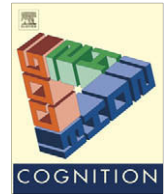




ELSEVIER

Contents lists available at ScienceDirect

## Cognition

journal homepage: [www.elsevier.com/locate/COGNIT](http://www.elsevier.com/locate/COGNIT)

## Developing PFC representations using reinforcement learning

Jeremy R. Reynolds<sup>a,\*</sup>, Randall C. O'Reilly<sup>b</sup>

<sup>a</sup> Department of Psychology, University of Denver, 2155 S. Race St., Denver, CO 80208, United States

<sup>b</sup> Department of Psychology, University of Colorado, United States

### ARTICLE INFO

#### Article history:

Received 18 October 2007

Revised 11 May 2009

Accepted 18 May 2009

#### Keywords:

PFC

Representation

Reinforcement learning

Functional organization

### ABSTRACT

From both functional and biological considerations, it is widely believed that action production, planning, and goal-oriented behaviors supported by the frontal cortex are organized hierarchically [Fuster (1991); Koehlin, E., Ody, C., & Kouneiher, F. (2003). *Neuroscience: The architecture of cognitive control in the human prefrontal cortex. Science, 424*, 1181–1184; Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt]. However, the nature of the different levels of the hierarchy remains unclear, and little attention has been paid to the origins of such a hierarchy. We address these issues through biologically-inspired computational models that develop representations through reinforcement learning. We explore several different factors in these models that might plausibly give rise to a hierarchical organization of representations within the PFC, including an initial connectivity hierarchy within PFC, a hierarchical set of connections between PFC and subcortical structures controlling it, and differential synaptic plasticity schedules. Simulation results indicate that architectural constraints contribute to the segregation of different types of representations, and that this segregation facilitates learning. These findings are consistent with the idea that there is a functional hierarchy in PFC, as captured in our earlier computational models of PFC function and a growing body of empirical data.

© 2009 Elsevier B.V. All rights reserved.

### 1. Introduction

The prefrontal cortex (PFC) plays a critical role in the execution of controlled behavior (Miller & Cohen, 2001). Many theories exist regarding the function of PFC (for reviews, see Stuss & Knight, 2002; Wood & Grafman, 2003), and this plethora of theories in part reflects our lack of understanding concerning the functional organization of the multiple anatomical areas that compose PFC (Duncan & Owen, 2000; Miller, 2000). This lack of understanding is not for a lack of trying; there have been a number of approaches to investigating this critical question. One approach has been to focus on various stimulus dimensions that have produced reliable dissociations in posterior areas. For example, researchers have hypothesized that the dorsal and ventral visual processing streams project

into PFC, and cause a spatial vs. object dissociation along a dorsal and ventral gradient (Goldman-Rakic, 1987; Haxby, Petit, Ungerleider, & Courtney, 2000; Wilson, Scalaidhe, & Goldman-Rakic, 1993). Although there appears to be some evidence suggesting that areas of PFC may be dissociated along these stimulus dimensions (Johnson, Raye, Mitchell, Greene, & Anderson, 2003), there are also studies that have not been able to identify such distinctions (i.e. Nystrom et al., 2000).

A second approach has focused on investigating whether different processes determine the functional organization of the PFC. In particular, a common distinction in the literature has suggested that ventrolateral PFC (BA 44/45) is responsible for maintaining and rehearsing information in working memory (WM) whereas dorsolateral PFC (BA 9/46) is responsible for the manipulation and monitoring of such information (D'Esposito, Postle, Ballard, & Lease, 1999; Haxby et al., 2000; Petrides, 2000). Although there have been studies supporting this claim, there have

\* Corresponding author. Tel.: +1 303 871 4622.

E-mail address: [jeremy.reynolds@psy.du.edu](mailto:jeremy.reynolds@psy.du.edu) (J.R. Reynolds).

been numerous studies demonstrating areas of dorsolateral PFC active under situations that do not require manipulation *prima facie* (Braver & Bongiolatti, 2002; Braver, Cohen, & Barch, 2002; Raye, Johnson, Mitchell, Reeder, & Greene, 2002). In addition to the hypothesis regarding dorsolateral and ventrolateral PFC, studies have also suggested that there is an additional gradient dissociating anterior PFC (BA 10) from more posterior areas of lateral PFC based on process, although the nature of that process is unclear. For example, data suggest that anterior PFC may process internally generated information (Christoff & Gabrieli, 2000; Christoff, Ream, Geddes, & Gabrieli, 2003), or that it may be associated with multi-tasking aspects of maintaining one goal in an active state while executing or scheduling a different one (Braver & Bongiolatti, 2002; Braver, Reynolds, & Donaldson, 2003; Koechlin, Basso, Pietrini, Panzer, & Grafman, 1999; Reynolds, West, & Braver, 2008). Finally, recent evidence has suggested that PFC may be hierarchically organized, such that more posterior, dorsal regions are intimately tied to determining the appropriate response, and more anterior regions process more abstract aspects of the task that are summed to inform the response (Koechlin, Ody, & Kounieher, 2003; Koechlin & Summerfield, 2007). Future research will be needed to determine whether these distinctions are isomorphic, or whether they capture something fundamentally different about cognition and PFC function.

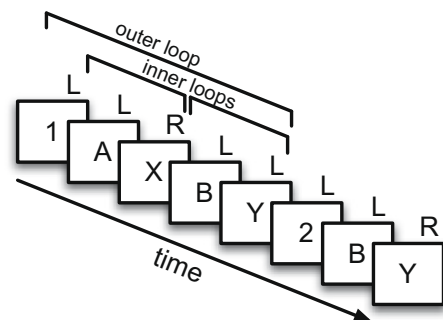
Despite the large number of theories and studies investigating the functional organization of PFC, there is not a satisfactory set of theories that is strongly supported by the available data. As mentioned above, one problem, particularly in the approach analyzing organization according to different processes, is that there do not appear to be general, grounded definitions of each proposed process. For example, What constitutes a manipulation in the context of the maintenance/manipulation distinction? Does simply updating the contents of working memory constitute a manipulation? Likewise, what does it mean to internally generate information? Any transformation upon externally presented information could be considered internally generated; where does one draw the line? Which specific computations are being performed that facilitate multi-task and sub-goal performance? Finally, what are the specific computations/dimensions that determine the aspect of the task that is being represented; (e.g. in Koechlin et al. (2003) tasks, what causes something to represent task context vs. episode context)?

One approach that can help address the ambiguity over the functional definition of various processes is computational modeling. To this end, the current simulations investigated the extent to which a biologically-based computational modeling framework can provide insight into the development of specialized neural representations in different PFC areas, with a focus on the hierarchical structuring of such representations (Fuster, 1991; Koechlin et al., 2003; Miller, Galanter, & Pribram, 1960). In particular, there are substantial gains to be had by hierarchically organized systems; systems organized in such a way frequently learn faster than systems that contain no such organization (Botvinick, 2008), and display substantially greater generalization (Botvinick, 2008; Lashley, 1951;

Miller et al., 1960). Effectively, a hierarchically organized system allows individuals to form abstractions (e.g. a plan to traverse a room to the opposite door) that can be applied in a variety of different contexts. Once discovered, these abstractions, or chunks, can then be applied to novel situations.

We explore these issues in the context of a relatively sophisticated task that has a hierarchical structure: the 1–2 AX-continuous performance task (Frank, Loughry, & O'Reilly, 2001). In this task, letters and numbers are presented sequentially over time, and participants must detect specific target sequences. The appropriate response to a particular sequence (such as A–X) is dependent on the most recently viewed number; thus, the cues (A's) and probes (X's) are nested hierarchically within an outer-loop of number information (see Fig. 1). For example, if a 1 was last seen, the target is A–X, but if a 2 was last seen, the target is B–Y (see Table 1). The computational model we employ here can successfully learn to perform this task (O'Reilly & Frank, 2006) using a reinforcement learning algorithm called PVLV (O'Reilly, Frank, Hazy, & Watz, 2007), coupled to a biologically-based model of the Basal Ganglia (BG) and PFC circuitry, all of which are detailed in Supplementary material.

Specifically, we build on previous work suggesting that the BG can dynamically and selectively gate the contents of PFC, and thus permit task-relevant information to be maintained in PFC while preventing extraneous information from interfering with performance (Frank et al., 2001; O'Reilly & Frank, 2006). A key property of this approach is that it allows for selective updating via the inclusion of parallel loops of connectivity between the BG and PFC, such that a particular set of neurons within the BG can control the updating of a particular set of neurons within the PFC (which we refer to as a stripe). This ability to selectively update some contents of WM while leaving other content intact is a process that is fundamental to hierarchical behavior, because the nature of the task representations have, by definition, different temporal dynamics. Larger goals are relevant over longer periods of time, and thus should not be updated once one sub-goal is com-



**Fig. 1.** The 1–2 AX-CPT task. Stimuli are presented one at a time in sequence. The participant responds by pressing the right key (R) to the target sequence; otherwise, a left key (L) is pressed. If the participant last saw a 1, then the target sequence is an A followed by an X. If a 2 was last seen, then the target is a B followed by a Y. The maintenance of the task stimuli (1 or 2) constitutes a temporal outer loop around multiple inner-loop memory updates required to detect the target sequence.

**Table 1**

The task paradigm performed by the model. The first row indicates each of the 4 outer-loop contexts (i.e. tasks). The second row indicates the inner-loop context cue nested within that task. For example, A and B in the last 2 columns appear in the context of the “2” outer-loop context, whereas R and S in the first two columns appear in the context of the “3” outer-loop context. The first column refers to the probes that may appear within each task. Within the table, empty fields mean those stimulus pairings do not appear, whereas T means that a target response is required on that combination of outer-loop, inner-loop, and probe and NT means that a non-target response is required. Numbers within parentheses refer to the relative percentage of the sequences within that particular outer-loop context (1, 2, 3 or 4). As can be seen from the table, the outer context is unnecessary in the 3 and 4 contexts, as an appropriate response can be determined purely from the inner cue and the probe.

	3		4		1		2	
	R	S	M	N	A	B	A	B
O	T (30%)	NT (20%)						
P	NT (20%)	NT (30%)						
Q			NT (30%)	NT (20%)				
T			NT (20%)	T (30%)				
X					T (30%)	NT (20%)	NT (30%)	NT(20%)
Y					NT (20%)	NT (30%)	NT(20%)	T (30%)

pleted and another is begun. The current framework provides an ideal avenue for exploring hierarchical behavior, because this ability to perform asynchronous updating has been established across several tasks, including the one explored in the current manuscript (O'Reilly & Frank, 2006).

While previous explorations have interrogated the ability of this type of network to learn and perform these tasks (O'Reilly & Frank, 2006), there has been little attention to the types of representations that are developed. Within the context of the task we investigated, outer-loop information (number) is available at the time at which the network processes inner-loop information (letter cues and probes). Thus, outer-loop information has the opportunity to guide and shape the developing inner-loop representations. We investigated the extent to which various manipulations of the network connectivity and other parameters would influence the degree to which the representation of these two types of information (inner-loop vs. outer-loop) were segregated or interdigitated.

The first simulation investigated whether a natural connectivity hierarchy would be sufficient to constrain the development of representations within PFC. This type of hierarchy has been associated with neuroanatomical data (Fuster, 1991, 2004; Petrides & Pandya, 2007), and has been shown to encourage the representation of stable context representations in previous computational models (Botvinick, 2007).

Additionally, we investigated the idea that a hierarchical structure allows for more generalized, flexible behavior by having higher-level areas dynamically contextualize and modulate the input/output mappings represented in lower-level systems (Botvinick, 2008; Lashley, 1951; Miller et al., 1960).

The third simulation is motivated by developmental data suggesting that (a) cognitive development seems to proceed in a hierarchical fashion, where lower-levels of cognitive processing are established prior to the development of more complex processing (Bunge & Zelazo, 2006; Halford, Wilson, & Phillips, 1999) and (b) that this behavioral timecourse could be mediated by developments in corresponding areas of PFC (Bunge & Zelazo, 2006; Gogtay et al., 2004; O'Donnell, Noseworthy, Levine, & Dennis, 2005). In order to model this type of developmental trajectory, we

included additional tasks in the training regimen. These tasks (called the 3–4 tasks) had exactly the same structure as the 1–2 AX task, but the stimulus contents of each inner loop were unique to each task (e.g. in the three task, R–P might be a target sequence, and those stimuli would never occur in any other outer-loop context, see Table 1). As such, the cue-probe pair uniquely determines the response, and there is no need to maintain or process the outer-loop information. In the context of these tasks, a developmental trajectory would suggest that one learns the easier, less complex tasks first, the representational content of which occupies resources. As one begins to learn more complicated tasks, the resources that were previously available are now occupied, and additional resources need to be utilized in order to solve the problem. The implementation of each approach is detailed in Section 2 below.

To foreshadow the results, we found that (a) anatomical constraints provide a strong constraint in the learning of such representations, and resulting hierarchical segregated representations facilitate learning, (b) outer-loop information is represented both in the form of stable representations across trials as well as conjunctive representations that dynamically update each trial, and (c) differential plasticity schedules have little influence on the development of segregated or hierarchical representations within the context of the various models implemented here.

## 2. Methods

### 2.1. Tasks

The model was trained to perform four different versions of the AX-CPT. In each of the versions of the AX-CPT modeled, the model was asked to make a target response to a particular probe (e.g. “X”) that follows a particular cue (e.g. “A”), and to make a non-target response after all other cue-probe combinations. Two of the versions followed this exact format; the only difference was the set of stimuli used for each task (see Table 1, columns 1–5). In the other two versions of the task, the exact same stimuli were used across the tasks, but the appropriate response mappings changed as a function of the task-level (henceforth called outer-loop) context (the number; see Table 1,

columns 6–9). Specifically, if the most recent task cue was a “1,” then the appropriate target sequence would be A–X, but if the most recent task cue was a “2,” then the appropriate target sequence would be B–Y (see Fig. 1). The probability of the cue–probe pairs was adjusted from the typical expectancy versions of the AX-CPT Braver et al., 2001 by decreasing the relative probability of the target sequences and increasing the relative probability of non-target trials that have no stimulus overlap with the target sequence (e.g. BY trials when AX is the target). The primary reason behind this change was so that the network would not be able to identify the current task by keeping track of the relative frequency of the AX and BY sequences. In this way, the stimulus frequencies provide no diagnostic information with respect to what task the model is supposed to be doing. Each task was equally likely to appear, and once the model encountered a task cue, the model would be asked to perform between 1 and 5 trials associated with that task prior to seeing a new task cue.

Each input and output was represented by its own localist representation. All other layers had distributed representations that were learned over the course of training (including the PFC, see below).

## 2.2. Modeling and training

The implemented models use the Leabra framework described in Appendix (O'Reilly, 1998, 2001; O'Reilly & Munakata, 2000), with the additional specialized prefrontal-cortex basal ganglia working memory (PBWM) mechanisms described in O'Reilly and Frank (2006). As a brief summary, the Leabra framework uses point neurons with excitatory, inhibitory, and leak conductances contributing to an integrated membrane potential, which is then thresholded and transformed via an  $x/(x + 1)$  sigmoidal function to produce a rate code output. Each layer uses a k-winners-take all (kWTA) function that computes an inhibitory conductance that keeps roughly the k most active units above firing threshold and keeps the rest below threshold. Units learn according to a combination of Hebbian, error-driven, and reinforcement learning, with the error-driven component computed using the generalized recirculation algorithm (GeneRec; O'Reilly, 1996), which computes backpropagation derivatives using two phases of activation settling. The cortical layers in the model use standard Leabra parameters and functionality, while the PBWM systems require some additional mechanisms to implement the DA modulation of Go/NoGo units, and toggling of PFC maintenance currents, as detailed in Supplementary material.

The base network architecture followed the organization depicted in Fig. 2, and it was largely consistent across simulation. The input, hidden, and output layers consisted of 16, 49, and 2 units, respectively. The PFC layer consisted of four stripes of 36 units each, while the Matrix layer consisted of four stripes of 28 units each (14 Go and 14 NoGo units). Implementation of the primary value-learned value (PVLV) system is detailed in Supplementary material. If two layers were connected, all units in one layer were connected with all units in the other, except for those connected to the PFC. The set of projections to the PFC was a target of manipulation within the Simulations, and the de-

tails regarding its various connectivity patterns are enumerated below.

## 2.3. Training procedure

The networks were trained to a similar criterion as in previous models using this task (O'Reilly & Frank, 2006; O'Reilly et al., 2007). Specifically, the networks had to complete 80 sequences of 1–5 trials with 0 errors (i.e. one epoch). Each epoch was constrained to include at least four trials of each type defined in Table 1. Although the stopping criterion is defined differently from previous published work with this task, the current criterion corresponds to approximately the same number of correct trials per outer-loop context as previously used criteria (current = 60 per outer-loop context, previous = mean of 62.5 per outer-loop context). Each network was trained 100 different times, with each training run initialized with different random seeds.

In order to collect behavioral and PFC activity data, learning was turned off after training, and each network was run through the tasks for an additional 240 sequences.

## 2.4. Data and analysis

### 2.4.1. Behavior/performance

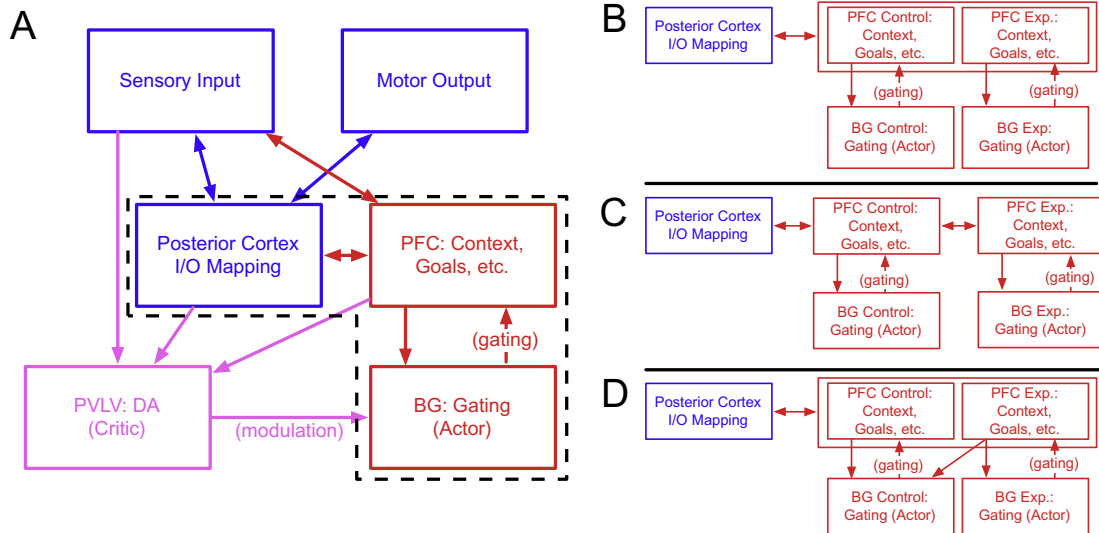
The behavior and performance of the networks were tested in two different ways. First, the rate of successful training and the time to criterion (number of epochs) was recorded for each replication, so that any manipulations could be tested to determine whether they influence learning. Second, we analyzed performance over the testing session in order to confirm that the networks were using active maintenance rather than weight-based memory. If weight-based memory were being utilized, then performance would greatly degrade during testing, as the only mechanism for memory is through active maintenance. Performance is reported for each type of task [i.e. low complexity (3–4) vs. high complexity (1–2)].

### 2.4.2. Representations

Due to the nature of our manipulations, we collapsed the four PFC stripes into two groups of two stripes. Within these two groups of stripes, a neurophysiology approach was taken to probe the types of representations the network was using to perform the task. This procedure involved using the activity patterns at the time of each probe in order to identify units demonstrating a significant response relative to baseline in any one of the four tasks ( $p < 0.001$ ), and then probing further to determine whether that response was selective to the task. This selectivity was determined by *t*-tests comparing the response to that particular task to each of the other tasks; if the response was different across all comparisons, it was said to be selective.

In order to interrogate differences between the two groups of units, an index was used to compare the proportion of each type of selective unit, controlling for the overall proportion of selective units:

$$\frac{p(\text{Exp}) - p(\text{Control})}{p(\text{Exp}) + p(\text{Control})} \quad (1)$$



**Fig. 2.** Network architectures for the simulations explored. Panel A reflects the overall structure of the network and its connections. The dashed box is expanded in panels B–D. Panel B reflects the structure of a typical network in this framework, with each PFC stripe governing the updating of itself. Panel C reflects the implementation of the first hypothesis. In this architecture, there is a different connectivity pattern associated with PFC; a subset of PFC units receive direct input from and provide direct output to hidden and output layers, while the rest of PFC only receives information from other areas of PFC. Panel D reflects the second hypothesis, in which the updating of one set of PFC units is governed both by its own activity as well as the activity of the other PFC units. “Control” and “Exp.” in panels B–D refer to the groups associated with the values in Eq. (1).

where  $p(\text{Control})$  is the proportion of selective units in the set of stripes with normal connectivity from the hidden layer and BG, and  $p(\text{Exp})$  is the proportion of selective units in the set of stripes with the experimental manipulation (see Fig. 2B–D). As reflected in Fig. 2B–D, the experimental group was defined for each simulation by either (a) replacing the projections between it and the hidden layer with projections to and from the other PFC group, (b) adding an additional projection to the BG stripes responsible for updating the control group of PFC units, or (c) delaying their learning. The experimental group always corresponds to the group of units that would be predicted to have more frequent outer-loop representations. The index places all measures on a  $-1$  to  $1$  scale in which  $-1$  means all selective units occur in the control group,  $1$  means all selective units occur in the experimental group, and  $0$  is an equal distribution of selective units across the groups. These ratios were then analyzed across multiple training runs to determine whether there were systematic differences in the type of representations that develop.

## 2.5. Simulations

The first two simulations investigate whether two similar, but distinct, computational constraints could potentially lead to the development of segregated representations: the ability of one set of units to directly influence the content of other units via direct biasing and the ability of one set of units to directly influence the timing of the updating signal associated with other units. The first simulation investigated whether a natural connectivity hierarchy would be sufficient to constrain the development of representations within PFC. This type of hierarchy has been associated with neuroanatomical data (Fuster,

1991, 2004; Petrides & Pandya, 2007), and has been shown to encourage the development of stable context representations in previous computational models (Botvinick, 2007; Paine & Tani, 2005). This hypothesis was implemented as shown in Fig. 2C, and is characterized by having bidirectional connections between the hidden layer and a group of stripes in the PFC layer, and bidirectional connectivity between these PFC stripes and the remaining PFC stripes. This is an explicit hierarchical structure similar to that used by Botvinick (2007) and Paine and Tani (2005), in which the stripes of PFC connected directly to the hidden layer are considered proximal the processing pathway, and as such, should come to represent information that is closer to the output pathway. Conversely, the PFC stripes connected only via the intermediate processing of PFC should reflect the apex of the processing pathway, and is predicted to represent the most abstract/temporally extended information. This network is labeled the PFC-hidden network. The representations developed by this network can then be compared to the representations developed in the context of a standard network in which all of PFC has access to information from the hidden layer, and does not direct connect with other areas of PFC (see Fig. 2B).

The second simulation investigated whether architectural changes allowing for higher-level areas to dynamically contextualize and modulate the input/output mappings represented in lower-level systems would encourage the development of segregated representations. This hypothesis was implemented as displayed in Fig. 2D; the experimental group of PFC units (see Fig. 2D), in addition to projecting to their corresponding BG stripes, also projected to the BG stripes associated with the group of control PFC units (labeled the PFC-BG network). This struc-



ture of projections was chosen for one primary reason. This projection formalizes the idea that anterior PFC has a strong direct influence over the timing of the updating process within lateral PFC. By influencing when lateral PFC is updated, the alternate area of PFC could likely shape the type of representations used by lateral PFC.

The third set of simulations interrogated whether different plasticity schedules could influence the development of segregated representations. This set of simulations was based on the observation that children learn to perform more simple tasks prior to learning more complex tasks (Halford, 1984, 1993; Robin & Holyoak, 1995), and further, that different areas of PFC tend to have different schedules of development (Shaw et al., 2008; Sowell, Thompson, Holmes, Jernigan, & Toga, 1999). The hypothesis is that if particular sets of PFC units are learning early, while the model is learning simple tasks, then those units would be pre-disposed to learn about the simple relationships in such tasks (e.g. between inner-loop cue-probe pairs). Once these tasks are learned, then if there are additional, non-committed units/stripes, then these non-committed units could be allocated to learning about additional sequential structure, namely about outer-loop task representations. In order to introduce this dynamic, the connections to and from the experimental group of PFC units did not learn until the less complex tasks achieved perfect performance for an epoch. Critically, the environment was constant over time, such the network was always attempting to learn all four tasks, but early in training, only half of its PFC resources were available to learn (a subsequent investigation investigated the role that a dynamic environment may play, but did not influence the results). While this procedure is not particularly realistic, it serves as a proxy for the natural development of the brain and cognition, in which it appears that young children learn to solve relatively simple problems prior to solving more complex problems (Halford, 1984, 1993), and that this cognitive development somewhat tracks the development of the cerebral cortex (Bunge & Zelazo, 2006).

## 2.6. Simulation results

### 2.6.1. Training

The only network that influenced the training success or training time was the PFC-hidden architecture (see Table 2). However, this difference in time to criterion appeared to be due largely to the number of stripes with access to hidden information, as a non-hierarchical net-

**Table 2**

Training information. The only network that produced different training statistics from the others was the network with constrained projections between the PFC and hidden layers (C). This network converged less frequently,  $\chi^2(1) = 18$ ,  $p < 0.001$ , and took longer to train when it successfully reached criterion,  $W = 1818$ ,  $p < 0.001$ . Letters in parentheses correspond to their component structure in Fig. 2.

	Proportion trained	Median time to criterion
Base (B)	100	66
PFC-hidden (C)	82***	296***
PFC-BG (D)	99	92
Plasticity	100	77.5

\*\*\*  $p < 0.001$ .

work with only two stripes demonstrated similar time to criterion (325 epochs,  $p = 0.9$ ).

### 2.6.2. Behavior

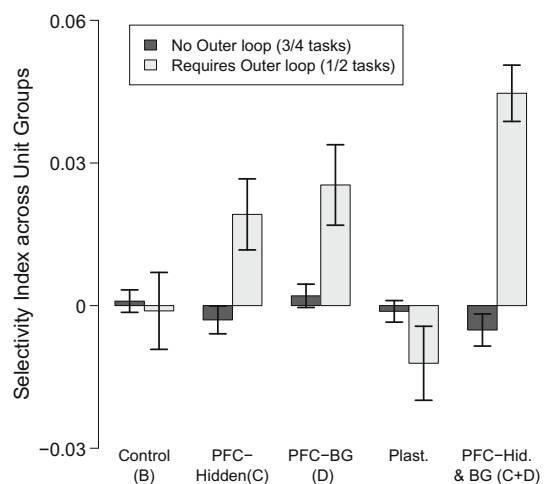
Across all networks, error rates during the testing phase were low, with the error rates in the 3 and 4 tasks being significantly lower than the error rates in the 1 and 2 tasks,  $F(1, 377) = 877$ ,  $p < 0.001$ ;  $\eta_p^2 = 0.7$ . This behavior did not fluctuate much as a function of the network architecture or plasticity schedule, as mean error rates in the 3–4 tasks were between 0.4% and 2.4%, and error rates in the 1–2 tasks were between 9.4% and 11.3%.

### 2.6.3. Selective representations

As expected, the selectivity index in the baseline network was not significantly different from 0 for either the complex 1–2 tasks or the less complex 3–4 tasks (collapsed:  $t(99) = 0.02$ ), indicating that there was no bias to develop representational specialization.

### 2.6.4. Constrained PFC-hidden layer projections

In contrast to baseline, the PFC-hidden network produced more segregated representations, such that the experimental group of units tended to code for outer-loop context more frequently than the control unit group, specifically when coding for information about the more complex, outer-loop tasks  $t(81) = 2.56$ ,  $p = 0.01$ , see Fig. 3. When compared directly to the control network, this network demonstrated a dissociation, such that the experimental group of units had a greater likelihood of repre-



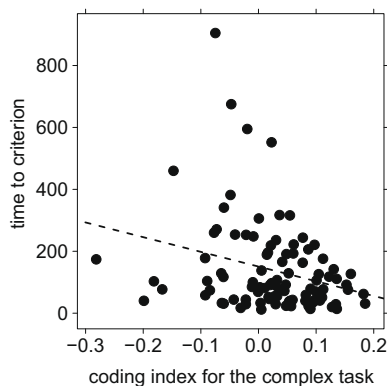
**Fig. 3.** Representational segregation for each type of task (low complexity = 3–4 tasks, high complexity = 1–2 tasks). Each set of bars corresponds to the difference between the control stripes and the experimental stripes for a different architecture or training regime. Dark bars correspond to the differences in type of representation for the less complex 3–4 tasks, whereas light bars correspond to differences in representation for the more complex 1–2 tasks. The first 3 groups correspond to the three different architectures presented in Fig. 2. The fourth set of bars correspond to the results when the control network is placed under a condition with developmental plasticity, such that the experimental stripes of PFC have a later onset of plasticity, and the fifth set of bars correspond to the results when both architectural constraints are included in the same model.

senting outer context in the more complex tasks  $F(1,162) = 4.84, p = 0.03$ .

Additionally, regression analyses indicated that there was a negative relationship between training time and the degree of representational segregation, specifically for the tasks requiring outer context ( $t(79) = -2.6, p = 0.01, \Delta R^2 = 0.08$ ). This relationship was different than that observed in the baseline network, where the degree of representational segregation was not related to training time (significant network  $\times$  index interaction  $t(177) = 3.6, p < 0.001$ ).

### 2.6.5. Hierarchical PFC-BG projections

Similar to the constrained PFC model, there was a relatively high degree of representational segregation within the model with hierarchical projections between the PFC and the BG, particularly for the complex tasks ( $t(98) = 3.0, p = 0.003$ , see Fig. 3). When compared directly to the control network, this network demonstrated a dissociation in developed responses, with the areas of PFC capable of directly influencing the updating of other areas of PFC having a greater likelihood of representing outer context  $F(1,197) = 4.5, p = 0.03$ . Similar to Simulation I, there was a negative relationship between time to train and the degree of representational segregation in the complex tasks, ( $t(96) = -2.8, p = 0.006, R^2 = 0.07$ , see Fig. 4). Visual inspection and diagnostic measures revealed heteroscedasticity in the model (see Fig. 4; significant Breusch–Pagan test:  $\chi^2(1) = 38, p < 0.001$ ), but the use of a heteroscedasticity consistent error term (Long & Ervin, 2000) revealed the relationship was still reliable, even after taking this violation into consideration ( $t(96) = -2.5, p = 0.01$ ). No quadratic trends were significant. Similar to Simulation I, this negative relationship was significantly different from the relationship identified in the control network (condition  $\times$  selectivity index interaction:  $t(194) = 3.8, p < 0.001$ ),  $\Delta R^2 = 0.07$ . In this particular example, there was no overall increase in training time relative to the control network (see Table 2), so this is a pure benefit associated with segregated representation.



**Fig. 4.** Relationship between representational specialization and training time for the PFC-BG network. An increased selectivity index was associated with faster training times. A similar pattern existed for the PFC-hidden network, but did not exist for the baseline or development networks.

### 2.6.6. Plasticity manipulation

Contrary to the earlier simulations, the simulation interrogating the role of plasticity did not produce a difference in representational structure ( $p = 0.12$ ), nor was the representational structure measure associated with training time ( $p = 0.8$ ).

### 2.6.7. Additional simulations

Each of the three experimental manipulations could be considered orthogonal to one another; one can perform the plasticity manipulation on the network with constrained PFC connections or the network with hierarchical projections to the BG. When all of the possible models were run and included in a 2 (PFC-hidden connectivity: either full or constrained)  $\times$  2 (PFC-BG connectivity: stripe specific vs. hierarchical)  $\times$  2 (plasticity: all learning vs. 1 delayed group)  $\times$  2 (complexity: low vs. high) ANOVA, the effects were additive and consistent with the simulations above (e.g. there was a complexity  $\times$  PFC-hidden connectivity interaction, and there was a complexity  $\times$  PFC-BG connectivity interaction, but no three-way interaction, see Fig. 3). The additive nature of the two significant effects indicates that both types of constraints may produce a similar tendency to segregate the nature of the representations learned, albeit by different underlying mechanisms. However, we only tested the additive nature of these effects when the organizational structure was completely overlapping. That is, we only interrogated the situation in which the most removed groups of PFC units are also the exact same units that can govern the updating of the intermediately located units. Future anatomical, empirical, and computational study will be needed to discern whether these effects occur when the overlap may not be quite so clear-cut.

### 2.6.8. Stable representations

In addition to coding selective outer-loop representations, we also differentiated between selective, stable representations and selective, dynamic representations. Representations were considered stable if they were selective for a particular outer-loop context, and also did NOT vary as a function of the inner-loop cues and probes that occurred within that outer-loop context. Representations were considered dynamic if they did fluctuate as a function of the inner-loop information.

The first thing to note regarding this distinction is that it highlights two distinct encoding strategies the network could adopt. In one case, it could simply learn to encode direct mappings of the stimuli, and to maintain some representation of that stimulus information in a constant state while updating other, independent pieces of information. This strategy would be characterized by stable units. On the other hand, the network could adopt a more dynamic updating policy in which it encodes and updates each cue it encounters based on the outer-loop information that is currently in memory. In this way, it could store a conjunction of the information regarding the cue and probe, and as such, would not necessarily need to maintain the outer-loop information in an isolated format. All networks predominantly utilized the later pattern. For example, given a particular outer-level context within the baseline

network, 85–95% of active units were task selective, depending on the task, which is consistent with values across other networks. However, less than 1% of these task-selective units were also stable. The percentage of stable units went up to approximately 3% for the hierarchical network if we specifically probe the tasks requiring outer loop contexts, but despite this increase, the predominant pattern of task encoding was that of a dynamic attractor, where outer-loop information was maintained across multiple patterns of activity that also reflected the task cue.

### 3. Discussion

The simulations presented here provide us with novel data that illustrate several points regarding PFC function and its role in the performance of sequential working memory tasks. Consistent with neuroanatomical data (Petrides & Pandya, 2007) and recent computational modeling work (Botvinick, 2007; Paine & Tani, 2005), anatomical constraints that impose a hierarchical structure within PFC resulted in segregated representational structure, with the most distal areas of PFC having higher likelihoods of representing outer-loop information. Additionally, anatomical constraints on the interactions between PFC and BG also produced such representational structure. While we also investigated the hypothesis regarding the utility of differential learning schedules, these manipulations resulted in minimal impact on the representational structure of the network. For those networks that promoted segregated representations, there was a unique benefit to training time, as those networks that had the strongest representational structure were also those that trained the fastest. Additionally, we probed the nature of the representations within these tasks, and demonstrated that all networks predominantly represented the task demands in a dynamic attractor in which each state was coded in terms of conjunctions between the outer-loop context and most recent inner-loop cue.

#### 3.1. Behavior

Each of the networks learned to perform the relevant tasks, and to perform such tasks, they relied upon the use of active maintenance processes subserved by interactions between the PFC and BG. Behavior during test was well above chance. However, there were clear differences in the ability of the networks to perform the two different types of tasks, as the networks were better at performing the tasks not requiring additional outer-loop information. This is not surprising given behavioral data from other paradigms, as there is substantial evidence that performance degrades as the number of dimensions an individual has to integrate increases (Christoff et al., 2001; Halford et al., 1999; Kroger et al., 2002).

##### 3.1.1. Architectural constraints

Imposing an architectural constraint on the connectivity of PFC clearly had an effect on the types of representations that were developed and the degree to which the different types of representations were segregated across

different tasks. The experimental networks produced stable representations of the outer-loop information more frequently (albeit, still a very small overall proportion), and tended to place such representations in the more remote areas of PFC. Further, it appeared as though such segregation was associated with faster learning within the context of the manipulated networks. This finding is similar to findings in the hierarchical reinforcement learning domain, particularly the options framework discussed below (Botvinick, Niv, & Barto, 2008; Sutton, Precup, & Singh, 1999).

It should be pointed out that the relationship between training time and selectivity index was not symmetric around 0 for either of the networks that demonstrated such a relationship. If the key to fast training was representation segregation (meaning that outer context representations tended to be represented in one particular group while not being represented in the other), then the relationship between this score and time to criterion should be symmetric around 0. However, it is very clear that this is not the case (adding an additional squared-term to the linear model accounted for no additional variance,  $F(1,95) < 1$ ). In these networks, it clearly matters what representation is placed in what set of stripes. It is useful for the network to place outer-loop context in the remote areas of PFC because the extra step of processing insulates it from updating signals originating from the input, and thus makes it less likely to update (and see this in that these networks have a higher chance of producing stable representations). It is useful for the network to place the outer context information in the areas of PFC that provide direct input to the updating mechanism of the other stripes, because it can then provide more direct input on the updating of the other stripes, and in such a way, mirror the demands of the task. Alternately, if outer context is placed in the areas that are being modulated by other areas, then the gain from the ability to scope the updating signal is somewhat lost, as these are relatively stable representations (as far as the task demands are concerned). Likewise, if outer-loop information is placed in the proximal areas of PFC, then the remote areas of PFC are unable to be updated with the relevant, dynamically changing inner-loop information, because they only have access to the static outer-loop information.

##### 3.1.2. Plasticity manipulation

This plasticity manipulation did not result in segregated, hierarchical representations. However, there are a number of potential reasons for this null effect. Based on our previous control simulations, it is not entirely surprising that there was no bias to segregate representations, as within that particular architecture, there is no computational pressure to form such types of representations. For the architecture used, it provides no additional benefit in terms of learning or performance, particularly when compared to the alternate networks. That being said, the additional models run were not influenced by the developmental training procedure, suggesting that what computational pressure may exist in the other model architectures did not provide the developmental procedure a way to bootstrap stronger segregated representations.



### 3.2. Relationship to hierarchical reinforcement learning approaches and other computational models of hierarchy

Other researchers have utilized reinforcement learning to approach the question of hierarchy (for reviews, see Botvinick, 2007; Botvinick et al., 2008). We focus on the relationship between our approach and an alternate approach in which temporally abstracted actions, referred to as *options*, facilitate learning by providing subgoals that can be attained prior to the achievement of some more distant reward (Sutton et al., 1999). At first glance, the approaches are quite similar in that they both use reinforcement learning techniques in order to accomplish temporally distant goals. As such, the current set of simulations, and the PBWM framework more generally, could be recast in the options framework, such that maintained outer-loop context could be considered an identifier of a particular *option* in that it has a relatively long-term outcome, and is being used to make more myopic decisions on a trial-by-trial basis. The maintenance of such information and its influence on ongoing processing would then be analogous to the utilization of such an option's policy (Botvinick et al., 2008).

However, a very clear difference between the two approaches arises when investigating the nature and discovery of component actions (also called subgoals). One problem that the *options* approach encounters is that the subgoals need to be identified and compiled, typically prior to task learning. A number of mechanisms for doing such identification have been suggested, ranging from analyzing paths through problem space for relevant statistical structure (Pickett & Barto, 2002), to using intrinsic motivation as a possible mechanism (Singh, Barto, & Chentanez, 2005). This is a problem that, in many ways, the PBWM system has been able to solve (at least in the restricted environments that it has been placed). Within the PBWM framework, the appropriate options and subgoal states are learned via the mechanism that governs the updating system of PFC (O'Reilly & Frank, 2006); to the extent that the network discovers and segregates the types of information relevant for long-term performance, then it could be said that appropriate subgoals and options are acquired. Without segregation of the long-term information, it is analogous to a model attempting to solve a complex, temporally extended problem with RL, without the benefit of the options framework. This interpretation is consistent with the identified relationship between training time and degree of representational segregation.

Previous models exploring how hierarchical representations may develop (Botvinick, 2007; Paine & Tani, 2005) have produced consistent results with those produced here. Namely, architectural constraints tend to be sufficient to bias the network into developing representations that segregated as a function of content and temporal abstraction. This consistency is in spite of huge differences in the actual tasks and algorithms used to investigate the question. Whereas previous approaches have varied between the purely computational approach of using genetic algorithms (Paine & Tani, 2005) and recurrent backpropagation (Botvinick, 2007), the current approach is an advance in that it uses a biologically-inspired algorithm to ask similar questions, and demonstrates that hierarchical

representations can be selected using such an algorithm. In addition to these models, other models have been used to interrogate hierarchy without regards to the development or origin of the representations. Specifically, previous models have focused on understanding and distinguishing the potential computational roles of various areas of PFC (Koechlin et al., 2003; Koechlin & Hyafil, 2007; Koechlin & Summerfield, 2007; Reynolds & Mozer, 2009). The approach utilized by Koechlin argues that different areas of PFC are governed by different types of information conveyed by stimuli, and that such information may be understood in the context of a hierarchy, whereas the approach taken by Dayan (2008) and Reynolds and Mozer (2009) argues that no explicit hierarchy is needed to elicit hierarchical behavior. Although these approaches do not focus on how such representations may develop, they provide constraints for future models – the ability to develop representations consistent with one or another framework could prove to be a powerful tool in discriminating between the various models.

### 3.3. Stable representations and conjunctive representations

The number of units that coded for conjunctions of various stimuli (e.g. A in the context of a 1) was quite large, and it was interesting that there was such a strong bias for the network to develop such representations. Although somewhat surprising, it is consistent with a range of neurophysiology data in which conjunctive codes have been identified in a number of paradigms, including working memory (Barone & Joseph, 1990; Rao, Rainer, & Miller, 1997), and task-switching paradigms (Wallis, Anderson, & Miller, 2001). It is possible that the development of more stable, abstract representations requires more extensive training of a particular sort (Rougier, Noelle, Braver, Cohen, & O'Reilly, 2005), and that such training allows for task-independent representations that can be utilized when learning about more complex tasks. It is quite possible that the tasks used in the current set of simulations is not ideal for generating or creating hierarchical representations; specifically, there is not notion of a particular sequence of behaviors that can be learned, “chunked,” and then applied in a novel situation (see Botvinick & Plaut (2004); Reynolds, Zacks, & Braver (2007) for alternate paradigms with exactly such a structure). As such, it is somewhat surprising that these tasks and constraints produced a measurable difference in the representational structure of the network at all. Further investigation will have to be performed in order to determine whether such true abstractions (either in the form of sequences or in the form of dimensional extractions) produce more stable or more specialized representations.

Despite this endeavor being in its infancy, it makes a strong prediction regarding the role of different areas of PFC in the performance of this sequential working memory task and novel variants. First, it suggests that this task may differ substantially from other empirical tasks used to prove the hierarchical organization of behavior (Badre & D'Esposito, 2008; Koechlin et al., 2003). The relatively small number of units that were statically coding for outer-loop context, combined with the large number of units

that were multiplexing outer-loop information and inner-loop information suggests that one may find much stronger dynamics on trial-by-trial effects, rather than in sustained maintenance across trials (Reynolds, 2005). Second, this bias towards dynamic attractors begs the question of what could increase the probability of forming a more static attractor? Although there are many possibilities, it seems like one potentially important and relevant variable is the timing between the inner-loop cue and probe information. Counter-intuitively, if you reduce this delay, there may not be enough time to re-encode the inner-loop cue as a conjunction with the outer-loop context. As such, manipulating this timing parameter may be sufficient to increase the maintenance of the outer-loop stimulus, and may account for differences in empirical investigations of this task (Reynolds, 2005) and other tasks where the outer-most piece of information is the only item presented prior to a multi-dimensional imperative stimulus (Badre & D'Esposito, 2008; Koehlin et al., 2003). Ongoing simulations and experiments are currently investigating this question.

Although the developmental plasticity hypothesis did not play out as predicted, there is still growing evidence that different regions of PFC mature at different rates (Brown et al., 2005; Shaw et al., 2008; Sowell et al., 1999), and there is likely some functional consequence of this (Bunge & Zelazo, 2006). While previous experimental approaches have suggested that there is a posterior–anterior gradient in terms of either rule complexity or rule abstraction (Badre & D'Esposito, 2008), recent data have suggested that the PFC may not develop in a strict posterior-to-anterior gradient. Specifically, it appears that posterior and anterior areas of PFC are the first to develop, with the areas in between developing later (Gogtay et al., 2004; Shaw et al., 2008). Additional investigation will be needed to understand how the differences in maturation schedule influence the nature of representation, and how these developmental trajectories may relate to current theories of the organization of PFC.

Despite our best intentions to concretely define what is meant by the maintenance of outer-loop information, there are several different variables that may influence whether information is considered outer-loop by the cognitive and neural system. For example, there are at least two differences between outer-loop information and inner-loop information in the current model, both of which are consistent with the notion of anterior PFC playing a high-level role within a hierarchy of goals. The first is that outer-loop information must be maintained while something else is updated. The second is that outer-loop information must be maintained for longer periods of time than inner-loop information. Finally, it is very possible that the cognitive system considers information to be outer-loop only when both of these constraints are met. The nature of the distinction between outer- and inner-loop information is a key question that is being investigated currently. Despite the potential concerns and limitations, the current set of simulations has allowed for a concrete, implemented definition regarding one particular dimension across which PFC may be organized, and provided a

biologically-inspired mechanism by which such an organization could develop.

The current set of simulations did not interrogate all of the potential dimensions of PFC organization. Rather, the current set of simulations should be viewed as an early step at using computational modeling techniques in order to probe the nature of such representations and differences between them. The current models could be extended in numerous ways to capture other potential dimensions. For example, there is growing evidence suggesting that left inferior PFC can be subdivided into regions that are differentially sensitive to semantic or phonological properties of stimuli (with more anterior areas of left inferior PFC being more closely associated with semantics; Poldrack et al., 1999). This could be viewed along a similar axis of abstraction as that investigated here, such that the fine grained muscle movements necessary to create utterances are required before verbal communication can develop. These muscle representations are likely used in the perception and comprehension of language (e.g. phonology; Rizzolatti & Arbib, 1998; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Once such a basis of movements is learned and used to, novel combinations of such movements can be generated and associated with an abstraction that represents the semantics of such a motion in a many-to-many mapping.

### 3.4. Conclusions

The current set of simulations provides convergent evidence that different types of representations within PFC can be developed by having specific kinds of architectural constraints. The segregation of these representations leads to faster learning, and as such, may provide some evidence for the computational pressures that govern the organization of PFC.

### Acknowledgements

The authors would like to thank Nicole Speer, Thomas Hazy, Seth Herd, and the rest of the Computational Cognitive Neuroscience Laboratory for helpful comments and suggestions. This research was supported in part by an NRSA post-doctoral training grant (1 F32 MH075300-01A2) and NIH R01 MH069597-01 (awarded to ROR).

### Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.cognition.2009.05.015](https://doi.org/10.1016/j.cognition.2009.05.015).

### References

- Badre, D., & D'Esposito, M. (2008). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Journal of Cognitive Neuroscience*, 19.
- Barone, P., & Joseph, J. P. (1990). Prefrontal cortex and spatial sequencing in macaque monkey. *Experimental Brain Research*, 78, 447–464.
- Botvinick, M., Niv, Y., & Barto, A. C. (2008). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition*, 113(3), 262–280.

- Botvinick, M., & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review*, *111*, 395–429.
- Botvinick, M. M. (2007). Multilevel structure in behaviour and in the brain: A model of fuster's hierarchy. *Philosophical Transactions of the Royal Society of London Series B – Biological Sciences*, *362*(1485), 1615–1626.
- Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences*, *12*(11), 201–208.
- Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., et al. (2001). Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology*, *130*, 746–763.
- Braver, T. S., & Bongiolatti, S. R. (2002). The role of frontopolar cortex in subgoal processing during working memory. *Neuroimage*, *15*, 523–536.
- Braver, T. S., Cohen, J. D., & Barch, D. M. (2002). The role of the prefrontal cortex in normal and disordered cognitive control: A cognitive neuroscience perspective. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 428–448). Oxford: Oxford University Press.
- Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron*, *39*, 713–726.
- Brown, T. T., Lugar, H. M., Coalson, R. S., Miezin, F. M., Petersen, S. E., & Schlaggar, B. L. (2005). Developmental changes in human cerebral functional organization for word generation. *Cerebral Cortex (New York, NY)*, *15*, 275–290.
- Bunge, S. A., & Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science*, *14*(3), 118–121.
- Christoff, K., & Gabrieli, J. D. E. (2000). The frontopolar cortex and human cognition: Evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychobiology*, *28*, 168–186.
- Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., et al. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *Neuroimage*, *14*(5), 1136–1119.
- Christoff, K., Ream, J. M., Geddes, L. P. T., & Gabrieli, J. D. E. (2003). Evaluating self-generated information: Anterior prefrontal contributions to human cognition. *Behavioral Neuroscience*, *117*(6), 1161–1168.
- Dayan, P. (2008). Bilinearity, rules, and prefrontal cortex. *Frontiers in Computational Neuroscience*, *1*(1).
- D'Esposito, M., Postle, B. R., Ballard, D., & Lease, J. (1999). Maintenance versus manipulation of information held in working memory: An event-related fmri study. *Brain and Cognition*, *41*, 66–86.
- Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*, 475–482.
- Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between the frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, and Behavioral Neuroscience*, *1*, 137–160.
- Fuster, J. M. (1991). Prefrontal cortex and the bridging of temporal gaps in the perception–action cycle. In A. Diamond (Ed.), *The development and neural bases of higher cognitive functions* (Vol. 608, pp. 318–336). New York: New York Academy of Science Press.
- Fuster, J. M. (2004). Upper processing stages of the perception–action cycle. *Trends in Cognitive Sciences*, *8*(4), 143–145.
- Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(21), 8174–8179.
- Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. *Handbook of Physiology – The Nervous System*, *5*, 373–417.
- Halford, G. S. (1984). Can young children integrate premises in transitivity and serial order tasks? *Cognitive Psychology*, *16*, 65–93.
- Halford, G. S. (1993). *Children's understanding: The development of mental models*. Hillsdale, NJ: Erlbaum.
- Halford, G. S., Wilson, W. H., & Phillips, S. (1999). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *The Behavioral and Brain Sciences*, *21*, 803.
- Haxby, J. V., Petit, L., Ungerleider, L. G., & Courtney, S. M. (2000). Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. *Neuroimage*, *11*, 380–391.
- Johnson, M. K., Raye, C. L., Mitchell, K. J., Greene, E. J., & Anderson, A. W. (2003). Fmri evidence for an organization of prefrontal cortex by both type of process and type of information. *Cerebral Cortex (New York, NY)*, *13*, 265–273.
- Koechlin, E., Basso, G., Pietrini, P., Panzer, S., & Grafman, J. (1999). The role of the anterior prefrontal cortex in human cognition. *Nature*, *399*, 148–151.
- Koechlin, E., & Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science (New York, NY)*, *318*(5850), 594–598.
- Koechlin, E., Ody, C., & Kouneiher, F. (2003). Neuroscience: The architecture of cognitive control in the human prefrontal cortex. *Science*, *424*, 1181–1184.
- Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, *11*(6), 229–235.
- Kroger, J. K., Sabb, F. W., Fales, C. L., Bookheimer, S. Y., Cohen, M. S., & Holyoak, K. J. (2002). Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: A parametric study of relational complexity. *Cerebral Cortex (New York, NY)*, *12*, 477–485.
- Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior: The hixon symposium* (pp. 112–136). New York: Wiley.
- Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, *54*, 217–224.
- Miller, E. K. (2000). The prefrontal cortex: No simple matter. *Neuroimage*, *11*, 447–450.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.
- Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt.
- Nystrom, L. E., Braver, T. S., Sabb, F. W., Delgado, M. R., Noll, D. C., & Cohen, J. D. (2000). Working memory for letters, shapes, and locations: Fmri evidence against stimulus-based regional organization in human prefrontal cortex. *Neuroimage*, *11*, 424–446.
- O'Donnell, S., Noseworth, M., Levine, B., & Dennis, M. (2005). Cortical thickness of the frontopolar area in typically developing children and adolescents. *Neuroimage*, *24*(4), 948–954.
- O'Reilly, R. C. (1996). Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Computation*, *8*(5), 895–938.
- O'Reilly, R. C. (1998). Six principles for biologically-based computational models of cortical cognition. *Trends in Cognitive Sciences*, *2*(11), 455–462.
- O'Reilly, R. C. (2001). Generalization in interactive networks: The benefits of inhibitory competition and Hebbian learning. *Neural Computation*, *13*, 1199–1242.
- 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*, 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*, 31–49.
- O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. Cambridge, MA: MIT Press.
- Paine, R. W., & Tani, J. (2005). How hierarchical control self-organizes in artificial adaptive systems. *Adaptive Behavior*, *13*(3), 211–225. doi:10.1177/105971230501300303.
- Petrides, M. (2000). The role of the mid-dorsolateral prefrontal cortex in working memory. *Experimental Brain Research*, *133*, 44.
- Petrides, M., & Pandya, D. N. (2007). Efferent association pathways from the rostral prefrontal cortex in the macaque monkey. *The Journal of Neuroscience*, *27*(43), 11573–11586.
- Pickett, M., & Barto, A. (2002). Policyblocks: An algorithm for creating useful macro-actions in reinforcement learning. In *ICML'02: Proceedings of the nineteenth international conference on machine learning* (pp. 506–513). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc..
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage*, *10*, 15–35.
- Rao, S. C., Rainer, G., & Miller, E. K. (1997). Integration of what and where in the primate prefrontal cortex. *Science*, *276*(5313), 821–824.
- Raye, C. L., Johnson, M. K., Mitchell, K. J., Reeder, J. A., & Greene, E. J. (2002). Neuroimaging a single thought: Dorsolateral pfc activity associated with refreshing just-activated information. *Neuroimage*, *15*, 447–453.

- Reynolds, J., West, R., & Braver, T. (2008). Distinct neural circuits support transient and sustained processes in prospective memory and working memory. *Cerebral Cortex (New York, NY)*. doi:10.1093/cercor/bhn164.
- Reynolds, J., Zacks, J., & Braver, T. (2007). A computational model of event segmentation from perceptual prediction. *Cognitive Science*, 31, 613–634.
- Reynolds, J. R. (2005). On the roles of duration and computational complexity in the recruitment of fronto-polar prefrontal cortex. Ph.D. Thesis. Washington University in St. Louis.
- Reynolds, J. R., & Mozer, M. C. (2009). Temporal dynamics of cognitive control. *Advances in Neural Information Processing Systems*.
- Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21, 188–194.
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3, 131–141.
- Robin, N., & Holyoak, K. J. (1995). Relational complexity and the functions of prefrontal cortex. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (1st ed., pp. 987–997). Cambridge, MA: MIT Press.
- Rougier, N. P., Noelle, D., Braver, T. S., Cohen, J. D., & O'Reilly, R. C. (2005). Prefrontal cortex and the flexibility of cognitive control: Rules without symbols. *Proceedings of the National Academy of Sciences*, 102(20), 7338–7343.
- Shaw, P., Kabani, N., Lerch, J., Eckstrand, K., Lenroot, R., Gogtay, N., et al. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *Journal of Neuroscience*, 28, 3586–3594.
- Singh, S., Barto, A., Chentanez, N. (2005). Intrinsically motivated reinforcement learning. In L. K. Saul, Y. Weiss, & L. Bottou (Eds.), *Advances in neural information processing systems 17: Proceedings of the 2004 conference* (pp. 1281–1288). Cambridge: MIT Press.
- Sowell, E. R., Thompson, P. M., Holmes, C. J., Jernigan, T. L., & Toga, A. W. (1999). In vivo evidence for post-adolescent brain maturation in frontal and striatal regions. *Nature Neuroscience*, 2(10), 859–861.
- Stuss, D. T., & Knight, R. T. (Eds.). (2002). *Principles of frontal lobe function*. New York: Oxford University Press.
- Sutton, R., Precup, D., & Singh, S. (1999). Between mdps and semi-mdps: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence*, 112(1–2), 181–211.
- Wallis, J. D., Anderson, K. C., & Miller, E. K. (2001). Single neurons in prefrontal cortex encode abstract rules. *Nature*, 411, 953–956.
- Wilson, F. A., Scalaidhe, S. P., & Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science (New York, NY)*, 260, 1955–1957.
- Wood, J. N., & Grafman, J. (2003). Human prefrontal cortex: Processing and representational perspectives. *Nature Reviews*, 4, 139–147.