Computational Neuroscientific Models of Working Memory

Thomas E. Hazy, Michael J. Frank, and Randall C. O'Reilly

1 Introduction

Originally coined by Newell and Simon (1956) in the context of computer science, the term *working memory* (WM) was introduced into Cognitive Psychology by G. A. Miller, Galanter, and Pribram (1960), who used it for the idea of holding goals and subgoals in mind in the service of planning and executing complex behaviors (Cowan, 2017). Since then the usage of the term has evolved in complex and nuanced ways such that Cowan (2017) could distinguish nine separate definitions currently in use by various researchers. For the work described in this chapter, the definition attributed to G. A. Miller et al. (1960) will be adopted (*Table 1: Glossary*).

Broadly speaking, there are two levels of computational working memory models: abstract cognitive-level models, and neurobiologically-based models, the latter of which are the primary focus of this chapter. These models are based on the discovery of persistent delay-period neuronal activity in the prefrontal cortex of nonhuman primates, in a variety of delayed-response tasks (e.g., Funahashi, Bruce, & Goldman-Rakic, 1989; Fuster & Alexander, 1971; Kubota & Niki, 1971). A central idea behind most of these models is that neural activity can be sustained through *mutual excitation*, where populations of interconnected neurons send each other excitatory activity in a self-perpetuating fashion (also described as *reverberant* or *recurrent* activity). Computationally, this corresponds to a stable *attractor* in a dynamical system: a state that remains constant over time once the system enters the vicinity of that state (known as the *attractor basin*) (see Barak & Tsodyks, 2014; X.-J. Wang, 2001, for reviews). This mechanism of working memory can be more specifically described as *robust active maintenance*, which is distinct from a more transitory form of continued neural activity in posterior cortex that can persist for a few hundreds of milliseconds, but is

---

Table 1: Glossary

Acronym / Term	Definition
WM	*Working Memory:* As used here, the set of cognitive processess used for holding goals and subgoals in mind in the service of planning and executing complex behaviors (after G. A. Miller et al., 1960 as attributed by Cowan, 2017).
1-2-AX	A hierarchical form of the AX-CPT in which the target sequence (AX vs. BY) is signaled by outer-loop cues (1 or 2).
ACT-R	*Adaptive Control of Thought – Rational:* A highly influential production system-based model of cognition developed by John Anderson and colleagues.
AX-CPT	A-then-X Continuous Performance Task: Subjects observe sequences of letters and have to respond correctly for the target sequence of an 'A' followed by and 'X'.
BPTT	*Back Propagation Through Time:* An extension of the backpropagation algorithm to RNNs
Backpropagation	The dominant learning algorithm used in connectionism and deep learning (Rumelhart, Hinton, & Williams, 1986)
BG	*Basal Ganglia:* A set of subcortical nuclei involved in modulating frontal cortical function including motor activity and executive function.
Connectionism	A very successful and highly influential approach to behavioral and, especially, cognitive modeling in Psychology that emerged in the 1980's and emphasized learning in neural networks.
Deep Learning	A general term for a growing number of neural network-based machine learning models that share the feature of having many different layers stacked hierarchially.
ID/ED	*Intradimensial/Extradimensional:* A dynamic categorization task switching task in which a block's operational rule switches either within a dimension (e.g., red vs. green) or extradimensionally (color vs. shape).
LSTM	*Long Short-Term Memory:* A highly influential recurrent neural network model developed by Juergen Schmidhuber and colleagues that introduced the idea of gating maintenance so as to protect it over long time periods.
ML	*Machine Learning:* A branch of computer science that deals with various forms of statistical learning. Roughly equivalent to artificial intelligence (AI)
N-back	A continuous performance task in which subjects must indicate when a currently displayed stimulus matches with one presented n-steps back. Typically $1 < n < 5$.
PBWM	*Prefrontal Cortex and Basal Ganglia Working Memory:* A neural network-based model of WM maintenance and updating that emphasized the role of the basal ganglia in gating items into active maintenance and updating them as appropriate (Hazy, Frank, & O'Reilly, 2007; O'Reilly & Frank, 2006).
Production System	A computer program typically used to provide a form of artificial intelligence. It is characterized by a set of *productions* or rules that pair states (IF part of the rule) with actions to be executed (THEN part of the rule).
PVLV	*Primary Value, Learned Value:* A neurobiologcially informed and constrained alternative to the temporal difference (TD) algorithm for generating reward prediction error (RPE) signals used to train the rest of a given network model.
RL	*Reinforcement Learning:* A branch of machine learning in which actions are learned by trial and error based only on scalar-valued feedback, i.e., good or bad.
RNN	*Recurrent Neural Network:* A category of neural network in which some subpopulation of the units feedback to excite themselves on sequential timesteps.
RPE	*Reward Prediction Error:* An error signal generated as the difference between actual received reward versus that that has come to be expected.
SRN	*Simple Recurrent Network:* An simple form of RNN that involves a direct copy of information from the prior time step to contextualize the current time step.
TD	*Temporal Differences:* The dominant RL algorithm for generating reward prediction error (RPE) signals used to train models.
Vector Rotation	A term used to describe the quantification of the changes in neural population activity that treats each unit as a single dimension in the high dimensional space corresponding to all recorded units. Thus, as the population activity changes over time it can be described as rotating in this high dimensional space.
WCST	*Wisconsin Card Sort Task:* Subjects match cards according to color or shape as defined by implicit rules that change periodically without instruction.

quickly overwritten by new stimuli (e.g., *distracters*).

Functionally, the ability to robustly maintain activity over time must also be complemented by an ability to rapidly update to encode new information into working memory, when such information is transiently present in the sensory input. These two demands are mutually contradictory, and the concept of *gating* has been introduced as a way to dynamically switch between robust maintenance versus rapid updating. The long-short-term-memory (LSTM) model (Hochreiter & Schmidhuber, 1997) introduced an abstract algorithm for multiple forms of gating (*maintenance* gating of new information into working memory, and *output* gating of maintained information from working memory), and various neurobiological mechanisms have been proposed to support gating, including the neuromodulator dopamine (Braver & Cohen, 2000; Durstewitz, Seamans, & Sejnowski, 2000; Seamans & Yang, 2004) and the *basal ganglia* (Dayan, 2007, 2008; Frank, Loughry, & O'Reilly, 2001; Frank & O'Reilly, 2006; Todd, Niv, & Cohen, 2008).

The neurobiologically-based approach has embraced empirical data from multiple species and levels of analysis to inform and constrain the models. At a systems and cognitive level of analysis, this work emphasizes the importance of working memory as a core component of higher cognitive function, including attention, cognitive control, decision-making, goal-directed behavior, and executive function (Baddeley, 1986; Baddeley & Hitch, 1974; Engle, Tuholski, Laughlin, & Conway, 1999; Friedman et al., 2006; Miyake et al., 2000). Machine learning algorithms (e.g., LSTM) are also an important source of inspiration for understanding the functional properties of such models, and learning more generally plays an important role in some of this work, to understand how complex cognitive functions can emerge from simpler neural machinery.

Sustained neural activity is essential for higher-level cognitive function, to enable consistent plans or goals to drive processing over the duration necessary to achieve desired outcomes. Mechanistically, actively-firing neurons in the prefrontal cortex can drive a *top-down biasing* of neurons in domain-specific posterior cortical areas, to focus their processing on task-relevant information (E. K. Miller & Cohen, 2001; O'Reilly, Braver, & Cohen, 1999). This is also known as *task-based attention*. The specific ability to maintain stable activity in the face of potentially distracting stimuli or thoughts has been an important feature of working memory in the cognitive literature (Baddeley & Hitch, 1974; Miyake & Shah, 1999), for example in the case of *complex working memory span* tasks, that require maintaining selected information in the face of ongoing complex cognitive processing.

The ability to plan or evaluate different possible future courses of action critically depends on this ability to maintain internal representations of these plans without the support of external stimuli. Indeed, based on the comparative development of frontal areas across species, the core working memory ability likely evolved to maintain affective goal states to guide behavior toward those goals, in frontal areas that correspond to ventral and medial areas in the primate brain (V. J. Brown & Bowman, 2002; Ongür & Price, 2000; O'Reilly, Russin, & Herd, 2019; Uylings, Groenewegen, & Kolb, 2003).

Table 2 includes specific examples of tasks and phenomena that have been modeled with this approach. For example, the *PBWM* model incorporates biologically-based mechanisms of frontal robust active maintenance, basal ganglia gating mechanisms, and learning mechanisms based on phasic dopamine, and can simulate a wide range of commonly studied working memory tasks including the 1-2-AX and phonological loop (O'Reilly & Frank, 2006), ID/ED dynamic categorization (O'Reilly, Noelle, Braver, & Cohen, 2002), WCST (Rougier & O'Reilly, 2002), N-back (e.g., Chatham et al., 2011), task switching, the Stroop task (Herd et al., 2014), hierarchical rule learning (Badre & Frank, 2012), and the reference-back-2 task (Rac-Lubashevsky & Frank, 2020).

This review of the field of neurobiologically-based working memory models focuses on the following central, open questions that characterize many of the important differences across existing models:

- From the gating perspective, what is the nature and scale of the neural substrate that is subject to gating modulation? The potential range here might extend from the gating of individual neurons at the most fine-grained end of the scale to the en-masse gating of the entire PFC by a global gating mechanism (e.g., the neuromodulator dopamine).

- What kinds of qualitatively different gating dynamics exist in the brain, and what are their respective neural substrates? Possibilities include: *input gating* (allowing sensory / bottom-up activation into prefrontal cortex), *maintenance gating* (updating new information into active maintenance), *forget gating* (removing, resetting active maintenance), and *output gating* (output of information from active maintenance).

- What is the temporal relationship between gating events and the maintenance period? For example, the gating of an item into robust maintenance could be a punctate event with the gate opening only transiently at the start, and then closing again. Alternatively, the gate could persist in an open state throughout the maintenance period, playing a critical role in sustaining the active maintenance.

- How static vs. dynamic are working memory representations over the maintenance period? Evidence for both relatively static, boxcar-like sustained activity, as well as various waxing-and-waning patterns of delay-period activity have been reported.

- What is the nature and source of working memory capacity limitations? Is capacity limited by something like a small number of discrete slots (Cowan, 2001; G. A. Miller, 1956), or is it more like a single shared resource (e.g., Ma, Husain, & Bays, 2014)?

- Can working memory representations provide a substrate for a form of content addressable memory in service of variable binding and transfer?

These questions also have numerous mutual interdependencies, such that a comprehensive theory needs to consider all of the issues interactively. Each of the above questions will be revisited in the *General Discussion* section that follows the model descriptions.

Although the focus is on the neurobiologically-oriented models here, there is an extensive literature on more abstract models that target human-level cognitive function specifically, and account for a range of behavioral data regarding the nature and limits to working memory capacity and the modalities involved (e.g., Logie, 2018; Oberauer et al., 2018a, 2018b; Vandierendonck, 2018). For additional background and reviews, interested readers are referred to other sources to learn about them (e.g., Adams, Nguyen, & Cowan, 2018; Burgess & Hitch, 2005). In addition, readers are encouraged to look at Oberauer et al. (2018a) for a compilation of benchmark human behavioral phenomena drawn from a wide swath of working memory tasks that a panel of researchers have deemed important for proposed models to address. These benchmarks constitute a kind of "psychophysics" of working memory: many different ways of probing the basic process of encoding and retrieving information over a relatively short interval, including: serial recall, free recall, complex span tasks, visual change detection, recognition, memory updating, and n-back.

The overall organization for the remainder of the chapter is as follows. First, the theoretical *Background* for many of the issues introduced here will be provided in following section. Then, models at different points on the spectrum articulated above are reviewed, considering how they might inform an understanding of the role of gating and whether there are qualitatively different forms of working memory systems or not. Finally, a synthetic summary of the basic ideas will be provided including a return to the motivating questions listed earlier.

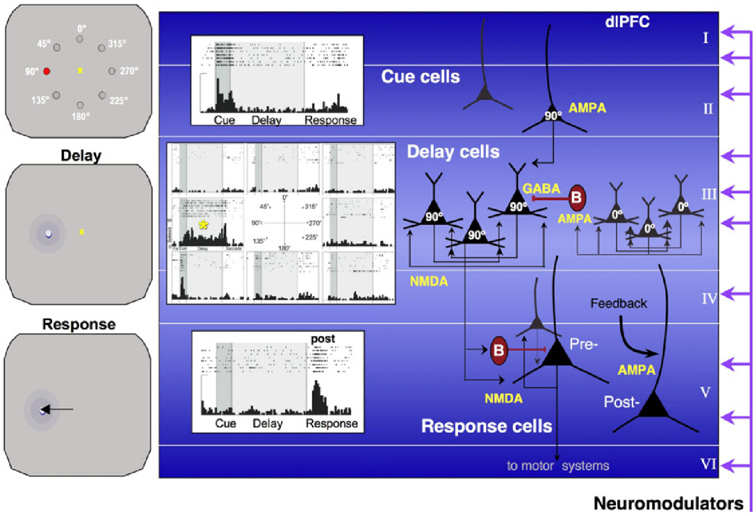

Figure 1: Detailed mapping of a standard occulomotor delayed response task onto patterns of neural activity across different lamina within the dorsolateral prefrontal cortex (dlPFC). Superficial layer (II) neurons receive bottom-up sensory inputs encoding the cued location for a delayed visual saccade, in this case, the red light at 90 degrees to the left of the central yellow fixation point. Specialized deep layer III neurons with extensive lateral recurrent connectivity, expressing both NMDA and GABA-B channels, provide the reverberant attractor dynamics to sustain the cue location over the delay period, during which time the animal must maintain central fixation. When the fixation cross dissappears, the animal is allowed to respond, and deep layer V output neurons drive the motor response, to saccade to the previously-cued target location. All aspects of this task are typically trained through reinforcement-based learning in a shaped fashion, such that the animal learns that reward only occurs when all steps are correctly performed. Figure adapted from Arnsten et al., (2012)

2 Background

The most central phenomenon for all neurobiological models of working memory is the sustained *delay period* firing of neurons in the prefrontal cortex (PFC) (e.g., Fuster & Alexander, 1971; Goldman-Rakic, 1995; Kubota & Niki, 1971; E. K. Miller & Desimone, 1994; Sommer & Wurtz, 2000). This phenomenon has been the subject of extensive computational modeling research, at multiple levels of analysis. The core ability for neural circuits to maintain a signal through the enduring firing of neurons has been extensively investigated through many variations on *attractor networks* (see Barak & Tsodyks, 2014; X.-J. Wang, 2001, for reviews). Specifically, neurons can maintain information over time through active firing sustained by a pattern of *mutual reciprocal excitation* (you pat my back and I'll pat yours, essentially). Although brief periods of self-sustained activity can be seen across much of the neocortex, the PFC seems clearly specialized in this regard (e.g., Funahashi et al., 1989; Fuster & Alexander, 1971; Goldman-Rakic, 1995; Kubota & Niki, 1971; E. K. Miller, Erickson, & Desimone, 1996; M. Wang et al., 2013). Thus, a critical question is: are there specialized neural mechanisms in the PFC that explain this ability?

Figure 1 from Arnsten, Wang, and Paspalas (2012) shows a widely-accepted framework for how these reverberant attractor dynamics operate within a standard oculomotor delayed response task to maintain the cue location during the delay period, enabling a delayed saccade to the cued location (J. W. Brown, Bullock, & Grossberg, 2004, developed an early system's level model with this structure, as discussed later). Specifically, a specialized population of deep layer 3 pyramidal neurons within the prefrontal cortex has been identified, which has extensive lateral, mutually excitatory (recurrent) connectivity (Kritzer & Goldman-Rakic, 1995; Y. Wang et al., 2006). This pattern of connectivity has undergone a prominent evolutionary expansion in primates (Elston, 2003; M. Wang et al., 2013), and has a high concentration of N-methyl-D-aspartate (NMDA) receptors which are important for stabilizing this reverberatory activity and contribute to its continued informational specificity. These receptors have a switch-like bistability, such that when they are activated they drive sustained excitatory currents that reinforce the activity of already-activated neurons. There are also important complementary bistable inhibitory GABA-B channels that prevent previously inactive neurons from becoming activated, which greatly enhances the robustness and stability of the attractor states (Sanders, Berends, Major, Goldman, & Lisman, 2013).

Several studies have shown that NMDA receptor blockade impairs working memory performance in multiple species (Krystal et al., 2005; Moghaddam & Adams, 1998; Roberts et al., 2010). A particularly elegant study by M. Wang et al. (2013) showed that the targeted administration of antagonists to NMDA, but not AMPA, in deep layer 3 pyramidal cells blocked persistent activity in monkey PFC and impaired performance on a spatial working memory task. These authors also showed that the NMDA receptors involved were phenotypically specialized to express high levels of the NR2B subunit.

The laminar specialization shown in Figure 1 makes sense according to standard patterns of cortical connectivity. Sensory inputs activate superficial layers directly and via layer 4, which then projects up to the superficial layers, and the subcortical output from the PFC arises from the deep layers, with the large layer 5b output neurons providing direct motor-level output (i.e., their axons constitute the pyramidal tract projections to the spinal cord). These layer 5b neurons also project to the basal ganglia and other subcortical targets. There is also a population of layer 6 corticothalamic (CT) neurons that project to the thalamus, which will be discussed below. In addition to driving output responses, the layer 5b output neurons also transmit both sensory input and sustained active maintenance signals, as revealed by the unambiguous recording of all of these firing patterns in identified layer 5b neurons (Sommer & Wurtz, 2000). This can arise from different patterns of projections from layer 2 and 3 neurons into layer 5b, and can be computationally useful in enabling all aspects of the PFC activity to be available to subcortical systems.

The issue of *gating* can be seen directly in the activation patterns illustrated in Figure 1. Specifically, what causes the layer 5b output neurons to only fire at the moment when a response should be initiated, and not sooner during the delay period? Furthermore, if the superficial layer neurons were always capable of updating the state of the layer 3 delay cells, irrelevant distracters would thus interrupt the working memory system, but a defining characteristic of working memory is its robustness in the face of such distractions. These questions are addressed in abstract, algorithmic terms by the LSTM model (Hochreiter & Schmidhuber, 1997), which has a *maintenance* gate that learns when to allow new information into working memory, and an *output* gate that learns when to allow information out of the working memory system. Both of these gates operate as a simple multiplicative factor on a precisely balanced, linear working memory cell that can perfectly maintain information indefinitely over time until further gated.

Thus, from a neurobiological perspective, a central question concerns the nature of possible neural mechanisms that could support these forms of gating. One early set of proposals focused on the neuromodulator *dopamine*, which affects virtually all aspects of the PFC circuitry, including NMDA and GABA-B receptors (Braver & Cohen, 2000; Durstewitz et al., 2000; Seamans & Yang, 2004). Specifically, transient changes in dopamine firing, driven by its synergistic role in reinforcement learning, could modulate the stability of activity dynamics in PFC, switching between robust maintenance and a more labile state where rapid updating is possible. However, such a mechanism would likely affect all of PFC at a time, due to the widespread nature of dopamine innervation, and the relative homogeneity of dopamine cell firing, making it difficult to *selectively* update some information while robustly maintaining other states. For hierarchical motor control and various standard working memory tasks, this ability to selectively update is essential.

Motivated by data on the extensive interconnectivity and functional relevance of the basal ganglia (BG) for frontal function (G. Alexander, DeLong, & Strick, 1986; R. G. Brown & Marsden, 1990; Graybiel, 1995; Middleton & Strick, 2000; Mink, 1996), a number of models have advanced the idea that the BG are well-positioned to provide this more selective gating function (Beiser & Houk, 1998; J. W. Brown et al., 2004; Dayan, 2007, 2008; Dominey & Arbib, 1992; Frank, 2005; Frank et al., 2001; Gruber, Dayan, Gutkin, & Solla, 2006; Houk, 2005; O'Reilly & Frank, 2006; Todd et al., 2008). Other work has directly addressed BG gating from a theoretical and empirical perspective (Chatham, Frank, & Badre, 2014; Dahlin, Neely, Larsson, Backman, & Nyberg, 2008; Voytek & Knight, 2010). Specifically, there are numerous parallel loops of circuitry between the frontal cortex and BG that could provide a more selective, focal gating signal, and the essential function of the BG is widely thought to be to disinhibit excitatory corticothalamic loops in frontal cortex. In the motor domain, this disinhibition is thought to drive the initiation of overt motor actions (Mink, 1996). Thus, by analogy, BG gating in higher-level PFC areas could drive the initiation of cognitive-level actions, including the updating of working memory representations.

With the above providing a relatively well-established foundation, the next section will motivate some of the more unresolved questions that different neurobiologically-based computational models have explored, which will then be reviewed in greater detail in the remainder of the chapter.

2.1 *The Nature of (BG) Gating and PFC Representations*

The nature of working memory gating at many different levels of analysis represents a huge space of unresolved questions, including the most basic question of whether gating is really even present in the first place. Some of these questions were highlighted in the introduction, including: the granularity over which gating might operate; which of the different kinds of gating (maintenance, output and others) might be active, and via which neural mechanisms; and how might gating dynamics relate to maintenance activation?

At the abstract computational level of analysis, there is an influential set of papers that showed how some working-memory like abilities could emerge in a basic type of recurrent neural network (RNN) without any form of gating mechanism (Botvinick & Plaut, 2004, 2006). Interestingly, these models focused on well-learned types of behavior, including highly-practiced task performance and immediate serial recall (e.g., repeating a phone number or other information you've just been told), and they took 100's of thousands of trials to learn. These models also lacked any strong form of specialized active maintenance mechanism, and instead learned to shape dynamically unfolding patterns of neural activity over time to systematically encode the relevant temporal structure.

To help situate these models within a larger functional taxonomy, the well-established dichotomy between *controlled* and *automatic* (habitual) processing in human behavior (Cohen, Dunbar, & McClelland, 1990; O'Reilly, Nair, Russin, & Herd, 2020; Shiffrin & Schneider, 1977) is particularly relevant. Controlled processing is specifically required in cases of novel or difficult cognitive tasks that require sustained attention and, typically, multiple cognitive steps. Paradigmatic examples include mental arithmetic, planning moves in a game of chess, and evaluating multiple potential aspects of a difficult decision-making problem. By contrast, automatic processing occurs for well-learned, often single-step cognitive operations, for example reading printed words. The widely-studied Stroop task demonstrates this distinction very clearly, where automatic word reading is unaffected by irrelevant ink colors, but less well-practiced color naming is strongly affected by conflicting color words (Dunbar & MacLeod, 1984; Stroop, 1935).

Thus, one could argue that the highly-trained, fine-grained, non-gated dynamics of recurrent neural networks capture the faster time-scale, automatized forms of behavior and cognition associated with well-learned tasks, which are thought to be supported by cortical networks in the parietal and lower-order frontal motor areas. In contrast, controlled processing may require strongly gated, more discrete, longer-time-scale dynamics supported by BG / PFC based models. The working memory contents in this latter case reflect plans, goals, and other more sustained forms of information, associated with dorsolateral PFC (dlPFC) and ventromedial PFC (vmPFC) areas. One can think of these controlled processing roles of the BG / PFC circuitry as longer-time-scale "outer loops" of cognitive function involved in maintaining and selecting task plans and goals, that organize the sequential order of actions and cognition over longer periods of time. Within these outer loops, "inner loops" of more automatic, well-learned cognitive steps and actions take place.

Thus, instead of representing a challenge to the importance of gating and specialized active maintenance mechanisms, the basic RNN models help to delineate the specific domain of relevance for these mechanisms, within the higher-level cognitive control / executive function domain, which is where at least some of these models have been specifically targeted.

Within the space of models with gating mechanisms, the question of *representational granularity* is of central importance. On one end of the spectrum is the LSTM model, which is typically used with each individual working memory unit having its own dedicated set of gating units. This produces a very fine-grained, diverse, and dynamic set of memory signals updating separately in many different ways over time. By contrast, the more biological models based on the constraints dictated by the BG / PFC system require a significantly more coarse-grained form of gating. Specifically, as reviewed below, biological data establish that there are orders of magnitude fewer gating neurons in the output nuclei of the BG, relative to neurons of the frontal cortex, meaning that relatively large aggregates of frontal neurons should share gating signals.

At the most coarse-grained end of the spectrum, the widely-used ACT-R computational modeling framework (Anderson & Lebiere, 1998; Stocco, Lebiere, & Anderson, 2010) features the BG as the central bottleneck that drives the sequence of production firing steps, according to the classical *production system* model of higher-level cognitive function. A *production* in this framework represents a single automatic inner-loop step of processing, such as adding together two single-digit numbers, retrieving a fact from declarative memory, or focusing attention on a particular element in a visual input display. Critically ACT-R requires that only a single such production can fire at any given time, producing a very coarse-grained form of gating (at least in the temporal domain), compared to models where many different gating signals can fire in parallel.

Interestingly, there is a nice convergence between the abstract, cognitive-level ACT-R framework and the more biologically-based BG-gating models (Jilk, Lebiere, O'Reilly, & Anderson, 2008), even though they were derived from very different starting points. The principle of BG-gating of PFC active maintenance is the hub that connects these frameworks most directly. Remarkably, based on purely behavioral considerations, the ACT-R framework converged on a production firing constraint of no-faster-than 50 msec, which directly matches the intrinsic oscillatory mode of the BG circuit (Bogacz, 2013; Courtemanche, Fujii, & Graybiel, 2003; Schmidt et al., 2019).

Another important angle on the representational question is in terms of how dynamic and high-dimensional working memory representations are over time and representational space? Several electrophysiological studies support the notion of *mixed selectivity* coding, where individual neurons have complex, high-dimensional response profiles relative to relevant task variables (Fusi, Miller, & Rigotti, 2016; Mante, Sussillo, Shenoy, & Newsome, 2013). The high dimensional aspects of mixed selectivity are recognized to be useful for flexibility in solving arbitrary tasks, but they come at the expense of generalizing to new stimuli within a dimension. On the other hand, a long history of studies also support a more discrete, lower-dimensional organization, with more discrete, "square wave" style temporal dynamics (Funahashi et al., 1989; Fuster & Alexander, 1971; Goldman-Rakic, 1995; Kubota & Niki, 1971; Sommer & Wurtz, 2000). These different temporal dynamics may interact with the representational organization of information as well, with more fluid, high-dimensional, mixed-selectivity coding associated with the more automatic processing, inner-loop end of the spectrum, and more discrete, square-wave dynamics associated with the more controlled, outer-loop end of the spectrum.

Ultimately, the computational models can only serve to raise and focus questions, and further empirical studies are required to more definitively answer these questions. For example, does the proposed distinction between more continuous, fine-grained, dynamical models and the more discrete, broader-scale gated models fit with direct contrasts between different levels of PFC and posterior cortex? Or, is it possible there is only one of these two types of mechanisms operating in the brain, supporting the whole scope of relevant time-scales and modes of cognitive function? And more specifically, for gating operating through the BG, how is this gating organized relative to representational content and neural structure, under the strong biological constraints that there are many fewer gating neurons in the BG relative to PFC neurons. Is there evidence for separate gating signals for different chunks of PFC, and what is the organization of these chunks if so?

At a more detailed, biological level, there are a number of questions about the neural mechanisms that could subserve different forms of gating (maintenance, output, etc). Based on the laminar organization of PFC (Figure 1), maintenance gating should preferentially affect the specialized deep layer 3 neurons (M. Wang et al., 2013), while output gating ultimately needs to affect the subcortically-projecting layer 5b output neurons (e.g., J. W. Brown et al., 2004; Harris & Shepherd, 2015; Larkum, Petro, Sachdev, & Muckli, 2018; Ramaswamy & Markram, 2015; Sommer & Wurtz, 2000). Interestingly, there are two different types of thalamic afferents to cortex, *core* vs. *matrix*, which may differentially impact these cortical layers (Clascá, Rubio-Garrido, & Jabaudon, 2012; Jones, 1998a, 1998b, 2007; Phillips et al., 2019), and could thus be involved in both forms of gating. Specifically, core-type thalamic projections target the central layers, including 3 and 4, while matrix-type preferentially target layer 1 where the apical tufts of pyramidal cells from layers 2, 3, and 5b reside, the thick tufts of subcortically-projecting layer 5b being particularly prominent (Harris & Shepherd, 2015; Larkum et al., 2018; Ramaswamy & Markram, 2015).

Furthermore, most areas of the frontal cortex receive input from at least two different thalamic nuclei, and both core- and matrix-type thalamic relay cells, with medial dorsal (MD) nucleus prominently sending core-type projections (Giguere & Goldman-Rakic, 1988), but also having matrix-type cells (Münkle, Waldvogel, & Faull, 2000; Phillips et al., 2019). On the other hand, certain ventral thalamic areas (VM, VA) predominantly send matrix-type (Kuramoto et al., 2009, 2015), while VL mostly sends core-type (Kuramoto et al., 2009). In addition, the basal ganglia output nuclei target the matrix-type ventral thalamic areas more densely and uniformly as compared to the more patchily covered MD (Ilinsky, Jouandet, & Goldman-Rakic, 1985; Kuramoto et al., 2009, 2015; Tanibuchi, Kitano, & Jinnai, 2009a).

Putting these biological data points together, the resulting hypothesis would be that BG-mediated effects on frontal cortex may be predominantly on the output-gating side (matrix type, targeting 5b output neurons), while corticothalamic pathways independent of the BG, predominantly via the MD, may drive PFC maintenance gating (core type, targeting layer 3). This is consistent with a growing body of empirical evidence supporting a role for the MD nucleus in both the maintenance (Tanibuchi, Kitano, & Jinnai, 2009b; Watanabe & Funahashi, 2012; Watanabe, Takeda, & Funahashi, 2009; Wyder, Massoglia, & Stanford, 2004) and updating (Rikhye, Gilra, & Halassa, 2018) of sustained PFC activity. While this idea remains relatively unexplored computationally, it nevertheless shows how neurobiologically-based models can usefully incorporate anatomical data to inform an understanding of the nature of the computations. It is also important to emphasize that output gating in one PFC area could then directly influence maintenance in other areas, and that the BG-driven gating could still result in sustained neural firing in targeted PFC areas, so it will likely require more detailed implemented computational models to really sort through the full implications and unique signatures of these different types of gating. Ideally, the predictions of such models could then be tested empirically, at which point some more definitive level of understanding could be established.

2.2 *Learning Mechanisms*

Another central question for the working memory system is how it ends up being “intelligent” enough to function as one of the core systems of generalized fluid intelligence, as cognitive-level theories and psychometric data suggest (Engle et al., 1999; Friedman et al., 2006; Miyake et al., 2000). Without a clear answer to this question, the PFC / BG working memory system ends up as a kind of unexplained *homunculus* — a “little person” inside the head that makes humans smart (Hazy, Frank, & O’Reilly, 2006; Hazy et al., 2007). One clear answer to this question is that the system *learns* how to strategically control the maintenance and updating of working memory over the protracted timecourse of PFC functional development.

As such, one of the intriguing features of the dopamine-based gating hypothesis (Braver & Cohen, 2000) was that it built the gating dynamics directly on top of an emerging understanding of phasic dopamine signaling in reinforcement learning (RL) (Montague, Dayan, & Sejnowski, 1996; Schultz, Dayan, & Montague, 1997), thus providing a direct connection to learning. Subsequent models based on BG gating also retained this connection to dopamine-based RL (Hazy et al., 2006, 2007; O’Reilly & Frank, 2006), operating directly within the BG where dopamine receptors are the most dense, and extensive evidence supports a critical role for dopamine in shaping learning in a manner directly compatible with these models (Collins & Frank, 2014;

Frank, 2005; Frank & O'Reilly, 2006; Gerfen & Surmeier, 2011; Moustafa, Sherman, & Frank, 2008).

These biologically-motivated uses of dopamine-based RL are broadly consistent with current machine-learning approaches that combine RL with deep learning networks (i.e., *Deep RL*), which have proven successful at learning to succeed at a variety of different competitive games including Atari video games, chess, and Go (e.g., Mnih et al., 2015). However, the LSTM gating model upon which Deep RL is based still relies on a form of error backpropagation that is difficult to reconcile with known biology (unlike simpler forms of backpropagation which do have a reasonable biological mapping; O'Reilly, 1996; Whittington & Bogacz, 2019). Overall, the direct connection between dopamine and motivated, goal-driven learning may be synergistic with the task-driven function of the PFC more generally, and together with its known biological basis, suggests it may be the more likely form of learning in these systems.

Also, the combination of RL with selectively updatable, actively maintained working memory representations can be exploited to produce a sort of inductive bias to use those representations in a way that can be co-opted under new task conditions, resulting in a form of out-of-distribution generalization or learning transfer (Bhandari & Badre, 2018; Collins & Frank, 2013, 2016; Frank & Badre, 2012; Kriete, Noelle, Cohen, & O'Reilly, 2013; Rougier, Noelle, Braver, Cohen, & O'Reilly, 2005; A. Williams & Phillips, 2020). Thus, there may be some connection with human-level symbolic-like processing abilities and these underlying neural systems (O'Reilly et al., 2014).

2.3 *Activity-Silent Working Memory*

Finally, although the focus here is mostly on the neural mechanism of sustained neural firing, considerable work has shown that the broad functionality attributed to working memory can also be supported by other neural mechanisms. For example, Braver and colleagues have championed the distinction between *proactive* vs. *reactive* cognitive control in which the former corresponds to sustained neural firing to span a temporal delay while the latter involves the temporary offline storage, e.g., in the hippocampus, and its retrieval later at the time in which the information is actually needed (e.g., Braver, Paxton, Locke, & Barch, 2009).

More recently, the potentially related idea of *activity-silent* working memory has gained considerable traction, based on the observation that neural activity is often quite variable during the delay interval, and sometimes seemingly even nonexistent (Stokes, 2015). Thus, perhaps *temporary* strengthening of recurrent synapses involved in WM could be contributing, consistent with the role of long-acting, intrinsic cellular mechanisms (e.g., O'Reilly & Frank, 2006; X.-J. Wang, 2001), specifically the recruitment of NMDA receptors shown to be critical for stabilizing reverberatory activity. It has also been proposed that activity-silent working memory reflects an optimization that PFC can use if it can get away with it, but not if manipulation of longer maintenance is needed (Masse, Yang, Song, Wang, & Freedman, 2019), which is consistent with the broader idea that the more demanding form of working memory supporting executive function may require sustained active maintenance, but more automatized forms may not.

Next, the following section will delve deeper into the ideas and questions raised here and in the *Introduction*, starting with a more detailed discussion of the abstract machine-learning level computational models, and then working down to more biologically-based models.

Table 2: Working Memory Models Covered

Model	**Salient Features**	**Key Results**
Active Maintenance - Persistent Cortical Activity		
Attractor-based	Corticocortical reverberant activity	Long time constants of NMDARs enable persistent activity (X.-J. Wang, 2001) Specialized NR2B NMDAR subunits critical to robust maintenance (M. Wang et al., 2013) (Nassar, Helmers, & Frank, 2018)
	Corticothalamocortical reverberatory activity	Mouse ALM (Guo et al., 2017)
Gating-Relevant (Machine Learning)		
AlphaStar (Deep Mind)	Deep RL, DCNN	Defeated human players at Starcraft II (Vinyals et al., 2019)
Botvinick-Plaut	SRN + BPTT	Immediate serial recall (Botvinick & Plaut, 2006)
Deep Q-Network	Deep RL	Learned to play a large suite of Atari games (Mnih et al., 2015)
LSTM	Multiple forms of fine-grained gating	(Hochreiter & Schmidhuber, 1997) (Gers, Schmidhuber, & Cummins, 2000) (Schmidhuber, Gers, & Eck, 2002)
Open AI Five	Deep RL (includes LSTM)	Team of five cooperating artificial agents defeated tournament-level human teams in Dota2 (https://openai.com/five)
	Combined Deep RL with supervised learning with sensory feedback signals	Learned facile manipulation using human-like robotic hand (Dactyl) (https://openai.com/blog/learning-dexterity/)
BG-Based Gating		
Beiser-Houk	i - Maintenance gating: reverberant corticothalamocortical activity ii - Transient disinhibition of thalamic relay cells switches them into a persistently active up state	Sequence learning (Beiser & Houk, 1998)
Dominey-Arbib	i - Maintenance gating: persistent suppression of BG output permits sustained corticothalamocortical reverberant activity.	i - Memory-guided saccades. (Dominey & Arbib, 1992)
	i - Input gating - BG selects between two presented potential targets	ii - Visuomotor discrimination for selective saccades (Arbib & Dominey, 1995; Dominey, Arbib, & Joseph, 1995)
FROST	i - Explicitly excludes a role for BG in the *initiation* of maintenance gating ii - Attentional, cortically-initiated maintenance feeds back to BG that then helps support it	Memory-guided action selection. (Ashby, Ell, Valentin, & Casale, 2005) Attentional effects on working memory capacity
Gruber et al.	Phasic dopamine trigger mechanism affects the bi-stability of cells in both BG and cortex	Initiation of WM maintenance; Prevention of drift for WM representations in continuous space. (Gruber et al., 2006)
PBWM	i - Intrinsic cellular maintenance mechanisms triggered by BG gating signals ii - Phasic dopamine signals train BG gating signals based on correct/incorrect outputs	1-2-AX, Phono loop (Hazy et al., 2007; O'Reilly & Frank, 2006), WCST (Rougier & O‘Reilly, 2002), N-back Chatham et al. (2011) task switching, the Stroop task (Herd et al., 2014), reference-back-2 task (Rac-Lubashevsky & Frank, 2020), and more...
Schroll et al.	Increased STN activity in response to salient stimuli transiently suppresses the thalamus and terminates reverberant corticothalamocortical activity.	WM memoranda updating (Schroll, Vitay, & Hamker, 2012)
TELOS	i - Division of labor between superficial cortical layers for maintenance and deep for output. ii - BG gating of maintenance signals in superficial cortical layers to deep layers for output	Output gating by BG of memory-guided saccades trained by RL (J. W. Brown et al., 2004)

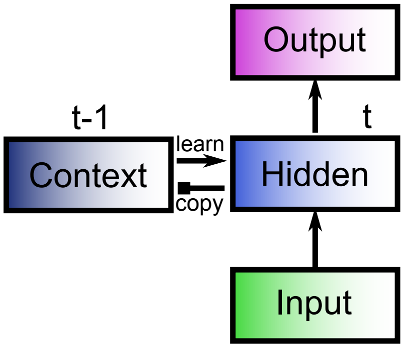

Figure 2: The simple recurrent network (SRN). The Context layer holds a copy of the prior (t-1) Hidden layer activation state, and the current Hidden layer has learnable synaptic weights that can adapt to incorporate this temporal context as needed to help learn the current Input / Output mapping. However, anything that is not needed on the current or few subsequent time steps will be rapidly forgotten: the system has a very limited effective memory span.

3 Recurrent Neural Networks, LSTM, and the Deep Learning Revolution

The machine learning / AI version of the classic attractor model of working memory involves *recurrent neural network (RNN)* models, which have some form of recurrent (reciprocal) connectivity, in contrast to the more predominant, simpler forms of neural networks that are purely *feedforward*. The *simple recurrent network (SRN)* (Cleeremans, Servan-Schreiber, & McClelland, 1989; Elman, 1990; Jordan, 1986) is a particularly simple version, based on a feedforward backpropagation network in which a copy of a layer's activation vector after each timestep is fed back into the network on the following timestep as an additional input, most typically involving the hidden layer feeding back into itself (Figure 2). This $t-1$ activation vector is input by a weight matrix that connects each $t-1$ unit with all of the hidden units at timestep t; that is, there is an all-to-all projection from a hidden layer to itself, offset by one timestep.

Thus, a hidden layer's previous activity state provides a continually updated and integrated temporal context input to itself at every timestep. Then, the recurrent weights conveying the $t-1$ information are updated after every timestep along with all the other network weights according to the standard backpropagation algorithm (Rumelhart et al., 1986). More recently, there is some indication that thalamocortical circuits in the posterior cortex might support something very similar to the SRN, which would be consistent with a more short-term role (O'Reilly, Russin, Zolfaghar, & Rohrlich, 2020).

Whereas learning in the SRN is limited to looking back a single timestep, a more general, powerful learning algorithm was also developed, known as *backpropagation through time (BPTT)* (R. J. Williams & Zipser, 1992), which can be understood as an "unrolling" of the multiple iteration timesteps of network

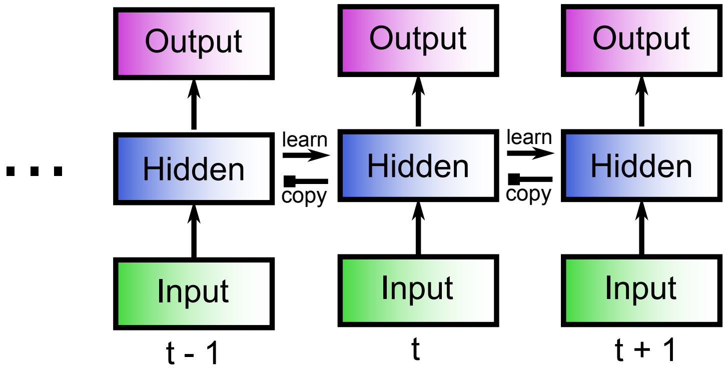

Figure 3: "Unrolling" an SRN for back-propagation through time (BPTT). As in the simple SRN a copy of the Hidden layer activation state is saved at the end of each timestep that sends learning weights to the current Hidden layer that can adapt to incorporate this temporal context as needed to help learn the current Input / Output mapping. All of these weights are then adapted after each timestep by the usual gradient descent back-propagation algorithm. Because of the veridical representation of each timestep's context the effective memory span of the system is extended.

processing constituting a particular sequence into an equivalent "spatialized" network to which standard back-propagation can be applied (Figure 3), with the critical factor being that in calculating the gradient-based contribution to the output error the recurrently connected hidden layer now has two descendent layers contributing to the calculation: the output layer on the *current* timestep as well as the hidden layer on the subsequent one. Although only a tiny part of the full BPTT algorithm as described in the Goodfellow, Bengio, and Courville (2016) text, Equation 1 shows how the BPTT computation of the gradient for a recurrently connected hidden layer depends on two descendent layers:

$$\nabla_{h^{(t)}} L = \left(\frac{\partial h^{(t+1)}}{\partial h^{(t)}} \right)^{\top} (\nabla_{h^{(t+1)}} L) + \left(\frac{\partial o^{(t)}}{\partial h^{(t)}} \right)^{\top} (\nabla_{o^{(t)}} L) \tag{1}$$

where $\nabla_{h^{(t)}} L$ and $\nabla_{o^{(t)}} L$ are the per timestep gradient contributions to the loss (error) function, L, of the hidden, h, and output, o, layers, respectively; and $\left(\frac{\partial h^{(t+1)}}{\partial h^{(t)}} \right)^{\top}$ and $\left(\frac{\partial o^{(t)}}{\partial h^{(t)}} \right)^{\top}$ are matrices of partial derivatives of the unit-by-unit changes in activity of descendent layers $h^{(t+1)}$ and $o^{(t)}$, respectively, with respect to the hidden layer activity on the reference timestep $h^{(t)}$. For further details on BPTT as well as the standard back-propagation algorithm itself, interested readers are referred to the excellent text by Goodfellow et al. (2016), and/or a very informative tutorial-level treatment by Werbos (1990), one of the original inventors of the back-propagation algorithm (Werbos, 1974).

The BPTT procedure can be combined with the SRN context copying method, and the combination can be quite powerful. Two important applications of this combined model (Botvinick & Plaut, 2004, 2006) provide a good illustration of the potential abilities of the more dynamic form of working memory, as explored next.

3.1 *The Botvinick-Plaut RNN Model*

The key contribution of the Botvinick and Plaut (2004) RNN model was to show that extensive backpropagation training enabled the model to develop a structured, hierarchical encoding of a well-learned task (preparing a cup of instant coffee or tea), which was robust to disruption in the sequence of events, and behaved similarly to humans overall. This model thus overcame the major limitation of a purely sequential *chaining* approach to sequence learning, which is that chaining is catastrophically brittle to any sort of disruption in the processing of a sequence, because every timestep is completely dependent on the state resulting from the prior one. Specifically, the extensive training enabled hierarchically-organized cross-step contingencies to be learned, overcoming the short-time-scale working memory properties of the SRN mechanism.

Subsequently, Botvinick and Plaut (2006) addressed the working memory domain more directly by adapting their model to reproduce many of the patterns of errors made by normal and impaired participants in a *serial recall* task in which four to-be-remembered items were presented in sequence (encoding stage) and the network was required to reproduce the items in the same order during a decoding stage. Like the coffee-making results, the core finding was that the network was again robust to disruption having learned representations that captured aspects of the hierarchical nature of the task on its own. Figure 4 shows how the hidden layer activation vector in this model evolves over the course of four encoding timesteps followed by four decoding timesteps. At each time point, the hidden layer population vector changes so as to best match its efferent weights to the output layer such that the output units decode the proper item in sequence.

What is responsible for this behavior? The answer is *learning* and the power of *distributed representations* (Hinton, McClelland, & Rumelhart, 1986). Consider the first recall timestep (second circled 1 starting from the left in Figure 4) during the training process. The context layer's population vector copied over from the previous timestep will correspond most to just-encoded stimulus 4. If the network's output on this timestep is wrong, the recurrent weights from the context to the current hidden layer will be weakened so that next time around a different output might be made. If correct, the recurrent weights will be strengthened, in particular, those weights coming from the context layer units that overlap with the activation vector

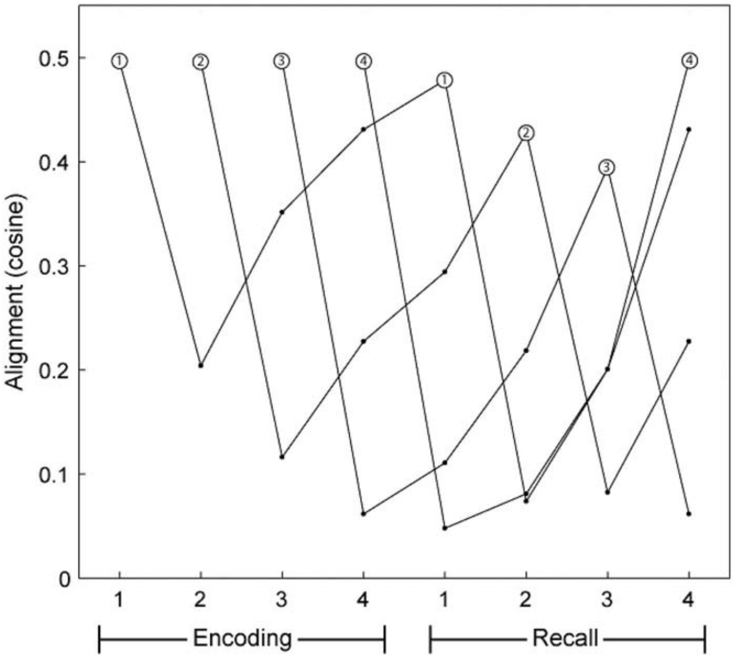

Figure 4: How RNNs convey information over time to make it available when needed. Data associated with four items are shown (circled numbers 1-4). Data points reflect a similarity measure between the population activity vector in the hidden layer and the corresponding weights that connect the hidden layer to the output. Following memorandum 1 as an example, note the high similarity value on the first (encoding) trial in which the network must output the identity of that item. However, on the next trial, the similarity drops precipitously when the second item is encoded (and output). Subsequently, the similarity measure for the first item gradually rises over further encoding trials until it again becomes highest on the fifth trial, which is the first decoding trial and the first item needs to be output again. One can think of the activity vector in the hidden layer as rotating over the sequence of trials such that each item to be recalled takes turns being the best match to the output weights. Figure from Botvinick & Plaut, 2006, Figure 5.

that most corresponds to stimulus 1.

Gradually, based on changes in the recurrent weights from the context layer (hidden at t-1), the current hidden activation vector will come to approach that corresponding to outputting stimulus 1. In this way, the population vector of the hidden layer comes to change systematically over subsequent timesteps in a way that allows for correct sequential outputs. This systematic change in the population vector activity is sometimes called “vector rotation” (see *Table 1: Glossary*). Thus, this evolution of the population vector along a trajectory that exposes representations only at the appropriate time is reminiscent of the dynamic population vector trajectories described in activity-silent and/or dynamically evolving working memory representations (e.g., Stokes, 2015; Stokes et al., 2013).

These models may best describe an *implicit* form of memory where the relevant information is deeply embedded in complex neural dynamics, which might be difficult for other systems to access in more gener-

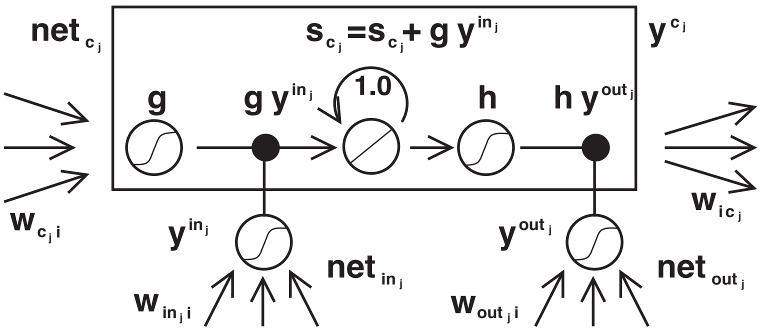

Figure 5: The LSTM memory cell (rectangle) with constant error carousel (CEC; circle with diagonal chord). See main text for explanation. From Hochreiter & Schmidhuber, 1997, Figure 1.

alizable, flexible ways. Furthermore, such dynamic temporally-evolving representations would not appear to be ideal for broadcasting a sustained plan of action, or desired goal state, over a relatively long period of time, to guide coordinated behavior across a wide range of different brain areas toward carrying plans and achieving goals. Indeed, most theories of conscious awareness emphasize that sustained stable activity over relatively long time periods (10's to 100's of milliseconds) is a necessary property (Lamme, 2006; Seth, Dienes, Cleeremans, Overgaard, & Pessoa, 2008), consistent with this overall idea that the kinds of memory associated with these rapidly rotating high-dimensional activity states would likely not be consciously accessible. This is consistent with the overall suggestion that the form of working memory supporting controlled processing is distinct from that supporting highly automated sequential behavior.

3.2 *Long Short-Term Memory and Gating*

Despite capturing many aspects of human behavior, the SRN / BPTT models remained strongly limited on their ability to span longer temporal delays, because each additional step back in time, which is equivalent to adding an additional hidden layer in the BPTT framework (Figure 3), results in another step of exponential decay of both the activations and the backpropagated learning signals (i.e., the "vanishing gradient" problem; Goodfellow et al., 2016). They also had difficulty filtering out the effects of distracters, and selectively updating to encode infrequent relevant items from a sequential stream. Furthermore, whatever flexibility and robustness they were able to exhibit required extensive training, and even then was relatively limited. To directly solve these problems, Schmidhuber and colleagues introduced dynamic, learned *gating* mechanisms in the long short-term memory (LSTM) model (Gers et al., 2000; Hochreiter & Schmidhuber, 1997; Schmidhuber et al., 2002).

The fundamental functional element in LSTM is the *memory cell* (the rectangular box in Figure 5). At the core of the memory cell is the *constant error carousel (CEC)*, which is effectively a unit having a linear activation function and a fixed self-recurrent connection of weight 1.0 (the circle with a diagonal chord at middle-bottom of the rectangle), which enables it to store activity states in veridical form over a potentially indefinite number of timesteps. By itself, however, the CEC would be constantly bouncing around under the influence of every input signal into it, and therefore the LSTM model added learnable *gating* units that preserve the CEC's current state when the gate is closed, and allow it to rapidly update when the gate is open. Thus, the CEC state s_{c_j} is updated at each timestep according to the following equation:

$$s_{c_j}(t) = s_{c_j}(t-1) + g(net_{c_j})(t))y^{in_j}(t) \tag{2}$$

where $s_{c_j}(t)$ is the CEC's activity state at timestep t; $g(net_{c_j}(t))$ is a nonlinear, squashing activation function with codomain 0 to 1; and $y^{in_j}(t)$ is the activation of the input gate function in_j (left circle beneath the rectangle with S-shape inside).

Furthermore, an output gate unit (right circle with S-shape) determines when the CEC activation is communicated to other neurons. Thus, the output of the memory cell, y^{c_j}, is computed at each timestep as follows:

$$y^{c_j}(t) = y^{out_j}(t)h(s_{c_j})(t)) \tag{3}$$

where $y^{c_j}(t)$ is the memory cell's output at each timestep; $y^{out_j}(t)$ is the activity of the output gate unit out_j; and $h(s_{c_j}(t))$ is a nonlinear function of the CEC's current state value, s_{c_j}..

With these gates in place, the LSTM can lock in and hold information for indefinitely long time periods, and learn to drive outputs at precise points in the future. Hochreiter and Schmidhuber (1997) adapted a real-time variant of the BPTT logic described by Robinson and Fallside (1987) for learning when to open and close these gates, as a function of overall task error. Critically, the input and output gates not only gate access in and out of the CEC state, they also serve to filter learning by gating the access of back-propagating error signals to the input ($w_{c_j i}$) and output ($w_{i c_j}$) weights of the whole memory cell (Figure 5), thereby shielding them from changing when the gate is closed.

Each LSTM memory cell is typically used as a single unit would be in a standard network, receiving full weighted synaptic inputs from lower layers, and sending outputs to higher layers. Although the original

LSTM paper envisioned the possibility of multiple CEC memory cells (and CECs) per set of gates, in practice this is rarely if ever used. As such, typical LSTM models exhibit similar kinds of complex, high-dimensional, rotation-like dynamics as the RNNs investigated by Botvinick and Plaut (2004, 2006), but with the significant advantage of being naturally biased to maintain information over time (instead of having to be explicitly trained to do so), and having the ability via gating of maintaining information in a relatively protected manner over long time intervals.

Schmidhuber and colleagues later added a *forget* gate (not included in Figure 5) to deal with an important problem that arises under conditions of continuous performance in which events (timesteps) are not grouped into discrete trials. The problem they identified was that their storage cells/carousels became saturated without the intermittent clearing (resetting to 0) that generally happens programmatically between discrete trials. Adding a forget gate unit allows the network to learn to clear storage cells adaptively (Gers et al., 2000). These forget gates are standard on most current LSTM implementations, and highlight the critical point that forgetting is really as important as remembering, from a signal-to-noise perspective: it is important to remove old, irrelevant information so that new, relevant information can naturally drive processing.

3.3 *Deep Reinforcement Learning*

With the explosion of deep learning over the last decade, it has turned out that the LSTM has become a workhorse for networks having a predictive, temporal contingency component. These are often still trained by traditional supervised backpropagation, but recently many deep learning researchers have started to train these LSTM-based deep networks with a version of RL such that it is only reward signals that are backpropagated in order to train the gating units controlling the LSTM units. This triple merger of deep convolutional neural networks, LSTMs, and reinforcement learning has become known as *deep reinforcement learning* and has spawned many impressive successes just in the last few years.

For example, Deep Q-Network, a Deep RL model, learned to play a large suite of Atari games in an end-to-end fashion, using only on-screen pixels as input and points from the game serving as a reward function (Mnih et al., 2015). However, the model was fairly brittle – e.g. if you move the paddle just two pixels in breakout, it fails to adapt (Kansky et al., 2017). Also, in 2017 a team of five cooperating artificial agents (Open AI Five) trained by deep RL defeated tournament-level human teams in a modified version of the Dota 2 virtual game (`https://openai.com/five/`). And, using the same algorithms as Open AI

Five, a different team combined deep RL with supervised learning on the sensory side (a deep convolutional neural network) to train a robotic hand (Dactyl) to manipulate a block in an impressively human-like way (`https://openai.com/blog/learning-dexterity/`). Finally, in 2019 DeepMind's AlphaStar used a combination of deep RL and supervised learning in a deep convolutional neural network to win at Starcraft II.

In summary, the LSTM model strongly suggests that dynamic gating of working memory has key computational benefits, but current LSTM models retain the more implicit form of dynamic, high-dimensional temporal dynamics of non-gated RNNs, and both are likely better models of implicit, highly automated task performance. A key limitation of these automated-task level models is their relative inflexibility, which contrasts strongly with the defining features of cognitive control and executive function, which is more closely associated with working memory in the cognitive neuroscience literature. Models of this latter domain will be examined next.

4 Gating: Models of Selective Updating

The computational-level insights about the benefits of dynamic, learnable gating in the LSTM algorithm converge with considerable biological data supporting the idea that the basal ganglia (BG) provides dynamic, learnable gating for PFC working memory activity. It has long been recognized that what most distinguishes the frontal cortex from more posterior areas is the additional involvement of the BG in modulating cortical activity. For motor cortex, this is reflected in the BG's generally accepted role in the selective gating of motor actions (e.g., Mink, 1996) and there is now a modern consensus that the BG are critically and analogously involved in cognitive functioning (R. G. Brown & Marsden, 1990; Dahlin et al., 2008; Frank, 2005; Frank & O'Reilly, 2006; Graybiel, 1995; Gruber et al., 2006; Houk, 2005; Middleton & Strick, 2000; Rac-Lubashevsky & Frank, 2020; Voytek & Knight, 2010).

Specifically, it has long been suggested that the same basic gating-like mechanisms operational in motor control may have been adapted during evolution to support cognitive functioning as well (e.g., Beiser & Houk, 1998; Middleton & Strick, 2000; Wickens, Alexander, & Miller, 1991) and there is now considerable empirical evidence suggesting that specific gating decisions made by the BG via thalamus can perform a maintenance gating function (Basso & Wurtz, 2002; Cole, Bagic, Kass, & Schneider, 2010; Hikosaka & Wurtz, 1983; McNab & Klingberg, 2008; Monchi, Petrides, Strafella, Worsley, & Doyon, 2006; Nyberg

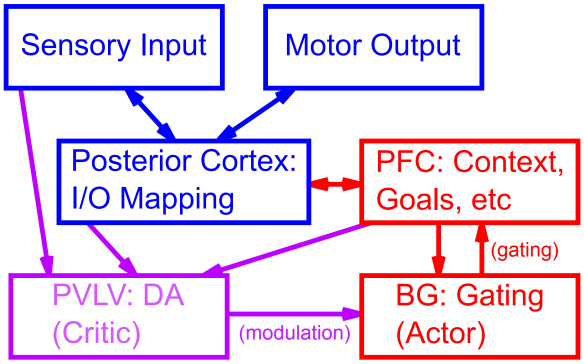

Figure 6: The basic PBWM framework illustrating the roles of the basal ganglia and PFC in working memory. Processed information from posterior areas (blue boxes) can be loaded into PFC for active maintenance under the control of gating by the BG (red boxes). Maintained information can in turn be used to bias processing in posterior areas. Learning in the BG uses phasic DA signals computed by the PVLV system (purple box; O'Reilly, Frank, Hazy et al., 2007; Mollick, Hazy, Krueger et al., 2020.

et al., 2009; Rikhye et al., 2018; Stelzel, Basten, Montag, Reuter, & Fiebach, 2010; Yehene, Meiran, & Soroker, 2008). This has led to a series of computational models based on the interaction of the PFC and BG, some of which will be reviewed here with a focus on the mechanisms each proposes with regard to working memory gating.

As noted in the introduction, these BG-based models tend to focus on longer time scales of action selection and cognitive control, with the general idea that the BG functions at a longer outer-loop time scale to help select the next course of action, and support the cognitive control and executive functions needed to organize behavior over these longer time scales. These ideas are consistent with the striking data from severe cases of Parkinsonism and other BG disorders, which result in a catatonic state with little to no voluntary, self-initiated action, as depicted in the movie *Awakenings* (starring Robert De Niro and Robin Williams). Thus, it is likely that these models describe entirely different phenomena compared to the automatic, habitual inner-loop level behavior characterized by the RNN models described above.

4.1 *The PBWM Framework*

The *PBWM* (prefrontal-cortex, basal-ganglia working memory) model was directly inspired by LSTM gating, combined with the extant BG biological data (Frank et al., 2001; Hazy et al., 2007; O'Reilly, 2006; O'Reilly & Frank, 2006) (Figure 6). PBWM assumes the basic sustained firing of PFC neurons as described above (supported by both recurrent excitatory loops and intrinsic mechanisms including NMDA channels), and shows how the BG disinhibition of PFC can drive the rapid updating of these sustained working memory representations. Specifically, as illustrated in Figure 7:

- Firing in the direct or *Go* pathway of the BG will disinhibit a select subset of one or a few of the excitatory thalamocortical loops in corresponding areas of PFC (called *stripes*), and this disinhibition should provide a sufficient jolt of extra excitation to open NMDA receptors, and trigger robust active maintenance. This notion of Go-gating for working memory updating is consistent with the characteristically sparse and episodic nature of much of BG signaling (G. E. Alexander, 1987; Kimura, Kato, & Shimazaki, 1990; Plenz & Wickens, 2010), and with the idea that BG is specifically engaged at the *initiation* of action.

- The *NoGo* pathway serves to oppose the Go pathway in the process of deciding whether to update individual stripes (Collins & Frank, 2014; Frank et al., 2001; O'Reilly, 2006; O'Reilly & Frank, 2006). In the PBWM model, if the NoGo pathway wins out in the competition between these two pathways, ongoing active maintenance continues in the associated PFC areas. This is in contrast to other possible models where the NoGo is seen as more directly inhibiting activity in the cortex (e.g., Arbib & Dominey, 1995; Ashby et al., 2005; Dominey et al., 1995; Dominey & Arbib, 1992; Mink, 1996; Schroll et al., 2012). In computational simulations, the ability of NoGo firing to protect ongoing active maintenance has proved valuable. Nevertheless, this is not a fully settled issue, and remains an important question for ongoing research. For example, D2 activity in the BG has been shown to suppress specific actions, induce NoGo learning, and affect updating and distractibility (Collins & Frank, 2014; Frank & O'Reilly, 2006; Hikida, Kimura, Wada, Funabiki, & Nakanishi, 2010; Kravitz, Tye, & Kreitzer, 2012; Yttri & Dudman, 2016; Zalocusky et al., 2016).

- Phasic dopamine signals generated by reward prediction errors serve to reinforce Go / NoGo decisions based on the relative value of reward outcomes.

- By enabling selective updating of different stripes where information can be encoded, a powerful form of role-filler variable binding (O'Reilly, 2006) and further levels of indirection Kriete, Mingus, Wyatte, Herd, and O'Reilly (2011) can be achieved, supporting systematic structure-sensitive cognitive processing (O'Reilly et al., 2014; Rougier et al., 2005).

A major focus of work in developing the PBWM model has been on how more biologically-realistic

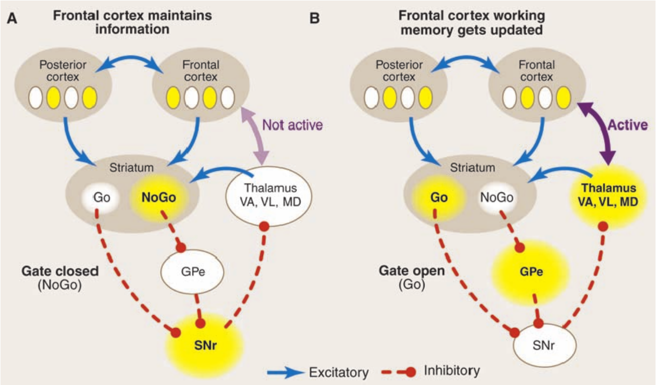

Figure 7: PBWM framework illustrating the roles of Go and NoGo pathways in the basal ganglia in the updating of working memory. A. When NoGo dominates in the BG, gating is prevented and information is maintained in PFC. B. When a Go is computed, the gate is opened and new information is loaded into PFC and then maintained.

learning mechanisms might be able to train the BG to learn to gate at appropriate points in time, to support effective cognitive function. Thus, instead of relying on the biologically-implausible BPTT algorithm as described above, PBWM uses well-established biological mechanisms of learning based on *phasic dopamine* neuromodulation. Specifically, reward-related phasic dopamine signaling provides an appropriate training signal for both the Go and NoGo pathways of the BG by virtue of the differential expression of dopamine D1 vs. D2 receptors in the two pathways, respectively (Frank, 2005; O'Reilly & Frank, 2006) (Figure 7). This directly implements Thorndike's *Law of Effect* logic: if gating leads to a better-than-expected outcome, reinforce that gating, and conversely, if gating leads to a worse-than-expected outcome, punish that gating.

A critical ongoing issue with this form of learning is the need to span potentially long temporal gaps between gating and subsequent outcomes (i.e., the *temporal credit assignment* problem). Whereas earlier versions of PBWM used a CS-like learning mechanism based on the working memory activity patterns themselves, more recent versions have explored the use of longer-lasting synaptic tags (Redondo & Morris, 2011), which can be initially activated by the gating activity but then modulated and effected by subsequent phasic dopamine signals. This produces an overall learning dynamic similar to the ACT-R version of reinforcement learning, which applies its reinforcement signal at the time of an outcome uniformly to all production firing (since the last outcome) leading up to that outcome (Stocco et al., 2010).

By incorporating a biologically based model of phasic dopamine signaling (PVLV model; Primary Value and Learned Value; Mollick et al., 2020; O'Reilly, Frank, Hazy, & Watz, 2007), PBWM has shown that many complex working memory tasks (including those with arbitrary numbers of intervening distractors) can be learned from trial-and-error experience using such a gating mechanism. These include the 1-2-AX and

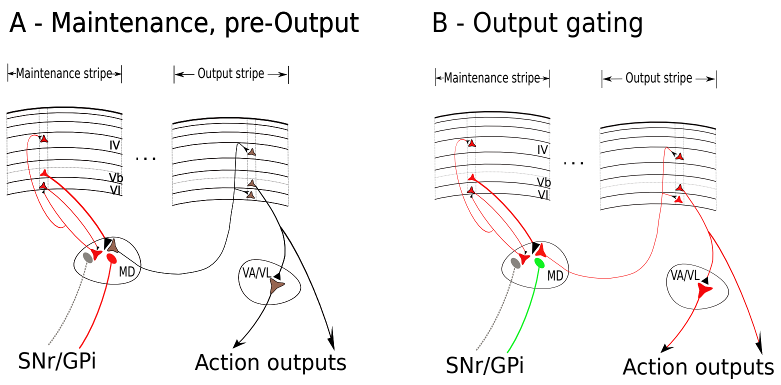

Figure 8: **Proposed division-of-labor between maintenance-specialized stripes and corresponding output-specialized stripes.** **A -** Maintenance stripe (left) in maintenance mode, with corticothalamocortical reverberant activity shown (red). Information from that stripe projects via layer Vb pyramidals to a thalamic relay cell for the corresponding output stripe (Type 2 corticothalamic projection; see text), but the BG gate is closed from tonic GPi/SNr inhibition so nothing happens yet (gray). **B -** Output gate opens due to Go signal-generated disinhibition of SNr/GPi output (green), triggering burst firing in the thalamic relay cell, which in turn activates the corresponding cortical stripe representation for the appropriate output. Projection from output stripe's layer Vb pyramidal cells then activates cortical and subcortical action/output areas, completing a handoff from maintenance to output. NOTE: Input stage of processing not relevant so left out. *Key:* MD = mediodorsal nucleus of the thalamus; VA,VL = ventral anterior, ventral lateral thalamic (motor) nuclei.

phonological loop (O'Reilly & Frank, 2006), ID/ED dynamic categorization (O'Reilly et al., 2002), WCST (Rougier & O'Reilly, 2002), N-back (e.g., Chatham et al., 2011), task switching, the Stroop task (Herd et al., 2014), hierarchical rule learning (Badre & Frank, 2012), and the reference-back-2 task (Rac-Lubashevsky & Frank, 2020).

In the original PBWM models, it was hypothesized that anatomical structures known as *stripes* (Levitt, Lewis, Yoshioka, & Lund, 1993) could be separately, selectively gateable regions, comprised of aggregates of cortical mini-columns, and correspond roughly to the *hypercolumns* described generally across a variety of different cortical areas (Mountcastle, 1997). However, it is not clear if this correspondence is strongly supported by extant data or not, as the relevant experiments have not been done. Nevertheless, there is some suggestive evidence of at least some degree of neighborhood consistency in the form of systematically ordered *iso-coding microcolumns* described by Rao, Williams, and Goldman-Rakic (1999), i.e., the equivalent of the mini-columns referred to above.

Another potential form of organization involves a distinction between neurons that fire well in advance of a later motor action (i.e., *preparatory* firing), versus those that fire at the time of the action (i.e., *output* or *action* firing). Different PFC neurons appear to be specialized according to these two different time domains, with an anatomical organization at least in the frontal eye fields (Sommer & Wurtz, 2000). More recent versions of the PBWM model have incorporated this distinction between preparatory (*maintenance*) gating, and *output* gating, which also maps well onto these distinct types of gating in the LSTM framework (Figure 8) (O'Reilly, Hazy, & Herd, 2016; O'Reilly, Munakata, Frank, Hazy, & Contributors, 2012). There are different learning and activation dynamics demands associated with these different forms of gating in the BG, which further supports the idea that they are supported by distinct sub-circuits within the overall system. Finally, there is a growing body of empirical data and theoretical analysis supporting the basic idea of a kind of maintenance vs. output organization in humans (e.g., Badre & Frank, 2012; Chatham & Badre, 2015; Chatham et al., 2014; Collins & Frank, 2013; Frank & Badre, 2012; Gayet, Paffen, & Van der Stigchel, 2013; Haith, Pakpoor, & Krakauer, 2016; Huang, Hazy, Herd, & O'Reilly, 2013; Kriete et al., 2013; van Moorselaar, Theeuwes, & Olivers, 2014).

In summary, PBWM captures the following core hypotheses in a biologically-based framework that, while significantly less computationally powerful than the full BPTT of LSTM, is nevertheless capable of learning executive function tasks that depend on sustained working memory:

- The basal ganglia gates active maintenance in the PFC, with phasic Go-pathway firing driving a rapid updating to encode new information, and opposing NoGo-pathway firing blocking this update and supporting continued maintenance (and not inhibiting it).
- This gating can be learned through phasic dopamine neuromodulation, via opposing effects of dopamine D1 and D2 receptors.
- BG gating affects many PFC neurons at once (those within the same "stripes"), and conversely there are many separable such stripes controlled by distinct BG gating signals (i.e., they are independently gatable), raising the important question as how these PFC neurons might be organized relative to their shared and distinct gating signals.
- There is evidence for separable maintenance vs. output gating, which have different learning and dynamic requirements in the PBWM model – more work could be done to investigate these issues empirically.

In the remainder of this section, various other models will be reviewed in the context of overall working memory and motor / cognitive control tasks, which have proposed different hypotheses about how the gating dynamics function. For example, in the PBWM framework BG gating works as a kind of spring-loaded gate in the sense that it serves only to initiate the maintenance process by a brief period of opening. The obvious alternative is for the BG to participate in the ongoing maintenance process by being the kind of gate that can stay open, in this case throughout the delay period. Several models have adopted versions of this idea for maintenance gating.

4.2 *Dominey-Arbib Model of Volitional Saccades*

Over a series of papers, Dominey and Arbib described a computational model of the saccade system that prominently included a working memory component for memory-guided saccades (Arbib & Dominey, 1995; Dominey et al., 1995; Dominey & Arbib, 1992). Based on then-extent electrophysiological data from primate frontal eye fields like that shown later in Figure 12, the Dominey-Arbib model included separate collections of memory-for-target and saccade-generating units (among a total of four unit types). Dominey and Arbib proposed a gating mechanism controlled by persistent suppression of BG output that acted permissively at the thalamus to sustain a corticothalamocortical loop of reverberant activity in their memory-coding cells over the delay period, a form of maintenance gating. Saccades were prevented during the delay by continued fixation and then permissively triggered by the removal of of the fixation stimulus at the end of the delay; thus, there was no distinct sense of output gating.

For a separate paradigm of visuomotor discrimination, in which subjects had to select between two simultaneously presented targets, Dominey and Arbib described a form of input gating performed by the BG that contributes to the selection between two targets (Arbib & Dominey, 1995; Dominey et al., 1995). Thus, the Dominey-Arbib model can be said to include versions of input and maintenance gating as defined here, but not output gating. The model is silent as to the cortical organization that might underlie the division-of-labor between these two kinds of processing.

4.3 *FROST Model of Ashby et al.*

An approach similar to that of Dominey and Arbib was taken by Ashby et al. (2005) in their FROST model (FROntal cortex, Striatum, and Thalamus). With regard to maintenance gating, an interesting and seemingly unique aspect of the FROST model is that it explicitly excludes a role for the BG in the *initiation* of maintenance gating, only its persistence. Citing data from Hikosaka, Sakamoto, and Usui (1989) showing sustained firing in striatal cells that starts only after the offset of the to-be-remembered stimulus, the authors

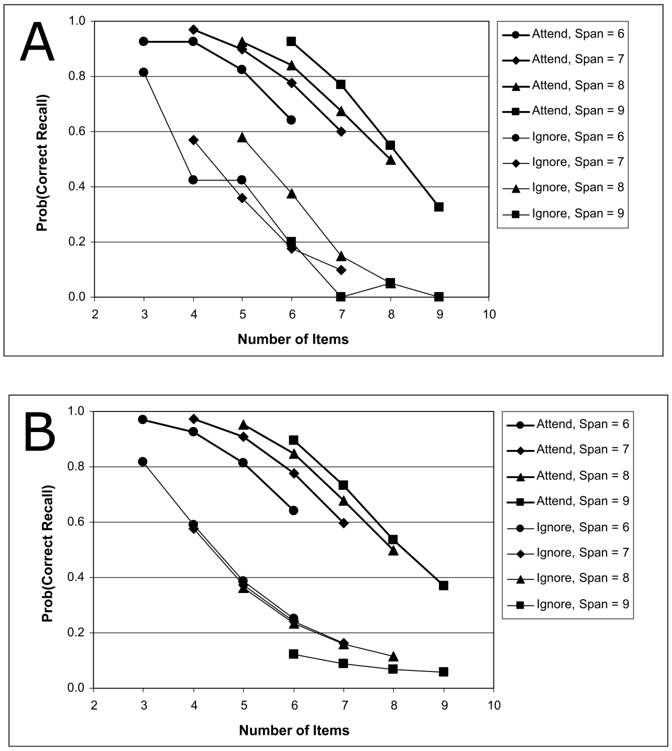

Figure 9: Results from the FROST model showing it captures the effects of attention and individual differences in working memory capacity as reported by Cowan et al., 1999. A - Empirical results. B - Model results. From Ashby, Ell, Valentin et al., 2005.

propose that the role of the BG is to allow maintenance activity already started in the cortex to recruit a loop of corticothalamocortical reverberant activity by activating striatal cells and thus disinhibiting the thalamus. No other kind of gating is mentioned including output gating.

Another distinguishing feature of the FROST model is that Ashby et al. (2005) explicitly attribute a role for *selective attention* in the cortical initiation of active maintenance and are able to account for attentional effects as well as individual differences in the pattern of measured working memory capacity reported by Cowan, Nugent, Elliott, Ponomarev, and Saults (1999). Figure 9 shows empirical results at the top (A) and FROST model results below (B) with the higher group of curves in each graph reflecting attentional effects and each individual curve a subject with differing measured working memory spans.

4.4 *Schroll et al. Model of BG*

Informed by considerable neurobiological detail regarding the BG, Schroll et al. (2012) developed a comprehensive model of BG function (Figure 10). Like the previous models in this section, their model implements

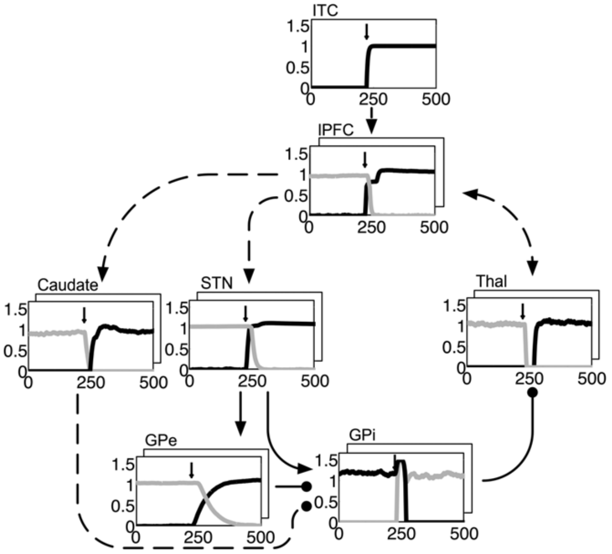

Figure 10: The model of Schroll, Vitay & Hamker, 2012. See main text for explanation. From Schroll, Vitay & Hamker, 2012, Figure 5.

maintenance gating as persistent activity in the striatum that permits continued reverberation of the corticothalamocortical loop. The subthalamic nucleus (STN) in their model exerts a strong excitatory tone on the output nuclei of the BG (GPi and SNr) and also itself receives widespread excitatory inputs from much of frontal cortex. The onset of a new relevant stimulus transiently increases STN, and therefore GPi and SNr, activity in a relatively global manner, thus transiently suppressing the thalamus and breaking the positive feedback loop of reverberatory corticothalamocortical activity, effectively clearing the current contents of working memory. This allows an updated memorandum to be stored. Because the input from STN to GPi and SNr is known to be relatively global, it is not clear, however, how this mechanism might be able to discriminate between to-be-stored items versus distracters. Similarly, it is not clear how such a mechanism might be able to selectively update only one out perhaps three or four currently maintained items.

4.5 *Beiser-Houk Model of Sequence Learning*

Two influential models have embraced something like a hybrid of the punctate and sustained versions of maintenance gating and may suggest some ways in which the two approaches might be synthesized. The sequence-production model of Beiser and Houk (1998) exploits unique biophysical characteristics of thalamic relay cells, which exhibit burst firing in response to BG-mediated disinhibition, which in turn activates the corticothalamocortical reverberatory activity. Although striatal activity was only transient so as to initiate

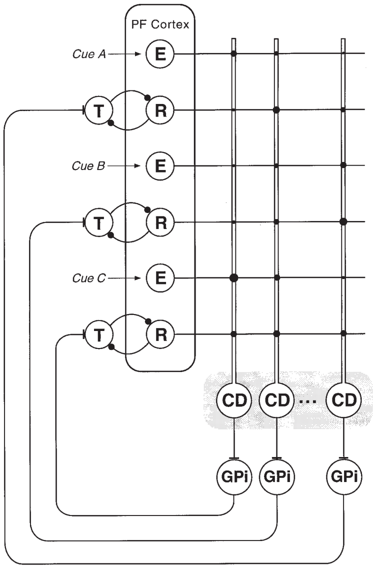

Figure 11: The model of Beiser & Houk, 1998. Three cortical-basal ganglionic loops are shown corresponding to three items (A,B,C). Active maintenance is a result of reverberatory activity in the corticothalamocortical recurrent loop (T—R), triggered by disinhibition at the thalamic relay cell (T) from a corresponding GPi cell. From Beiser & Houk, 1998, Figure 2.

maintenance-gating in their simulations, they also described instances in which sustained firing throughout the delay also followed the initial maintenance-triggering activity. Although not directly relevant for their model, this could provide a bridge to the sustained activity models described above. In addition, this model was able to reproduce a significant number of sequences, based purely on random initial connectivity without any learning, suggesting that these burst-firing dynamics may provide a useful general-purpose sequencing mechanism.

4.6 *Gruber et al. Model of Dopamine-Modulated Gating*

The model of Gruber et al. (2006) also primarily relies on a trigger-like form of maintenance gating by the BG, but also had follow-on permissive role over the full maintenance period. In this model, phasic dopamine affects the bi-stability characteristics of cells in both cortex and striatum, triggering an upstate among MSNs which in turn triggers a variably stable attractor state in the cortex. A small amount of persistent striatal activity could stabilize cortical representations even in a continuous space by holding open the gate at the thalamus for the initialized spatial location, thereby preventing noise-induced drift that is otherwise problematic for continuous line-attractor models.

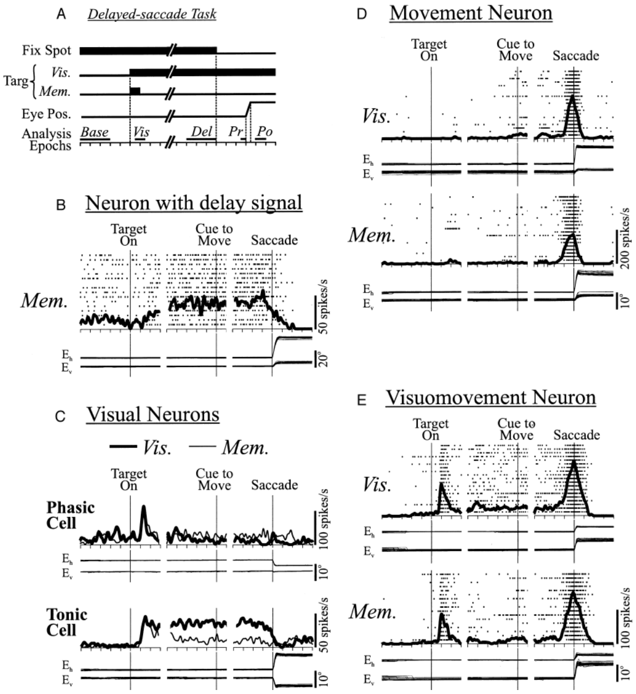

Figure 12: **Layer 5 projecting cells of FEF showing heterogeneous firing patterns suggesting different roles for input vs. output processing.** Histograms and activity rate curves for individual cells recorded from the frontal eye fields (FEF) during a visually-guided and memory-guided saccades **A** Schematic for both tasks. **B** Delay period cell. Histogram (background dots) and curve of activity rate for an individual cell recorded in the frontal eye fields (FEF) during a delayed saccade task. The target stimulus is only on briefly at the beginning of the trial. This cell maintained its activity during the delay so as to enable other cells to generate a correct saccade at the end of the trial. **C** Visual only cells **D** Movement only cells. **E** Visuomovement cells showingactivity during both the visual and movement time epochs. Adapted from Sommer & Wurtz, 2000, Figure 2. Permission pending.

4.7 *Brown, Bullock & Grossberg TELOS Model*

Informed by the same kind of monkey electrophysiological data as had guided Dominey and Arbib's work (e.g., see Figure 12), J. W. Brown et al. (2004) developed a detailed model (TELOS) to account for the results from many different saccade paradigms, in particular addressing the tension between voluntary (top-down generated) and involuntary (bottom-up) saccades. Most relevant to the issue of working memory and maintenance gating are two aspects of the authors' treatment of the memory-guided saccade case:

- J. W. Brown et al. (2004) explicitly mapped the categories of FEF cells exhibiting differential patterns of responses to the cortical laminae of the FEF: input-responsive to middle cortical laminae (roughly layer 4); memory-coding to superficial laminae (2, 3, 5a); and saccade-generating to layer 5b specifically. Thus, their story about delayed responding was that the superficial layers maintained a memory of the target location over the delay, while the large subcortically-projecting pyramidal cells of layer

5b were activated at the appropriate time to generate the saccade (see Figure 1 for a diagram).

- In terms of BG-mediated gating, TELOS seems to have been the first neurobiologically informed model to describe a form of output gating in which the BG served to open a gate that allowed the layer 5b cells to get active at the appropriate time to generate the saccade. Both input processing and the initiation and maintenance of sustained firing during the delay were treated as more-or-less automatic processes without involvement of the BG.

Thus, while the models discussed earlier in this section have included a role for the BG in some form for maintenance gating (Dominey-Arbib also included input gating), TELOS included only a role for BG in output gating.

With regard to the authors' mapping of maintenance to the superficial cortical laminae and output to the deep 5b cells, an apparent problem with this account is that the data from Sommer and Wurtz (2000) (Figure 12) unequivocally demonstrated that all varieties of activity signals, including memory-cell signals, are transmitted to the superior colliculus during the delay and these signals can *only* be coming from subcortically-projecting layer 5b pyramidal cells. Given that a TELOS-like laminar specialization is consistent with a considerable other data as discussed in the *Background*, it will be important to reconcile these two seemingly contradictory data sets with it being likely that some combination of both interlaminar and intercolumnar divisions-of-labor are both involved.

It is now well-established that layer 5b pyramidals are not homogeneous and can be subdivided into multiple subtypes according to both morphology (Fries, 1984; Leichnetz, Spencer, Hardy, & Astruc, 1981) and, critically, differing subcortical targets (Economo et al., 2018; Harris & Shepherd, 2015; Hattox & Nelson, 2007; Ramaswamy & Markram, 2015; Winnubst et al., 2019). Thus, the functional effects of output gating depend on *which* of the 5b subtypes are getting gated and their corresponding subcortical targets. Although not directly addressed as such in this context by J. W. Brown et al. (2004), their model does adopt a functional distinction between 5a and 5b subtypes (which are also morphologically distinct), both of which are likely to project to the superior colliculus, but only 5b is hypothesized to be output-gated by the BG. Thus, a straightforward reconciliation is to suggest that the 5a neurons convey input and maintenance signals from other lamina in an ungated fashion, while the 5b are output-gated by the BG, to drive overt responses such as saccades.

This account is consistent with several details from Sommer and Wurtz (2000) and earlier anatomical data (Fries, 1984; Leichnetz et al., 1981), suggesting a diversity of morphologies within layer 5 cells that project to the colliculus, and that the movement cells specifically identified by Sommer and Wurtz (2000) were indeed the largest and fastest-conducting cells, consistent with the 5b profile. Furthermore, although Sommer and Wurtz (2000) identified a topographic bias in the locations of motor output vs other cell types at the most extreme lateral edge of the FEF, there was substantial intermingling of these cell types throughout most of the extent of the FEF, consistent with the laminar specialization model, and not a stronger topographic segregation of cells across different regions of FEF.

4.8 *Embracing Diversity*

In summary, a diverse range of different ideas have been explored across many different neurobiologically-oriented models developed by several different research groups, but at least there is a general consensus around the idea that frontal cortex is critical for active maintenance of working memory states over time, and that the basal ganglia likely plays some kind of role in driving a gating-like modulation of these frontal activity states. As discussed earlier, there is evidence that multiple different thalamic circuits may modulate the PFC, with potentially different characteristic patterns of connectivity and targets, in addition to differential patterns of connectivity with the BG. There are a growing number of empirical studies using advanced neuroscience techniques to determine the properties and functions of these circuits, the results of which should directly inform the further development of computational models. Thus, the field may be poised for a new wave of “second generation” models that incorporate this new data, and may end up adopting different subsets of the overall mechanisms across the existing set of models reviewed above.

5 General Discussion

This chapter reviewed some of the seminal computational models of working memory in the context of higher cognitive function overall. In particular, the development of LSTMs was used to motivate the computational requirements for maintenance and output gating. The authors’ own gating-focused PBWM framework was also highlighted and compared with several other models through the lens of basal ganglia-mediated gating. Below are summarized some of the tentative conclusions that might be drawn with regard to the motivating questions presented in the *Introduction*.

5.1 *Representational Scale of Independently-Gatable Units*

All of the neurobiologically-motivated models reviewed in this chapter employ, at least implicitly, some version of separate channels for separate items, although the PBWM framework is perhaps the most explicit by mapping these channels onto the biological feature of "stripes". Interestingly, the adoption of the LSTM framework by the AI community has evolved in such a way that gating functions at the individual unit level, which is at the extreme fine-grained end of the granularity scale. It would nevertheless be interesting to more systematically explore this gating granularity dimension in these models, because it likely has not yet been explored, and the biological constraints strongly suggest that, at least for BG-mediated gating, there are many PFC neurons per gating signal.

The relevant biological data is as follows. Originally, G. Alexander et al. (1986) described five largely independent, closed loops connecting specific regions of frontal cortex with themselves and running through the BG. Since then, numerous studies have established that the connectivity between the cortex and the BG has both closed loop and open loop qualities (e.g., Haber, 2003; Haber & Knutson, 2010; Joel & Weiner, 2000), and that the closed loop aspect can be observed at a much more fine-grained level than the original five loops (Ferry, Öngür, An, & Price, 2000; Flaherty & Graybiel, 1993a, 1993b; Graybiel, Flaherty, & Gimenez-Amaya, 1991; Haber, 2003), including in humans (Choi, Yeo, & Buckner, 2012; Jung et al., 2014; Pauli, O'Reilly, Yarkoni, & Wager, 2016). This raises the critical question of just how fine-grained this closed loop connectivity might be, because that could serve as a kind of lower bound on the neuroanatomical and representational scope of individually BG-gateable units in terms of working memory updating.

The strongest constraint comes from the fact that there are many fewer neurons in the output pathway of the BG, the GPi / SNr, than in the corresponding areas of frontal cortex that are affected by BG gating signals. A reasonable, perhaps conservative, estimate is that roughly 5 billion (35%) of the 14 billion pyramidal cells in the human brain reside in the frontal cortex (Pakkenberg & Gundersen, 1997). Meanwhile, a reasonable, possibly generous, estimate for the total number of cells in the output nuclei (GPi and SNr) of the BG is approximately 740,000 in humans (GPi: 352,000; SNr (non-dopamine): 288,000) (Hardman et al., 2002). Thus, there are approximately 6,750 frontal pyramidal cells downstream for each BG output cell. Furthermore, because each isocoding minicolumn has 70 or so pyramidal cells, this implies that there are on the order of 100 cortical minicolumns downstream for each BG output cell, a ratio that is likely to be a lower bound. Based on this back-of-the-envelope calculation, as well as the known thalamocortical connectivity patterns, it seems clear that the gating of individual pyramidal cells, or even individual minicolumns, is virtually impossible.

5.2 *Working Memory Capacity Limitations*

Another possible source of constraints on the scope of working memory gating and overall representational organization comes from studies attempting to determine the origin and nature of capacity limitations in working memory. George Miller (1956) famously showed that working memory appears to be limited to holding only 7 plus-or-minus 2 items at a time. Does that magic number somehow reveal how many independently-gatable working memory states there are? If so, it would suggest a much coarser-grained form of gating than the most fine-grained end of the spectrum possible according to the GPi / SNr bottleneck, which is certainly a possibility: many individual GPi / SNr neurons could work together to drive gating for larger swaths of PFC. However, further research suggests that this capacity constraint can apply separately to many different representational domains (verbal vs. visual vs. numerical vs. spatial etc) and is actually more like 4 items than 7 (Cowan, 2001, 2011; Luck & Vogel, 1997, 2013; Zhang & Luck, 2008) as, for example, when digit span is tested with unpredictable reporting points, where rehearsal and chunking strategies are less able to contribute to performance (Cowan, 2001). More recently, it has been recognized that differences in measured memory span may also be complicated by a variable contribution of rapid learning effects (Cowan, 2019).

It is difficult to know how many such representational domains there are, but for example if there were 70 GPi / SNr neurons per gating unit, and 4 gating units per domain, that would amount to a total of approximately 2,640 different such domains, which might be a reasonable number considering the entire scope of information coded by the frontal cortex. Again, these are just rough order-of-magnitude calculations, and it is unlikely that the brain would be crisply organized in this way (i.e., there is likely to be partial overlap and different subsets activated in different situations, etc).

In contrast to this more "slot-based" analysis, a body of research has found that the precision of memory varies as a function of the memory load and similarity between visual stimuli (Bays, Catalao, & Husain, 2009; Bays & Husain, 2008; Ma et al., 2014; Wilken & Ma, 2004), and that increased precision for one item comes at the expense of other co-maintained representations (Gorgoraptis, Catalao, Bays, & Husain, 2011; Pertzov, Bays, Joseph, & Husain, 2013). Thus, this view holds that, instead of a fixed number of slots, working memory capacity might be better conceived as a single shared resource that can be flexibly allocated between multiple items (e.g., Ma et al., 2014).

The attractor model, augmented with lateral inhibitory connections, can potentially reconcile this slots vs. resources debate (e.g., Fukuda, Vogel, Mayr, & Awh, 2010; Nassar et al., 2018; Wei, Wang, & Wang, 2012). Wei et al. (2012) showed how the representation of multiple items in a shared neural population exhibits characteristics of both continuous resource sharing and discretized items in that only a limited number of "bump attractors" can co-exist in a single population without colliding (merging), and that the strength and fidelity of each bump representation is diminished the more items there are that are retained. Nassar et al. (2018) showed that by adding a center-surround pattern of lateral excitation-inhibition to the Wei et al. (2012) network they could further account for additional aspects of the precision vs. recall tradeoff by positing a chunking-like mechanism that serves to combine features of similar value across items (e.g., treating various shades of red as a single feature value) and that the benefits of such a representational strategy seemed to asymptote at a partitioning of the feature space of about four categories.

It would seem at least theoretically possible that discrete gating slots might make different predictions from these attractor models, and that some particular combination of these two models might provide a more comprehensive account – this would be a good target for future research.

5.3 *Variable Binding and Transfer*

The combination of reinforcement learning with selectively updatable, actively maintained working memory representations enables a form of role-filler style variable binding that supports flexible working memory function. Information can be encoded into different functionally-defined "slots" of working memory, and then retrieved according to the relevant functional category, independent (at least to some extent) of the detailed content (O'Reilly, 2006). In addition, the combination can be exploited to produce a sort of inductive bias to use those representations in a way that can be co-opted under new task conditions, a form of out-of-distribution generalization or learning transfer. Examples of this kind of learning transfer are Bhandari and Badre (2018); Collins and Frank (2013); Frank and Badre (2012); Kriete et al. (2013); Rougier et al. (2005); A. Williams and Phillips (2020).

The Stocco et al. (2010) model of the BG, based on the ACT-R architecture, provides a particularly powerful form of flexible BG gating that supports the arbitrary *routing* of information from one part of the brain to another, like a system bus in a standard computer architecture. However, an important constraint on such a model is the very small size of the GPi / SNr bottleneck through which all BG output flows — it is not clear if there is sufficient capacity there to directly route much detailed content through the BG itself. Instead, it may make more sense to think of the BG as selecting the relevant brain areas through indirect effects of gating on the frontal cortex, which in turn can provide top-down attentional gain modulation on the relevant brain areas, and then the information is routed through much higher capacity corticocortical pathways between these areas. Nevertheless, the principle that the BG may be important for flexible, controlled processing is much more consistent with a wide range of data compared to the older notion that it is the locus of habitual responding (O'Reilly, Nair, et al., 2020).

5.4 *Nature and Kinds of Gating*

Across many neurobiologically-oriented models developed by several different research groups there has emerged a remarkable consensus that the BG plays *some* kind of role in gating activity in the PFC, even while there is considerable diversity in ideas for exactly what this role is, among the set of functionally-defined types of gating supported by the abstract LSTM model (Gers et al., 2000; Hochreiter & Schmidhuber, 1997). Some argue that it is important for maintenance gating of new information into PFC, while others argue for a more specific role in output-gating of information out of working memory, while yet others advocate both roles. As discussed above, a wide range of neuroscience data can be brought to bear on addressing this question, and while definitive answers are not yet available, there is some indication that the BG is likely to be more specifically involved in output-gating, via matrix-type thalamic projections, versus maintenance gating, which is supported by core-type thalamic pathways. Hopefully, the considerable empirical work going on in this area will soon provide more definitive answers to these important questions.

Another ongoing question concerns the degree to which the BG gating signal functions in a more punctate way to initiate a corresponding effect in the PFC, versus participating in a more sustaining regulation of cortical activity throughout the delay period. There seems to be strong empirical evidence for both punctate and sustained maintenance signals in the striatum and BG output nuclei. At this point it seems the most likely case is that there are multiple BG-mediated contributions, including a punctate initiating event, an ongoing permissive component that supports ongoing corticothalamocortical reverberatory activity, and possibly even a punctate terminating or clearing event in some cases.

5.5 *Static Vs. Dynamic Working Memory Representations*

There seems to be compelling evidence for both boxcar-like sustained activity as well as various waxing-and-waning patterns of activity during working memory delay periods. Thus, it is hard to avoid the conclusion that both patterns of activity must contribute to working memory-like processing. Assuming this to be the case, an important challenge for future work will be to better characterize the circumstances under which different activity patterns tend to predominate in order to better understand the contributions of each. One obvious contribution to the apparently conflicting stories is that the sustained activity story is generally older and comes from single-cell recording data, while the dynamic, variable activity story is generally based on much more recent data and comes from population-based recording data. Thus, at least some of the difference in the two stories is likely a matter of methodologies and researcher emphasis.

One intriguing possibility is that the sustained activity may be more prevalent during the early stages of learning any particular task when controlled processing is thought to be most necessary, while the less metabolically costly, dynamic trajectory pattern may become increasingly prevalent as learning proceeds and performance transitions to a more automatic mode of processing, perhaps approaching something like that captured by RNN models such as described by Botvinick and Plaut (2006).

6 Conclusion

The last several decades have seen a great deal of progress in understanding the neurobiological mechanisms underlying working memory and there is now extensive evidence in support of the basic idea that the PFC and BG function as an integrated system with the BG is performing something like a gating function for controlling cognition as well as motor action, including determining when working memory is updated in PFC. In particular, the BG seems to participate in initiating and/or maintaining a robust form of persistent activity in the PFC as well as in controlling downstream access to working memory contents via the similar process of output gating. Nonetheless, much of the story remains to be worked out including many of the specific details involved and how the transition from controlled to automatic processing may evolve over repeated experience through continuous learning.

Acknowledgements

Supported by ONR grants N00014-07-1-0651 and N00014-03-1-0428, and NIH grants MH069597 DEFD ICArUS

References

Adams, E. J., Nguyen, A. T., & Cowan, N. (2018, July). Theories of working memory: Differences in definition, degree of modularity, role of attention, and purpose. *Language, Speech, and Hearing Services in Schools*, *49*(3), 340–355. doi: 10.1044/2018_LSHSS-17-0114

Alexander, G., DeLong, M., & Strick, P. (1986, May). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, *9*, 357–381.

Alexander, G. E. (1987, November). Selective neuronal discharge in monkey putamen reflects intended direction of planned limb movements. *Experimental Brain Research*, *67*, 623–634.

Anderson, J. R., & Lebiere, C. (1998). *The Atomic Components of Thought.* (First ed.). Mahwah, NJ: Lawrence Erlbaum Associated, Publishers.

Arbib, M. A., & Dominey, P. F. (1995, December). Modeling the roles of basal ganglia in timing and sequencing saccadic eye movements. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of Information Processing in the Basal Ganglia* (pp. 149–162). Cambridge, MA: MIT Press.

Arnsten, A. F. T., Wang, M. J., & Paspalas, C. D. (2012, October). Neuromodulation of thought: Flexibilities and vulnerabilities in prefrontal cortical network synapses. *Neuron*, *76*(1), 223–239. doi: 10.1016/j.neuron.2012.08.038

Ashby, F. G., Ell, S. W., Valentin, V. V., & Casale, M. B. (2005, November). FROST: A distributed neurocomputational model of working memory maintenance. *Journal of Cognitive Neuroscience*, *17*(11), 1728–1743. doi: 10.1162/089892905774589271

Baddeley, A. D. (1986). *Working Memory*. New York: Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974, December). Working memory. In G. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. VIII, pp. 47–89). New York: Academic Press.

Badre, D., & Frank, M. J. (2012, March). Mechanisms of hierarchical reinforcement learning in cortico-striatal circuits 2: Evidence from FMRI. *Cerebral cortex*, *22*(3).

Barak, O., & Tsodyks, M. (2014, April). Working models of working memory. *Current Opinion in Neurobiology*, *25*, 20–24. doi: 10.1016/j.conb.2013.10.008

Basso, M. A. ., & Wurtz, R. H. (2002, March). Neuronal activity in substantia nigra pars reticulata during target selection. *Journal of Neuroscience*, *22*(5), 1883–1894.

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009, September). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10), 7–7. doi: 10.1167/9.10.7

Bays, P. M., & Husain, M. (2008, August). Dynamic shifts of limited working memory resources in human vision. *Science (New York, N.Y.)*, *321*(5890), 851–854. doi: 10.1126/science.1158023

Beiser, D. G., & Houk, J. C. (1998, July). Model of cortical-basal ganglionic processing: Encoding the serial order of sensory events. *Journal of Neurophysiology*, *79*, 3168–3188.

Bhandari, A., & Badre, D. (2018, March). Learning and transfer of working memory gating policies. *Cognition*, *172*, 89–100. doi: 10.1016/j.cognition.2017.12.001

Bogacz, R. (2013). Basal Ganglia: Beta Oscillations. In D. Jaeger & R. Jung (Eds.), *Encyclopedia of Computational Neuroscience* (pp. 1–5). New York, NY: Springer. doi: 10.1007/978-1-4614-7320-6_82-1

Botvinick, M. M., & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological review*, *111*(2), 395.

Botvinick, M. M., & Plaut, D. C. (2006, April). Short-term memory for serial order: A recurrent neural network model. *Psychological Review*, *113*, 201–233.

Braver, T. S., & Cohen, J. D. (2000, December). On the Control of Control: The Role of Dopamine in Regulating Prefrontal Function and Working Memory. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 713–737). Cambridge, MA: MIT Press.

Braver, T. S., Paxton, J. L., Locke, H. S., & Barch, D. M. (2009, May). Flexible neural mechanisms of cognitive control within human prefrontal cortex. *Proceedings of the National Academy of Sciences USA*, *106*(18), 7351–7356.

Brown, J. W., Bullock, D., & Grossberg, S. (2004, April). How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Networks*, *17*, 471–510.

Brown, R. G., & Marsden, C. D. (1990, February). Cognitive function in Parkinson's disease: From description to theory. *Trends in Neurosciences*, *13*, 21–29.

Brown, V. J., & Bowman, E. M. (2002, December). Rodent models of prefrontal cortical function. *Trends* *in Neurosciences*, *25*, 340–343.

Burgess, N., & Hitch, G. (2005, November). Computational models of working memory: Putting long-term memory into context. *Trends in Cognitive Sciences*, *9*(11), 535–541. doi: 10.1016/j.tics.2005.09.011

Chatham, C. H., & Badre, D. (2015, February). Multiple gates on working memory. *Current Opinion in Behavioral Sciences*, *1*, 23–31. doi: 10.1016/j.cobeha.2014.08.001

Chatham, C. H., Frank, M., & Badre, D. (2014, January). Corticostriatal output gating during selection from working memory. *Neuron*, *81*(4), 930–942.

Chatham, C. H., Herd, S. A., Brant, A. M., Hazy, T. E., Miyake, A., O'Reilly, R. C., & Friedman, N. P. (2011, November). From an executive network to executive control: A computational model of the n-back task. *Journal of Cognitive Neuroscience*, *23*, 3598–3619.

Choi, E. Y., Yeo, B. T. T., & Buckner, R. L. (2012, July). The organization of the human striatum estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *108*(8), 2242–2263. doi: 10.1152/jn.00270.2012

Clascá, F., Rubio-Garrido, P., & Jabaudon, D. (2012). Unveiling the diversity of thalamocortical neuron subtypes. *European Journal of Neuroscience*, *35*(10), 1524–1532. doi: 10.1111/j.1460-9568.2012.08033.x

Cleeremans, A., Servan-Schreiber, D., & McClelland, J. L. (1989, January). Finite State Automata and Simple Recurrent Networks. *Neural Computation*, *1*(3), 372–381.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990, January). On the Control of Automatic Processes: A Parallel Distributed Processing Model of the Stroop Effect. *Psychological Review*, *97*(3), 332–361.

Cole, M. W., Bagic, A., Kass, R., & Schneider, W. (2010, October). Prefrontal dynamics underlying rapid instructed task learning reverse with practice. *The Journal of Neuroscience*, *30*(42), 14245–14254.

Collins, A. G. E., & Frank, M. J. (2013, January). Cognitive control over learning: Creating, clustering, and generalizing task-set structure. *Psychological Review*, *120*(1), 190–229.

Collins, A. G. E., & Frank, M. J. (2014, July). Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychological Review*, *121*(3), 337–366.

Collins, A. G. E., & Frank, M. J. (2016, January). Surprise! Dopamine signals mix action, value and error. *Nature Neuroscience*, *19*(1), 3–5. doi: 10.1038/nn.4207

Courtemanche, R., Fujii, N., & Graybiel, A. M. (2003, December). Synchronous, focally modulated betaband oscillations characterize local field potential activity in the striatum of awake behaving monkeys. *Journal of Neuroscience*, *23*(37), 11741–11752.

Cowan, N. (2001, August). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185.

Cowan, N. (2011, May). The focus of attention as observed in visual working memory tasks: Making sense of competing claims. *Neuropsychologia*, *49*(6), 1401–1406. doi: 10.1016/j.neuropsychologia.2011.01.035

Cowan, N. (2017, August). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*(4), 1158–1170. doi: 10.3758/s13423-016-1191-6

Cowan, N. (2019, August). Short-term memory based on activated long-term memory: A review in response to Norris (2017). *Psychological Bulletin*, *145*(8), 822–847. doi: 10.1037/bul0000199

Cowan, N., Nugent, L. D., Elliott, E. M., Ponomarev, I., & Saults, J. S. (1999, November). The role of attention in the development of short-term memory: Age differences in the verbal span of apprehension. *Child development*, *70*(5), 1082–1097.

Dahlin, E., Neely, A. S., Larsson, A., Backman, L., & Nyberg, L. (2008, June). Transfer of learning after updating training mediated by the striatum. *Science*, *320*(5882), 1510–1512.

Dayan, P. (2007). Bilinearity, rules, and prefrontal cortex. *Frontiers in Computational Neuroscience*, *1*(1), 1–14.

Dayan, P. (2008, December). Simple substrates for complex cognition. *Frontiers in Neuroscience*, *2*(2), 255.

Dominey, P. F., Arbib, M., & Joseph, J.-P. (1995, July). A model of corticostriatal plasticity for learning oculomotor associations and sequences. *Journal of Cognitive Neuroscience*, *7*(3), 311–336. doi: 10.1162/jocn.1995.7.3.311

Dominey, P. F., & Arbib, M. A. (1992, January). Cortico-subcortical model for generation of spatially accurate sequential saccades. *Cerebral Cortex*, *2*, 153–175.

Dunbar, K., & MacLeod, C. M. (1984, November). A horse race of a different color: Stroop interference patterns with transformed words. *Journal of Experimental Psychology. Human Perception and Performance*, *10*, 622–639.

Durstewitz, D., Seamans, J. K., & Sejnowski, T. J. (2000, January). Neurocomputational models of working memory. *Nature Neuroscience*, *3 supp*, 1184–1191.

Economo, M. N., Viswanathan, S., Tasic, B., Bas, E., Winnubst, J., Menon, V., … Svoboda, K. (2018, November). Distinct descending motor cortex pathways and their roles in movement. *Nature*, *563*(7729), 79–84. doi: 10.1038/s41586-018-0642-9

Elman, J. L. (1990, January). Finding structure in time. *Cognitive Science*, *14*(2), 179–211.

Elston, G. N. (2003). Cortex, cognition and the cell: New insights into the pyramidal neuron and prefrontal function. *Cerebral Cortex*, *13*(11), 1124–1138.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999, November). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology. General*, *128*, 309–331.

Ferry, A. T., Öngür, D., An, X., & Price, J. L. (2000). Prefrontal cortical projections to the striatum in macaque monkeys: Evidence for an organization related to prefrontal networks. *Journal of Comparative Neurology*, *425*(3), 447–470.

Flaherty, A. W., & Graybiel, A. M. (1993a, August). Output architecture of the primate putamen. *Journal of Neuroscience*, *13*(8), 3222–3237.

Flaherty, A. W., & Graybiel, A. M. (1993b, March). Two input systems for body representations in the primate striatal matrix: Experimental evidence in the squirrel monkey. *Journal of Neuroscience*, *13*(3), 1120–1137.

Frank, M. J. (2005, January). When and when not to use your subthalamic nucleus: Lessons from a computational model of the basal ganglia. In A. K. Seth, T. J. Prescott, & J. J. Bryson (Eds.), *Modelling Natural Action Selection: Proceedings of an International Workshop* (pp. 53–60). Sussex: AISB.

Frank, M. J., & Badre, D. (2012, March). Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: Computational analysis. *Cerebral Cortex*, *22*(3), 509–526.

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001, January). Interactions between the frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, and Behavioral Neuroscience*, *1*, 137–160.

Frank, M. J., & O'Reilly, R. C. (2006, June). A mechanistic account of striatal dopamine function in human cognition: Psychopharmacological studies with cabergoline and haloperidol. *Behavioral Neuroscience*, *120*, 497–517.

Friedman, N., Miyake, A., Corley, R., Young, S., Defries, J., & Hewitt, J. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172–179.

Fries, W. (1984). Cortical projections to the superior colliculus in the macaque monkey: A retrograde study using horseradish peroxidase. *Journal of Comparative Neurology*, *230*(1), 55–76. doi: 10.1002/cne.902300106

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010, October). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, *17*(5), 673–679. doi: 10.3758/17.5.673

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989, April). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, *61*(2), 331–349.

Fusi, S., Miller, E. K., & Rigotti, M. (2016, April). Why neurons mix: High dimensionality for higher cognition. *Current Opinion in Neurobiology*, *37*, 66–74. doi: 10.1016/j.conb.2016.01.010

Fuster, J. M., & Alexander, G. E. (1971, January). Neuron activity related to short-term memory. *Science*, *173*, 652–654.

Gayet, S., Paffen, C. L. E., & Van der Stigchel, S. (2013, December). Information matching the content of visual working memory is prioritized for conscious access. *Psychological Science*, *24*(12), 2472–2480. doi: 10.1177/0956797613495882

Gerfen, C. R., & Surmeier, D. J. (2011). Modulation of striatal projection systems by dopamine. *Annual Review of Neuroscience*, *34*, 441–466.

Gers, F. A., Schmidhuber, J., & Cummins, F. (2000, November). Learning to forget: Continual prediction with LSTM. *Neural Computation*, *12*, 2451–2471.

Giguere, M., & Goldman-Rakic, P. S. (1988). Mediodorsal nucleus: Areal, laminar, and tangential distribution of afferents and efferents in the frontal lobe of rhesus monkeys. *Journal of Comparative Neurology*, *277*(2), 195–213. doi: 10.1002/cne.902770204

Goldman-Rakic, P. S. (1995, March). Cellular basis of working memory. *Neuron*, *14*(3), 477–485.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. Cambridge, Massachusetts: The MIT Press.

Gorgoraptis, N., Catalao, R. F. G., Bays, P. M., & Husain, M. (2011, June). Dynamic updating of working memory resources for visual objects. *Journal of Neuroscience*, *31*(23), 8502–8511. doi: 10.1523/JNEUROSCI.0208-11.2011

Graybiel, A. M. (1995). Building action repertoires: Memory and learning functions of the basal ganglia. *Current Opinion in Neurobiology*, *5*(6), 733–741.

Graybiel, A. M., Flaherty, A. W., & Gimenez-Amaya, J. M. (1991). Striosomes and matrisomes. In G. Bernardi, M. B. Carpenter, G. Di Chiara, M. Morelli, & P. Stanzione (Eds.), *The Basal Ganglia III: Proceedings of the Third Triennial Meeting of the International Basal Ganglia Society, held June 10-13, 1989, in Cagliari, Italy* (pp. 3–12). Plenum Press.

Gruber, A. J., Dayan, P., Gutkin, B. S., & Solla, S. A. (2006, May). Dopamine modulation in the basal ganglia locks the gate to working memory. *Journal of Computational Neuroscience*, *20*(2), 153–166.

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

Haber, S. N. (2003, January). The primate basal ganglia: Parallel and integrative networks. *Journal of Chemical Neuroanatomy*, *26*(4), 317–330.

Haber, S. N., & Knutson, B. (2010, January). The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology*, *35*.

Haith, A. M., Pakpoor, J., & Krakauer, J. W. (2016, March). Independence of movement preparation and movement initiation. *Journal of Neuroscience*, *36*(10), 3007–3015. doi: 10.1523/JNEUROSCI.3245-15.2016

Hardman, C. D., Henderson, J. M., Finkelstein, D. I., Horne, M. K., Paxinos, G., & Halliday, G. M. (2002, April). Comparison of the basal ganglia in rats, marmosets, macaques, baboons, and humans: Volume and neuronal number for the output, internal relay, and striatal modulating nuclei. *The Journal of Comparative Neurology*, *445*(3), 238–255.

Harris, K. D., & Shepherd, G. M. G. (2015, February). The neocortical circuit: Themes and variations. *Nature Neuroscience*, *18*(2), 170–181. doi: 10.1038/nn.3917

Hattox, A. M., & Nelson, S. B. (2007, December). Layer V neurons in mouse cortex projecting to different targets have distinct physiological properties. *Journal of Neurophysiology*, *98*, 3330–3340.

Hazy, T. E., Frank, M. J., & O'Reilly, R. C. (2006, April). Banishing the homunculus: Making working memory work. *Neuroscience*, *139*, 105–118.

Hazy, T. E., Frank, M. J., & O'Reilly, R. C. (2007, August). Towards an executive without a homunculus: Computational models of the prefrontal cortex/basal ganglia system. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1485), 1601–1613.

Herd, S. A., O'Reilly, R. C., Hazy, T. E., Chatham, C. H., Brant, A. M., & Friedman, N. P. (2014, April). A neural network model of individual differences in task switching abilities. *Neuropsychologia*.

Hikida, T., Kimura, K., Wada, N., Funabiki, K., & Nakanishi, S. (2010). Distinct Roles of Synaptic Transmission in Direct and Indirect Striatal Pathways to Reward and Aversive Behavior. *Neuron*, *66*, 896–907.

Hikosaka, O., Sakamoto, M., & Usui, S. (1989, July). Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *Journal of Neurophysiology*, *61*(4), 814–832.

Hikosaka, O., & Wurtz, R. H. (1983, August). Visual and oculomotor functions of monkey substantia nigra pars reticulata. III. Memory-contingent visual and saccade responses. *Journal of Neurophysiology*, *49*(5), 1268–1284.

Hinton, G. E., McClelland, J. L., & Rumelhart, D. E. (1986, January). Distributed Representations. In D. E. Rumelhart, J. L. McClelland, & P. R. Group (Eds.), *Parallel Distributed Processing. Volume 1: Foundations* (pp. 77–109). Cambridge, MA: MIT Press.

Hochreiter, S., & Schmidhuber, J. (1997, January). Long Short-Term Memory. *Neural Computation*, *9*, 1735–1780.

Houk, J. C. (2005, June). Agents of the mind. *Biol Cybern*, *92*(6), 427–437.

Huang, T.-R., Hazy, T. E., Herd, S. A., & O'Reilly, R. C. (2013, June). Assembling old tricks for new tasks: A neural model of instructional learning and control. *Journal of Cognitive Neuroscience*, *25*(6), 843–851.

Ilinsky, I. A., Jouandet, M. L., & Goldman-Rakic, P. S. (1985). Organization of the nigrothalamocortical system in the rhesus monkey. *Journal of Comparative Neurology*, *236*(3), 315–330. doi: 10.1002/cne.902360304

Jilk, D., Lebiere, C., O'Reilly, R. C., & Anderson, J. (2008, September). SAL: An explicitly pluralistic cognitive architecture. *Journal of Experimental & Theoretical Artificial Intelligence*, *20*(3), 197–218.

Joel, D., & Weiner, I. (2000, April). The connections of the dopaminergic system with the striatum in rats and primates: An analysis with respect to the functional and compartmental organization of the striatum. *Neuroscience*, *96*, 451–474.

Jones, E. G. (1998a). A new view of specific and nonspecific thalamocortical connections. *Advances in Neurology*, *77*, 49–71.

Jones, E. G. (1998b, April). Viewpoint: The core and matrix of thalamic organization. *Neuroscience*, *85*(2), 331–345. doi: 10.1016/S0306-4522(97)00581-2

Jones, E. G. (2007). *The Thalamus* (Second ed., Vol. 2 volume set). Cambridge: Cambridge University Press.

Jordan, M. I. (1986). Attractor Dynamics and Parallelism in a Connectionist Sequential Machine. In *Proceedings of the 8th Confererence of the Cognitive Science Society* (pp. 531–546). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jung, W. H., Jang, J. H., Park, J. W., Kim, E., Goo, E.-H., Im, O.-S., & Kwon, J. S. (2014, September). Unravelling the Intrinsic Functional Organization of the Human Striatum: A Parcellation and Connectivity Study Based on Resting-State fMRI. *PLOS ONE*, *9*(9), e106768. doi: 10.1371/journal.pone.0106768

Kansky, K., Silver, T., Mély, D. A., Eldawy, M., Lázaro-Gredilla, M., Lou, X., … George, D. (2017, August). Schema Networks: Zero-shot transfer with a generative causal model of intuitive physics. *arXiv:1706.04317 [cs]*.

Kimura, M., Kato, M., & Shimazaki, H. (1990). Physiological properties of projection neurons in the monkey striatum to the globus pallidus. *Experimental Brain Research*, *82*(3), 672–676. doi: 10.1007/bf00228811

Kravitz, A. V., Tye, L. D., & Kreitzer, A. C. (2012, April). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nature Neuroscience*, *15*(6), 816–818.

Kriete, T., Mingus, B., Wyatte, D., Herd, S., & O'Reilly, R. C. (2011). Learned Figure-Ground Processing in an Embodied Brain Model. *Under review at the Psychological Research*.

Kriete, T., Noelle, D. C., Cohen, J. D., & O'Reilly, R. C. (2013, October). Indirection and symbol-like processing in the prefrontal cortex and basal ganglia. *Proceedings of the National Academy of Sciences U.S.A.*, *110*(41), 16390–16395.

Kritzer, M. F., & Goldman-Rakic, P. S. (1995, August). Intrinsic circuit organization of the major layers and sublayers of the dorsolateral prefrontal cortex in the rhesus monkey. *Journal of Comparative Neurology*, *359*(1), 131–143.

Krystal, J. H., Abi-Saab, W., Perry, E., D'Souza, D. C., Liu, N., Gueorguieva, R., … Breier, A. (2005, April). Preliminary evidence of attenuation of the disruptive effects of the NMDA glutamate receptor antagonist, ketamine, on working memory by pretreatment with the group II metabotropic glutamate receptor agonist, LY354740, in healthy human subjects. *Psychopharmacology*, *179*(1), 303–309. doi: 10.1007/s00213-004-1982-8

Kubota, K., & Niki, H. (1971, September). Prefrontal cortical unit activity and delayed alternation performance in monkeys. *Journal of Neurophysiology*, *34*(3), 337–347.

Kuramoto, E., Furuta, T., Nakamura, K. C., Unzai, T., Hioki, H., & Kaneko, T. (2009, September). Two types of thalamocortical projections from the motor thalamic nuclei of the rat: A single neuron-tracing study using viral vectors. *Cerebral cortex*, *19*(9), 2065–2077.

Kuramoto, E., Ohno, S., Furuta, T., Unzai, T., Tanaka, Y. R., Hioki, H., & Kaneko, T. (2015, January). Ventral medial nucleus neurons send thalamocortical afferents more widely and more preferentially to layer 1 than neurons of the ventral anterior–ventral lateral nuclear complex in the rat. *Cerebral Cortex*, *25*(1), 221–235. doi: 10.1093/cercor/bht216

Lamme, V. A. F. (2006, January). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, *10*(11), 494–501. doi: 10.1016/j.tics.2006.09.001

Larkum, M. E., Petro, L. S., Sachdev, R. N. S., & Muckli, L. (2018). A perspective on cortical layering and layer-spanning neuronal elements. *Frontiers in Neuroanatomy*, *12*. doi: 10.3389/fnana.2018.00056

Leichnetz, G. R., Spencer, R. F., Hardy, S. G., & Astruc, J. (1981). The prefrontal corticotectal projection in the monkey; an anterograde and retrograde horseradish peroxidase study. *Neuroscience*, *6*(6), 1023–1041.

Levitt, J. B., Lewis, D. A., Yoshioka, T., & Lund, J. S. (1993). Topography of Pyramidal Neuron Intrinsic Connections in Macaque Monkey Prefrontal Cortex (Areas 9 & 46). *Journal of Comparative Neurology*, *338*, 360–376.

Logie, R. H. (2018, September). Scientific advance and theory integration in working memory: Comment on Oberauer et al. (2018). *Psychological Bulletin; Washington*, *144*(9), 959.

Luck, S. J., & Vogel, E. K. (1997, December). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279.

Luck, S. J., & Vogel, E. K. (2013, August). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391–400. doi: 10.1016/j.tics.2013.06.006

Ma, W. J., Husain, M., & Bays, P. M. (2014, March). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356. doi: 10.1038/nn.3655

Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013, November). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, *503*(7474), 78–84. doi: 10.1038/nature12742

Masse, N. Y., Yang, G. R., Song, H. F., Wang, X.-J., & Freedman, D. J. (2019, July). Circuit mechanisms for the maintenance and manipulation of information in working memory. *Nature Neuroscience*, *22*(7), 1159–1167. doi: 10.1038/s41593-019-0414-3

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

Middleton, F. A., & Strick, P. L. (2000, May). Basal ganglia output and cognition: Evidence from anatomical, behavioral, and clinical studies. *Brain and Cognition*, *42*(2), 183–200.

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

Miller, E. K., & Desimone, R. (1994, February). Parallel neuronal mechanisms for short-term memory. *Science (New York, N.Y.)*, *263*, 520–522.

Miller, E. K., Erickson, C. A., & Desimone, R. (1996, August). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, *16*(16), 5154–5167.

Miller, G. A. (1956). *The magical number seven, plus or minus two: Some limits on our capacity for processing information.* (Vol. 101). Indiana : Bobbs-Merrill.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt.

Mink, J. W. (1996, March). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology*, *50*(4), 381–425.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000, September). The unity and diversity of executive functions and their contributions to complex ”Frontal Lobe” tasks: A latent variable analysis. *Cognitive psychology*, *41*, 49–100.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control.* New York: Cambridge University Press.

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

Moghaddam, B., & Adams, B. W. (1998, August). Reversal of phencyclidine effects by a group II metabotropic glutamate receptor agonist in rats. *Science*, *281*(5381), 1349–1352. doi: 10.1126/science.281.5381.1349

Mollick, J. A., Hazy, T. E., Krueger, K. A., Nair, A., Mackie, P., Herd, S. A., & O'Reilly, R. C. (2020, November). A systems-neuroscience model of phasic dopamine. *Psychological Review*, *127*(6), 972–

1021. doi: 10.1037/rev0000199

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

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

Mountcastle, V. B. (1997, June). The columnar organization of the neocortex. *Brain*, *120*(Pt 4), 701–722.

Moustafa, A. A., Sherman, S. J., & Frank, M. J. (2008, January). A dopaminergic basis for working memory, learning, and attentional shifting in Parkinson's Disease. *Neuropsychologia*, *46*, 3144–3156.

Münkle, M. C., Waldvogel, H. J., & Faull, R. L. M. (2000, July). The distribution of calbindin, calretinin and parvalbumin immunoreactivity in the human thalamus. *Journal of Chemical Neuroanatomy*, *19*(3), 155–173. doi: 10.1016/S0891-0618(00)00060-0

Nassar, M. R., Helmers, J. C., & Frank, M. J. (2018, July). Chunking as a rational strategy for lossy data compression in visual working memory. *Psychological Review*, *125*(4), 486–511. doi: 10.1037/rev0000101

Newell, A., & Simon, H. (1956, September). The logic theory machine–A complex information processing system. *IRE Transactions on Information Theory*, *2*(3), 61–79. doi: 10.1109/TIT.1956.1056797

Nyberg, L., Andersson, M., Forsgren, L., Jakobsson-Mo, S., Larsson, A., Marklund, P., ... Backman, L. (2009, July). Striatal dopamine D2 binding is related to frontal BOLD response during updating of long-term memory representations. *NeuroImage*, *46*(4), 1194–1199.

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., ... Ward, G. (2018a, September). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*(9), 885–958. doi: http://dx.doi.org.colorado.idm.oclc.org/10.1037/bul0000153

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., ... Ward, G. (2018b, September). Benchmarks provide common ground for model development: Reply to Logie (2018) and Vandierendonck (2018). *Psychological Bulletin*, *144*(9), 972–977. doi: http://dx.doi.org.colorado.idm.oclc.org/10.1037/bul0000165

Ongür, D., & Price, J. L. (2000, April). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral Cortex*, *10*(3), 206–219.

O'Reilly, R. C. (1996, January). 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. (2006, October). Biologically based computational models of high-level cognition. *Science*, *314*(5796), 91–94.

O'Reilly, R. C., Braver, T. S., & Cohen, J. D. (1999, January). A Biologically Based Computational Model of Working Memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control.* (pp. 375–411). New York: Cambridge University Press.

O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, *18*(2), 283–328.

O'Reilly, R. C., Frank, M. J., Hazy, T. E., & Watz, B. (2007). PVLV: The primary value and learned value Pavlovian learning algorithm. *Behavioral Neuroscience*, *121*(1), 31–49.

O'Reilly, R. C., Hazy, T. E., & Herd, S. A. (2016). The Leabra cognitive architecture: How to play 20 principles with nature and win! In S. Chipman (Ed.), *Oxford handbook of cognitive science.* Oxford University Press.

O'Reilly, R. C., Munakata, Y., Frank, M. J., Hazy, T. E., & Contributors. (2012). *Computational Cognitive Neuroscience*. Wiki Book, 1st Edition, URL: http://ccnbook.colorado.edu.

O'Reilly, R. C., Nair, A., Russin, J. L., & Herd, S. A. (2020, March). How Sequential Interactive Processing Within Frontostriatal Loops Supports a Continuum of Habitual to Controlled Processing. *Frontiers in Psychology*, *11*, 380. doi: 10.3389/fpsyg.2020.00380

O'Reilly, R. C., Noelle, D. C., Braver, T. S., & Cohen, J. D. (2002, February). Prefrontal cortex and dynamic categorization tasks: Representational organization and neuromodulatory control. *Cerebral cortex (New York, N.Y. : 1991)*, *12*, 246–257.

O'Reilly, R. C., Petrov, A. A., Cohen, J. D., Lebiere, C. J., Herd, S. A., & Kriete, T. (2014). How Limited Systematicity Emerges: A Computational Cognitive Neuroscience Approach. In I. P. Calvo & J. Symons (Eds.), *The architecture of cognition: Rethinking Fodor and Pylyshyn*^{1}*s Systematicity Challenge.* Cambridge, MA: MIT Press.

O'Reilly, R. C., Russin, J. L., & Herd, S. A. (2019, January). Computational models of motivated frontal function. In M. D'Esposito & J. Grafman (Eds.), *Handbook of Clinical Neurology* (Vol. 163, pp. 317–332). Amsterdam Oxford New York: Elservier.

O'Reilly, R. C., Russin, J. L., Zolfaghar, M., & Rohrlich, J. (2020, June). Deep Predictive Learning in Neocortex and Pulvinar. *arXiv:2006.14800 [q-bio]*.

Pakkenberg, B., & Gundersen, H. J. (1997, July). Neocortical neuron number in humans: Effect of sex and age. *Journal of Comparative Neurology*, *384*(2), 312–320.

Pauli, W. M., O'Reilly, R. C., Yarkoni, T., & Wager, T. D. (2016, February). Regional specialization within the human striatum for diverse psychological functions. *Proceedings of the National Academy of Sciences*, *113*(7), 1907–1912. doi: 10.1073/pnas.1507610113

Pertzov, Y., Bays, P. M., Joseph, S., & Husain, M. (2013, October). Rapid forgetting prevented by retrospective attention cues. *Journal of Experimental Psychology. Human Perception and Performance*, *39*(5), 1224–1231. doi: 10.1037/a0030947

Phillips, J. W., Schulmann, A., Hara, E., Winnubst, J., Liu, C., Valakh, V., … Hantman, A. W. (2019, November). A repeated molecular architecture across thalamic pathways. *Nature Neuroscience*, *22*(11), 1925–1935. doi: 10.1038/s41593-019-0483-3

Plenz, D., & Wickens, J. R. (2010). The striatal skeleton: Medium spiny projection neurons and their lateral connections. In H. Steiner & K. Y. Tseng (Eds.), *Handbook of Basal Ganglia Structure and Function* (pp. 99–112). USA: Academic Press.

Rac-Lubashevsky, R., & Frank, M. J. (2020, December). Analogous computations in working memory input, output and motor gating: Electrophysiological and computational modeling evidence. *bioRxiv*, 2020.12.21.423791. doi: 10.1101/2020.12.21.423791

Ramaswamy, S., & Markram, H. (2015). Anatomy and physiology of the thick-tufted layer 5 pyramidal neuron. *Frontiers in Cellular Neuroscience*, *9*. doi: 10.3389/fncel.2015.00233

Rao, S. G., Williams, G. V., & Goldman-Rakic, P. S. (1999, May). Isodirectional tuning of adjacent interneurons and pyramidal cells during working memory: Evidence for microcolumnar organization in PFC. *Journal of Neurophysiology*, *81*(4), 1903–1916.

Redondo, R. L., & Morris, R. G. M. (2011, January). Making memories last: The synaptic tagging and capture hypothesis. *Nature Reviews Neuroscience*, *12*(1), 17–30. doi: 10.1038/nrn2963

Rikhye, R. V., Gilra, A., & Halassa, M. M. (2018, December). Thalamic regulation of switching between cortical representations enables cognitive flexibility. *Nature Neuroscience*, *21*(12), 1753–1763. doi: 10.1038/s41593-018-0269-z

Roberts, B. M., Shaffer, C. L., Seymour, P. A., Schmidt, C. J., Williams, G. V., & Castner, S. A. (2010, March). Glycine transporter inhibition reverses ketamine-induced working memory deficits. *NeuroReport*, *21*(5), 390–394. doi: 10.1097/WNR.0b013e3283381a4e

Robinson, A. J., & Fallside, F. (1987, January). *The utility driven dynamic error propagation network* (Tech. Rep. No. CUED/F-INFENG/TR.1). Cambridge: Cambridge University Engineering Department.

Rougier, N. P., Noelle, D., Braver, T. S., Cohen, J. D., & O'Reilly, R. C. (2005, January). Prefrontal Cortex and the Flexibility of Cognitive Control: Rules Without Symbols. *Proceedings of the National Academy of Sciences*, *102*(20), 7338–7343.

Rougier, N. P., & O'Reilly, R. C. (2002, January). Learning representations in a gated prefrontal cortex model of dynamic task switching. *Cognitive Science*, *26*, 503–520.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986, January). Learning representations by back-propagating errors. *Nature*, *323*(9), 533–536.

Sanders, H., Berends, M., Major, G., Goldman, M. S., & Lisman, J. E. (2013, January). NMDA and GABAB (KIR) Conductances: The "Perfect Couple" for Bistability. *Journal of Neuroscience*, *33*(2), 424–429. doi: 10.1523/JNEUROSCI.1854-12.2013

Schmidhuber, J., Gers, F., & Eck, D. (2002, August). Learning nonregular languages: A comparison of simple recurrent networks and LSTM. *Neural Computation*, *14*(9), 2039–2042.

Schmidt, R., Ruiz, M. H., Kilavik, B. E., Lundqvist, M., Starr, P. A., & Aron, A. R. (2019, October). Beta oscillations in working memory, executive control of movement and thought, and sensorimotor function. *Journal of Neuroscience*, *39*(42), 8231–8238. doi: 10.1523/JNEUROSCI.1163-19.2019

Schroll, H., Vitay, J., & Hamker, F. H. (2012, February). Working memory and response selection: A computational account of interactions among cortico-basalganglio-thalamic loops. *Neural Networks*, *26*, 59–74. doi: 10.1016/j.neunet.2011.10.008

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

Seamans, J. K., & Yang, C. R. (2004, September). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Progress in Neurobiology*, *74*(1), 1–57.

Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008, August). Measuring consciousness: Relating behavioural and neurophysiological approaches. *Trends in Cognitive Sciences*, *12*(8), 314–321. doi: 10.1016/j.tics.2008.04.008

Shiffrin, R. M., & Schneider, W. (1977, January). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127–190.

Sommer, M. A., & Wurtz, R. H. (2000, April). Composition and topographic organization of signals sent from the frontal eye field to the superior colliculus. *Journal of Neurophysiology*, *83*(4), 1979–2001.

Stelzel, C., Basten, U., Montag, C., Reuter, M., & Fiebach, C. J. (2010, October). Frontostriatal involvement in task switching depends on genetic differences in D2 receptor density. *Journal of Neuroscience*, *30*(42), 14205–14212.

Stocco, A., Lebiere, C., & Anderson, J. (2010). Conditional routing of information to the cortex: A model of the basal ganglia's role in cognitive coordination. *Psychological Review*, *117*, 541–574.

Stokes, M. G. (2015, July). 'Activity-silent' working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences*, *19*(7), 394–405. doi: 10.1016/j.tics.2015.05.004

Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., & Duncan, J. (2013, April). Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, *78*(2), 364–375. doi: 10.1016/j.neuron.2013.01.039

Stroop, J. R. (1935, January). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Tanibuchi, I., Kitano, H., & Jinnai, K. (2009a, November). Substantia nigra output to prefrontal cortex via thalamus in monkeys. I. Electrophysiological identification of thalamic relay neurons. *Journal of Neurophysiology*, *102*(5), 2933–2945.

Tanibuchi, I., Kitano, H., & Jinnai, K. (2009b, November). Substantia nigra output to prefrontal cortex via thalamus in monkeys. II. Activity of thalamic relay neurons in delayed conditional go/no-go discrimination task. *Journal of Neurophysiology*, *102*(5116), 2946–2954.

Todd, M. T., Niv, Y., & Cohen, J. D. (2008). Learning to use Working Memory in Partially Observable Environments through Dopaminergic Reinforcement. In D. Koller (Ed.), *Advances in Neural Information Processing Systems (NIPS) 21* (Vol. 21). Curran Associates.

Uylings, H., Groenewegen, H., & Kolb, B. (2003). Do rats have a prefrontal cortex? *Behavioural Brain Research*, *146*(1-2), 3–17.

van Moorselaar, D., Theeuwes, J., & Olivers, C. N. L. (2014, August). In competition for the attentional template: Can multiple items within visual working memory guide attention? *Journal of Experimental Psychology. Human Perception and Performance*, *40*(4), 1450–1464. doi: 10.1037/a0036229

Vandierendonck, A. (2018, September). Working memory benchmarks—A missed opportunity: Comment on Oberauer et al. (2018). *Psychological Bulletin*, *144*(9), 963–971. doi: http://dx.doi.org.colorado

.idm.oclc.org/10.1037/bul0000159

Voytek, B., & Knight, R. T. (2010, October). Prefrontal cortex and basal ganglia contributions to visual working memory. *Proceedings of the National Academy of Sciences USA*, *107*(42), 18167–18172.

Wang, M., Yang, Y., Wang, C.-J., Gamo, N. J., Jin, L. E., Mazer, J. A., … Arnsten, A. F. T. (2013, February). NMDA receptors subserve persistent neuronal firing during working memory in dorsolateral prefrontal cortex. *Neuron*, *77*(4), 736–749. doi: 10.1016/j.neuron.2012.12.032

Wang, X.-J. (2001, August). Synaptic reverberation underlying mnemonic persistent activity. *Trends in Neurosciences*, *24*(8), 455–463.

Wang, Y., Markram, H., Goodman, P. H., Berger, T. K., Ma, J., & Goldman-Rakic, P. S. (2006, April). Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nature Neuroscience*, *9*(4), 534–542.

Watanabe, Y., & Funahashi, S. (2012, January). Thalamic mediodorsal nucleus and working memory. *Neuroscience & Biobehavioral Reviews*, *36*(1), 134–142. doi: 10.1016/j.neubiorev.2011.05.003

Watanabe, Y., Takeda, K., & Funahashi, S. (2009, June). Population vector analysis of primate mediodorsal thalamic activity during oculomotor delayed-response performance. *Cerebral Cortex*, *19*, 1313–1321.

Wei, Z., Wang, X.-J., & Wang, D.-H. (2012, August). From distributed resources to limited slots in multiple-item working memory: A spiking network model with normalization. *Journal of Neuroscience*, *32*(33), 11228–11240.

Werbos, P. (1974). *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences* (Unpublished doctoral dissertation). Harvard University.

Werbos, P. (1990, October). Backpropagation through time: What it does and how to do it. *Proceedings of the IEEE*, *78*(10), 1550–1560. doi: 10.1109/5.58337

Whittington, J. C. R., & Bogacz, R. (2019, March). Theories of error back-propagation in the brain. *Trends in Cognitive Sciences*, *23*(3), 235–250. doi: 10.1016/j.tics.2018.12.005

Wickens, J. R., Alexander, M. E., & Miller, R. (1991, May). Two dynamic modes of striatal function under dopaminergic-cholinergic control: Simulation and analysis of a model. *Synapse (New York, N.Y.)*, *8*(1), 1–12. doi: 10.1002/syn.890080102

Wilken, P., & Ma, W. J. (2004, December). A detection theory account of change detection. *Journal of Vision*, *4*(12), 1120–1135. doi: 10.1167/4.12.11

Williams, A., & Phillips, J. (2020, April). Transfer reinforcement learning using output-gated working memory. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(02), 1324–1331. doi: 10.1609/aaai.v34i02.5488

Williams, R. J., & Zipser, D. (1992, January). Gradient-based learning algorithms for recurrent networks and their computational complexity. In Y. Chauvin & D. E. Rumelhart (Eds.), *Backpropagation: Theory, Architectures and Applications.* Hillsdale, NJ: Erlbaum.

Winnubst, J., Bas, E., Ferreira, T. A., Wu, Z., Economo, M. N., Edson, P., … Chandrashekar, J. (2019, September). Reconstruction of 1,000 projection neurons reveals new cell types and organization of long-range connectivity in the mouse brain. *Cell*, *179*(1), 268-281.e13. doi: 10.1016/j.cell.2019.07.042

Wyder, M. T., Massoglia, D. P., & Stanford, T. R. (2004, June). Contextual modulation of central thalamic delay-period activity: Representation of visual and saccadic goals. *Journal of Neurophysiology*, *91*(6), 2628–2648.

Yehene, E., Meiran, N., & Soroker, N. (2008, June). Basal ganglia play a unique role in task switching within the frontal-subcortical circuits: Evidence from patients with focal lesions. *Journal of Cognitive Neuroscience*, *20*, 1079–1093.

Yttri, E. A., & Dudman, J. T. (2016, May). Opponent and bidirectional control of movement velocity in the basal ganglia. *Nature*, *533*(7603), 402–406. doi: 10.1038/nature17639

Zalocusky, K. A., Ramakrishnan, C., Lerner, T. N., Davidson, T. J., Knutson, B., & Deisseroth, K. (2016, March). Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. *Nature*, *531*(7596), 642-646. doi: 10.1038/nature17400

Zhang, W., & Luck, S. J. (2008, May). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235.