From an Executive Network to Executive Control: A Computational Model of the *n*-back Task

Christopher H. Chatham, Seth A. Herd, Angela M. Brant,
Thomas E. Hazy, Akira Miyake, Randy O'Reilly,
and Naomi P. Friedman

Abstract

■ A paradigmatic test of executive control, the *n*-back task, is known to recruit a widely distributed parietal, frontal, and striatal "executive network," and is thought to require an equally wide array of executive functions. The mapping of functions onto substrates in such a complex task presents a significant challenge to any theoretical framework for executive control. To address this challenge, we developed a biologically constrained model of the *n*-back task that emergently develops the ability to appropriately gate, bind, and maintain information in working memory in the course of learning to perform the task. Furthermore, the model is sensitive to proactive interference in ways that match findings from neuroimaging and shows a U-shaped performance curve after manipulation of prefrontal dopaminergic mechanisms similar to that observed in studies of genetic polymorphisms and pharmacological manipulations. Our model represents a formal computational link between anatomical, functional neuroimaging, genetic, behavioral, and theoretical levels of analysis in the study of executive control. In addition, the model specifies one way in which the pFC, BG, parietal, and sensory cortices may learn to cooperate and give rise to executive control. ■

INTRODUCTION

Goal-directed behaviors are enabled by executive functions that help stop prepotent responses, resolve interference, update working memory, shift mental sets, and coordinate multiple tasks (e.g., Friedman & Miyake, 2004; Logie, Cocchini, Delia Sala, & Baddeley, 2004; Salthouse, Atkinson, & Berish, 2003; Miyake et al., 2000). Such broad categories of executive function can be fractionated into lower-level component processes. For example, working memory updating tasks require storing information, gating information into and out of working memory, tracking serial order, and selective attention. These processes may in turn be mapped to diverse parietal, frontal, and striatal substrates (e.g., Wager & Smith, 2003), posing a many-to-many problem in mapping executive functions to their neural substrates. A paradigmatic example of this many-to-many mapping problem is the *n*-back task (e.g., Kirchner, 1958). The main purpose of this article is to elucidate the mechanistic basis of this complex task using a biologically constrained computational model.

The *n*-back Task

In the *n*-back task, subjects identify over consecutive trials whether the current stimulus matches a stimulus presented in *n* trials previously. At the cognitive level, this task is thought to involve numerous executive processes: active maintenance of the last *n* items; updating of new items so that they can be actively maintained; rapid binding of items to their serial order so that responses are based on the match between the current item and the *n*-back item and not between items matching at a non-*n* lag; and resolution of any proactive interference arising from non-*n* lag items. At the biological level, neuroimaging, pharmacological, and genetic polymorphism studies indicate that *n*-back performance is associated with a distributed network of parietal, frontal, and striatal sites (Tsuchida & Fellows, 2009; Owen, McMillan, Laird, & Bullmore, 2005; Olesen, Westerberg, & Klingberg, 2003) and dopaminergic mechanisms (Apud & Weinberger, 2007; Tan et al., 2007; Meyer-Lindenberg et al., 2006; Aalto, Brück, Laine, Någren, & Rinne, 2005; Goldberg et al., 2003; Mattay et al., 2003; Egan et al., 2001). This cognitive and neurobiological complexity makes the *n*-back task a useful test case for formal accounts of how executive functions arise from their neural substrates.

One feature of the *n*-back task makes it especially appropriate for this undertaking: It appears to require rapid binding of stimuli to representations of their serial order. Symbolic cognitive models (e.g., ACT-R) fulfill this requirement through the use of propositional representations and explicit variables and have yielded a working *n*-back model (Juvina & Taatgen, 2007). However, the brain operates on the basis of distributed representations and slowly adapting synaptic connections. The difficulty in reconciling

University of Colorado

this distributed, slowly adapting neural substrate with the n-back's rapid binding requirements could explain the absence of more biologically constrained models of this task. Here we present a model that overcomes this challenge and is capable of learning the n-back without a mechanism specifically implemented for symbolic binding.

Model Architecture

Our model is rooted in the biologically plausible prefrontal BG working memory (PBWM) architecture (Hazy, Frank, & O'Reilly, 2006, 2007, 2010; O'Reilly & Frank, 2006). PBWM's essential principle is that task-relevant information can be maintained in pFC and help guide successful task performance by a process of biased competition (Desimone & Duncan, 1995); the reward signals resulting from successful task performance, in the form of phasic dopamine (DA), can then train the BG through reinforcement learning to send an "updating signal" for gating new information into pFC. PBWM, thus, integrates numerous ideas from computational neuroscience, implementing reinforcement learning in terms of phasic DA via the primary value–learned value (PVLV) mechanism (Hazy et al., 2010; O'Reilly & Frank, 2006), resolving the stability–flexibility dilemma (Goschke, 2000) with flexible gating mechanisms and yielding biased competition via prefrontal representations stabilized through recurrent connectivity and tonic DA (e.g., Cohen, Braver, & Brown, 2002). The architecture supporting these interactions is schematically illustrated in Figure 1.

Our implementation of PBWM, depicted in Figure 2, is most closely based on the PBWM model of the phonological loop (O'Reilly & Frank, 2006). As in that previous work, the model receives input from a layer in which 1 of 10 different units is activated on every trial of the task, each unit corresponding to a different stimulus. The network's response on every trial is indicated by the patterns of activa-

Figure 2. (A) The PBWM architecture includes units based on the pFC and BG, including ventral and dorsal striatum, grouped into "stripes" (the visible subgroups within prefrontal and striatal layers). Input is provided to the model about the identity of the current stimulus and its serial order; the model is required to produce a manual output about whether the current stimulus matches that presented n trials previously and a verbal output corresponding to the identity of the stimulus presented n trials previously. (B) The parietal layers represent the serial order of successive trials in terms of n, using a graded and compressive code based on the mean and variance observed in the tuning curves of rank order sensitive neurons in the horizontal segment of the IPS.

tion across two output layers: a "verbal output" layer with 10 units, corresponding to each of the input stimuli indicating the network's best guess as to the n-back stimulus, and a "manual output" layer with 2 units, corresponding to match and nonmatch responses, indicating the network's best guess as to whether the current stimulus matches the n-back stimulus. (Some n-back tasks require subjects only to indicate whether there is a match between the current and n-back stimulus and not the actual identity of the n-back stimulus; we included both output requirements in our model because subjects are likely to keep the identity of the n-back stimulus identity in memory regardless of the precise variant of n-back they are performing.) Finally, a "posterior cortex" layer of 100 units is bidirectionally connected with each of these layers and provides a substrate for biased competition to take place. We refrain from identifying this layer with a particular neocortical area, as it contains no special mechanisms that might be thought to differentiate it from many areas of neocortex.

Superimposed on this structure are the core components of PBWM. These components include prefrontal layers organized into stripes, consistent with the functional macrocolumns observed in the monkey pFC (e.g., Rao, Williams, & Goldman-Rakic, 1999; Pucak, Levitt, Lund, & Lewis, 1996; Levitt, Lewis, Yoshioka, & Lund, 1993). The units constituting these stripes are unique relative to all other units in two ways: They are recurrently self-connected,

Figure 1. Schematic illustration of core PBWM architecture, in which prefrontal context representations of relevant prior information and current goals bias the sensory motor mappings that are learned by posterior cortical "hidden" layers. The prefrontal context representations are updated via dynamic gating by the BG. These gating functions are learned by the BG on the basis of input from the PVLV system, which provides modulatory dopaminergic input depending on the reward value of the actions performed by the BG.

and they contain an excitatory hysteresis current. When combined, these features enable persistent, self-sustaining patterns of activity. The resulting temporally stable patterns of activity are gated on a stripe-specific basis by a set of corresponding stripes in a BG "matrix" layer, modeled after the medium spiny projection neurons of the striatal matrix. Each stripe in the matrix (stripes are represented as visible subgroups within the BG layers in Figure 2) contains go and no-go units that are spatially intermixed (as they are biologically). Go units correspond to the direct pathway of the striatum, and no-go units correspond to the indirect pathway. As such, go units have a disinhibitory effect on corticothalamic gating, thereby allowing working memory to be updated; no-go units have an inhibitory effect on corticothalamic gating, thereby helping to keep the contents of working memory the same despite new incoming information.

Learning Algorithms in the Model

Central to the PBWM architecture is the use of the PVLV algorithm, which can be seen as a biologically plausible implementation of traditional temporal difference reinforcement learning (Hazy et al., 2010; O'Reilly, Frank, Hazy, & Watz, 2007). The PVLV algorithm is used specifically and selectively to train the go and no-go units of the striatum. Ultimately, PVLV trains go units to fire in response to stimuli that predict reward (and which might therefore be updated into working memory), whereas no-go units learn to fire when stimuli do not predict anything more rewarding than the information currently represented in working memory. In conjunction with the prefrontal layers, PBWM implements mechanisms that at a higher level of analysis can enable basic executive functions like active maintenance and gating (e.g., O'Reilly & Frank, 2006).[1]

The other components of the model are all trained with a standard Hebbian learning rule and an error-driven learning rule (O'Reilly & Munakata, 2000). The end result of this combination of learning rules and PVLV is that, by the end of training, networks learned to fire primarily go units in certain stripes in the BG, such that the particular stripes activated depend on the activity patterns in other layers. This stripe-specific go firing within the BG updates corresponding stripes in pFC with information currently present in the input layer. BG stripes that are not used for a given trial fire primarily no-go units, resulting in the preserved maintenance of information from preceding trials. Finally, pFC activity representing this important maintained information biases the posterior layer, which in turn biases the verbal and manual output layers. These connection weights are incrementally refined via Hebbian and error-driven learning so that they are most likely to produce the correct verbal and manual outputs (see Appendices I and II for additional details).

Serial Order Representations of the Model

Interestingly, the model autonomously learns to take advantage of the stripe-specific gating possible within the PBWM

architecture so as to solve to the variable binding problem posed by the n-back task (Juvina & Taatgen, 2007). With training, the BG sends an increasingly differentiated gating signal such that the pFC can learn to maintain items in different stripes conditional on their serial order. This increased specificity of gating enables distribution of the task's mnemonic demands across multiple stripes and solves the rapid binding problem by allowing the model to autonomously bind representations of items to their serial order (e.g., O'Reilly, Busby, & Soto, 2003).

One crucial addition to this standard PBWM architecture is the parietal layer, which represents the serial order of successive stimuli using a graded and compressive code (in which representations are distributed and increasingly similar to one another as the serial order of the current stimulus increases; depicted in Figure 2B for serial orders 1, 2, and 3, respectively). The localization of such a serial order representation to parietal cortex is consistent with previous models (Botvinick & Watanabe, 2007; Botvinick & Plaut, 2006), with electrophysiology and neuroimaging of serial order representation in the intraparietal sulcus (IPS; Marshuetz, Reuter-Lorenz, Smith, Jonides, & Noll, 2006; Nieder, Diester, & Tudusciuc, 2006; Marshuetz, Smith, Jonides, Degutis, & Chenevert, 2000) and with the IPS activity observed across an n-back meta-analysis (Owen et al., 2005). Recent evidence suggests that working memory contents are encoded as a function of their ordinal position in the sequence of to-be-remembered items (Van Dijck & Fias, 2011), consistent with our use of a parietally based serial order mechanism to satisfy the working memory updating demands of the n-back task. Thus, the serial order representations in our model are different from the representations expected to support processing of other attributes (e.g., color or shape), in that they are explicitly based on the known tuning curves of neurons coding for serial order in the IPS.

Importantly, we implemented serial order representations not as a continuous number line that stretches to the number of trials, but as a periodic repeat of item positions. For example, in the 2-back task, the serial order representations alternate between 1 and 2, whereas in the 3-back task, they repeatedly cycle through 1, 2, and 3. This periodicity of the serial order representations is imposed by fiat or "prescribed." Although we return to this issue in the discussion, it is a difficult and outstanding problem of how such serial order representations and their dynamics might be learned. We abstract over this difficulty here. This nonetheless leaves much to be solved: The model must still autonomously learn that these serial order representations are important to bind them to the items presented on each trial and to update and maintain this information appropriately.

Organization of the Current Paper

The results of our simulations are outlined as follows. After describing the details of the model and the way in which

the *n*-back task and its instructions were presented to the network, we demonstrate the capacity of the network to replicate hallmark findings from the *n*-back literature, spanning multiple levels of analysis (behavioral, hemodynamic, and genetic). We then quantitatively analyze the model's prefrontal and striatal functioning to support an expository description of the model's functioning. We next discuss how our simulations inform cognitive theorizing about how executive functions like active maintenance, gating, and the resolution of proactive interference may emerge from a highly interactive fronto-parieto-striatal circuit. Finally, we describe how our model provides not only a computationally explicit example of the prefrontal–parietal interactions commonly observed in the considerable neuroimaging literature on this and other executive control tasks (de Frias et al., 2010; Tsuchida & Fellows, 2009; Owen et al., 2005; Egan et al., 2003; Olesen et al., 2003) but also how it leads to new theoretical insights and untested predictions.

METHODS

Implementation

To illustrate the biologically constrained nature of our model, here we briefly review the Leabra framework (O'Reilly, 2001). This framework simulates neural processing in terms of interconnected units, each of which has a membrane potential determined by separate excitatory, inhibitory, and leak conductances. Fluctuations in the resulting membrane potential are thresholded and transformed to yield a rate-coded output that contributes to the excitatory conductance of all other units to which a particular unit is connected in proportion to the connection weight. Connection and bias weights are initially randomized but are shaped over the course of training according to Hebbian, reward-driven, and biologically realistic error-driven learning rules (see below). Units are grouped into layers that undergo a *k*-winners-take-all (kWTA) function for simulating the influence of local inhibitory interneurons. These biologically inspired mechanisms have been used in over 40 models to capture a variety of detailed phenomena (e.g., O'Reilly & Munakata, 2000), indicating that these simple biological mechanisms can yield human-like performance in a number of domains.

In addition to the PBWM implementation (see Appendix I, and depiction in Figure 2A), sequential order was represented via the scaled log-normal function (Botvinick & Watanabe, 2007):

$$R_\rho(r) = \exp\left(-\frac{(\ln r - \ln \rho)^2}{2\sigma^2}\right)$$

wherein $R_\rho(r)$ is the activation level of the ρth unit in the layer on a trial with rank r and σ is a parameter determining the relative specificity of each unit to its preferred rank. We

elaborated on this scheme by convolving the activation levels specified by this scaled-log normal function with Gaussian variance conforming to that empirically observed in the IPS (Nieder et al., 2006); this convolution is included here as an additional biological constraint on the parietal layer and can be observed as "noise" in the activation dynamics depicted in the parietal layer of Figure 2B. In the model, this parietal layer is interconnected both with the pFC and (strongly) with the BG, consistent with the known anatomy of humans and other primates (Fernández-Miranda et al., 2008; Yeterian & Pandya, 1993) and with the functionally interconnected parieto-fronto-striatal network commonly observed in neuroimaging studies of the executive functions.

Each named layer of the model contains features that uniquely associate its layers with the identified brain regions. For example, prefrontal layers are unique because of recurrent connections and an excitatory hysteresis current, as well as the stripe organization connected with a parallel stripe organization in striatal layers; parietal layers are unique because of the graded and compressive activation dynamics imposed there; striatal layers are unique because of their DA-driven reinforcement learning. The posterior layer is distinct because it contains none of the unique features above, but only the more general mechanisms implemented by Leabra and thought to apply to neocortex in general. Moreover, the connectivity among these layers is based on known neurobiology (Hazy et al., 2006, 2007, 2010; O'Reilly & Frank, 2006).

Training and Testing

All models were run in batches of 25 networks, and each network was initialized with random patterns of connection weights. To compare performance on the 2- and 3-back tasks, we employed networks with 12 pFC stripes so that the same networks were capable of learning both tasks, as the 3-back task seemed to require more working memory "capacity" than the 2-back task. For all other analyses, we used a scaled-down model consisting of only six stripes, both to speed training time and make detailed analyses of network behavior more tractable.

Training on the 2- and 3-back tasks consisted of activating 1 of 10 possible input units and the corresponding distributed representation of serial order in the parietal layer (each trial corresponds to one of the three serial orders illustrated in Figure 2B). "Lure" trials, in which the current stimulus matched a previous stimulus at a non-*n* lag, were allowed to occur. "Recent" lure trials are those where the current stimulus matches the $n - 1$ stimulus; "Nonrecent" lure trials are those where the current stimulus matches a preceding stimulus with a lag larger than *n*.

Human subjects are instructed on the value of *n* for each *n*-back task they perform. In our simulations, the network was informed of the value of *n* by way of a small, probabilistic bias to replace stimuli occurring at values of *n*. This bias was implemented in the 2-back task by increasing the

Figure 3. A schematic example of a trained model's inputs, outputs, and "hidden" layer activations on the 2-back task. Trial 1: The network is presented with the input A and a parietal representation corresponding to serial order 1. The three leftmost units for the striatum have learned to fire on trials with this serial order and, therefore, gate the stimulus "A" into the corresponding units in pFC, which has learned to represent "A." This conjunction of the item "A" in the stripe that has learned to represent information from serial order "1" produces a bound representation that can be termed "A1." Finally, the network has learned to produce the verbal output corresponding to the 2-back item (not applicable [n/a] here, because this is the first trial), and the manual output corresponding to "nonmatch," because the current item does not match the item presented 2-back. Trial 2: The network is presented with input "D" and serial order 2 is represented in the parietal layer; the right most units in the striatum fire for this serial order, and therefore gate the stimulus "D" into the corresponding pFC units, producing a bound representation that can be termed "D2." Trials 3–6: New stimuli are presented, the parietal layer continues to count off the serial order of the current stimulus, and the striatal layer continues to fire at the appropriate times, thereby updating pFC with the current stimulus in the correct set of units. The network produces nonmatch responses for all trials except Trial 4, which is a match trial.

activity level of the go units in the matrix layer on a random 10% of the trials in which they had not fired on the previous trial. Similarly, in the 3-back task, the activity level of those units was increased on a random 10% of the trials in which they had not fired on the previous two trials. This probabilistic bias yields a proportion of trials in which the pFC is updated with a periodicity of n. For this bias to yield good performance, the network must not only perform correctly on the few trials where this probabilistic updating occurs but must also generalize that behavior across all trials and stimuli.

For testing, the patterns of activity in the verbal and manual output layers were recorded after those activity patterns had stabilized or a maximum number of cycles had occurred (here we use the Leabra default of 60 cycles). The most active output unit was considered the network's response, and this output was compared with the correct output for computing the error statistics described in Results. Networks were trained in epochs of 500 trials each until the network was tested to perform above 80% correct in terms of both its verbal and manual outputs for seven consecutive epochs. This performance criterion allows networks to develop individual differences in the range of those observed in humans: Some networks will perform substantially better than 80% correct by the end of training, whereas others may have a shallower learning curve. For all analyses except those pertaining to learning across the entire course of training, network behavior is tested during the final 10% of training.

For individual differences analyses, three batches of 25 networks were run with variations in the gain of prefrontal units (a proxy for tonic prefrontal DA) but the same 25

random seeds were used to initialize weights across each batch to ensure comparability across model runs. Generalization was assessed in terms of the verbal responses in a distinct batch of 25 networks on a randomly selected set of 10 trial sequences; these 10 trial sequences had been entirely omitted from the training set. For example, the sequence A1X2B1 might have been excluded from the training set for the 2-back network, where the intervening "X" stimulus could have been any of the possible stimuli.

The activation dynamic resulting from training is schematically illustrated in Figure 3 for the 2-back task. The first trial is a nonmatch trial with input stimulus "A" and serial order "1" (in Results, this type of trial is represented with the phrase "A1"). A subset of BG stripes fire (the leftmost three BG units in Figure 3), resulting in maintenance of stimulus "A" within a corresponding subset of pFC stripes (the left-most three pFC units in Figure 3). On the following trial, a different subset of BG stripes fire, resulting in the maintenance of the next stimulus ("D") within the corresponding new subset of pFC stripes. This two-part activation dynamic repeats across all subsequent trials but is illustrated for several trials in Figure 3 for clarity, including a recent lure trial, a non-recent lure trial, and a match trial. Three-part activation dynamics emerge in networks trained to perform the 3-back task, such that pFC, parietal, and BG layers have three distinct activation states (as opposed to the two distinct states illustrated in Figure 3). The only remaining difference in 3-back is that the correct verbal and manual outputs correspond to matches between the item presented currently and that presented *three* trials previously in the 3-back.

RESULTS AND DISCUSSION

The Model Captures Benchmark Findings in the *n*-back Literature

Our model was capable of de novo learning of both the 2-back and 3-back tasks, without an underlying symbolic variable system for performing rapid binding. This learning was not rote, in that all networks generalized to untrained sequences at a rate significantly above chance $t(1, 24) = 17.9, p < .001$ for 2-back and $t(1, 24) = 12.9, p < .005$ for 3-back. As described below, the model also captured numerous benchmark features of human performance in the *n*-back task.

One hallmark finding in the *n*-back literature is reduced accuracy as *n* increases from 2 to 3. The model also showed this pattern, such that 2-back accuracy was higher than 3-back accuracy, $F(1, 24) = 10.54, p < .005$, as shown in Figure 4A. This result arises from two features of the *n*-back: Relative to 2-back, 3-back requires an additional item be maintained by the prefrontal layers; also, 3-back involves a less reliable signal of a current item's serial order, owing to the logarithmic compression of the parietal layer. These two constraints jointly produce lower performance on (and also slower learning of) the 3-back task, because they diminish the ability of the network to appropriately bind an item to its serial order and to maintain this binding over subsequent trials.

A second benchmark finding in the *n*-back literature is that performance is sensitive to the presence of lures—items that match a preceding item but not at the critical *n* lag. The model also captures this phenomenon, such that accuracy was significantly lower for recent lures than non-recent lures (Figure 4B) in both the 2-back, $F(1, 24) = 77.2, p < .001$, and the 3-back, $F(1, 24) = 15.8, p = .001$. This effect reflects interference caused by items in the input, which match items maintained in memory, albeit with a different temporal order, thereby yielding a tendency for the network to inappropriately detect a match on lure trials. Moreover, accuracy is particularly low on recent lure trials $(n - 1)$, reflecting proactive interference, because the prefrontal layers are more likely to represent items with lags

less than *n* than items with lags greater than *n* (the latter are task irrelevant); thus, the network is more prone to erroneously detect matches in the former case.

One counterintuitive result from the *n*-back literature is that the effect of $n - 1$ lures, relative to the effect of nonrecent lures (i.e., lures at positions $> n$), is reduced as *n* moves from 2- to 3-back (Oberauer, 2005). Although this effect is counterintuitive—one might expect that the cost of lure trials on accuracy would increase proportionally with overall difficulty—the model reproduced the observed result (Figure 4B; $F(1, 24) = 18.14, p < .001$). Consistent with the model's functioning, this effect reflects the fact that proactive interference arising from a match between the current item and maintained items is diluted when more items are being simultaneously maintained, as in the 3-back task.

Neuroimaging studies of this kind of proactive interference reveal a larger hemodynamic response in the lateral pFC to recent relative to nonrecent lures (Jonides & Nee, 2006; Badre & Wagner, 2006; Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998). The hemodynamic response is thought to reflect metabolic demands; furthermore, 50%–80% of the brain's energy consumption reflects the input and output activity of its neurons (Buzsáki, Kaila, & Raichle, 2007). As an approximation of this metabolic demand, we calculated a proxy hemodynamic response by summing the net input to each unit in pFC, with each unit's contribution to the sum weighted by its net output. Consistent with extant neuroimaging data on proactive interference, our simulated hemodynamic response was markedly increased in prefrontal layers during recent lures, relative to nonrecent lures or targets, $t(24) = 11.35, p < .0005$ and $t(24) = 5.01, p < .001$, respectively (see Figure 5). This result was not due solely to simulated excitatory neurotransmission: The same pattern was observed in terms of net inhibitory input (see Appendix I for details about inhibitory currents in Leabra), consistent with theories of inhibitory contributions to the hemodynamic response (Buzsáki et al., 2007) and with the involvement of inhibition in resolving proactive interference (Jonides et al., 1998).

Figure 4. (A) The model reproduces the benchmark result of lower accuracy on 3-back than 2-back. (B) The model shows reduced accuracy on recent $(n - 1)$ lures, relative to both nonrecent lures $(>n)$ and match trials. In addition, the relative difference of these trial types is smaller in the 3-back task than the 2-back task, consistent with human data.
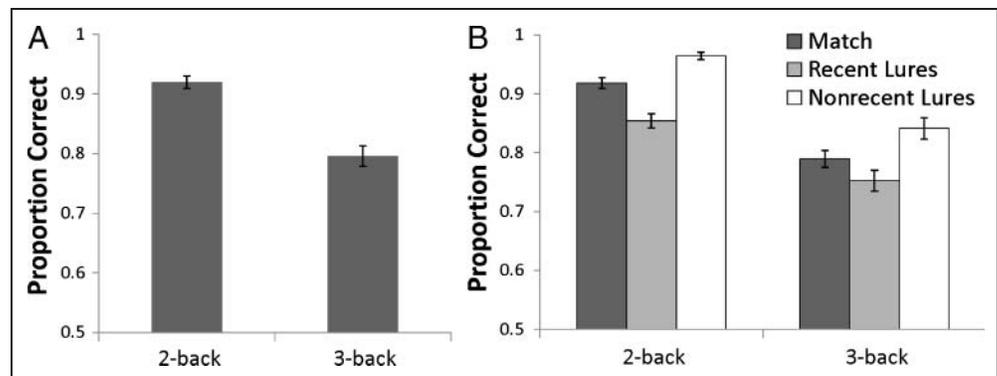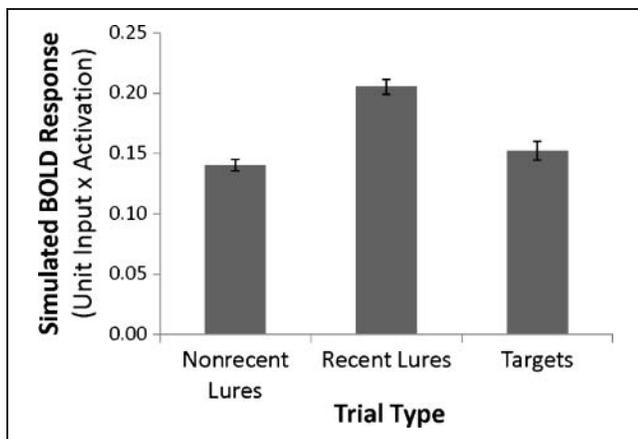
Figure 5. Recent lures were associated with a greater simulated hemodynamic response than nonrecent lures and targets, where the hemodynamic response is simulated as the weighted average of unit inputs and unit activations in the pFC layers.

The Model Captures Individual Differences in Human *n*-back Performance

In addition to capturing the above hallmark phenomena in the *n*-back task, we also tested whether the model captures individual differences. One source of individual differences is genetic variation related to dopaminergic functioning, such as the Val158Met polymorphism in the gene coding for catechol-O-methyl transferase (COMT), the principal enzyme that degrades DA in the pFC (Boulton & Eisenhofer, 1998). The low efficiency variant (the met allele) yields a large net reduction in prefrontal DA metabolism relative to carriers of the higher efficiency val allele (Chen et al., 2004; Männistö & Kaakkola, 1999). This differing efficacy results in a higher tonic level of prefrontal DA in met carriers (Bilder, Volavka, Lachman, & Grace, 2004).

Consistent with the hypothesized inverted U-shaped curve relating prefrontal DA levels to executive control (see Mattay et al., 2003), homozygotes for the val allele perform worse on the *n*-back than met carriers, either in terms of performance (e.g., Goldberg et al., 2003) or efficiency (i.e., neural activation required to achieve the same level of performance; e.g., Egan et al., 2001). Additionally, met carriers perform worse following pharmacological manipulations thought to increase prefrontal DA levels, such as administration of amphetamine (Mattay et al., 2003). Some recent studies suggest that the effect of the Val158Met polymorphism on *n*-back performance is weak, if it exists at all (e.g., Barnett, Scoriels, & Munafo, 2008). However, in practice any conclusion about the influence of COMT polymorphisms is complicated by other unmeasured and confounding genetic differences that may also distinguish val and met carriers (e.g., linkage disequilibrium or functional epistasis; Tan et al., 2007; Meyer-Lindenberg et al., 2006). Biologically constrained computational modeling can offer clarity to this situation as a way of testing the underlying hypothesis that extremes in prefrontal DA

should be associated with worse performance when all other factors are held constant.

Higher extracellular DA levels are frequently thought to increase the gain in individual pyramidal cells' activation function so as to make strongly active cells more active—an excitatory effect—and weakly active cells less active—an inhibitory effect (Cohen et al., 2002). The net result is an increase in signal-to-noise ratio for pFC as a whole (Winterer et al., 2006; Stefanis et al., 2005; Durstewitz, Seamans, & Sejnowski, 2000). In this way, individual differences at the Val158Met locus of the COMT gene might be hypothesized to produce differences in the relative sharpness of active representations in the pFC. All things being equal, sharper, sparser representations will promote faster processing and more robust maintenance in the pFC areas (O'Reilly & Munakata, 2000). Thus, to mimic the putative effects of individual differences in COMT function, we trained models to perform the 2-back task under variations in tonic DA's aforementioned (and most widely hypothesized) influence on pFC: signal-to-noise ratio (Winterer et al., 2006; Stefanis et al., 2005; Cohen et al., 2002; Durstewitz et al., 2000). Specifically, we increased the gain of the sigmoidal activation function on the units in the prefrontal layers from the default value (from 400 to 600). The gain was also decreased from the default value as a proxy for reduced levels of prefrontal DA (from 400 to 100).

Results of the simulations indicate that, although none of these variations in prefrontal DA precluded learning of the 2-back task to criterion, the final levels of performance reached by these networks after training conformed to the expected U-shaped curve, $F(1, 24) = 5.25$, $p = .027$ and $F(1, 24) = 9.09$, $p = .006$ for target and lure trial accuracy, respectively (see Figure 6). In our model, the U-shaped curve arises by rebalancing the flexibility
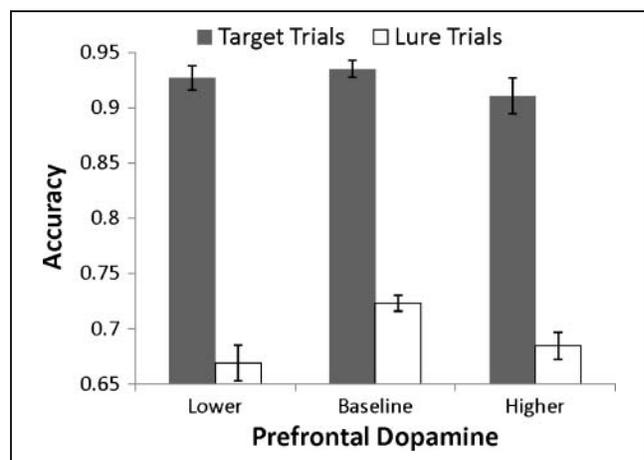
Figure 6. As a proxy for the effects of the polymorphisms in the COMT gene, we manipulated the effects of DA in the prefrontal layers of the model. This manipulation revealed an inverted U-shaped curve relating DA levels to performance, consistent with the hypothesized effects of varying DA levels in pFC.

stability tradeoff: With low gain/tonic DA, prefrontal representations are somewhat unstable, but with high gain/tonic DA, prefrontal representations become somewhat difficult to update. Note that the individual differences resulting from pFC gain are not unique to n-back or PBWM; other models positing similar DA effects in pFC may exhibit similar results (e.g., Chadderdon & Sporns, 2006; Deco, 2006; Tagamets & Horwitz, 2000), and as such, this result represents an important point of convergence across multiple formalisms.

Another source of individual differences in the n-back relates to the influence of control strategies and response bias on behavioral performance. Juvina and Taatgen (2007) showed that subjects encouraged to use a high-control strategy in this task—that is, to rely on active maintenance as opposed mere familiarity—show a positive correlation between accuracy on recent lure trials and on target n-back trials. In contrast, subjects encouraged to use a low-control familiarity strategy demonstrate a negative correlation between these trial types. Because our model includes only the mechanisms thought to be involved in high-control strategies, the model should also show this positive correlation. Indeed, we observed a robust positive correlation, $r(73) = .44, p < .0005$, between lure and target accuracy in the 2-back task (Figure 7) performed by models that varied in their pFC gain parameters (as described in the previous paragraph). This positive correlation arises because networks differing in prefrontal gain consequently also differ in their ability to update and maintain information in working memory—abilities that support performance on both lure and target trials alike.[2]

Inside the n-back: How Gating, Binding, and Resolution of Proactive Interference Occur

As described above, our model captures numerous empirical phenomena from the n-back task. Crucially, this good match to empirical data is enabled not by the explicit fitting of parameters but rather by the types of representations that develop through learning in PBWM. These representations can be readily understood as instantiations of the very executive functions hypothesized to be crucial for n-back performance: the need to flexibly update working memory, to bind stimulus representations to representations of serial order, and to manage proactive interference. Below, we demonstrate how these functions are accomplished using quantitative analysis of the model's learning trajectories.
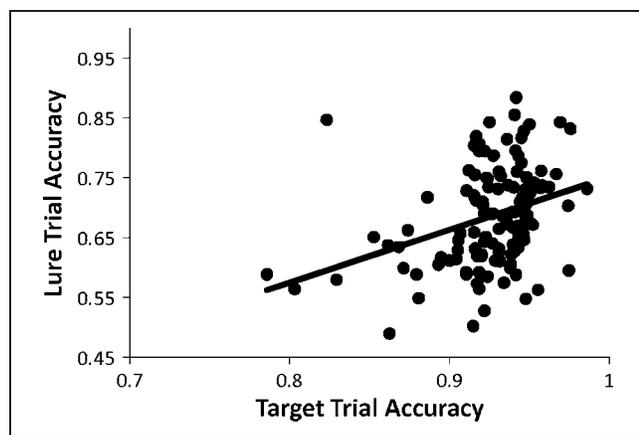
Gating

One principal executive function important for the n-back task is working memory updating; our model reveals what form this updating may take as a result of the striatal reinforcement learning mechanisms implemented in our model. In particular, the striatal layers learn to maximize reinforcement by firing differentially in terms of the serial order of each stimulus (i.e., 1, 2, or 3) instead of stimulus identity (i.e., A, B, C, etc.). This policy develops because it is supported by network connectivity (such that parietal layers project particularly strongly to striatal layers) but also because it maximizes reinforcement. Had striatal layers learn to fire differentially based on stimulus identity information (i.e., A, B, C, etc.), then it would be up to prefrontal layers to learn to represent whether each of those stimuli had been seen one, two, or three trials ago or not at all, as would most commonly be the case. Because it is less efficient to use limited prefrontal resources to represent stimuli that have not been recently experienced than to specifically represent those stimuli that have been seen recently, the latter updating policy is what emerges naturally through reinforcement learning.

The order-based gating striatal policy can be seen in how the activity patterns of these layers become more discrete with respect to serial order as training progresses. Formally, this change can be quantified as a reduction in entropy (such that lower entropy reflects greater certainty in which striatal units will be activated by a particular serial order) over the course of learning, as illustrated in Figure 8A. Thus, increasingly distinguishable neural patterns in the BG occur for distinct serial orders as training progresses, thereby yielding an order-specific gating signal.

The importance of this reduction in entropy can be seen in its relationship to performance. Although all networks ultimately reached approximately the same level of updating ability (i.e., near-zero entropy by the end of training), differences in accuracy on the task at that final point could be predicted based on the history of striatal entropy. That is, networks that were less error prone at the end of training showed no difference in striatal entropy at that late point but rather lower striatal entropy only early in training (as illustrated in Figure 8B). This effect occurs because the separation of items occurring with different serial orders to different prefrontal stripes is essential for two subsequent

Figure 7. The model captures the individual differences in the relationship of lure and target trial accuracy observed empirically when subjects are encouraged to adopt the same strategy as adopted by our model.
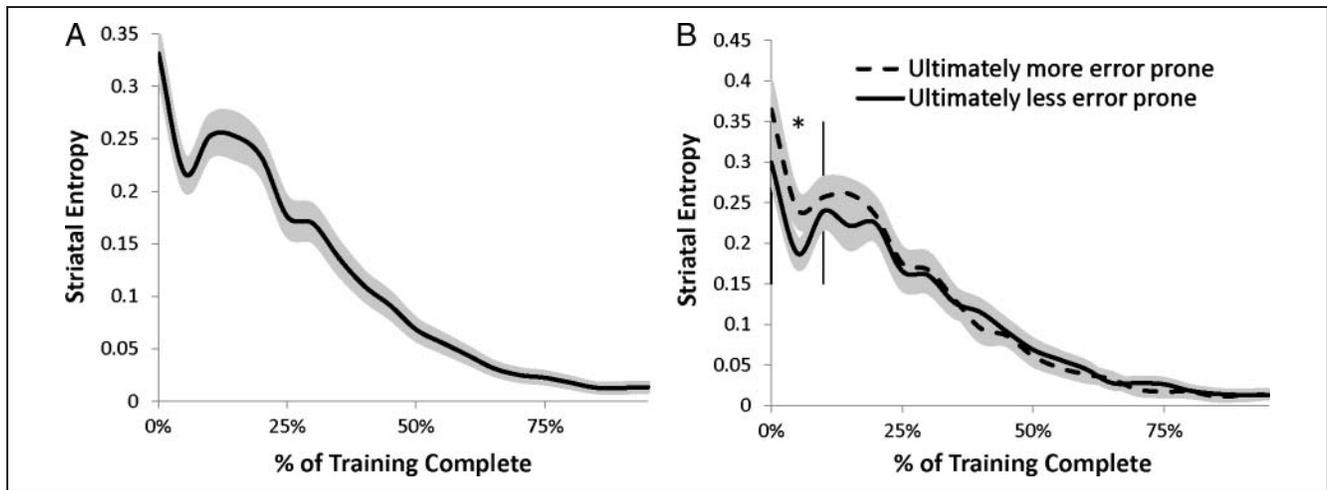
Figure 8. (A) The model learns to appropriate gate information into working memory by developing increasingly discrete firing patterns in the striatum over the course of training, here visualized in terms of reductions in entropy. (B) Individual differences in the ultimate post-training performance of models across runs can be predicted based on the reduction in striatal entropy much earlier in training: networks that ultimately commit less errors following training (solid vs. dotted lines) show significantly more (*$p < .05$) discrete patterns of firing between 0 and 10% of the total training time (vertical bars). Shaded regions represent *SEM* for each time point.

developments: the differentiation of items by prefrontal stripes, and the active maintenance of this information to resolve proactive interference. Networks that achieve earlier reductions in striatal entropy have a "head start" in these subsequent and slow refinements, each discussed in turn below.

Binding

A stable, order-specific gating signal is a prerequisite for prefrontal units to learn to differentiate items occurring with different serial orders—that is, binding, an important process for *n*-back performance (Badre & Wagner, 2006; Oberauer, 2005). The binding that occurs in our model is distinct from that typically occurring in connectionist models, which relies on coactivation of shared features. Our model binds items to serial order in a fundamentally different way, as described below.

First, the order-specific gating policy developed by striatal reinforcement learning mechanisms exposes particular pFC stripes to stimuli occurring with a particular serial order and other pFC stripes to stimuli occurring with other serial orders. By itself, this order-specific gating policy does not suffice for binding; the network must also be able to differentiate between the stimuli of any given serial order (e.g., to differentiate an A of serial order 1 from a B of serial order 1). Because this stimulus identity information is not provided by firing in striatal layers (which is serial order based), the prefrontal layers must discriminate stimulus identity on the basis of posterior representations. Ultimately, this discrimination is accomplished through a process of representational differentiation supported by Hebbian and error-driven learning,[3] such that each prefrontal stripe will learn to discriminate all items occurring with the serial order that that stripe is preferentially exposed

to (via the order-based striatal gating policy). This progressive differentiation both within and across stripes in pFC reflects the network emergently learning to bind items to their serial order.

The nature of the resulting bound representations can be quantified in terms of the Euclidean distance between prefrontal activity patterns across the various items and order combinations. This high-dimensional analysis can be illustrated with a cluster plot in which the *y* axis represents item by order combinations, horizontal lines represent the Euclidean distances between clusters, and cluster membership is indicated by vertical lines. We performed this type of hierarchical cluster analysis on the activity patterns, both before and after training, of one pFC stripe that learned to code for items appearing with one serial order (Figure 9A and B) and for a different pFC stripe that learned to code for items appearing with a different serial order (Figure 9C and D). The resulting figure reveals an initially haphazard pattern of representational similarity across items by order combinations (represented as a letter followed by a number; e.g., "D1"; Figure 9, left). After learning, this disorganization resolves into a highly structured representational scheme (Figure 9, right) in which all items occurring with a nonpreferred serial order for a given pFC stripe are highly similar, as indicated by very short horizontal lines linking the items into large clusters. In contrast, items of a preferred serial order become much more differentiated, as indicated by the increasingly pairwise clusters.

Thus, reinforcement learning mechanisms drive an order-based gating policy, whereas Hebbian/error-driven mechanisms support representational differentiation within particular serial orders. These processes jointly give rise to the active maintenance of item–context bindings in our model, such that items are bound to their context in terms of which pFC stripe they are gated into. Our model further suggests that this binding may occur through a sensitivity of

the pFC to serial order; indeed, empirical evidence suggests the pFC encodes information about serial order (Amiez & Petrides, 2007), and our model demonstrates how such sensitivity might emerge.

Proactive Interference

As the prefrontal-striatal circuit learns to gate and actively maintain information about items and their serial order, the network must also learn to resolve proactive interference from lure trials. In essence, the network must produce responses based on the match or mismatch between the current stimulus and information that was updated n-trials previously, while avoiding responses that would be based on any mapping between the current stimulus and the information updated and maintained from non-n lag lure trials. This constraint is at its core a selection problem: The prefrontal stripe with stimulus identity information for the current trial's serial order—and not stripes with information from other serial orders—must convey this information to the verbal output and posterior cortical layers so that corresponding match/nonmatch outputs can be activated.

The network learns to solve this selection problem through two mechanisms that emerge over learning. First, the stimulus identity information relevant to the current

trial's serial order biases the verbal output unit that corresponds to that stimulus's identity, as learned in the weights connecting that prefrontal stripe and the verbal output layer. Second, the posterior cortex acts as a kind of comparator, such that error-driven and Hebbian learning mechanisms craft a set of weights in the posterior cortical layer to detect matches between the stimulus input and verbal output layers and activate the appropriate manual output (see Appendix II for more details). The source of lure errors is, therefore, multicausal: Some errors (approximately 25%) reflect inappropriate detection of input–output matches by the posterior cortical area (i.e., the verbal output is correct and does not match the current stimulus, but the manual output nevertheless indicates a match response). Other errors (approximately 75%) reflect item confusion in the prefrontal layers as a result of interference from current stimuli, ultimately leading to the biasing of the incorrect unit in the verbal output layer (Appendix II provides a detailed analysis of recent lure errors in the 2-back task, which provides no evidence for the interpretation that lure errors arise because of incorrect gating on previous trials).

Figure 10 illustrates that this selection problem is solved relatively slowly over the course of training, with more rapid reductions in the proportion of errors that occur

Figure 9. Cluster plots reflect the Euclidean distance (indicated by the length of horizontal lines) between every item (indicated by letters along the y axis) and serial order (indicated by numbers along the y axis); thus, if the path from one item to another requires a large amount of horizontal travel, then the representations of those items are relatively distinct. (A) One prefrontal stripe shows an initially haphazard pattern of representational similarity across items, as indicated by the lack of systematic clustering between items and their order. (B) After training, the same prefrontal stripe illustrated in A develops a highly structured representation, by collapsing across all items of serial order 2 (top half of cluster plot) but differentiating among all items of serial order 1 (as indicated by the large horizontal lines separating each item; bottom half, enclosed by rounded rectangle). This stripe is preferentially tuned to code items of serial order 1. (C) A different pFC stripe also shows initially haphazard representational similarity. (D) After training this stripe shows a different pattern than that illustrated in B, in that it collapses equally across all items of serial order 1 (top half of cluster plot) but increasingly differentiates every item occurring with the other serial order (bottom half, enclosed by rounded rectangle).
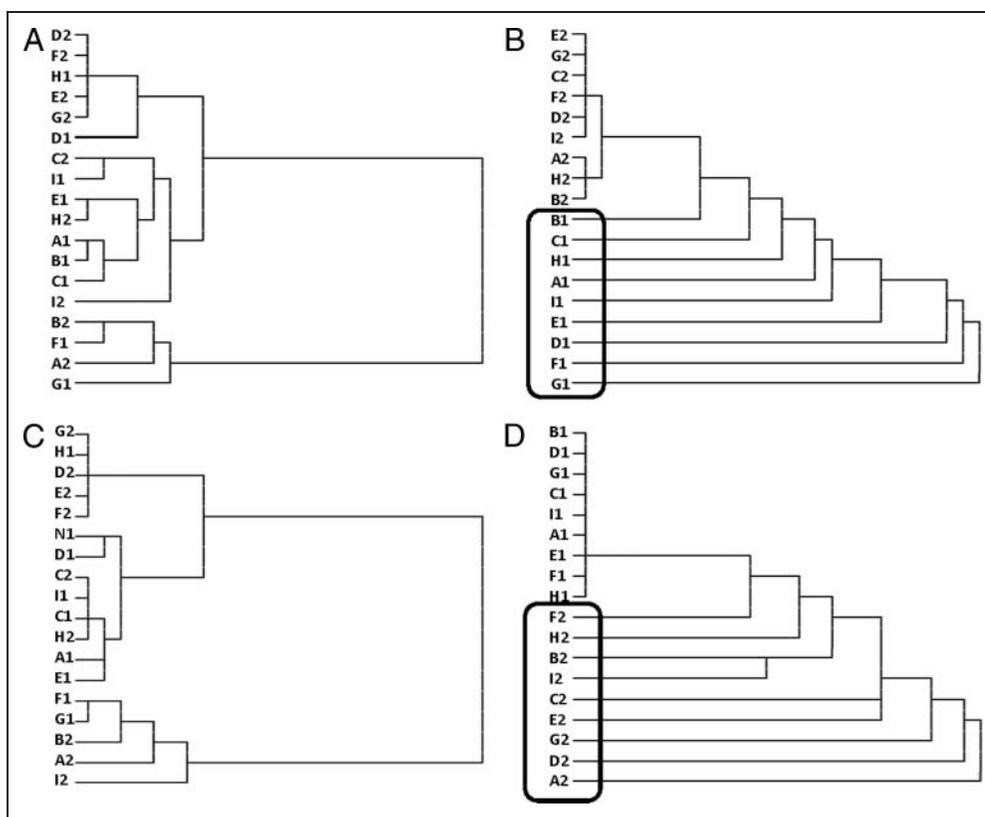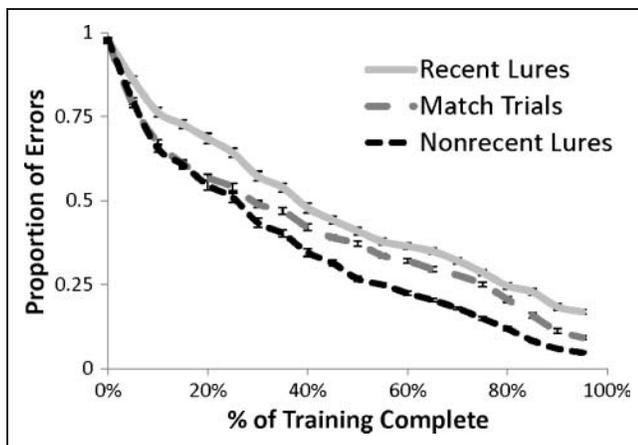
Figure 10. Performance on recent lures trials undergoes a shallower learning curve than performance on all other trial types, reflecting a more rapid reduction in error rate on trials that do not require the resolution of proactive interference (match and nonrecent lure trials as compared with recent lures).

on nonrecent lures and match trials than on recent lure trials, as well as a lower asymptotic error rate on those trials. We, thus, observed a relatively protracted development of resistance to interference, which is consistent with new evidence on developmental trajectories in the *n*-back (Schleepen & Jonkman, 2010). Our model shows this protracted development because of an interdependency between the resolution of proactive interference and other executive functions: Gating, maintenance and binding control processes supported by reinforcement learning (in the case of gating) as well as Hebbian and error-driven learning must first construct a relatively stable state before those representations can be incrementally refined to reduce proactive interference through additional Hebbian and error-driven learning.

GENERAL DISCUSSION

Here we report a biologically based model of the parieto-fronto-striatal system that learns to perform the *n*-back task, emergently producing representations that support executive functions like gating, binding, and the resolution of proactive interference. The model's acquisition of these functions enables its close match to empirical data, including behavioral, genetic, and neuroimaging findings, without the need for fine-grained tuning of the model's underlying parameters. Specifically, the updating of working memory is accomplished as the BG learns to provide a gating signal that is increasingly differentiated by an item's serial order, quantified above in terms of entropy. Active maintenance occurs as the pFC learns to bind items and their serial order, quantified above via hierarchical cluster analysis. Finally, proactive interference resulting from recent lure trials affects these prefrontal representations via increases in net input, unit activation, and inhibitory neuro-

transmission, consistent with the increased BOLD response observed in pFC during proactive interference (Badre & Wagner, 2006; Jonides et al., 1998). The model also captures the effect of individual differences in prefrontal tonic DA, individual differences observed when humans use the same type of high-control strategy implemented by our model, as well as decreased overall accuracy and an increase in the relative accuracy of recent lures as *n* moves from 2 to 3.

Our model integrates previous work to use an order representation in PBWM (as in the phonological loop model developed by O'Reilly & Frank, 2006) by more firmly rooting it in the biology of the IPS (as in Botvinick & Watanabe, 2007) and thereby differentiating the serial order signal from representations that might be used for dimensions like color or form. This work extends the phonological loop model in two important ways. First, the *n*-back task differs from the serial recall performed by these models in that serial order representations must now do "double duty"—simultaneously supporting the recall of old information as well as the storage of new information. Second, our model addresses the lack of an explicit external frame of reference for the representation of serial order by positing an internally generated one, such that there is a periodicity relation rather than a strictly serial relation between successive stimuli. Although the importance of a self-imposed periodicity function for our model actually leads to testable predictions, how these representations might autonomously develop nonetheless remains an important and unsolved problem.

Insights

Our model also leads to a number of theoretical insights. First, previous literature suggests that executive functions show unity and diversity at both behavioral and genetic levels (e.g., Friedman et al., 2008), but our model may challenge modular interpretations of this unity and diversity. Executive functions emerge here from an integrated parieto-fronto-striatal circuit instead of discrete mechanisms involved in only some executive tasks (cf. Cooper & Davelaar, 2010). Instead, a more emergent view of unity and diversity may enable a better match to neural mechanisms. Our model also indicates this emergent view will need to include frontal, striatal, and parietal areas, at the minimum. Executive functions are often discussed in terms of frontal, parieto-frontal (e.g., Corbetta, Patel, & Shulman, 2008), or fronto-striatal (O'Reilly & Frank, 2006) substrates, but theoretical accounts of frontal, parieto-frontal, or fronto-striatal interactions may be substantially incomplete without considering all three parts to the larger, integrated network. For example, although executive control might be considered relatively distinct from abilities like serial order processing, our model indicates the neural mechanisms supporting behavior across these domains may be related; this relationship should be explicitly considered in determining the role of parietal cortex in updating tasks.

However, it is also likely that other areas of parietal cortex contribute to performance in ways that do not selectively relate to time or order representations (Collette et al., 2005), and our model does not encapsulate the only form of parieto-frontal interaction. These anatomically and functionally diverse regions (e.g., Rushworth, Behrens, & Johansen-Berg, 2006) are likely to support multiple forms of processing. Thus, other computational and theoretical models of fronto-parietal function (Edin et al., 2009; Corbetta et al., 2008; Corbetta & Shulman, 2002) may describe aspects of the parietal cortex not captured by our model, and vice versa.

The model also provides insight into the ability to resolve proactive interference. Our simulations suggest that the increased hemodynamic response observed during recent lures may reflect the presence of interference and not a separate control process recruited to resolve interference. Instead, proactive interference resolution unfolds as an emergent consequence of the network's learning in general. That is, proactive interference resolution depends on striatal gating becoming increasingly specific to serial orders (which reduces interference across different serial orders that involved presentation of the same item) and increased representational differentiation among items of the same serial order within pFC (which reduces interference across different items presented with the same serial order). However, proxy hemodynamic increases to recent lures also reflect that a similar preceding item is being strongly maintained and that the current stimulus is being fully processed. These multiple facets of the hemodynamic response to proactive interference may explain the seemingly paradoxical findings that activation of the lateral pFC positively correlates with fluid intelligence (Gray, Chabris, & Braver, 2003), which presumably relies on strong maintenance and full processing of stimuli, but also positively correlates with behavioral indices of proactive interference (Nee, Jonides, & Berman, 2007). Our model, thus, offers one explanation for these apparent contradictions in the current empirical literature.

Predictions and Extensions

Our *n*-back model also leads to new testable predictions. First, because the models rely on a periodic serial order representation, 2-back accuracy should be differentially disrupted if subjects must simultaneously complete a task that requires a different periodicity of serial order representations (e.g., a three-movement spatial tapping task relative to a two-movement one). Indeed, serial order may be important for precisely this type of motor control (Salinas, 2009).

Second, the neural substrates of self-imposed periodicity should be identifiable with fMRI, using regressors whose onsets correspond to a periodicity of *n*. Striatal activation should show the same parametric variation with *n* as has been previously observed in the cortex: Our model

predicts these areas form a highly interconnected circuit modulated by memory demands. Moreover, representational similarity analysis or other multivoxel pattern analysis methods might reveal the same striatal hemodynamics reported here in terms of representational differentiation patterns.

Third, to the extent that humans are capable of good performance on $n > 3$-back tasks, they may recruit additional mechanisms, such as the hippocampal complex (HPC), to compensate for the logarithmically compressed nature of serial order representations in the IPS. Indeed, the HPC has only been inconsistently observed during performance of 2- and 3-back tasks (de Frias et al., 2010; Egan et al., 2003), and other accounts of *n*-back might predict HPC involvement only insofar as subjects adopt a low-control (i.e., familiarity based) strategy (Juvina & Taatgen, 2007). In contrast, the current model predicts high-control strategies will involve additional mechanisms not modeled here, when $n > 3$.

This third set of predictions suggests several possible extensions of the model to capture different strategies and training effects. The parietal layer is an important constraint on the ability of networks to perform adequately on $n > 3$-back tasks, because its tuning curves become increasingly compressed at higher serial orders. Parietal serial order representations simply become too compressed at high levels of *n* to support discrete representations. Nonetheless, because humans are apparently capable of learning $n > 3$ back tasks with training, one extension to our model would be a top–down projection to this area from the pFC. Over training, the network might learn to support increasingly discrete serial order representations using a top–down biasing signal (e.g., Edin et al., 2009). Our model might also be extended to capture the HPC mechanisms possibly used by subjects adopting a familiarity-based strategy. Conceptually similar mechanisms are used in a symbolic model of the *n*-back task (Juvina & Taatgen, 2007), such that a "time tagging" system is integrated with a familiarity system that relies on declarative memory. Different control strategies are then simulated in terms of whether the time tags are actively maintained (as in our current model) or retrieved only when familiarity is detected. With the appropriate biological extensions, our model might capture these and more *n*-back phenomena.

Our model may be relevant to the burgeoning field of executive functions training, in which the *n*-back is playing a prominent role. For example, *n*-back performance improves following training on the letter memory task (Dahlin, Neely, Larsson, Bäckman, & Nyberg, 2008). Our model is also capable of performing the letter memory task, and the types of executive functions that emerge in our model from its training on letter memory are extremely similar to those reported here. However, in the current report our models were trained only on the *n*-back task; clearly, human performance in any task relies on a longer and more varied history of experience than the training we provided to our model. Future work will pretraining

models on a larger variety of more elemental cognitive tasks and test transfer effects.

Conclusions and Future Directions

Follow-up work is ongoing, including the more complete modeling of these and other neural structures with a role in executive functioning and a more elaborate mapping of this type of model to behavior in other executive function tasks. Indeed, the PBWM framework used here can model a number of other tasks, and the overlap among these models may reveal the computational origin of the unity and diversity of executive functions (Friedman et al., 2008; Miyake et al., 2000). The current work represents a first step in that direction by specifying a formal computational link between anatomical connectivity studies demonstrating a highly interconnected parieto-fronto-striatal network with studies of genetic polymorphisms with individual differences at the behavioral level and with theoretical accounts of the executive functions important for working memory updating tasks.

APPENDIX I

The Leabra framework used for implementing the model is described in detail in O'Reilly (2001) and O'Reilly and Munakata (2000) and summarized here. This framework has been used in over 40 different models in O'Reilly and Munakata (2000) and a number of other research models. The current model, therefore, represents an extension to a systematic modeling framework using standardized mechanisms. (The model can be obtained by emailing the corresponding author.)

Pseudocode

The pseudocode for Leabra is given here, showing exactly how the pieces of the algorithm described in more detail in the subsequent sections fit together.

For each event:

1. Iterate over minus (−), plus (+), and update (++) phases of settling for each event.

   (a) At start of settling:

      i. For non-pFC/BG units, initialize state variables (activation, v m, etc.).
      ii. Apply external patterns (clamp input in minus, input and output, external reward based on minus-phase outputs).

   (b) During each cycle of settling, for all nonclamped units:

      i. Compute excitatory net input ($g_e(t)$ or η_j, Equation 2) (Equation 21 for SNr/Thal units).

      ii. For striatum go/no-go units in ++ phase, compute additional excitatory and inhibitory currents based on DA inputs from SNc (Equation 20).
      iii. Compute kWTA inhibition for each layer, based on

         A. Sort units into two groups based on g.
         B. If basic, find k and $(k + 1)$th highest; if average-based, compute average of $1 \rightarrow k$ and $k + 1 \rightarrow n$.
         C. Set inhibitory conductance g_i from g.

      iv. Compute point neuron activation, combining excitatory input and inhibition.

   (c) After settling, for all units:

      i. Record final settling activations by phase.
      ii. At the end of + and ++ phases, toggle pFC maintenance currents for stripes with SNr/Thal act > threshold (.1).

2. After these phases, update the weights (based on linear current weight values):

   (a) For all non-BG connections, compute error-driven weight changes (Equation 8) with soft weight bounding (Equation 9), Hebbian weight changes from plus-phase activations (Equation 7), and overall net weight change as weighted sum of error-driven and Hebbian (Equation 10).
   (b) For PV units, weight changes are given by delta rule computed as difference between plus phase external reward value and minus phase expected rewards (Equation 11).
   (c) For LV units, only change weights (using Equation 13) if PV expectation > θ_{pv} or external reward–punishment actually delivered.
   (d) For striatum units, weight change is the delta rule on DA-modulated second-plus phase activations minus unmodulated plus phase acts (Equation 19).
   (e) Increment the weights according to net weight change.

Point Neuron Activation Function

Leabra uses a *point neuron* activation function that models the electrophysiological properties of real neurons while simplifying their geometry to a single point. The membrane potential V_m is updated as a function of ionic conductances g with reversal (driving) potentials E as follows:

$$\Delta V_m(t) = \tau \sum_c (g_c(t) \bar{g}_c (E_c) - V_m(t)) \qquad (1)$$

with three channels (c) corresponding to excitatory input e, leak current l, and inhibitory input i. Following electrophysiological convention, the overall conductance is decomposed into a time-varying component $g_c(t)$, computed as a function of the dynamic state of the network, and a

constant g_c that controls the relative influence of the different conductances.

The excitatory net input/conductance $g_e(t)$ or η_j is computed as the proportion of open excitatory channels as a function of sending activations times the weight values:

$$\eta_j = g_e(t) = (x_i w_{ij}) = \frac{1}{n} \sum_i x_i w_{ij} \qquad (2)$$

The inhibitory conductance is computed via the kWTA function described in the next section, and leak is a constant. Activation communicated to other cells (y_j) is a thresholded (θ) sigmoidal function of the membrane potential with gain parameter γ:

$$y_j(t) = \frac{1}{\left(1 + \frac{1}{\gamma[V_m(t) - \theta]_+}\right)} \qquad (3)$$

where $[x]_+$ is a threshold function that returns 0 if $x < 0$ and x if $x > 0$. Note that if it returns 0, we assume $y_j(t) = 0$, to avoid dividing by 0. To produce a less discontinuous deterministic function with a softer threshold, the function is convolved with a Gaussian noise kernel ($\mu = 0$, $\sigma = 0.005$), which reflects the intrinsic processing noise of biological neurons:

$$y_j^*(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z^2}{2\sigma^2}} y_j(z-x) dz \qquad (4)$$

wherein x represents the $[V_m(t) - \theta]_+$ value and $y_j(x)$ is the noise-convolved activation for that value. In the simulation, this function is implemented using a numerical lookup table.

kWTA Inhibition

Leabra uses a kWTA function to achieve inhibitory competition among units within a layer (area). The kWTA function computes a uniform level of inhibitory current g_i for all units in the layer, such that the $(k + 1)$th most excited unit within a layer is generally below its firing threshold, whereas the kth is typically above threshold:

$$g_i = g_{k+1}^\theta + q\left(g_k^\theta - g_{k+1}^\theta\right) \qquad (5)$$

wherein $0 < q < 1$ (0.25 default used here) is a parameter for setting the inhibition between the upper bound of g_k^θ and the the lower bound of g_{k+1}^θ. These boundary inhibition values are computed as a function of the level of inhibition necessary to keep a unit right at threshold:

$$g_i^\theta = \frac{g_e^* \overline{g}_e (E_e - \theta) + g_l \overline{g}_l (E_l - \theta)}{\theta - E_i} \qquad (6)$$

wherein g_e^* is the excitatory net input without the bias weight contribution—this allows the bias weights to override the kWTA constraint.

In the basic version of the kWTA function, which is relatively rigid about the kWTA constraint and is therefore used for output layers, g_k^θ and g_{k+1}^θ are set to the threshold inhibition value for the kth and $(k + 1)$th value for the top most excited units, respectively. In the average-based kWTA version used here, g_k^θ is the average g_i^θ value for the top k most excited units and g_{k+1}^θ is the average of g_i^θ for the remaining $n - k$ units. This version allows for more flexibility in the actual number of units active depending on the nature of the activation distribution in the layer.

Hebbian and Error-driven Learning

Leabra uses a combination of error-driven and Hebbian learning. Error-driven learning in Leabra is the symmetric midpoint version of the GeneRec algorithm (O'Reilly & Munakata, 2000), which is functionally equivalent to contrastive Hebbian learning. The network settles in two distinct phases, an expectation (minus) phase where the network's produces an output and an outcome (plus) phase where the target output is experienced. The network then computes the difference of a pre- and postsynaptic activation product between these two phases. For Hebbian learning, Leabra uses essentially the same learning rule used in competitive learning, which can be seen as a variant of the Oja normalization. The error-driven and Hebbian learning components are combined additively at each connection to produce a net weight change.

The equation for the Hebbian weight change is:

$$\Delta_{\text{hebb}} w_{ij} = x_i^+ y_j^+ - y_j^+ w_{ij} = y_j^+ \left(x_i^+ - w_{ij}\right) \qquad (7)$$

and for error-driven learning using contrastive Hebbian learning:

$$\Delta_{\text{err}} w_{ij} = \left(x_i^+ y_j^+\right) - \left(x_i^- y_j^-\right) \qquad (8)$$

which is subject to a soft-weight bounding to keep within the 0–1 range:

$$\Delta_{\text{sberr}} w_{ij} = [\Delta_{\text{err}}] + (1 - w_{ij}) + [\Delta_{\text{err}}] - w_{ij} \qquad (9)$$

The two terms are then combined additively with a normalized mixing constant k_{hebb}:

$$\Delta w_{ij} = \in [k_{\text{hebb}}(\Delta_{\text{hebb}}) + (1 - k_{\text{hebb}})(\Delta_{\text{sberr}})] \qquad (10)$$

PVLV Equations

See Hazy et al. (2010), O'Reilly et al. (2007), and O'Reilly and Frank (2006) for further details on the PVLV system. We assume that time is discretized into steps that

correspond to environmental events (e.g., the presentation of a CS or US). All of the following equations operate on variables that are a function of the *current* time step *t*—we omit the *t* in the notation because it would be redundant. PVLV is composed of two systems, PV (primary value) and LV (learned value), each of which in turn are composed of two subsystems (excitatory and inhibitory). Thus, there are four main value representation layers in PVLV (PV$_e$, PV$_i$, LV$_e$, LV$_i$), which then drive the DA (DA) layers (VTA/SNc). There are several changes in the algorithm from this previous work (most notably the inclusion of the PV$_r$ and novelty value (NV) systems; see Learning Rules below). These changes are efforts to increase the biological plausibility of the system (e.g., removing synaptic depression) and will be discussed in detail in a future work. The simulations and results described in this article were only performed using the PVLV system described here; the changes to the algorithms described here were developed completely independently.

Value Representations

The PVLV layers use standard Leabra activation and kWTA dynamics as described above, with the following modifications. They have a three-unit distributed representation of the scalar values they encode, where the units have preferred values of (0, 0.5, 1). The overall value represented by the layer is the weighted average of the unit's activation times its preferred value, and this decoded average is displayed visually in the first unit in the layer. The activation function of these units is a "noisy" linear function (i.e., without the $x/(x + 1)$ nonlinearity to produce a linear value representation but still convolved with Gaussian noise to soften the threshold, as for the standard units; Equation 4), with gain $\gamma = 220$, noise variance $\sigma = 0.01$, and a lower threshold $\theta = 0.17$. The *k* for kWTA (average based) is 1, and the *q* value is 0.9 (instead of the default of 0.6 in other layers). These values were obtained by optimizing the match for value represented with varying frequencies of 0–1 reinforcement (e.g., the value should be close to 0.4 when the layer is trained with 40% of 1 values and 60% of 0 values). Note that having different units for different values, instead of the typical use of a single unit with linear activations, allows much more complex mappings to be learned. For example, units representing high values can have completely different patterns of weights than those encoding low values, whereas a single unit is constrained by virtue of having one set of weights to have a monotonic mapping onto scalar values.

Learning Rules

The PV$_e$ layer does not learn and is always just clamped to reflect any received reward value (*r*). By default we use a value of 0 to reflect negative feedback, 0.50 for no feedback, and 1 for positive feedback (the scale is arbitrary). The PV$_i$ layer units (y_j) are trained at every point in time to produce an expectation for the amount of reward that will be received at that time. In the minus phase of a given trial, the units settle to a distributed value representation based on sensory inputs. This results in unit activations y_j^-, and an overall weighted average value across these units denoted PV$_i$. In the plus phase, the unit activations (y_j^+) are clamped to represent the actual reward *r* (a.k.a., PV$_e$). The weights (w_{ij}) into each PV$_i$ unit from sending units with plus-phase activations x_i^+, are updated using the delta rule between the two phases of PV$_i$ unit activation states:

$$\Delta w_{ij} = \in (y_j^+ - y_j^-)x_i^+ \tag{11}$$

This is equivalent to saying that the US/reward drives a pattern of activation over the PV$_i$ units, which then learn to activate this pattern based on sensory inputs. In addition to the PV$_e$ and PV$_i$ layers there is an additional PV$_r$ layer that is associated with learning about reward detection. This system learns the same way as the PV$_i$ system, but has a slower learning rate for weight decreases relative to increases. The LV$_e$ and LV$_i$ layers learn in much the same way as the PV$_i$ layer (Equation 11), except that the PV system filters the training of the LV values, such that they only learn from actual reward outcomes or when reward is expected by the PV$_r$ system, and not when no rewards are present or expected. This condition is as follows:

$$PV_{filter} = \min(PV_r, PV_i) < \theta_{min} - \max(PV_r, PV_i) > \theta_{max} \tag{12}$$

$$\Delta w_i = \begin{cases} \in \left(y_j^+ - y_j^-\right)x_i^+ & \text{if } PV_{filter} \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

wherein θ_{min} is a lower threshold (0.20 by default), below which negative feedback is indicated and θ_{max} is an upper threshold (0.80), above which positive feedback is indicated (otherwise, no feedback is indicated). Biologically, this filtering requires that the LV systems be driven directly by primary rewards (which is reasonable and required by the basic learning rule anyway) and that they learn from DA dips driven by high PV$_r$ expectations of reward that are not met. The only difference between the LV$_e$ and LV$_i$ systems is the learning rate $\in$, which is 0.05 for LV$_e$ and 0.001 for LV$_i$. Thus, the inhibitory LV$_i$ system serves as a slowly integrating inhibitory cancellation mechanism for the rapidly adapting excitatory LV$_e$ system.

Finally, the NV layer signals stimulus novelty and produces DA bursts for novel stimuli, which slowly decay in magnitude as a stimulus becomes familiar. The habituation for this system is simply:

$$\Delta w_i = -NVx_i \tag{14}$$

The PV, LV, and NV distributed value representations drive the DA layer (VTA/SNc) activations in terms of

the difference between the excitatory and inhibitory terms for each. Thus, there is a PV delta, an LV delta, and an NV delta:

$$\delta_{pv} = PV_e - PV_i \tag{15}$$

$$\delta_{lv} = LV_e - LV_i \tag{16}$$

$$\delta_{nv} = NV \tag{17}$$

The DA system integrates each of these inputs, using a temporal derivative computation to only produce brief bursts or dips relative to a baseline level of activation (this is the primary difference from the synaptic depression mechanism used in the earlier published version). The key issue is when to use each of the above values: If primary rewards are present or expected but not present, then the PV system dominates, and otherwise, LV + NV drive it. With the differences in learning rate between LV_e (fast) and LV_i (slow), the LV delta signal reflects recent deviations from expectations and not the raw expectations themselves, just as the PV delta reflects deviations from expectations about primary reward values. This is essential for learning to converge and stabilize when the network has mastered the task. These two delta signals need to be combined to provide an overall DA delta value, as reflected in the firing of the VTA and SNc units. One sensible way of doing so is to have the PV system dominate at the time of primary rewards, whereas the LV system dominates otherwise by using the same PV-based filtering as holds in the LV learning rule:

$$\delta = \begin{cases} \left(\delta_{pv}^{t} - \delta_{pv}^{(t-1)}\right) & \\ \left(\delta_{lv}^{t} - \delta_{lv}^{(t-1)}\right) + \left(\delta_{nv}^{t} - \delta_{nv}^{(t-1)}\right) & \text{otherwise.} \end{cases} \quad if\ PV_{filter} \tag{18}$$

Special Basal Ganglia Mechanisms

Striatal Learning Function

Each stripe (group of units) in the striatum layer is divided into go versus no-go in an alternating fashion. The DA input from the SNc modulates these unit activations in the update phase by providing extra excitatory current to go and extra inhibitory current to the no-go units in proportion to the positive magnitude of the DA signal and vice versa for negative DA magnitude. This reflects the opposing influences of DA on these neurons (Frank, 2005). This update phase DA signal reflects the PVLV system's evaluation of pFC updates produced by gating signals in the plus phase. Learning on weights into the go/no-go units is based on the activation delta between the update (++) and plus phases:

$$\Delta w_i = \epsilon\, x_i(y^{++} - y^{+}) \tag{19}$$

To reflect the finding that DA modulation has a contrast-enhancing function in the striatum (Frank, 2005; Nicola, Surmeier, & Malenka, 2000; Hernández-López, Bargas, Surmeier, Reyes, & Galarraga, 1997) and to produce more of a credit assignment effect in learning, the DA modulation is partially a function of the previous plus phase activation state:

$$g_e = \gamma[da]_+ y^+ + (1 - \gamma)[da]_+ \tag{20}$$

where $0 < \gamma < 1$ controls the degree of contrast enhancement (0.5 is used in all simulations), $[da]_+$ is the positive magnitude of the DA signal (0 if negative), y_+ is the plus-phase unit activation, and g_e is the extra excitatory current produced by the da (for go units). A similar equation is used for extra inhibition (g_i) from negative da ($[da]_-$) for go units and vice versa for no-go units.

SNrThal Units

The SNrThal units provide a simplified version of the SNr/GPe/Thalamus layers. They receive a net input that reflects the normalized go/no-go activations in the corresponding striatum stripe:

$$\eta_j = \left[\frac{\sum go - \sum no - go}{\sum go + \sum no - go}\right]_+ \tag{21}$$

(where [ ]+ indicates that only the positive part is taken; when there is more no-go than go, the net input is 0). This net input then drives standard Leabra point neuron activation dynamics, with kWTA inhibitory competition dynamics that cause stripes to compete to update pFC. This dynamic is consistent with the notion that competition/selection takes place primarily in the smaller GP/SNr areas and not much in the much larger striatum (e.g., Mink, 1996). The resulting SNrThal activation then provides the gating update signal to pFC: If the corresponding SNrThal unit is active (above a minimum threshold; 0.1), then active maintenance currents in pFC are toggled.

This SNrThal activation also multiplies the per-stripe DA signal from the SNc:

$$\delta_j = snr_j\delta \tag{22}$$

where snr_j is the snr unit's activation for stripe j and δ is the global DA signal.

Random Go Firing

The PBWM system only learns after go firing, so if it never fires go; it can never learn to improve performance. One simple solution is to induce go firing if a go has not fired after some threshold number of trials. However, this threshold would have to be either task specific or set very

high, because it would effectively limit the maximum maintenance duration of pFC (because by updating pFC, the go firing results in loss of currently maintained information). Therefore, we have adopted a somewhat more sophisticated mechanism that keeps track of the average DA value present when each stripe fires a go:

$$\overline{da}_k = \overline{da}_k + \epsilon\left(da_k - \overline{da}_k\right) \qquad (23)$$

If this value is <0 and a stripe has not fired go within one or two trials (in the 2-back and 3-back, respectively), a random go firing is triggered with some probability (.1). We also compare the relative per-stripe DA averages, if the per-stripe DA average is low but above 0, and one stripe's da_k is 0.05 below the average of that of the other stripes:

$$\text{if}\left(\overline{da}_k < 0.1\right) \text{ and } \left(\overline{da}_k - \langle\overline{da}\rangle < -.05\right); \text{go} \qquad (24)$$

a random go is triggered, again with some probability (.1). Finally, we also fire random go in all stripes with some very low baseline probability (.0001) to encourage exploration.

When a random go fires, we set the SNrThal unit activation to be above go threshold, and we apply a positive DA signal to the corresponding striatal stripe, so that it has an opportunity to learn to fire for this input pattern on its own in the future.

pFC Maintenance

pFC active maintenance is supported in part by excitatory ionic conductances that are toggled by go firing from the SNrThal layers. This is implemented with an extra excitatory ion channel in the basic V_m update Equation 1. This channel has a conductance value of 0.5 when active. See Frank, Loughry, and O'Reilly (2001) for further discussion of this kind of maintenance mechanism. The first opportunity to toggle pFC maintenance occurs at the end of the first plus phase and then again at the end of the second plus phase (third phase of settling). Thus, a complete update can be triggered by two gos in a row, and it is almost always the case that if a go fires the first time, it will fire the next, because striatum firing is primarily driven by sensory inputs, which remain constant.

APPENDIX II

Computations Supporting Manual Output: Match versus Nonmatch Decision

As noted in the main text, the network must learn to produce not only the correct verbal output (corresponding to the *n*-back item) but also a manual output (corresponding to whether the current item matches or does not match the *n*-back item). This match versus nonmatch decision can be computed by the network simply by comparing the activation patterns in the input with those in the verbal output layer and pFC. Indeed, it is precisely this form of "coincidence detection" that is accomplished by the posterior cortical layer.

We confirmed that "coincidence detection" between the verbal output layer and stimulus input layer was the underlying computation performed by the posterior cortical layers as follows. First, we examined those units in the posterior cortical layer that received strong projections from corresponding units in the input and verbal output layers (e.g., large weights from the "A" stimulus in both layers, or from the "B" stimulus in both layers, as indicated by a positive correlation of weights from these two layers). We found that these units projected disproportionately strongly to the target output response than to the nontarget response, relative to those posterior cortical units that do not show this correspondence of weights (e.g., strong weights from the "A" stimulus in the input layer but weak weights from the "A" stimulus in the verbal output layer): $(F(1, 98) = 4.482, p < .05)$. Thus, the target manual output is driven largely by those posterior cortical units that are themselves strongly activated by matches between the input and verbal output layers.

As such, the match–nonmatch decision relies not only on coincidence detection mechanisms but also on mechanisms supporting activation of the correct verbal output—a requirement fulfilled by the connectivity of striatal areas with parietal areas, which trigger the gating of prefrontal information into the verbal output layer. Thus, the match–nonmatch response can be seen as a cumulative result of the network's behavior in total, although it is most directly supported by weight-based computations occurring in the posterior cortical layer.

The Underlying Source of Recent Lure Errors

Our approach to identify the source of recent lure errors was to examine in detail the performance of one network performing the 2-back task over the final seven epochs of training. We first determined that the match–nonmatch decision was typically being performed correctly by the posterior cortical layer (i.e., detecting matches between the recalled verbal output and the stimulus present in the input). Only 25% of recent lure errors reflected a failure to respond to matches–mismatches between the (correct) verbal output and stimulus input layers. That is, the prefrontal layers recalled the correct information to the verbal output layer, but the posterior cortical layer incorrectly responded as though this information matched the information presented in the input.

Nonetheless, approximately 75% of recent lure errors reflected recall of the 1-back instead of the 2-back item in the verbal output layer. To determine whether item confusion within the relevant prefrontal stripe was to blame for this incorrect recall, we examined the representational differentiation among items in prefrontal layers that were gated on a particular trial, using a similar cluster plot analysis as presented in the main text. In particular, we recorded

the activations in a prefrontal stripe with a preferred serial order of 1 across the final seven epochs of training. For each unit in this stripe, we averaged activations across correct trials and incorrect trials separately, conditional on the verbal output for that trial and the current trial's serial order. Finally, we constructed separate cluster plots for correct and incorrect trials to visualize the differentiation of prefrontal representations of each item × order combination.

This analysis indicated that stripes demonstrated good differentiation among items of the preferred serial order on correct recent lure trials (Figure A1A), but a much more haphazard pattern of representational differentiation on incorrect recent lure trials (Figure A1B). Thus, recent lure

Figure A1. Cluster plots of prefrontal activations on correct (A) and incorrect (B) recent lure trials. On correct trials, different items of the preferred serial order for this stripe are well differentiated in prefrontal activations, as indicated by the relatively long paths interconnecting items of serial order one (A1, B1, C1, etc.). Items of a dispreferred serial order are not as well differentiated, as indicated by relatively shorter paths interconnecting those items (A2, B2, C2, etc.). In contrast, on incorrect recent lure trials (B), there is a much more haphazard pattern of differentiation among items in terms of prefrontal activations, indicative of item confusion.

errors are associated with increased item confusion within the prefrontal layers.

In principle, this item confusion within the prefrontal layers could arise from a gating error. That is, this prefrontal stripe may have been gated inappropriately on the current trial or some other recent trial and been exposed to items of a dispreferred serial order. In this case, poor differentiation of items would reflect that this stripe had been updated with information that it was poorly suited to represent. However, we found no appreciable differences in the striatal activations between correct and incorrect trials—neither on the trial where the incorrect verbal output was provided, nor on either of the two preceding trials. Thus, prefrontal stripes were gated similarly on incorrect and correct recent lure trials, as well as on the trials immediately preceding them, indicating that gating errors are not a source of the item confusion occurring on incorrect recent lure trials.

If not because of gating, what could be the source of the item confusion occurring on incorrect recent lure trials? We found that the haphazard pattern of representational differentiation in prefrontal activation states on incorrect recent lure trials—that is, item confusion—was paralleled by haphazard patterns of net input to prefrontal layers on incorrect recent lure trials. Whereas net input to prefrontal layers was substantially different in terms of whether the current trial was of serial order 1 or 2 on correct recent lure trials (Figure A2A), incorrect recent lure trials showed much more similar net input to prefrontal layers across trials of serial orders 1 and 2 (Figure A2B). This result indicates that recent lure errors arise from an instability of prefrontal activation states independent of gating: The clean separation between representations of items of different serial orders is corrupted on incorrect recent lure trials, both in terms of prefrontal activations and net input to prefrontal layers.

We conducted further analyses of representational differentiation on the trial preceding incorrect and correct recent lure trials but found no appreciable differences in the prefrontal representations on the trials preceding recent lure errors relative to the representations on the trials preceding correct rejections of recent lures. This similarity indicates that the corruption of prefrontal representations on incorrect recent lure trials is due to the recent lure itself and not to a corruption of the representation of the 2-back stimulus occurring before the recent lure. Our model, thus, indicates that recent lure errors occur because of a lack of stability of prefrontal representations to interference arising from the recent lure itself.

In summary, these analyses suggested that recent lure errors did not arise because of gating problems but rather because of nonrobust representations in pFC that were susceptible to interference from incoming stimuli. These particular representations may have been susceptible to interference from lures to the extent that they were similar to the 1-back stimulus, perhaps as a result of Hebbian learning in the sequences leading up to recent lure errors.

Figure A2. Cluster plots of prefrontal net input on correct (A) and incorrect (B) lure trials. Although net input to prefrontal layers on correct recent lure trials (A) was reliably differentiated in terms of the serial order of the current trial (as indicated by the large cluster separating items of serial order 2 from those of serial order 1), net input to prefrontal layers on incorrect recent lure trials (B) was not reliably differentiated according to serial order. Thus, the representational differentiation among items of different serial orders in terms of activation states (Figure A1A and B) was paralleled by differentiation among items of different serial orders in terms of net input.

Hebbian and Error-driven Computations Contributing to Item Differentiation in the Prefrontal Layers

The binding of items to context in our *n*-back model relies on two principle developments: the development of an order-based striatal gating signal as a result of reinforcement learning and the increasing prefrontal differentiation of items occurring with a preferred serial order as a result of Hebbian and error-driven learning. As discussed in the main text, the order-based gating policy develops as a result of reinforcement learning because it is supported by strong connectivity between the parietal and striatal layers, but also because it maximizes reinforcement relative to alternative gating policies.

In contrast, increasing representational differentiation in the prefrontal layers develops via Hebbian and error-driven learning processes over repeated training experiences. To see why Hebbian and error-driven learning lead naturally to this kind of representational differentiation, consider an incorrect trial on the 2-back task, where the serial order-based gating policy had correctly updated a prefrontal stripe with the "A" stimulus presented two trials previously, but the prefrontal representation of this "A" stimulus is not yet sufficiently distinct from its representation of other stimuli. This indistinct prefrontal representation may bias the posterior cortical and verbal output layers such that the "B" unit in the verbal output layer is ultimately activated instead of the correct "A" unit. Thus, there will be a resulting difference in activation states between the incorrect answer (produced during Leabra's minus phase, as described in Appendix I) and the correct answer (produced during Leabra's plus phase, also described in Appendix I). This difference will lead to an error-driven learning signal that changes specifically those weights—from the prefrontal layer that was gated on this trial to the verbal output and posterior cortical layer with which the prefrontal layers are connected—that served to conflate the "B" and "A" stimuli. In addition, Hebbian learning will further strengthen connections among those (correct) units that are simultaneously activated in Leabra's plus phase. Iterative learning of this type eventually converges to yield prefrontal representations that maximally distinguish the stimuli that any given stripe must represent, so that such errors are not produced. Thus, because each stripe eventually contains representations of items occurring with only one particular serial order (because of the order-based gating policy learned by the striatum), error-driven and Hebbian learning only ever train stripes to maximally distinguish those stimuli of that preferred serial order.

Notes

1. We note that the dense interconnectivity of the PBWM architecture is based on known neurobiology, and is therefore taken as a given in the current attempt to map from neurobiology to executive function. Previous work has identified that the intact striatal and prefrontal mechanisms of PBWM are necessary for good serial recall performance (O'Reilly & Frank, 2006), of which the *n*-back is a particularly demanding variant.

2. We suggest this effect is intrinsic to the model's emergent behavior, and not merely epiphenomenal, for two reasons. First, a more likely result would have been a negative correlation between the accuracy on lure and target trials, owing to the fact that Leabra involves the learning of "bias weights" which might produce the widely-observed tradeoff between hit and false alarm rate in target detection tasks. Second, this correlation was specific to networks in the trained state; no significant correlation between lure and target accuracy was observed following the first epoch of training ($r = .10$, ns).

3. Appendix II describes how error-driven and Hebbian learning cooperate to support representational differentiation.

REFERENCES

Aalto, S., Brück, A., Laine, M., Någren, K., & Rinne, J. O. (2005). Frontal and temporal dopamine release during working memory and attention tasks in healthy humans: A positron emission tomography study using the high-affinity dopamine D2 receptor ligand [11C]FLB 457. *Journal of Neuroscience, 25,* 2471–2477.

Amiez, C., & Petrides, M. (2007). Selective involvement of the mid-dorsolateral prefrontal cortex in the coding of the serial order of visual stimuli in working memory. *Proceedings of the National Academy of Sciences, U.S.A., 104,* 13786–13791.

Apud, J. A., & Weinberger, D. R. (2007). Treatment of cognitive deficits associated with schizophrenia: Potential role of catechol-*O*-methyltransferase inhibitors. *CNS Drugs, 21,* 535–557.

Badre, D., & Wagner, A. D. (2006). Computational and neurobiological mechanisms underlying cognitive flexibility. *Proceedings of the National Academy of Sciences, U.S.A., 103,* 7186–7191.

Barnett, J. H., Scoriels, L., & Munafo, M. R. (2008). Meta-analysis of the cognitive effects of the catechol-*O*-methyltransferase gene Val158/108Met polymorphism. *Biological Psychiatry, 64,* 137–144.

Bilder, R. M., Volavka, J., Lachman, H. M., & Grace, A. A. (2004). The catechol-*O*-methyltransferase polymorphism: Relations to the tonic-phasic dopamine hypothesis and neuropsychiatric phenotypes. *Neuropsychopharmacology, 29,* 1943–1961.

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

Botvinick, M., & Watanabe, T. (2007). From numerosity to ordinal rank: A gain-field model of serial order representation in cortical working memory. *Journal of Neuroscience, 27,* 8636–8642.

Boulton, A. A., & Eisenhofer, G. (1998). Catecholamine metabolism: From molecular understanding to clinical diagnosis and treatment. *Advances in Pharmacology, 42,* 273–292.

Buzsáki, G., Kaila, K., & Raichle, M. (2007). Inhibition and brain work. *Neuron, 56,* 771–783.

Chadderdon, G. L., & Sporns, O. (2006). A large-scale neurocomputational model of task-oriented behavior selection and working memory in prefrontal cortex. *Journal of Cognitive Neuroscience, 18,* 242–257.

Chen, J., Lipska, B. K., Halim, N., Ma, Q. D., Matsumoto, M., Melhem, S., et al. (2004). Functional analysis of genetic variation in catechol-*O*-methyltransferase (COMT): Effects on mRNA, protein, and enzyme activity in postmortem human brain. *American Journal of Human Genetics, 75,* 807–821.

Cohen, J. D., Braver, T. S., & Brown, J. W. (2002). Computational perspectives on dopamine function in prefrontal cortex. *Current Opinion in Neurobiology, 12,* 223–229.

Collette, F., Van Der, L. M., Laureys, S., Delfiore, G., Degueldre, C., Luxen, A., et al. (2005). Exploring the unity and diversity of the neural substrates of executive functioning. *Human Brain Mapping, 25,* 409–423.

Cooper, R. P., & Davelaar, E. J. (2010). Modelling the correlation between two putative inhibition tasks: A simulation approach. In D. D. Salvucci & G. Gunzelmann (Eds.), *Proceedings of the 10th International Conference on Cognitive Modeling* (pp. 31–36). Philadelphia, PA: Drexel University.

Corbetta, M., Patel, G. H., & Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron, 58,* 306–324.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3,* 215–229.

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

de Frias, C. M., Marklund, P., Eriksson, E., Larsson, A., Oman, L., Annerbrink, K., et al. (2010). Influence of COMT gene polymorphism on fMRI-assessed sustained and transient activity during a working memory task. *Journal of Cognitive Neuroscience, 22,* 1614–1622.

Deco, G. (2006). A dynamical model of event-related fMRI signals in prefrontal cortex: Predictions for schizophrenia. *Pharmacopsychiatry, 39,* 65–67.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of visual selective attention. *Annual Review of Neuroscience, 18,* 192–222.

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

Edin, F., Klingberg, T., Johansson, P., McNab, F., Tegnér, J., & Compte, A. (2009). Mechanism for top–down control of working memory capacity. *Proceedings of the National Academy of Sciences, U.S.A., 106,* 6802–6807.

Egan, M. F., Goldberg, T. E., Kolachana, B. S., Callicott, J. H., Mazzanti, C. M., Straub, R. E., et al. (2001). Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proceedings of the National Academy of Sciences, U.S.A., 98,* 6917–6922.

Egan, M. F., Kojima, M., Callicott, J. H., Goldberg, T. E., Kolachana, B. S., Bertolino, A., et al. (2003). The BDNF val66met polymorphism affects activity dependent secretion of BDNF and human memory and hippocampal function. *Cell, 112,* 257–269.

Fernández-Miranda, J. C., Rhoton, A. L., Alvarez-Linera, J., Kakizawa, Y., Choi, C., & de Oliveira, E. P. (2008). Three-dimensional microsurgical and tractographic anatomy of the white matter of the human brain. *Neurosurgery, 62,* 989–1026.

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and non-medicated Parkinsonism. *Journal of Cognitive Neuroscience, 17,* 51–72.

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.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent variable analysis. *Journal of Experimental Psychology: General, 133,* 101–135.

Friedman, N. P., Miyake, A., Young, S. E., Defries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences

in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General, 137,* 201–225.

Goldberg, T. E., Egan, M. F., Gscheidle, T., Coppola, R., Weickert, T., Kolachana, B. S., et al. (2003). Executive subprocesses in working memory: Relationship to catechol-*O*-methyltransferase Val158Met genotype and schizophrenia. *Archives of General Psychiatry, 60,* 889–896.

Goschke, T. (2000). Involuntary persistence and intentional reconfiguration in task-set switching. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 331–355). Cambridge, MA: MIT Press.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6,* 316–322.

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

Hazy, T. E., Frank, M. J., & O'Reilly, R. C. (2007). 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,* 1601–1613.

Hazy, T. E., Frank, M. J., & O'Reilly, R. C. (2010). Neural mechanisms of acquired phasic dopamine responses in learning. *Neuroscience & Biobehavioral Reviews, 34,* 701–720.

Hernández-López, S., Bargas, J., Surmeier, D. J., Reyes, A., & Galarraga, E. (1997). D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca2+ conductance. *Journal of Neuroscience, 17,* 3334–3342.

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience, 139,* 181–193.

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proceedings of the National Academy of Sciences, U.S.A., 95,* 8410–8413.

Juvina, I., & Taatgen, N. A. (2007). Modeling control strategies in the N-back task. In *Proceedings of the 8th International Conference on Cognitive Modeling* (pp. 73–78). New York, NY: Psychology Press.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology, 55,* 352–358.

Levitt, J. B., Lewis, D. A., Yoshioka, T., & Lund, J. S. (1993). Topograhy of pyramidal neuron intrinsic connections in macaque monkey prefrontal cortex (Areas 9 & 46). *Journal of Comparative Neurology, 338,* 360–376.

Logie, R. H., Cocchini, G., Delia Sala, S., & Baddeley, A. D. (2004). Is there a specific executive capacity for dual task coordination? Evidence from Alzheimer's disease. *Neuropsychology, 18,* 504–513.

Männistö, P. T., & Kaakkola, S. (1999). Catechol-*O*-methyltransferase (COMT): Biochemistry, molecular biology, pharmacology, and clinical efficacy of the new selective COMT inhibitors. *Pharmacology Review, 51,* 593–628.

Marshuetz, C., Reuter-Lorenz, P. A., Smith, E. E., Jonides, J., & Noll, D. C. (2006). Working memory for order and the parietal cortex: An event-related fMRI study. *Neuroscience, 139,* 311–316.

Marshuetz, C., Smith, E. E., Jonides, J., Degutis, J., & Chenevert, T. L. (2000). Order information in working memory: fMRI evidence for parietal and prefrontal mechanisms. *Journal of Cognitive Neuroscience, 12,* 130–144.

Mattay, V. S., Goldberg, T. E., Fera, F., Hariri, A. R., Tessitore, A., Egan, M. F., et al. (2003). Catechol *O*-methyltransferase val158-met genotype and individual variation in the brain response to amphetamine. *Proceedings of the National Academy of Sciences, U.S.A., 100,* 6186–6191.

Meyer-Lindenberg, A., Nichols, T., Callicott, J. H., Ding, J., Kolachana, B., Buckholtz, J., et al. (2006). Impact of complex genetic variation in COMT on human brain function. *Molecular Psychiatry, 11,* 867–877.

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

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

Nee, D. E., Jonides, J., & Berman, M. G. (2007). Neural mechanisms of proactive interference-resolution. *Neuroimage, 38,* 740–751.

Nicola, S. M., Surmeier, D. J., & Malenka, R. C. (2000). Dopaminergic modulation of neuronal excitability in the striatum and nucleus accumbens. *Annual Review of Neuroscience, 23,* 185–215.

Nieder, A., Diester, I., & Tudusciuc, O. (2006). Temporal and spatial enumeration processes in the primate parietal cortex. *Science, 313,* 1431–1435.

Oberauer, K. (2005). Binding and inhibition in working memory: Individual and age differences in short-term recognition. *Journal of Experimental Psychology: General, 134,* 368–387.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2003). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience, 7,* 75–79.

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., Busby, R. S., & Soto, R. (2003). Three forms of binding and their neural substrates: Alternatives to temporal synchrony. In A. Cleeremans (Ed.), *The unity of consciousness: Binding, integration, and dissociation* (pp. 168–192). Oxford: Oxford University Press.

O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: A computational model of learning in the frontal 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.

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). *n*-Back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping, 25,* 46–59.

Pucak, M. L., Levitt, J. B., Lund, J. S., & Lewis, D. A. (1996). Patterns of intrinsic and associational circuitry in monkey prefrontal cortex. *Journal of Comparative Neurology, 376,* 614–630.

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

Rushworth, M. F., Behrens, T. E., & Johansen-Berg, H. (2006). Connection patterns distinguish 3 regions of human parietal cortex. *Cerebral Cortex, 16,* 1418–1430.

Salinas, E. (2009). Rank-order-selective neurons form a temporal basis set for the generation of motor sequences. *Journal of Neuroscience, 29,* 4369–4380.

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General, 132,* 566–594.

Schleepen, T. J., & Jonkman, L. M. (2010). The development of non-spatial working memory capacity during childhood and adolescence and the role of interference control: An n-back task study. *Developmental Neuropsychology, 35,* 37–56.

Stefanis, N. C., van Os, J., Avramopoulos, D., Smyrnis, N., Evdokimidis, I., & Stefanis, C. N. (2005). Effect of COMT Val158Met polymorphism on the Continuous Performance Test, Identical Pairs Version: Tuning rather than improving performance. *American Journal of Psychiatry, 162,* 1752–1754.

Tagamets, M. A., & Horwitz, B. (2000). A model of working memory: Bridging the gap between electrophysiology and human brain imaging. *Neural Networks, 13,* 941–952.

Tan, H. Y., Chen, Q., Goldberg, T. E., Mattay, V. S., Meyer-Lindenberg, A., Weinberger, D. R., et al. (2007). Catechol-*O*-methyltransferase Val158Met modulation of prefrontal–parietal-striatal brain systems during arithmetic and temporal transformations in working memory. *Journal of Neuroscience, 49,* 13393–13401.

Tsuchida, A., & Fellows, L. K. (2009). Lesion evidence that two distinct regions within prefrontal cortex are critical for *n*-back performance in humans. *Journal of Cognitive Neuroscience, 21,* 2263–2275.

Van Dijck, J., & Fias, W. (2011). A working memory account for spatial-numerical associations. *Cognition, 119,* 114–119.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: A meta-analysis. *Cognitive, Affective and Behavioral Neuroscience, 3,* 255–274.

Winterer, G., Musso, F., Vucurevic, G., Stoeter, P., Konrad, A., Seker, B., et al. (2006). COMT genotype predicts BOLD signal and noise characteristics in prefrontal circuits. *Neuroimage, 32,* 1722–1732.

Yeterian, E. H., & Pandya, D. N. (1993). Striatal connections of the parietal association cortices in rhesus monkeys. *Journal of Comparative Neurology, 332,* 175–197.