Opening the gate to working memory

David Badre[a,b,1]

[a]Department of Cognitive, Linguistic, and Psychological Sciences and [b]Brown Institute for Brain Science, Brown University, Providence, RI 02912

In a changing and complex world, we require the ability to sustain pursuit of a goal, while also updating our goals as the need arises. For example, we are capable of rapidly shifting tasks, such as stopping work on a paper to have an involved discussion with a colleague. However, we are also able to avoid distracting cues, such as ignoring the unread e-mail messages incrementing in our inbox while we write. The mechanisms by which the brain achieves an adaptive balance between flexibility and stability remain the basis of much current investigation in cognitive neuroscience. In PNAS, D'Ardenne et al. (1) provide evidence for two key pieces to this puzzle. They demonstrate that the dorsolateral prefrontal cortex (DLPFC) is necessary for flexible updating of contexts to control behavior and that activity in cells of the dopaminergic midbrain may signal an updating response.

One solution to the flexibility vs. stability paradox has been to hypothesize a working-memory "gate." Such models assume that working memory, supported by PFC, maintains contextual information to modulate thought and action (2). Critical to such a system is a mechanism by which useful contextual information is updated into working memory and distracting information is kept out; in other words, a working-memory gate. When the gate is open, available information can enter working memory. When the gate is closed, the current contents of working memory are sustained, while irrelevant information is kept out.

Having separate maintenance and gating mechanisms is computationally efficient (3). And, although there are proposals of general form recurrent networks with diverse neural responses (akin to those in PFC) that can exhibit self-gating behavior (4), the most influential models of working-memory updating assume separate gating and maintenance mechanisms (5, 6). Specifically, the PFC is proposed to support noise-resistant maintenance, such as via recurrence in a network of active neurons (7). Gating is achieved by delivery of a brief stimulation that can elicit hysteresis in the resting network or make an active ensemble available to new inputs.

A prominent example of such a mechanism was proposed by Braver and Cohen (5). In their model, encounter with contextual information elicits phasic delivery of the neurotransmitter dopamine (DA) from the ventral tegmental area (VTA) and substantia nigra (SN) to the PFC. Thus, DA acts as a gating signal, timing when the PFC accepts new input. In this model, the SN/VTA learns when to gate from the positive or negative outcomes that follow gating.

Several aspects of the DA-updating hypothesis are supported empirically. For

> ## D'Ardenne et al. provide an important step forward in our understanding of working-memory updating.

example, it is established that the PFC is necessary for controlled behavior (2, 8) and that PFC neurons exhibit delay period activity (9, 10) and selectivity to task rules (11). Similarly, evidence from functional MRI (fMRI) has implicated DLPFC in gating responses (12, 13).

Moreover, working-memory performance has been linked to DA. Midbrain DA cells fire transiently to cues during working-memory encoding (14). In PFC, extracellular DA levels increase during working-memory tasks (15), and performance suffers following DA depletion (16). In general, higher DA levels in PFC drive stability over distractibility (17). However, no evidence shows that DLPFC is necessary for flexible updating of context information triggered by phasic responses in SN/VTA. Using a combined-methods approach, D'Ardenne et al. (1) provide evidence to fill this fundamental gap.

In three experiments, human participants saw pairs of letters presented sequentially and responded on the basis of the identity of the letters. The first letter of each pair was the cue and the second letter was the probe. For some trials, the appropriate response entirely depended on the probe identity. These "context-independent" trials did not require maintenance of the cue as a context in working memory. In other trials, the appropriate response to a probe was conditioned on the identity of the preceding cue. Thus, in these "context-dependent" trials, the cue was encoded as a context.

Following this logic, comparison of context-dependent vs. context-independent cues provided a means of testing working-memory updating. First, in fMRI, greater bilateral DLPFC activation accompanied context-dependent vs. context-independent cues. Next, single-pulse transcranial magnetic stimulation (TMS) was applied to each participant's DLPFC in a locus that showed that participant's maximal updating effect. When the TMS pulse was applied to right DLPFC 150 ms after onset of the cue, subsequent performance was disrupted in context-dependent trials. This timing aligns with a previous event-related potential study of context updating (18). Disruption was not observed following stimulation of left DLPFC or stimulation of right DLPFC at offsets earlier than 150 ms or when encoding of the context was not required (context-independent trials). Thus, right DLPFC is necessary for context updating.

In a third experiment, D'Ardenne et al. (1) provide evidence for a link between SN/VTA and DLPFC updating. Using a high-resolution fMRI technique, greater activation for context-dependent than for context-independent conditions was again observed in right DLPFC, but also in the SN/VTA. Moreover, the activation in SN/VTA correlated with behavior and with the activation in right DLPFC. Thus, these observations provide support for the dopamine-updating hypothesis in that not only is SN/VTA responsive to the presentation of stimuli that predict reward, but also its activity is coordinated with frontal regions necessary for updating.

D'Ardenne et al. (1) provide an important step forward in our understanding of working-memory updating. Of course, an interesting result also raises a number of open questions. For example, (i) Is updating global or selective? (ii) Are there multiple gating mechanisms? (iii) What is the relationship between gating and the functional organization of frontal cortex? (iv) What is the reason for the right laterality of the effects observed by D'Ardenne et al. (1)? (v) What is the nature of the interaction between episodic and working-memory during updating? And (vi) what

is the relationship between the updating function studied here and working-memory capacity? In what follows, the first two of these questions are elaborated further, as they relate to the nature of updating.

Is Working-Memory Updating Global or Selective?

A potential limitation of assuming a global updating signal, such as that arising from SN/VTA (5), is that it updates everything in working memory. There are no means to selectively hold one contextual representation in working memory while updating a second one. However, selective updating of this type is likely adaptive in a number of circumstances, and so it is important to know whether PFC gating can be selective or whether tasks requiring selective updating must be solved in other ways (such as by relying on episodic memory).

At least one class of computational models, such as those of O'Reilly and Frank (6), highlights the striatum in delivering selective gating inputs to PFC. These models assume that, through elaborated frontostriatal-thalamic loops (19), striatum can update recurrent activity in organized ensembles of frontal neurons, while information in other ensembles is sustained. DA and SN/VTA play an im-

portant role in these models, as well. However, here SN/VTA is crucial for training the striatal gating responses. Thus, these models separate learning and gating into separate components.

The results of D'Ardenne et al. (1) are consistent with both global and selective gating models. However, several other recent results have highlighted the importance of striatum in flexibility and stability in working memory (20). Also, recent fMRI data have provided support for selective gating in separate regions of frontal cortex based on context abstraction, implicating striatum but also parietal cortex in gating (21). Thus, it remains open whether gating within working memory is selective or global and what systems might support one or the other.

Are There Multiple Gates on Working Memory?

Gating can involve more than just deciding what information to allow into working memory. Indeed, it may be the case that we also want to select which maintained representation is permitted to have an influence over behavior. Thus, it may be useful to distinguish gathering contextual information, "input gating," from allowing the information resident in working

memory to influence behavior, "output gating." Given its advantages, some computational models of cognitive control include input and output gating (6, 22). For example, in a model of abstract rule learning, output gating allowed a hierarchical corticostriatal network to gain a learning advantage over a nonhierarchical network and better matched human participants (23). By contrast, a hierarchical network that leveraged input gating did not show this benefit (also see ref. 24). Notably, this nested output gating model (23) predicted an organization of frontostriatal connectivity that has been partly supported by fMRI (25) and by diffusion tractography (26). However, the empirical evidence in brain and behavior for output gating of working memory is primarily indirect (27, 28). Thus, it remains open whether output gating is supported by a similar DA gating system to that of input gating.

D'Ardenne et al. (1) provide important evidence regarding the neural substrates of working-memory updating and the importance of interactions between DLPFC and the SN/VTA for this function. These results provide a basis for new investigations into the neural mechanisms of flexible, goal-directed behavior.

1. D'Ardenne K, et al. (2012) Role of prefrontal cortex and the midbrain dopamine system in working memory updating. *Proc Natl Acad Sci USA* 109:19900–19909.
2. Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202.
3. Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780.
4. Rigotti M, Ben Dayan Rubin D, Wang XJ, Fusi S (2010) Internal representation of task rules by recurrent dynamics: The importance of the diversity of neural responses. *Front Comput Neurosci* 4:24.
5. Braver TS, Cohen JD (2000) On the control of control: The role of dopamine in regulating prefrontal function and working memory. *Attention and Performance XVIII*, eds Monsell S, Driver J (MIT Press, London), pp 713–737.
6. O'Reilly RC, Frank MJ (2006) Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput* 18(2):283–328.
7. Durstewitz D, Seamans JK, Sejnowski TJ (2000) Neurocomputational models of working memory. *Nat Neurosci* 3(Suppl):1184–1191.
8. Stuss DT, Benson DF (1987) The frontal lobes and control of cognition and memory. *The Frontal Lobes Revisited*, ed Perecman E (IRBN Press, New York), pp 141–158.
9. Fuster JM, Alexander GE (1971) Neuron activity related to short-term memory. *Science* 173(3997):652–654.
10. Goldman-Rakic PS (1995) Cellular basis of working memory. *Neuron* 14(3):477–485.

11. Wallis JD, Anderson KC, Miller EK (2001) Single neurons in prefrontal cortex encode abstract rules. *Nature* 411(6840):953–956.
12. Postle BR (2005) Delay-period activity in the prefrontal cortex: One function is sensory gating. *J Cogn Neurosci* 17(11):1679–1690.
13. Nee DE, Brown JW (2012) Dissociable frontal-striatal and frontal-parietal networks involved in updating hierarchical contexts in working memory. *Cereb Cortex*, in press.
14. Schultz W, Apicella P, Ljungberg T (1993) Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J Neurosci* 13(3):900–913.
15. Watanabe M, Kodama T, Hikosaka K (1997) Increase of extracellular dopamine in primate prefrontal cortex during a working memory task. *J Neurophysiol* 78(5):2795–2798.
16. Brozoski TJ, Brown RM, Rosvold HE, Goldman PS (1979) Cognitive deficit caused by regional depletion of dopamine in prefrontal cortex of rhesus monkey. *Science* 205(4409):929–932.
17. Cools R (2008) Role of dopamine in the motivational and cognitive control of behavior. *Neuroscientist* 14(4):381–395.
18. Lenartowicz A, Escobedo-Quiroz R, Cohen JD (2010) Updating of context in working memory: An event-related potential study. *Cogn Affect Behav Neurosci* 10(2):298–315.
19. Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits link-

ing basal ganglia and cortex. *Annu Rev Neurosci* 9:357–381.
20. Cools R (2011) Dopaminergic control of the striatum for high-level cognition. *Curr Opin Neurobiol* 21(3):402–407.
21. Nee DE, Brown JW (2012) Rostral-caudal gradients of abstraction revealed by multi-variate pattern analysis of working memory. *Neuroimage* 63(3):1285–1294.
22. Kriete T, Noelle DC (2011) Generalization benefits of output gating in a model of prefrontal cortex. *Connect Sci* 23(2):119–129.
23. Frank MJ, Badre D (2012) Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: Computational analysis. *Cereb Cortex* 22(3):509–526.
24. Reynolds JR, O'Reilly RC (2009) Developing PFC representations using reinforcement learning. *Cognition* 113(3):281–292.
25. Badre D, Frank MJ (2012) Mechanisms of hierarchical reinforcement learning in cortico-striatal circuits 2: Evidence from FMRI. *Cereb Cortex* 22(3):527–536.
26. Verstynen TD, Badre D, Jarbo K, Schneider W (2012) Microstructural organizational patterns in the human corticostriatal system. *J Neurophysiol* 107(11):2984–2995.
27. Farrell S, Lelièvre A (2012) The dynamics of access to groups in working memory. *J Exp Psychol Learn Mem Cogn* 38(6):1659–1674.
28. Kiyonaga A, Egner T, Soto D (2012) Cognitive control over working memory biases of selection. *Psychon Bull Rev* 19(4):639–646.