser, D. G. (1994). Adaptive critics and the basal ganglia. In *Models of information processing in the basal ganglia* (p. 215-232).

Hubel, D. H., & Wiesel, T. N. (1961). Integrative action in the cat's lateral geniculate body. *J Physiol*, *155*, 385–398.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol*, *160*, 106–154.

Hunnicutt, B. J., Jongbloets, B. C., Birdsong, W. T., Gertz, K. J., Zhong, H., & Mao, T. (2016). A comprehensive excitatory input map of the striatum reveals novel functional organization. *Elife*, *5*.

Hélie, S., Ell, S. W., & Ashby, F. G. (2015). Learning robust cortico-cortical associations with the basal ganglia: an integrative review. *Cortex*, *64*, 123–135.

Ikemoto, S., & Panksepp, J. (1999). The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Res Brain Res Rev*, *31*(1), 6–41.

Jacobs, D. S., & Moghaddam, B. (2020). Prefrontal Cortex Representation of Learning of Punishment Probability During Reward-Motivated Actions. *J Neurosci*, *40*(26), 5063–5077.

Jiang, H., & Kim, H. F. (2018). Anatomical Inputs From the Sensory and Value Structures to the Tail of the Rat Striatum. *Front Neuroanat*, *12*, 30.

Jones, E. G. (Ed.). (1985). *The thalamus*. Springer US. Retrieved from `https://doi.org/10.1007/978-1-4615-1749-8` doi: 10.1007/978-1-4615-1749-8

Jung, H., Ju, J., Jung, M., & Kim, J. (2018). Less-forgetful learning for domain expansion in deep neural networks. In *Aaai conference on artificial intelligence.* Retrieved from `https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/17073`

Kemker, R., & Kanan, C. (2018). Fearnet: Brain-inspired model for incremental learning. In *International conference on learning representations.* Retrieved from `https://openreview.net/forum?id=SJ1Xmf-Rb`

Kemker, R., McClure, M., Abitino, A., Hayes, T., & Kanan, C. (2018). Measuring catastrophic forgetting in neural networks. In *Aaai conference on artificial intelligence.* Retrieved from `https://aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/16410`

Ketz, N., Morkonda, S. G., & O'Reilly, R. C. (2013). Theta coordinated error-driven learning in the hippocampus. *PLOS Computational Biology*, *9*(6), 1-9. Retrieved from `https://doi.org/10.1371/journal.pcbi.1003067` doi: 10.1371/journal.pcbi.1003067

Kim, C., Johnson, N. F., Cilles, S. E., & Gold, B. T. (2011). Common and distinct mechanisms of cognitive flexibility in prefrontal cortex. *J Neurosci*, *31*(13), 4771–4779.

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., … Hadsell, R. (2017). Overcoming catastrophic forgetting in neural networks. *Proc Natl Acad Sci U S A*, *114*(13), 3521–3526.

Kirkwood, A., Rioult, M. C., & Bear, M. F. (1996). Experience-dependent modification of synaptic plasticity in visual cortex. *Nature*, *381*(6582), 526–528.

Kornfeld, J., Januszewski, M., Schubert, P., Jain, V., Denk, W., & Fee, M. (2020). An anatomical substrate of credit assignment in reinforcement learning. *bioRxiv*. Retrieved from `https://www.biorxiv.org/content/early/2020/02/19/2020.02.18.954354` doi: 10.1101/2020.02.18.954354

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (Vol. 25). Curran Associates, Inc. Retrieved from `https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf`

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What Learning Systems do Intelligent Agents Need? Complementary Learning Systems Theory Updated. *Trends Cogn Sci*, *20*(7), 512–534.

Kusmierz, L., Isomura, T., & Toyoizumi, T. (2017). Learning with three factors: modulating Hebbian plasticity with errors. *Curr Opin Neurobiol*, *46*, 170–177.

Lanciego, J. L., Luquin, N., & Obeso, J. A. (2012). Functional neuroanatomy of the basal ganglia. *Cold Spring Harb Perspect Med*, *2*(12), a009621.

Lapish, C. C., Kroener, S., Durstewitz, D., Lavin, A., & Seamans, J. K. (2007). The ability of the mesocortical dopamine system to operate in distinct temporal modes. *Psychopharmacology (Berl)*, *191*(3), 609–625.

Lewkowicz, D. J. (2014). Early experience and multisensory perceptual narrowing. *Dev Psychobiol*, *56*(2), 292–315.

Li, Z., & Hoiem, D. (2018). Learning without forgetting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(12), 2935-2947. doi: 10.1109/TPAMI.2017.2773081

Lien, A. D., & Scanziani, M. (2018). Cortical direction selectivity emerges at convergence of thalamic synapses. *Nature*, *558*(7708), 80–86.

Lillicrap, T. P., Cownden, D., Tweed, D. B., & Akerman, C. J. (2016). Random synaptic feedback weights support error backpropagation for deep learning. *Nat Commun*, *7*, 13276.

Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Backpropagation and the brain. *Nat Rev Neurosci*, *21*(6), 335–346.

Liu, Y. H., Smith, S., Mihalas, S., Shea-Brown, E., & Sümbül, U. (2020). A solution to temporal credit assignment using cell-type-specific modulatory signals. *bioRxiv*. Retrieved from `https://www.biorxiv.org/content/early/2020/11/23/2020.11.22.393504` doi: 10.1101/2020.11.22.393504

Makino, H., Hwang, E. J., Hedrick, N. G., & Komiyama, T. (2016). Circuit Mechanisms of Sensorimotor Learning. *Neuron*, *92*(4), 705–721.

Maltoni, D., & Lomonaco, V. (2019). Continuous learning in single-incremental-task scenarios. *Neural Netw*, *116*, 56–73.

Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, *503*(7474), 78–84.

Marton, T. F., Seifikar, H., Luongo, F. J., Lee, A. T., & Sohal, V. S. (2018). Roles of Prefrontal Cortex and Mediodorsal Thalamus in Task Engagement and Behavioral Flexibility. *J Neurosci*, *38*(10), 2569–2578.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev*, *102*(3), 419–457.

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bower (Ed.), (Vol. 24, p. 109-165). Academic Press. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0079742108605368` doi: https://doi.org/10.1016/S0079-7421(08)60536-8

McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nat Neurosci*, *11*(1), 103–107.

Mehaffey, W. H., & Doupe, A. J. (2015). Naturalistic stimulation drives opposing heterosynaptic plasticity at two inputs to songbird cortex. *Nat Neurosci*, *18*(9), 1272–1280.

Miller, E. K. (2000). The prefontral cortex and cognitive control. *Nature Reviews Neuroscience*, *1*(1), 59–65. Retrieved from `https://doi.org/10.1038/35036228` doi: 10.1038/35036228

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu Rev Neurosci*, *24*, 167–202.

Minsky, M. (1961). Steps toward artificial intelligence. *Proceedings of the IRE*, *49*(1), 8-30. doi: 10.1109/JRPROC.1961.287775

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., … Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533.

Monchi, O., Petrides, M., Strafella, A. P., Worsley, K. J., & Doyon, J. (2006). Functional role of the basal ganglia in the planning and execution of actions. *Ann Neurol*, *59*(2), 257–264.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci*, *16*(5), 1936–1947.

Morris, G., Nevet, A., Arkadir, D., Vaadia, E., & Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future action. *Nat Neurosci*, *9*(8), 1057–1063.

Mukherjee, A., Bajwa, N., Lam, N. H., Porrero, C., Clasca, F., & Halassa, M. M. (2020). Variation of connectivity across exemplar sensory and associative thalamocortical loops in the mouse. *Elife*, *9*.

Mukherjee, A., Lam, N. H., Wimmer, R. D., & Halassa, M. M. (2021). Thalamic circuits for independent control of prefrontal signal and noise. *Nature*, *600*(7887), 100–104.

Murray, J. D., Bernacchia, A., Freedman, D. J., Romo, R., Wallis, J. D., Cai, X., … Wang, X. J. (2014). A hierarchy of intrinsic timescales across primate cortex. *Nat Neurosci*, *17*(12), 1661–1663.

Murray, M. M., Lewkowicz, D. J., Amedi, A., & Wallace, M. T. (2016). Multisensory Processes: A Balancing Act across the Lifespan. *Trends Neurosci*, *39*(8), 567–579.

Nakajima, M., Schmitt, L. I., & Halassa, M. M. (2019). Prefrontal Cortex Regulates Sensory Filtering through a Basal Ganglia-to-Thalamus Pathway. *Neuron*, *103*(3), 445–458.

Nambu, A. (2011). Somatotopic organization of the primate Basal Ganglia. *Front Neuroanat*, *5*, 26.

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154. Retrieved from `https://doi.org/10.1016/j.jmp.2008.12.005` doi: 10.1016/j.jmp.2008.12.005

O'Reilly, R. C. (1996). Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Computation*, *8*(5), 895-938. doi: 10.1162/neco.1996.8.5.895

O'Reilly, R. C., Russin, J. L., Zolfaghar, M., & Rohrlich, J. (2021). *Deep predictive learning in neocortex and pulvinar.*

Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., & Wermter, S. (2019). Continual lifelong learning with neural networks: A review. *Neural Netw*, *113*, 54–71.

Perrin, E., & Venance, L. (2019). Bridging the gap between striatal plasticity and learning. *Curr Opin Neurobiol*, *54*, 104–112.

Peters, A. J., Fabre, J. M. J., Steinmetz, N. A., Harris, K. D., & Carandini, M. (2021). Striatal activity topographically reflects cortical activity. *Nature*.

Petersen, C. C. H. (2019). Sensorimotor processing in the rodent barrel cortex. *Nat Rev Neurosci*, *20*(9), 533–546.

Phillips, J. M., Kambi, N. A., & Saalmann, Y. B. (2016). A Subcortical Pathway for Rapid, Goal-Driven, Attentional Filtering. *Trends Neurosci*, *39*(2), 49–51.

Power, J. D., & Schlaggar, B. L. (2017). Neural plasticity across the lifespan. *Wiley Interdiscip Rev Dev Biol*, *6*(1).

Rakic, P. (2009). Evolution of the neocortex: a perspective from developmental biology. *Nat Rev Neurosci*, *10*(10), 724–735.

Reinagel, P., Godwin, D., Sherman, S. M., & Koch, C. (1999). Encoding of visual information by LGN bursts. *J Neurophysiol*, *81*(5), 2558–2569.

Richards, B. A., & Lillicrap, T. P. (2019). Dendritic solutions to the credit assignment problem. *Curr Opin Neurobiol*, *54*, 28–36.

Rikhye, R. V., Gilra, A., & Halassa, M. M. (2018). Thalamic regulation of switching between cortical representations enables cognitive flexibility. *Nat Neurosci*, *21*(12), 1753–1763.

Roelfsema, P. R., & Holtmaat, A. (2018). Control of synaptic plasticity in deep cortical networks. *Nat Rev Neurosci*, *19*(3), 166–180.

Roelfsema, P. R., & van Ooyen, A. (2005). Attention-gated reinforcement learning of internal representations for classification. *Neural Comput*, *17*(10), 2176–2214.

Roesch, M. R., Calu, D. J., & Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci*, *10*(12), 1615–1624.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. Retrieved from `https://doi.org/10.1038/323533a0` doi: 10.1038/323533a0

Rusu, A. A., Rabinowitz, N. C., Desjardins, G., Soyer, H., Kirkpatrick, J., Kavukcuoglu, K., … Hadsell, R. (2016). Progressive neural networks. *CoRR*, *abs/1606.04671*. Retrieved from `http://arxiv.org/abs/1606.04671`

Saalmann, Y. B., & Kastner, S. (2015). The cognitive thalamus. *Front Syst Neurosci*, *9*, 39.

Sacramento, J., Ponte Costa, R., Bengio, Y., & Senn, W. (2018). Dendritic cortical microcircuits approximate the backpropagation algorithm. In *Advances in neural information processing systems* (Vol. 31, pp. 8735–8746). Curran Associates, Inc. Retrieved from `https://proceedings.neurips.cc/paper/2018/file/1dc3a89d0d440ba31729b0ba74b93a33-Paper.pdf`

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

Schiess, M., Urbanczik, R., & Senn, W. (2016). Somato-dendritic synaptic plasticity and error-backpropagation in active dendrites. *PLOS Computational Biology*, *12*(2), 1-18. Retrieved from `https://doi.org/10.1371/journal.pcbi.1004638` doi: 10.1371/journal.pcbi.1004638

Schmitt, L. I., Wimmer, R. D., Nakajima, M., Happ, M., Mofakham, S., & Halassa, M. M. (2017). Thalamic amplification of cortical connectivity sustains attentional control. *Nature*, *545*(7653), 219–223.

Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., … Silver, D. (2020). Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature*, *588*(7839), 604–609.

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

Seamans, J. K., & Robbins, T. W. (2010). Dopamine modulation of the prefrontal cortex and cognitive function. In *The dopamine receptors* (pp. 373–398). Totowa, NJ: Humana Press. Retrieved from `https://doi.org/10.1007/978-1-60327-333-6_14` doi: 10.1007/978-1-60327-333-6_14

Seo, M., Lee, E., & Averbeck, B. B. (2012). Action selection and action value in frontal-striatal circuits. *Neuron*, *74*(5), 947–960.

Sherman, S. M., & Guillery, R. W. (2005). *Exploring the thalamus and its role in cortical function, second edition* (Hardcover ed.). The MIT Press.

Sherman, S. M., & Spear, P. D. (1982). Organization of visual pathways in normal and visually deprived cats. *Physiol Rev*, *62*(2), 738–855.

Shin, H., Lee, J. K., Kim, J., & Kim, J. (2017). Continual learning with deep generative replay. In *Advances in neural information processing systems* (Vol. 30). Curran Associates, Inc. Retrieved from `https://proceedings.neurips.cc/paper/2017/file/0efbe98067c6c73dba1250d2beaa81f9-Paper.pdf`

Shine, J. M. (2021). The thalamus integrates the macrosystems of the brain to facilitate complex, adaptive brain network dynamics. *Prog Neurobiol*, *199*, 101951.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., … Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), 484–489.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., … Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature*, *550*(7676), 354–359.

Singer, W., Sejnowski, T., & Rakic, P. (2019). *The neocortex*. MIT Press. Retrieved from `https://books.google.com/books?id=aL60DwAAQBAJ`

Sohn, H., Meirhaeghe, N., Rajalingham, R., & Jazayeri, M. (2021). A Network Perspective on Sensorimotor Learning. *Trends Neurosci*, *44*(3), 170–181.

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

Soliveri, P., Brown, R. G., Jahanshahi, M., Caraceni, T., & Marsden, C. D. (1997). Learning manual pursuit tracking skills in patients with Parkinson’s disease. *Brain*, *120 (Pt 8)*, 1325–1337.

Suri, R. E., & Schultz, W. (1999). A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience*, *91*(3), 871–890.

Sutton, R., & Barto, A. (2018). *Reinforcement learning: An introduction*. MIT Press. Retrieved from `https://books.google.com/books?id=sWV0DwAAQBAJ`

Sutton, R. S., & Barto, A. G. (1990). Time-derivative models of pavlovian reinforcement. In *Learning and computational neuroscience: Foundations of adaptive networks* (pp. 497–537). MIT Press.

Tanaka, M. (2007). Cognitive signals in the primate motor thalamus predict saccade timing. *J Neurosci*, *27*(44), 12109–12118.

Tesileanu, T., Olveczky, B., & Balasubramanian, V. (2017). Rules and mechanisms for efficient two-stage learning in neural circuits. *Elife*, *6*.

Thomas-Ollivier, V., Reymann, J. M., Le Moal, S., Schück, S., Lieury, A., & Allain, H. (1999). Procedural memory in recent-onset Parkinson's disease. *Dement Geriatr Cogn Disord*, *10*(2), 172–180.

Thorndike, E. (2017). *Animal intelligence: Experimental studies*. Taylor & Francis. Retrieved from `https://books.google.com/books?id=1_hADwAAQBAJ`

Tsutsui, K., Hosokawa, T., Yamada, M., & Iijima, T. (2016). Representation of Functional Category in the Monkey Prefrontal Cortex and Its Rule-Dependent Use for Behavioral Selection. *J Neurosci*, *36*(10), 3038–3048.

Usrey, W. M., Alonso, J. M., & Reid, R. C. (2000). Synaptic interactions between thalamic inputs to simple cells in cat visual cortex. *J Neurosci*, *20*(14), 5461–5467.

Voytek, B., & Knight, R. T. (2010). Prefrontal cortex and basal ganglia contributions to visual working memory. *Proc Natl Acad Sci U S A*, *107*(42), 18167–18172.

Wang, J. X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., … Botvinick, M. (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nat Neurosci*, *21*(6), 860–868.

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

Whittington, J. C. R., & Bogacz, R. (2019). Theories of Error Back-Propagation in the Brain. *Trends Cogn Sci*, *23*(3), 235–250.

Wickens, J. R., & Kotter, R. (1994). Cellular Models of Reinforcement. In *Models of Information Processing in the Basal Ganglia.* The MIT Press. Retrieved from `https://doi.org/10.7551/mitpress/4708.003.0017` doi: 10.7551/mitpress/4708.003.0017

Wimmer, R. D., Schmitt, L. I., Davidson, T. J., Nakajima, M., Deisseroth, K., & Halassa, M. M. (2015). Thalamic control of sensory selection in divided attention. *Nature*, *526*(7575), 705–709.

Wolff, M., & Vann, S. D. (2019). The Cognitive Thalamus as a Gateway to Mental Representations. *J Neurosci*, *39*(1), 3–14.

Xiao, T., Zhang, J., Yang, K., Peng, Y., & Zhang, Z. (2014). Error-driven incremental learning in deep convolutional neural network for large-scale image classification. In *Acm multimedia* (ACM Multimedia ed.).

Yamins, D. L., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc Natl Acad Sci U S A*, *111*(23), 8619–8624.

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

Zenke, F., & Ganguli, S. (2018). SuperSpike: Supervised Learning in Multilayer Spiking Neural Networks. *Neural Comput*, *30*(6), 1514–1541.

Zenke, F., Gerstner, W., & Ganguli, S. (2017). The temporal paradox of Hebbian learning and homeostatic plasticity. *Curr Opin Neurobiol*, *43*, 166–176.

Zenke, F., Poole, B., & Ganguli, S. (2017). Continual learning through synaptic intelligence. In *Proceedings of the 34th international conference on machine learning* (Vol. 70, pp. 3987–3995). Retrieved from `http://proceedings.mlr.press/v70/zenke17a.html`

Zhou, H., Schafer, R. J., & Desimone, R. (2016). Pulvinar-Cortex Interactions in Vision and Attention. *Neuron*, *89*(1), 209–220.