



This article is part of the topic “Best of Papers from the 2016 International Conference on Cognitive Modeling,” David Reitter and Frank E. Ritter (Topic Editors). For a full listing of topic papers, see: <http://onlinelibrary.wiley.com/doi/10.1111/tops.2017.9.issue-1/issuetoc>.

## The Effects of Guanfacine and Phenylephrine on a Spiking Neuron Model of Working Memory

Peter Duggins, Terrence C. Stewart, Xuan Choo, Chris Eliasmith

*Centre for Theoretical Neuroscience, University of Waterloo*

Received 19 October 2016; received in revised form 5 November 2016; accepted 15 November 2016

---

### Abstract

We use a spiking neural network model of working memory (WM) capable of performing the spatial delayed response task (DRT) to investigate two drugs that affect WM: guanfacine (GFC) and phenylephrine (PHE). In this model, the loss of information over time results from changes in the spiking neural activity through recurrent connections. We reproduce the standard forgetting curve and then show that this curve changes in the presence of GFC and PHE, whose application is simulated by manipulating functional, neural, and biophysical properties of the model. In particular, applying GFC causes increased activity in neurons that are sensitive to the information currently being remembered, while applying PHE leads to decreased activity in these same neurons. Interestingly, these differential effects emerge from network-level interactions because GFC and PHE affect all neurons equally. We compare our model to both electrophysiological data from neurons in monkey dorsolateral prefrontal cortex and to behavioral evidence from monkeys performing the DRT.

*Keywords:* Working memory; Delayed response task; Neural Engineering Framework; Pharmacology; ADHD

---

### 1. Introduction

Working memory (WM) is a central component of cognitive systems that is required for temporary information storage during the execution of complex tasks. WM can be

---

Correspondence should be sent to Peter Duggins, Centre for Theoretical Neuroscience, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada. E-mail: [psipeter@gmail.com](mailto:psipeter@gmail.com)

impaired by a variety of mental disorder, including attention-deficit-hyperactivity disorder (ADHD) (Scahill et al., 2014). Because WM is biologically realized in networks of neurons, one goal for researchers studying WM is to understand how populations of spiking neurons implement information storage and retrieval in the brain, and how these neurobiological processes are disrupted by WM disorders (Avery, Franowicz, Studholme, van Dyck, & Arnsten, 2000). Although computational models are well suited to this task, existing WM models rarely provide both biological detail and a functional architecture capable of generating behavioral predictions. For example, models such as CoJACK (Dancy, Ritter, Berry, & Klein, 2015) and Gunzelmann, Gross, Gluck, and Dinges (2009) are concerned with how high-level cognitive abilities like mental arithmetic, perception, and tactical planning relate to low-level details like caffeine or sleep loss, but must implement these low-level details through the models' symbolic plans and production rules rather than through the underlying neurobiological substrate. ACT-R/ $\phi$  (Ritter et al., 2012) also investigates low-level details (e.g., epinephrine levels) using a mathematical model of physiology, but it does not yet simulate neurons explicitly. On the other hand, the Human Brain Project (Markram et al., 2015) simulates cortical microcircuits with unprecedented biological accuracy, but it lacks a theoretical framework that relates model activity to high-level cognitive abilities like perception, decision-making, and WM. New theories and models are needed to unify these approaches and characterize the complex relationships between the pharmacological, neurobiological, and cognitive aspects of WM.

In this paper, we present a spiking neural network model of WM and action selection applied to a mnemonic cognitive test, the spatial delayed response task (DRT). Our model captures a broad set of low- and high-level features: It includes enough biophysical detail to simulate the underlying causes of mental disorders and typical interventions (e.g., drugs, neural stimulation, etc.), enough neural detail to respect biological constraints and produce data that can be externally validated, and enough functional detail to provide a conceptual description of WM systems and their disorder-induced deficits. We first describe the biological and computational basis of WM using the Neural Engineering Framework (NEF) (Eliasmith & Anderson, 2003), a general method for building cognitive models from spiking neurons. We then present our model, describing how it extends previous WM models and performs the DRT by storing, retrieving, and forgetting information. To introduce the relevant pharmacology, we examine two WM drugs, guanfacine (GFC) and phenylephrine (PHE), from a functional, neural, and biophysical perspective, hypothesize about why GFC alleviates WM deficits produced by ADHD while PHE exacerbates them, then construct drug simulations that reproduce the effects of GFC and PHE at three different levels of analysis. We find that these drug simulations alter the model's electrophysiology and DRT performance in a manner that aligns closely with empirical data from monkeys. We conclude by discussing how these results consolidate our understanding of WM disorders and proposing biophysical and anatomical extensions to the model.

## 2. Background

WM is at least partly realized in the prefrontal cortex (PFC), a brain region whose prominent size in highly evolved primates suggests its importance in complex cognitive tasks that require a flexible mental workspace. The PFC represents information that is temporarily held in mind, used to guide behavior and decision-making, and is thought to be maintained through recurrent excitatory connections between neurons with similar tuning properties (Goldman-Rakic, 1995). Computationally, this recurrence realizes an extended temporal integration that preserves the represented item without external stimulation (Singh & Eliasmith, 2006). Therefore, the core requirement in a neural model of WM is that a population of neurons can maintain its state over time. That is, given a brief input, the internal connectivity should cause the neural activity pattern that results from that input to persist after the input has stopped. This persistence will not be perfect—over time the neural activity will drift away from its initial value.

However, this population of neurons cannot maintain *any* possible pattern of firing: We expect there to be correlations in the structure of this neural activity. Indeed, it has become common to analyze neural activity in WM areas (and elsewhere in the brain) by performing dimensionality reduction through techniques such as jPCA (Shenoy, Sahani, & Churchland, 2013). These approaches characterize the underlying patterns of correlation between the spiking neurons, identifying a lower dimensional subspace that the neural activity represents. That is, rather than treating each neuron independently, we assume there is some vector  $x$  that is being represented by the population of neurons. The dimensionality of this vector is much smaller than the number of neurons, meaning that the information is redundantly encoded across these neurons. In particular, each neuron  $i$  will have some particular vector  $e_i$  for which that neuron fires most strongly (these are often known as “preferred direction vectors” or “encoders” and have been widely used as a useful way of characterizing cortical activity (e.g., Georgopoulos, Kalaska, Caminiti, & Massey, 1982). We can consider the total overall current going into a neuron to be proportional to  $e_i \cdot x$  (the similarity between  $x$  and the preferred vector  $e_i$ ). To produce a variety of tuning curves and firing rates that matches those in PFC, we randomly choose a gain  $\alpha_i$  and bias current  $\beta_i$  for each neuron, resulting in a total input current of  $\alpha_i e_i \cdot x + \beta_i$ . This current can be fed into any neuron model, but here we simply use the standard leaky integrate-and-fire (LIF) model.

Given that the neural spiking activity encodes some vector  $x$ , it should be possible to recover that information by observing the spikes. The simplest method is to “decode” this spiking information via a weighted sum of the spikes, such that  $\hat{x}(t) = \sum_i a_i(t) d_i \times h(t)$ , where  $a_i(t)$  is the spiking activity of the  $i$ th neuron,  $h(t)$  is the shape of the post-synaptic current caused by the spikes, and  $d_i$  is the weighting factor for each neuron. The decoder (i.e.,  $d_i$ ) values can be found by performing a least-squares optimization that minimizes the difference between  $x$  (the original vector) and  $\hat{x}$  (the vector recovered by observing the spiking activity). This method of characterizing neural representation is the first principle of the NEF.

Now that we have defined how a population of neurons can represent a value  $x$ , we can construct recurrent connections within this population such that the neural activity continues to represent  $x$  over time. To realize such a WM, we must find recurrent connection weights that stabilize dynamical neural activity, regardless of the value  $x$  being represented. Using the third principle of the NEF, this can be characterized as another least-squares minimization problem: Previous work has shown that the optimal weights from neuron  $i$  to neuron  $j$  are  $w_{ij} = \alpha_j e_j \cdot d_i$  (Eliasmith & Anderson, 2003). The result is a population of spiking neurons that maintains its activity over time and has been the basis of multiple WM models (Choo & Eliasmith, 2010; Singh & Eliasmith, 2006).

To simulate the WM population, we let the input vector  $x$  be two-dimensional, where the first dimension is the value to be remembered, and the second dimension is the amount of time it has been remembered for. Empirical and modeling evidence are consistent with the claim that PFC neurons explicitly encode the passage of time (Bekolay, Laubach, & Eliasmith, 2014; Lewis & Miall, 2006; Singh & Eliasmith, 2006). For example, some PFC neurons start firing only after a given amount of time has passed, while others gradually decrease their firing rate over time (Romo, Brody, Hernández, & Lemus, 1999). These “positive monotonic” and “negative monotonic” neurons can be thought of as neurons that are sensitive to both the value being represented and the amount of time the memory has been held; in other words, these are *spatial mnemonic* neurons whose  $e_i$  values are large for both the first and second dimension. Other neurons may only be sensitive to one or the other dimension (i.e., would have small  $e_i$  values for one of those two dimensions). This variability in  $e_i$  matches well to the observed variability in WM tuning curves (Singh & Eliasmith, 2006).

### 3. Model description

A standard behavioral test of working memory is called the spatial DRT. In this task, a monkey fixates on a point in the center of the screen, and then it is briefly presented a visual cue on the left or right (cue period, 1 s). The cue is removed and then comes a delay period (2, 4, 6, or 8 s), during which the monkey has to represent and maintain the cue’s location in working memory. After the delay period, the monkey recalls the cue’s location and responds by pressing a button on the left or right.

We extend the NEF models described above (Choo & Eliasmith, 2010; Singh & Eliasmith, 2006) to perform this task, using the architecture shown in Fig. 1. The cue’s location is represented by a value  $cue \in \{-1, +1\}$  and is the first dimension of the input vector  $x$ . This value is fed as a stimulus into the WM population, causing each of its  $N = 1,000$  neurons to spike with frequency determined by the similarity between its preferred vector  $e_i$  and the represented value  $x$ . This stimulus is applied for the duration of the cue period (1 s) then removed; after this, the memory must be maintained by activity fed back through the WM recurrent connections.

We introduce two sources of *instability* to simulate forgetting during the delay period. First, external noise approximating the stochastic variability found in the brain is injected

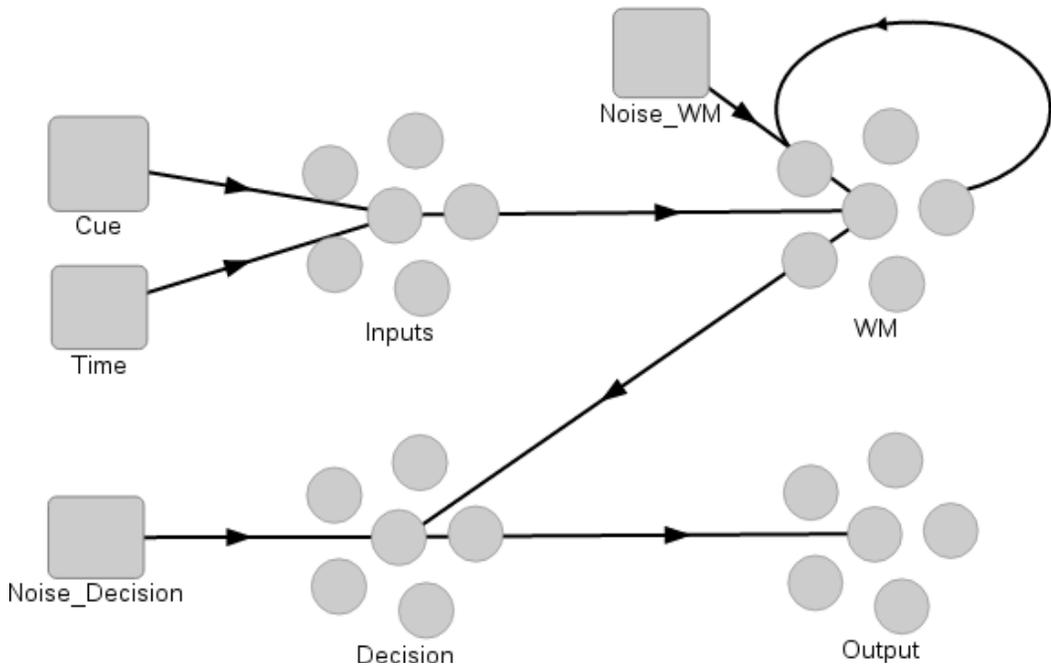


Fig. 1. Schematic of the spiking neuron model of the delayed response task. Circles represent neural populations and boxes represent nodes that output vectors. Parameters:  $dt = 0.001$  s;  $neurons_{WM} = 1,000$ ,  $neurons_{In} = 100$ ,  $neurons_{Dec} = 100$ ;  $cue(t) = \{-1.0, +1.0\}$  for  $t < 1.0$ ,  $time(t) = 0.4$  for  $t > 1.0$ ;  $noise_{wm} = N(0, 0.005)$ ,  $noise_{dec} = 0.3$ ; exponential synaptic time constants  $\tau_{wm} = 0.1$  s, a value consistent with NMDA-type glutamate receptors in PFC.

into WM neurons using a bias current through the `Noise_WM` node. Second, the constant passage of time, encoded as the second dimension of the input vector  $x$ , steadily increases the firing rate of WM neurons until they saturate. Once a significant portion of the neurons saturate, decoding the cue value from the population's activity becomes noisy and inaccurate. Fig. 2 shows that, without these instabilities, the information stored in WM is stable for a very long time (minutes to hours), but when they are present, the information decays over tens of seconds, consistent with decay rates of human WM (Choo & Eliasmith, 2010).

To produce a response, the model must access the stored value and produce one of two  $outputs \in \{-1, +1\}$ . A decision population attempts to decode the cue information by taking the neural activity of the WM neurons and computing their weighted sum, giving an estimate of the original value ( $\hat{x}(t) = \sum_i a_i(t) d_i \times h(t)$ ). Because a neural mechanism to convert this value into a decision will include some degree of variability, we add normally distributed noise to this estimate. If the result is above zero, we interpret this as the model giving the +1 response, and if it is below zero, we interpret it as giving the -1 response.

There are four important free parameters in this model: The ramp magnitude controls the rate of interference due to elapsed time while the working memory noise interferes

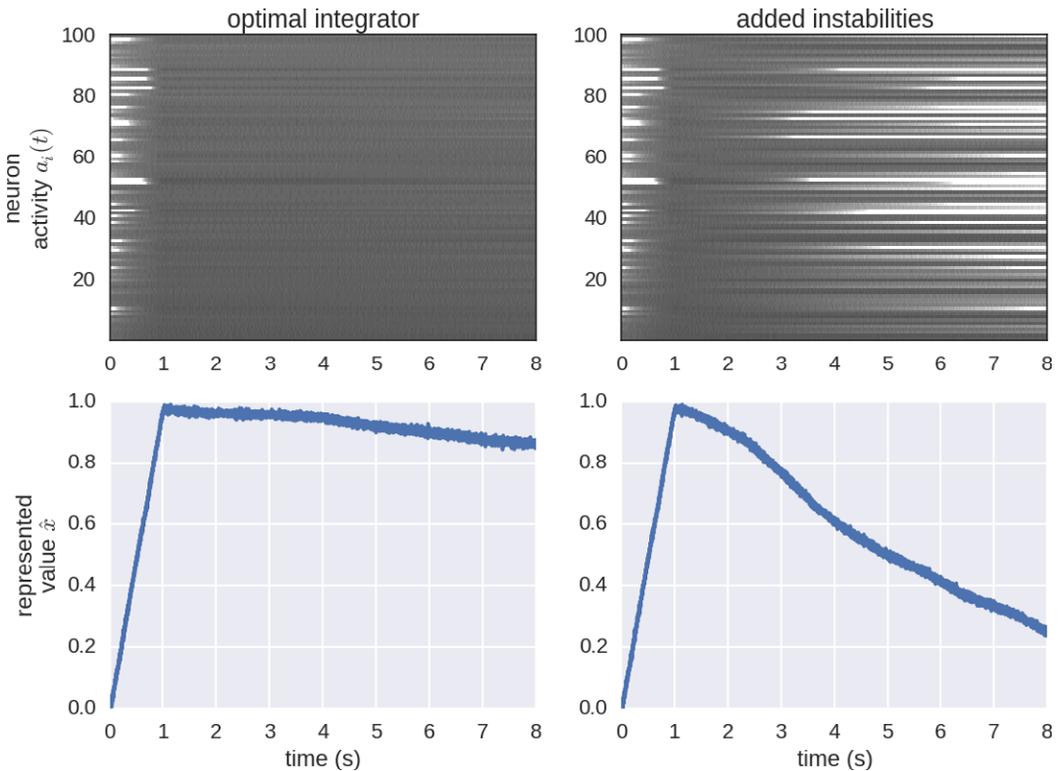


Fig. 2. Top: spike rasters from 100 WM neurons (out of 1,000). Bottom: the represented value is computed from the spiking activity with  $\hat{x} = \sum_i a_i(t) d_i \times h(t)$ . With added noise and temporal ramp, the neural integrator is more unstable and decays towards zero, causing forgetting of the represented information.

with the cue representation; the decision noise controls the accuracy of the decision procedure; and the misperception probability gives the likelihood that the model fails to perceive the cue in the first place ( $cue = 0$ ). We set the two noise parameters to biologically plausible values,  $\sigma_{wm} = 0.005$  and  $\sigma_d = 0.3$ , and then fine-tuned the misperception and ramp until we reproduced the baseline forgetting curve (see Fig. 5),  $misperceive = 0.1$  and  $ramp = 0.4$ . The model is available on GitHub ([https://github.com/psipeter/drugs\\_and\\_working\\_memory](https://github.com/psipeter/drugs_and_working_memory)).

#### 4. Simulating guanfacine and phenylephrine

The stable representation of items stored in WM is sensitive to the synaptic connections of intra-PFC loops and the biochemical environment of PFC neurons. Impairments in the dopamine and norepinephrine system are closely associated with WM disorders such as ADHD (Arnsten & Lombroso, 2000; Chandler, Waterhouse, & Gao, 2014), and the drugs used to treat them target these impaired systems

biophysically (Avery et al., 2000; Scahill et al., 2014). Specifically, drugs prescribed for ADHD affect PFC neurons that express Hyperpolarization-activated Cyclic Nucleotide-gated (HCN) ion channels (Franowicz et al., 2002). HCN channels are located on neurons' dendritic spines and are open at rest, shunting synaptic input by permitting nonspecific cations to flow out of the cell, as shown in Fig. 3. These channels control the excitability of pyramidal neurons by modulating dendritic summation and the cells' resting potentials (Magee, 1999; Poolos, Migliore, & Johnston, 2002); when the neuromodulator norepinephrine binds to the  $\alpha$ 2A-adrenoreceptor ( $\alpha$ 2A-AR), it activates a cAMP-mediated intracellular signalling cascade that ultimately closes HCN channels. The result is reduced shunting and increased excitability of the neuron.

The drugs guanfacine (GFC) and phenylephrine (PHE) are an agonist and an antagonist of the  $\alpha$ 2A-AR, respectively; GFC is prescribed to alleviate WM deficits in patients with ADHD (Scahill et al., 2014), while PHE reproduces many of the disorder's symptoms (Arnsten & Leslie, 1991; Levy, 2008). A study by Wang et al. (2007) showed that GFC increased (and a compound similar to PHE decreased) the firing rate of PFC neurons with weak mnemonic tuning in the direction of the cue presented in the DRT, while having no effect on cells tuned in the opposite direction; see Fig. 4. These results are consistent with monkeys' increased (decreased) performance on DRT when injected with GFC and PHE (Mao, Arnsten, & Li, 1999; Ramos, Stark, Verduzco, van Dyck, & Arnsten, 2006); see Fig. 5. We hypothesize that GFC raises the firing rate of neurons with cue-aligned encoders, slowing the decay of information stored in the PFC neural integrator and increasing performance on the DRT.

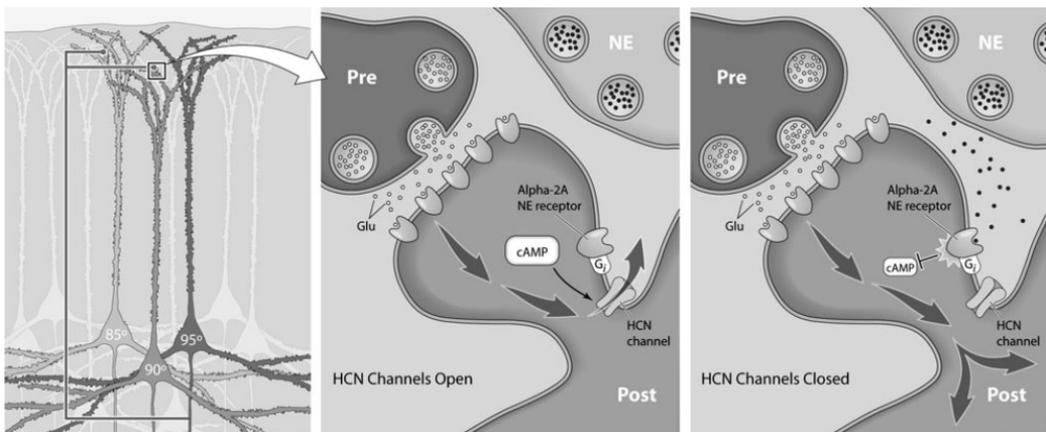


Fig. 3. Diagram of guanfacine's (GFC) biophysical interactions with prefrontal cortex neurons. EPSCs induced by presynaptic glutamate release are shunted from dendritic spines via open Hyperpolarization-activated Cyclic Nucleotide-gated channels, leading to minimal postsynaptic potentiation. When norepinephrine or its agonist GFC binds the  $\alpha$ 2A-AR, HCN channels close, increasing the efficacy of cortical inputs. Image reproduced from Wang et al. (2007) (with permission).

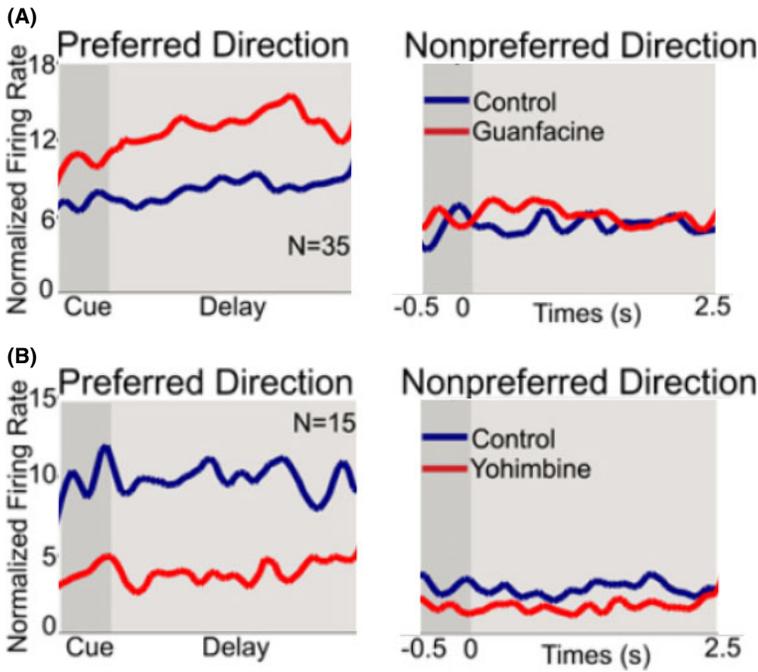


Fig. 4. Guanfacine increases (Yohimbine decreases) the activity of neurons that are spatially tuned to the cue’s location in the delayed response task (DRT) but have a negligible effect on nonpreferred direction neurons. Data were obtained from neurons in area 46 of dorsolateral prefrontal cortex that displayed spatial tuning during control conditions. Guanfacine and Yohimbine were applied iontophoretically to the monkeys during DRT performance (Wang et al., 2007). Yohimbine and PHE are both antagonists for the  $\alpha$ 2A-AR, so we expect their neural and behavioral effects to be quite similar.

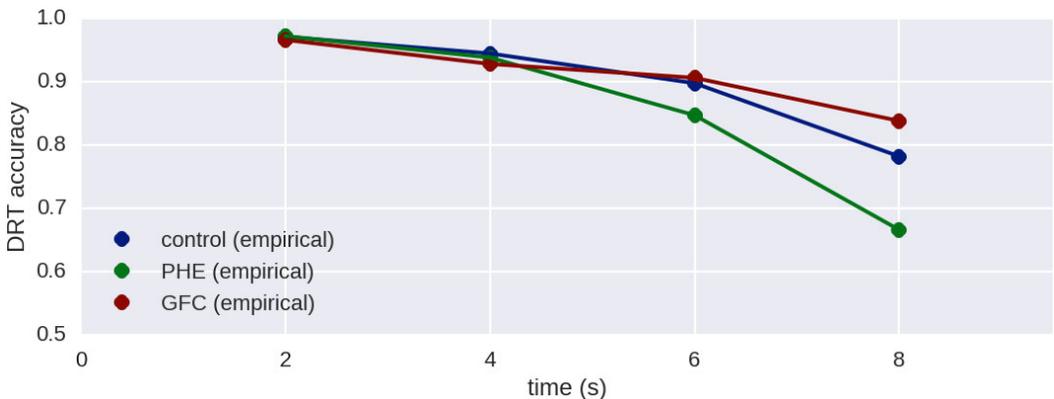


Fig. 5. Delayed response task (DRT) accuracy versus delay period length in monkeys injected with saline, guanfacine (GFC), and phenylephrine (PHE) (Mao et al., 1999). The outlier datapoint, GFC at  $t = 4$ , probably arises from the small sample size of the dataset: A single (unique) monkey was used for each experimental condition, though each line represents  $N = 800$ – $1,200$  DRT trials from that animal. Errors in the original data were negligible, so they are not plotted here.

## 5. Results

We performed three drug simulations, each of which approximates the effects of GFC and PHE on the model at a different scale. We began at the highest level, looking at how the drugs functionally alter forgetting rates in WM by manipulating the recurrent connection weights. Next, we investigated how injecting a constant current into all WM neurons biases their resting states, changing their firing rates and the population's ability to maintain information. Finally, we examined the underlying causes of these firing rate changes by manipulating the neurons' biophysical properties to approximate the effects of  $\alpha$ 2A-AR (in)activation.

### 5.1. Functional simulation

To simulate the high-level, functional effects of GFC and PHE on working memory, we multiplied the weights in the WM recurrent connection by a constant  $k_f$ , with the expectation that  $k_f > 1$  would increase feedback and promote remembering, whereas  $k_f < 1$  would increase decay and promote forgetting. Under normal conditions, as the model forgets the original stimulus, the value of  $\hat{c}ue$  decoded from the WM neurons' spiking activity decays exponentially. When we increased the strength of the recurrent connection ( $k_f = 1.03$ ), a higher value of  $\hat{x}$  was fed back as input to the WM population, increasing the firing rate of cue-aligned neurons and more strongly encoding the cue's location. As shown in Fig. 6 (top), the cue representation rose and its exponential decay fell compared to control. This made it easier for the decision procedure to distinguish the decoded cue location from noise, which shifted the "forgetting curve" up; see Fig. 6 (bottom). Conversely, weakening the recurrent connection ( $k_f = 0.985$ ) increased the decay rate and shifted the forgetting curve down. The model's response qualitatively matches the forgetting curves of monkeys injected with these drugs (Mao et al., 1999). Reported results were averaged over  $N = 1,000$  model realizations with randomized cues, neuron properties, and noise.

### 5.2. Neural simulation

Although the functional simulation is conceptually simple and produces a decent empirical match, it is not biologically realistic; GFC and PHE do not transform the physical synaptic connections between neurons. Our hypothesis is that these drugs alter the firing rate of PFC neurons in a way that later manifests functionally as improved or impaired forgetting. To test this, we introduced a global increase (decrease) in somatic current to all WM neurons:  $I_{GFC} = 0.5$  and  $I_{PHE} = -0.2$ , realized as the mean value of noise input from the `Noise_WM` node. Importantly, even though Wang et al. (2007) showed that, *in vivo*, an increase in activity was only observed for neurons whose preferred direction was aligned with the stimulus being remembered, we *do not* apply this extra current only to those neurons. This is because there is no direct mechanism by

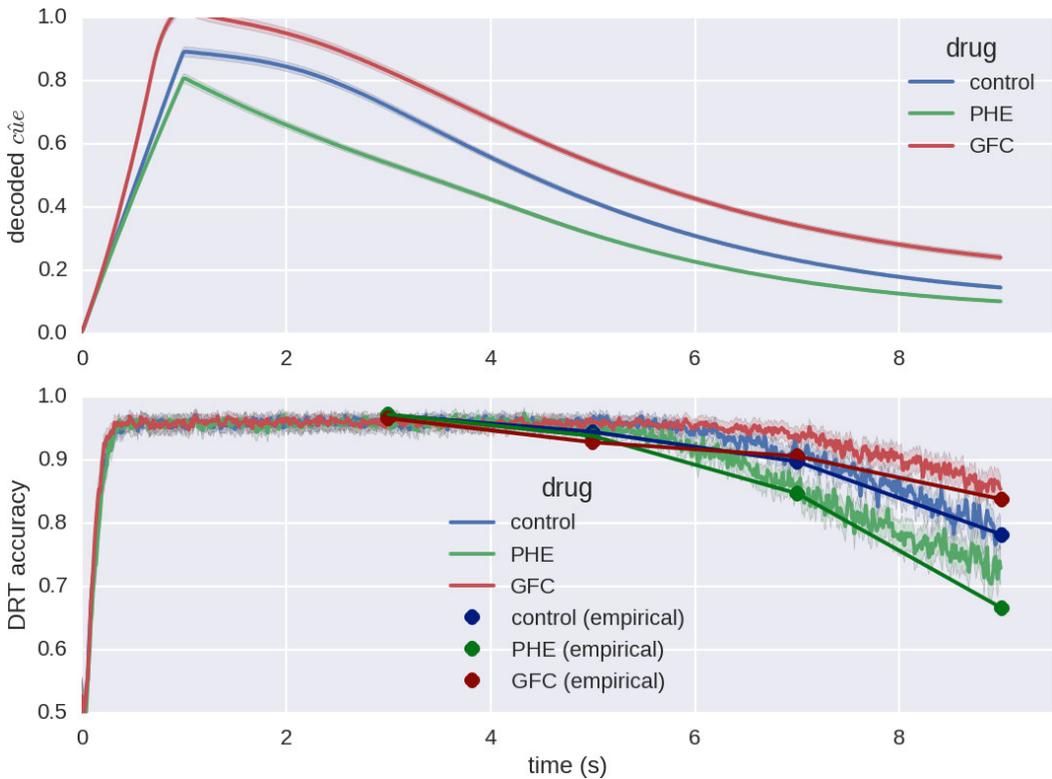


Fig. 6. Model data for the functional drug simulation. Top: decoded *cue* value from spiking neural activity in the working memory (WM) population. Bottom: the “forgetting curve,” or percentage of correct classifications by the decision procedure over  $N = 1,000$  trials as a function of delay period length. Both for monkeys and the model, accuracy decreases steadily from 2 to 6 s, then drops sharply at 8 s when the value stored in the WM become indistinguishable from zero to the model’s noisy decision procedure. Consistent with behavioral data from monkeys performing the delayed response task (DRT), applied guanfacine (GFC) increases task accuracy while phenylephrine (PHE) decreases it. Gray regions represent 95% confidence intervals and sub-100% accuracy before  $t = 3$  s arises from the model’s misperception of the cue.

which GFC or PHE could affect only those neurons that are actively encoding information. Rather, we apply the simulated drug effect to *all* the neurons in the WM model. While this seems counter-intuitive, the network effects of the recurrent connections are sufficient to cause the differential response observed by Mao et al. (1999).

Fig. 7 shows the normalized firing rate of neurons before and after the simulated application of GFC and PHE. As with the empirical data, the neural drug simulation for GFC increased (PHE decreased) the firing rate of simulated preferred-direction neurons while having little effect on neurons in the nonpreferred direction. This differential activation of preferred direction neurons, in turn, allowed the integrator to maintain a coherent representation of the cue’s location for a longer duration, shifting the forgetting curve up. These electrophysiological and behavioral results are consistent both with our functional drug simulation and with empirical data; see Fig. 8.

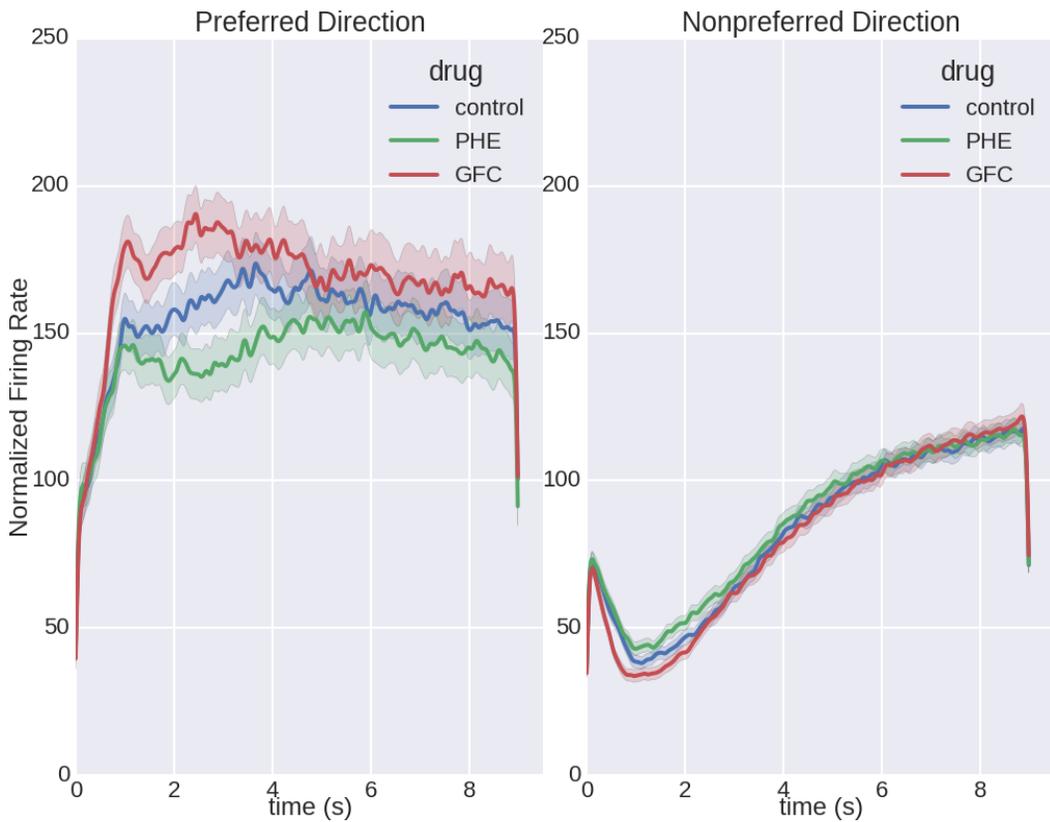


Fig. 7. Firing rate of simulated neurons with encoders in the preferred versus nonpreferred directions in response to injected current. Spikes were generated from  $n = 0-10$  neurons per trial for  $N = 1,000$  trials and then smoothed using Gaussian convolution with  $\sigma = 0.005$  every 0.2 s. We plotted model neurons that were tuned to the preferred direction during control conditions, as per their hypothesized importance in representing the cue's location during the delay period. Wang et al. (2007) failed to provide a precise definition of “weak spatial mnemonic tuning” or their procedure for choosing such neurons, so we selected model neurons based on the magnitude of their encoders ( $0.3 < |e| < 0.6$ ). They also did not discuss their method of calculating “normalized firing rate,” so we did not attempt to fine-tune our working memory neurons' properties to match the absolute rates reported in Fig. 4—only the differences induced by the drugs.

### 5.3. Biophysical simulation

In our final experiment, we approximated GFC and PHE at the biophysical level by altering the inherent properties of model neurons. At rest, HCN channels allow positive ions to flow into the cell, so closing HCN effectively induces a negative current, lowering the resting membrane potential. We modeled this effect by lowering the bias current  $\beta_i$  of each LIF neuron in the WM population. Additionally, closing HCN channels modulates neurons' dendritic summation such that small, desynchronized dendritic spikes more strongly influence the somatic membrane potential. This effectively increases neurons'

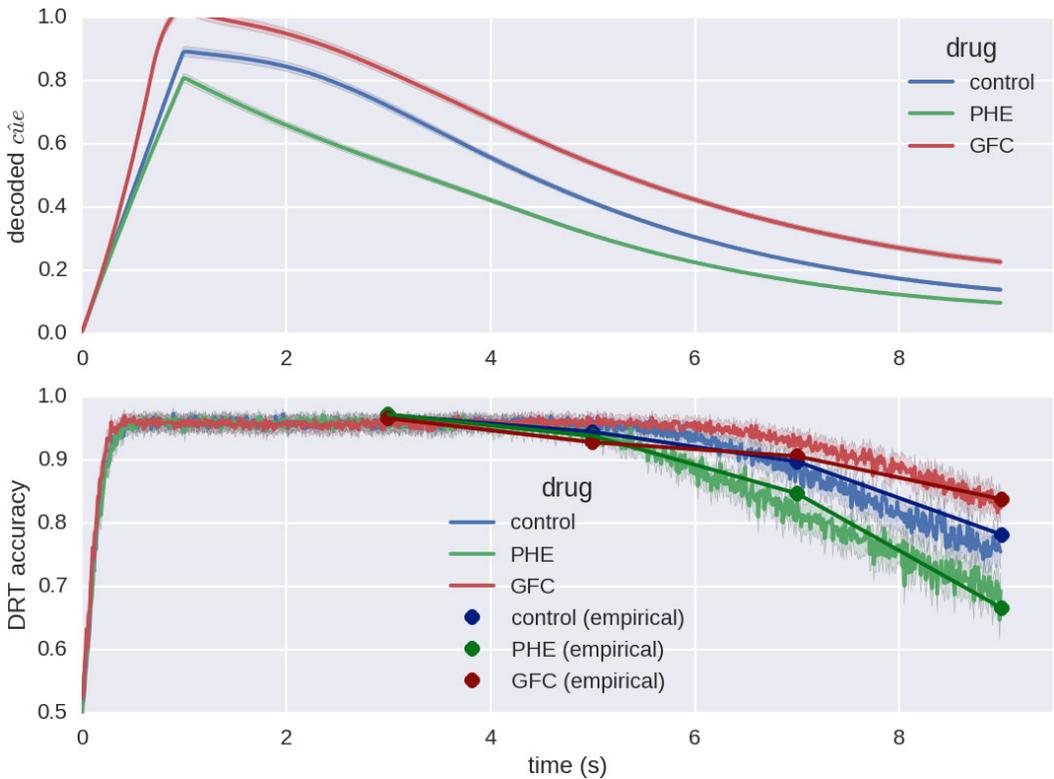


Fig. 8. Cue representation and forgetting curve for the neural drug simulation.

response to a given synaptic input, which we modeled by increasing the gain  $\alpha_i$  of each neuron. This simulation was validated using a study in mice (Nolan et al., 2004) that showed that closing HCN channels decreased neurons' resting membrane potentials and increased their gains in the subthreshold regime; see Fig. 9.

After initializing the neural model to perform least-squares optimal integration (i.e., distributing neurons' initial gains, biases, encoders, and decoders), we multiplied the gains of all WM neurons by a constant  $k_\alpha = 1.05$  for GFC ( $k_\alpha = 0.99$  for PHE), and the biases of all WM neurons by  $k_\beta = 0.95$  ( $k_\beta = 1.02$ ). The impact of these multiplications on the firing rate of example LIF neurons is depicted in Fig. 10; with a constant input current in the physiologically relevant ranges, increases (decreases) in gains overwhelm decreases (increases) in biases, meaning that simulated GFC should increase the overall activity of recurrently connected WM neurons. Fig. 11 confirms that this biophysical simulation reproduces the empirical drug-induced change in PFC neurons' activities. Again, this simulation was applied to all neurons in the WM population, so the network effects from the recurrent connection are responsible for the differential response of preferred versus nonpreferred direction neurons. The biophysical intervention also altered cue encodings in WM and shifted the forgetting curve in a manner consistent with the behavioral data and the previous drug simulations; see Fig. 12.

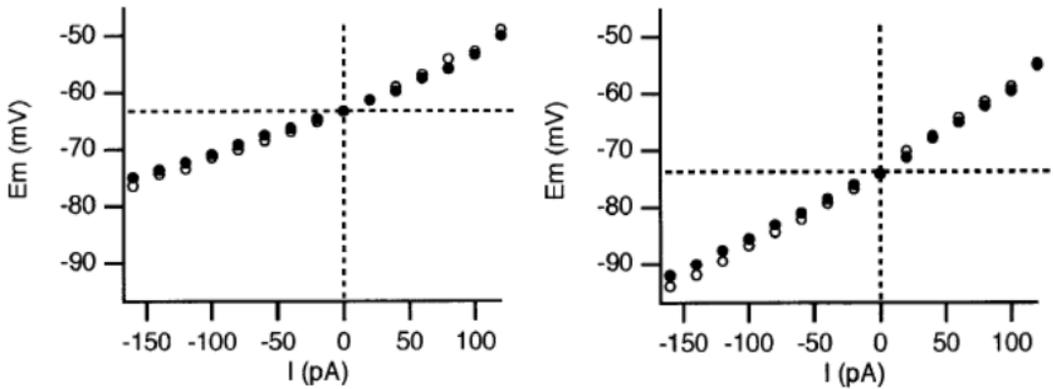


Fig. 9. Subthreshold resting membrane potential as a function of applied current for normal mice (left) versus mice that have had Hyperpolarization-activated Cyclic Nucleotide-gated (HCN) channel genes artificially inactivated (right). Closing HCN channels lowers the neuron’s resting potential (lower value of  $E_m$  at  $I = 0$ ) while increasing the neuron’s response to subsequent input (higher slope of  $E_m$  vs.  $I$ ). Image reproduced from Nolan et al. (2004) (with permission).

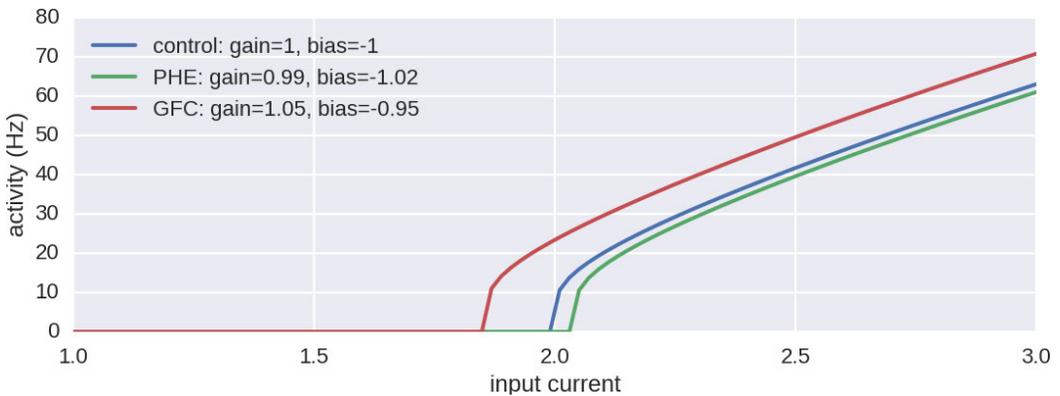


Fig. 10. Tuning curve of a sample leaky integrate-and-fire neuron in response to the biophysical simulation.

Finally, we checked the robustness of our results by replacing the decision procedure with a full basal ganglia network. This network (Stewart & Eliasmith, 2011; Stewart, Choo, & Eliasmith, 2010) has previously been used to simulate several cognitive tasks that require the model to read information stored in working memory, such as action selection and procedure following, and its structure and parameters are biologically plausible. We observed that this replacement changed the shape of the forgetting curve slightly, such that the model had higher accuracy for delay periods of 6 s or less, but a sharper dropoff in accuracy afterward; see Fig. 13. However, this curve still shifts up and down with the application of simulated GFC and PHE and aligns reasonably with the empirical forgetting curve.

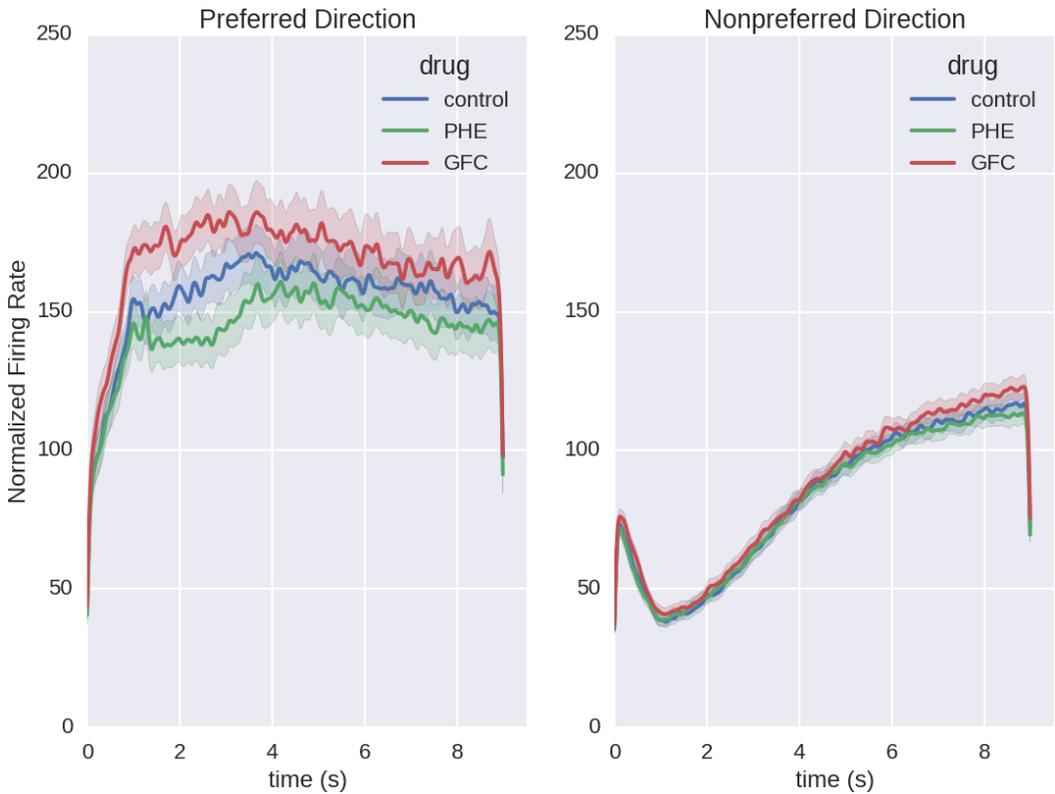


Fig. 11. Firing rate of simulated neurons for the biophysical drug simulation.

## 6. Discussion and conclusion

In this paper, we presented a spiking neuron model of WM and the DRT, and then used this model to investigate the underlying causes of WM disorders and their treatments through the simulated application of GFC and PHE. The model extends classical works on WM dynamics (Brunel & Wang, 2001) by incorporating the NEF, an approach that allows for (a) the principled encoding and decoding of information in large-scale spiking neural networks, and (b) the manipulation of these networks at levels ranging from the biophysical to the functional. We investigated the interactions between WM and two drugs that reduce and enhance WM deficits in ADHD, showing that these interactions could be explained from a functional, neural, and biophysical perspective. We demonstrated that three distinct drug simulations, each computationally realizing one of these perspectives by perturbing a different part of the model, all produce surprisingly similar, and empirically accurate, effects on electrophysiology and task performance. This result unifies these seemingly disparate descriptions of the drugs' interaction with WM systems and was robust to the chosen decision procedure.

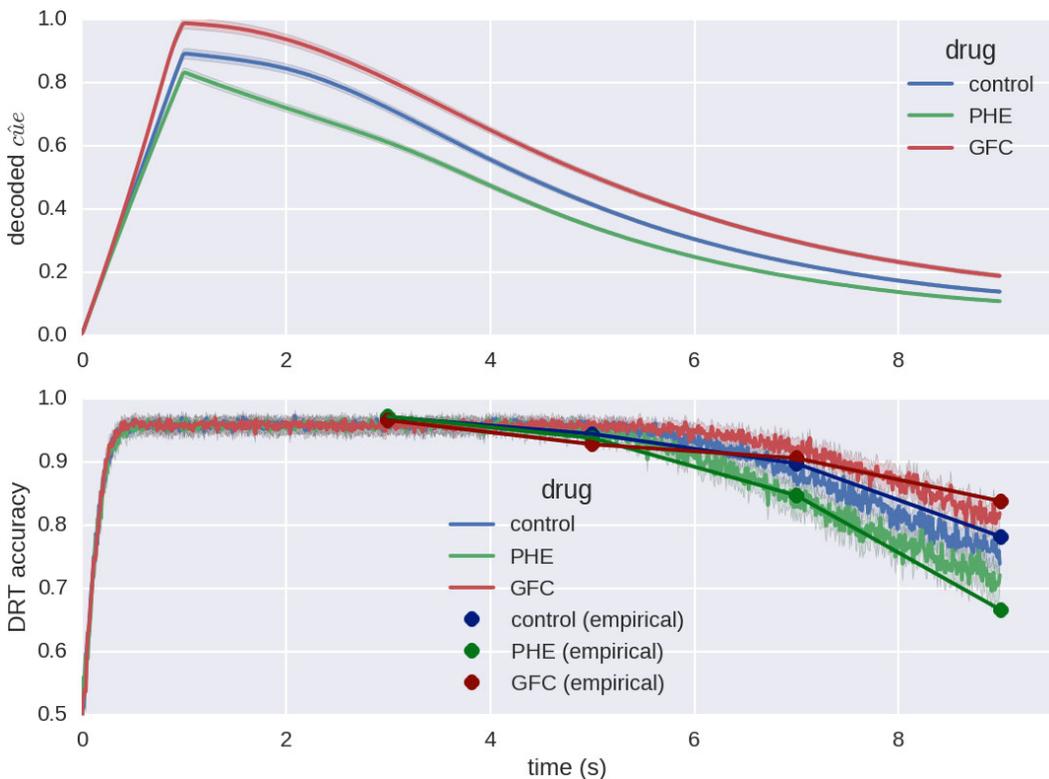


Fig. 12. Cue representation and forgetting curve for the biophysical drug simulation.

The fact that our model successfully reproduces changes in WM retention times via these three different levels of simulation gives strong evidence that the mechanisms being investigated here (i.e., the strength of synaptic connections within a population of recurrently connected spiking neurons) form the underlying basis of WM in mammals. Though further work is needed to confirm and explore this result, we believe it provides a concrete grounding for high-level WM theories and lays the foundation for future investigations into other aspects of WM. For example, Choo and Eliasmith (2010) use a similar model (without the GFC or PHE interactions) to model serial recall accuracy over lists of different lengths, providing a neural explanation for the primacy and recency effects. Because these models are implemented using compatible (and biologically plausible) spiking neuron models, they can easily be integrated into larger unified models.

Future work will address several simplifying assumptions made in this study. First, we had to approximate the effects of HCN channel opening and closing on LIF point neurons; this was possible largely because existing work had previously classified the relationship between GFC application and firing rate (Wang et al., 2007). Though our approximations were successful in reproducing the empirical data, replacing LIF neurons with biologically detailed neurons that include explicit ion channels (that can be closed or opened by drug interactions) would expand the range of biochemical processes we

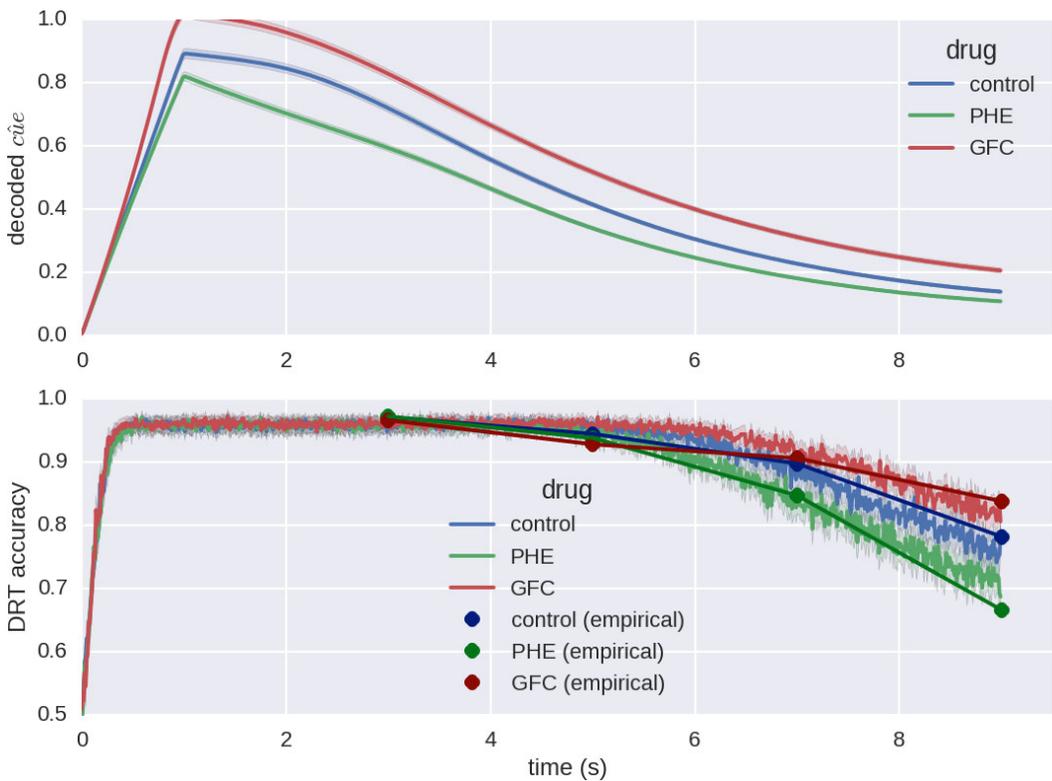


Fig. 13. Cue representation and forgetting curve for the biophysical drug simulation with the original decision population replaced by a full basal ganglia network.

could simulate without detailed foreknowledge. To progress in this direction, we are currently integrating the NEURON simulation package and the Hodgkin–Huxley-type neuron models developed by Bahl, Stemmler, Herz, and Roth (2012) and the Human Brain Project (Markram et al., 2015) with the NEF-style modeling performed here.

Second, while our model focused on the representational and dynamic aspects of WM, the processes by which information is placed in, and retrieved from, WM are equally important for its implementation in unified cognitive systems. Adding additional cognitive modules to the model would greatly expand the range of cognitive tasks that we could simulate, as well as present new neural targets for drugs that affect different aspects of cognition. For example, dopamine (D1) receptors are present both in PFC and hippocampus, and abnormal neurotransmitter/receptor levels have been implicated in WM deficits related to Parkinson’s and schizophrenia (Goldman-Rakic, 1995). Many of these systems have already been built using the NEF and applying the above methods for detailed drug simulation would be straightforward (Stewart & Eliasmith, 2011). In future work, we plan to implement these extensions on the world’s largest functional brain model, SPAUN (Eliasmith et al., 2012), in pursuit of a deeper understanding of the neural basis, psychological dysfunction, and pharmaceutical modulation of working memory.

## Acknowledgments

This work was supported by CFI and OIT infrastructure funding as well as the Canada Research Chairs program, NSERC Discovery grant 261453, ONR grant N000141310419, AFOSR grant FA8655-13-1-3084, and OGS graduate funding.

## References

- Arnsten, A., & Leslie, F. (1991). Behavioral and receptor binding analysis of the  $\alpha$  2-adrenergic agonist, 5-bromo-6 [2-imidazoline-2-yl amino] quinoxaline (uk-14304): Evidence for cognitive enhancement at an  $\alpha$  2-adrenoceptor subtype. *Neuropharmacology*, *30*(12), 1279–1289.
- Arnsten, A. F., & Lombroso, P. J. (2000). Genetics of childhood disorders: Xviii. adhd, part 2: Norepinephrine has a critical modulatory influence on prefrontal cortical function. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*(9), 1201–1203.
- Avery, R. A., Franowicz, J. S., Studholme, C., van Dyck, C. H., & Arnsten, A. F. (2000). The alpha-2a-adrenoceptor agonist, guanfacine, increases regional cerebral blood flow in dorsolateral prefrontal cortex of monkeys performing a spatial working memory task. *Neuropsychopharmacology*, *23*(3), 240–249.
- Bahl, A., Stemmler, M. B., Herz, A. V., & Roth, A. (2012). Automated optimization of a reduced layer 5 pyramidal cell model based on experimental data. *Journal of Neuroscience Methods*, *210*(1), 22–34.
- Bekolay, T., Laubach, M., & Eliasmith, C. (2014). A spiking neural integrator model of the adaptive control of action by the medial prefrontal cortex. *The Journal of Neuroscience*, *34*(5), 1892–1902.
- Brunel, N., & Wang, X.-J. (2001). Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *Journal of Computational Neuroscience*, *11*(1), 63–85.
- Chandler, D. J., Waterhouse, B. D., & Gao, W.-J. (2014). New perspectives on catecholaminergic regulation of executive circuits: evidence for independent modulation of prefrontal functions by midbrain dopaminergic and noradrenergic neurons. *Frontiers in Neural Circuits*, *8*, 1–10.
- Choo, F.-X., & Eliasmith, C. (2010). A spiking neuron model of serial-order recall. In S. Ohlsson, & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 2188–2193). Austin, TX: Cognitive Science Society.
- Dancy, C. L., Ritter, F. E., Berry, K. A., & Klein, L. C. (2015). Using a cognitive architecture with a physiological substrate to represent effects of a psychological stressor on cognition. *Computational and Mathematical Organization Theory*, *21*(1), 90–114.
- Eliasmith, C., & Anderson, C. H. (2003). *Neural engineering: Computation, representation, and dynamics in neurobiological systems*. Cambridge, MA: MIT Press.
- Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., & Rasmussen, D. (2012). A large-scale model of the functioning brain. *Science*, *338*(6111), 1202–1205.
- Franowicz, J. S., Kessler, L. E., Borja, C. M. D., Kobilka, B. K., Limbird, L. E., & Arnsten, A. F. (2002). Mutation of the  $\alpha$ 2a-adrenoceptor impairs working memory performance and annuls cognitive enhancement by guanfacine. *The Journal of Neuroscience*, *22*(19), 8771–8777.
- Georgopoulos, A. P., Kalaska, J. F., Caminiti, R., & Massey, J. T. (1982). On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *The Journal of Neuroscience*, *2*(11), 1527–1537.
- Goldman-Rakic, P. (1995). Cellular basis of working memory. *Neuron*, *14*(3), 477–485.
- Gunzelmann, G., Gross, J. B., Gluck, K. A., & Dinges, D. F. (2009). Sleep deprivation and sustained attention performance: Integrating mathematical and cognitive modeling. *Cognitive Science*, *33*(5), 880–910.
- Levy, F. (2008). Pharmacological and therapeutic directions in adhd: Specificity in the pfc. *Behavioral and Brain Functions*, *4*(1), 1.

- Lewis, P. A., & Miall, R. C. (2006). Remembering the time: A continuous clock. *Trends in Cognitive Sciences*, 10(9), 401–406.
- Magee, J. C. (1999). Dendritic Ih normalizes temporal summation in hippocampal CA1 neurons. *Nature Neuroscience*, 2(6), 508–514.
- Mao, Z.-M., Arnsten, A. F., & Li, B.-M. (1999). Local infusion of an  $\alpha$ -1 adrenergic agonist into the prefrontal cortex impairs spatial working memory performance in monkeys. *Biological Psychiatry*, 46 (9), 1259–1265.
- Markram, H., Muller, E., Ramaswamy, S., Reimann, M. W., Abdellah, M., Sanchez, C. A., ... et al. (2015). Reconstruction and simulation of neocortical microcircuitry. *Cell*, 163(2), 456–492.
- Nolan, M. F., Malleret, G., Dudman, J. T., Buhl, D. L., Santoro, B., Gibbs, E., ... et al. (2004). A behavioral role for dendritic integration: HCN1 channels constrain spatial memory and plasticity at inputs to distal dendrites of CA1 pyramidal neurons. *Cell*, 119(5), 719–732.
- Poolos, N. P., Migliore, M., & Johnston, D. (2002). Pharmacological upregulation of h-channels reduces the excitability of pyramidal neuron dendrites. *Nature Neuroscience*, 5(8), 767–774.
- Ramos, B. P., Stark, D., Verduzco, L., van Dyck, C. H., & Arnsten, A. F. (2006).  $\alpha$ 2a-adrenoceptor stimulation improves prefrontal cortical regulation of behavior through inhibition of cAMP signaling in aging animals. *Learning & Memory*, 13(6), 770–776.
- Ritter, F. E., Bittner, J. L., Kase, S. E., Everts, R., Pedrotti, M., & Busetta, P. (2012). Cojack: A high-level cognitive architecture with demonstrations of moderators, variability, and implications for situation awareness. *Biologically Inspired Cognitive Architectures*, 1, 2–13.
- Romo, R., Brody, C. D., Hernández, A., & Lemus, L. (1999). Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature*, 399(6735), 470–473.
- Scahill, L., Chappell, P. B., Kim, Y. S., Schultz, R. T., Katsovich, L., Shepherd, E., ... Leckman, J. F. (2014). A placebo-controlled study of guanfacine in the treatment of children with tic disorders and attention deficit hyperactivity disorder. *American Journal of Psychiatry*, 158(7), 1067–1074.
- Shenoy, K. V., Sahani, M., & Churchland, M. M. (2013). Cortical control of arm movements: a dynamical systems perspective. *Neuroscience*, 36(1), 337.
- Singh, R., & Eliasmith, C. (2006). Higher-dimensional neurons explain the tuning and dynamics of working memory cells. *The Journal of Neuroscience*, 26(14), 3667–3678.
- Stewart, T. C., & Eliasmith, C. (2011). Neural cognitive modelling: A biologically constrained spiking neuron model of the tower of hanoi task. In S. Ohlsson, & R. Cattrambone (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 656–661). Austin, TX: Cognitive Science Society.
- Stewart, T. C., Choo, X., & Eliasmith, C. (2010). Dynamic behaviour of a spiking model of action selection in the basal ganglia. In D. D. Salvucci, & G. Gunzelmann (Eds.), *Proceedings of the 10th international conference on cognitive modeling* (pp. 235–240). Philadelphia, PA: Drexel University.
- Wang, M., Ramos, B. P., Paspalas, C. D., Shu, Y., Simen, A., Duque, A., ... et al. (2007).  $\alpha$ 2a-adrenoceptors strengthen working memory networks by inhibiting cAMP-HCN channel signaling in prefrontal cortex. *Cell*, 129(2), 397–410.