



## How cognitive theory guides neuroscience



Michael J. Frank\*, David Badre

Cognitive, Linguistic & Psychological Sciences, Brown Institute for Brain Science, United States

### ARTICLE INFO

#### Article history:

Available online 8 December 2014

#### Keywords:

Neuroscience  
Computational models  
Memory  
Cognitive control  
Decision making

### ABSTRACT

The field of cognitive science studies latent, unobservable cognitive processes that generate observable behaviors. Similarly, cognitive neuroscience attempts to link latent cognitive processes with the neural mechanisms that generate them. Although neural processes are partially observable (with imaging and electrophysiology), it would be a mistake to 'skip' the cognitive level and pursue a purely neuroscientific enterprise to studying behavior. In fact, virtually all of the major advances in understanding the neural basis of behavior over the last century have relied fundamentally on principles of cognition for guiding the appropriate measurements, manipulations, tasks, and interpretations. We provide several examples from the domains of episodic memory, working memory and cognitive control, and decision making in which cognitive theorizing and prior experimentation has been essential in guiding neuroscientific investigations and discoveries.

© 2014 Elsevier B.V. All rights reserved.

### 1. Introduction

Mental operations emerge from interactions among large populations of neurons and interconnected brain systems. Neuroscientists leverage the principles of physical reductionism and reconstructionism to explain not only the role of individual elements (neurons, ion channels, receptors, etc.) but also how these interact in a dynamical system with emergent properties that drive cognition and behavior. Causal manipulations of underlying circuits (with lesions, pharmacology, optogenetics, etc.) allow researchers to study the mechanisms required for cognition, by observing predictable and selective changes in relevant cognitive measures. This characterization might lead one to think that neuroscience experts can go along their merry way discovering the principles that explain the mind without the help from cognitive scientists or cognitive theory since one level encompasses or 'explains'

the other. But *Cognition* readers can stop writhing in their seats; of course, this is not the case.

In this article we elaborate concrete examples articulating how principles of cognition – in particular, computational tradeoffs identified by studying functional requirements at the cognitive level – have, and will continue to be, instrumental in guiding neuroscientific discoveries. Neuroscience is rapidly accumulating a wealth of data at multiple levels ranging from molecules to cells to circuits to systems. However, in the absence of cognitive theory, this effort runs the risk of mere "stamp collecting", or the tendency to catalog the phenomena of the brain without gaining understanding or explanation. It follows, then, that many of the most influential findings in neuroscience have been understood within the functional context of cognitive theory. We focus on three examples: episodic memory, working memory and cognitive control, and decision making. In each case, we articulate how cognitive theory has set the stage to constrain measurements and manipulations which have advanced the neuroscientific enterprise. Thus, our primary focus in this review concerns how cognition has influenced neuroscience. The converse case, namely the influence that neuroscience can have on cognitive

\* Corresponding author.

E-mail addresses: [Michael\\_Frank@Brown.edu](mailto:Michael_Frank@Brown.edu) (M.J. Frank), [David\\_Badre@Brown.edu](mailto:David_Badre@Brown.edu) (D. Badre).

theory, is an important topic that we have each dealt with in detail elsewhere (see Chatham, Badre, & Badre, *in press-a*; Frank, *in press*, both of which emphasize the role of modeling endeavors that bridge across levels of analysis). However, for some examples, we also briefly note how reciprocally taking neuroscientific constraints into account has validated or refined cognitive models.

## 2. Hippocampus and functional tradeoffs in memory

The hippocampal formation has long been a focus of neuroscientists investigating its distinguishing anatomical and electrophysiological properties. Importantly, however, the progress of neuroscientific study of the hippocampus has been closely and continuously intertwined with cognitive theory regarding its widely celebrated role in episodic memory.

Henry Molaison, the famous patient H.M., had widespread hippocampal damage and exhibited profound episodic memory deficits, characterized by anterograde and retrograde amnesia (Scoville & Milner, 1957). But H.M.'s case was particularly compelling because of what he was still capable of learning. For example, he could acquire and retain complex motor skills, all while having no explicit memory of ever having performed these tasks. These results provided the strongest evidence to that time for the existence of multiple memory systems. However, these early investigations arose in a prevailing context of cognitive theory that already hinted at the existence of distinct forms of knowledge. Indeed, in her seminal paper on motor skill learning by H.M., Corkin (1968) motivated the investigation with "observations in normal man" that motor and other forms of memory were distinct, explicitly citing distinctions drawn in cognitive psychology between visual and kinesthetic codes (Posner, 1966, 1967) and verbal versus non-verbal forms of memory (e.g., McGeoch & Melton, 1929). This theoretical framing of H.M. grounded in cognitive theory led to a generation of investigations by neuropsychologists and cognitive neuroscientists studying multiple memory systems and their neural underpinnings (Cohen, Poldrack, & Eichenbaum, 1997; Squire, 1992), and influenced synaptic physiologists attempting to uncover the cellular basis of learning and memory in the hippocampus (Bliss & Lomo, 1973).

Cognitive theory has not only framed and motivated new neuroscience investigations, it also provides a functional level of analysis that motivates a deeper investigation as to why the brain evolved to support distinct systems. For example, computational cognitive modeling has indicated that multiple memory systems may be required to confront the functional tradeoffs between memory processes required to remember "where did I park my car today?" versus "what has been on average the best place to park my car?" (O'Reilly and Norman, 2002). The former question requires keeping overlapping memories separate, so as to not mistake yesterday's parking spot for today's, whereas the latter question requires an integration of many previous parking experiences into a coherent representation linked to its average value (McClelland, McNaughton & O'Reilly, 1995; Norman & O'Reilly, 2003; O'Reilly and Rudy, 2001). A single system

would have difficulty handling both of these functions, leading neuroscientists to conduct studies providing evidence that the hippocampus supports distinct memories for overlapping events, whereas the cortex and basal ganglia are well suited to represent similarities among these memories and to integrate their reward values across time. This complementary learning systems (CLS) perspective accounts for existing lesion studies (Myers et al., 2003; Squire & Knowlton, 1995) and motivated experiments involving pharmacological manipulations differentially affecting these systems, and imaging to identify their neural signatures, combined with the requisite cognitive manipulations for uncovering their dissociable effects (Curran, DeBuse, Woroch, & Hishman, 2006; Frank, O'Reilly, & Curran, 2006; Huffman & Stark, 2014). Finally, the seminal work of Tolman, who suggested that rats use cognitive representations to map space and plan behavioral actions, and Hull, who argued for habit-like stimulus-response learning are both encompassed within the CLS framework and have directly informed modern neuroscientific investigations showing that these strategies co-exist and tradeoff against each other between distinct hippocampal and striatal networks in rodents (Johnson & Redish, 2007; Packard & McGaugh, 1996; van der Meer, Johnson, Schmitzer-Torbert, & Redish, 2010) and humans (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Poldrack & Packard, 2003). Nevertheless, whether these systems truly compete or collaborate seems to depend on task demands (Sadeh, Shohamy, Levy, Reggev, & Maril, 2011) and further cognitive theory may be useful to resolve this controversy.

The 2014 Nobel prize in physiology and medicine was awarded to three neuroscientists (John O'Keefe, May-Britt Moser and Edvard Moser) for their work on the hippocampus. The prize was awarded for the discoveries of hippocampal "place cells" that fire when an animal visits a particular location, and "grid cells" that fire in a grid-like fashion across multiple locations in an environment with a particular spatial frequency. Importantly, the impact of these discoveries, and what elevated their influence to the highest levels in science, was their straightforward relationship to cognitive theory regarding cognitive maps. This relationship was first proposed in the seminal book by O'Keefe and Nadel (1978). Notably, this book takes as its starting point two detailed chapters focusing on cognitive theory regarding the nature of space, physical and psychological, and its ubiquitous role in memory, followed by a review of the principles and studies of navigation. Only after more than 100 pages do the authors begin to discuss the anatomy and physiology of the hippocampus in the context of this literature. To the present day, theorizing regarding the hippocampus and the role of place and grid cells has focused on spatial codes and their ability to bind separate elements, provide linking contexts, and naturally encode relationships among distinct features. Indeed, this general property of binding to location (spatial or virtual) is fundamental to many functions, and has led to a broadened view of the function of the hippocampus beyond declarative memory, as highlighted in a recent special issue of *Journal of Experimental Psychology: General* (Vol. 142, No. 4) devoted to the topic.

Cognitive scientists have also developed models of memory at varying levels of abstraction that point to functional tradeoffs between memory processes that constrain processing in hippocampus. One such trade off concerns how a memory system that rapidly encodes and retrieves individual episodes knows whether to treat partly overlapping events as distinct so as to be stored as separate memories ("pattern separation"), or to identify that an experienced event is similar enough to a previous episode that it should be used as a cue to recall other aspects of that event ("pattern completion"). This distinction was first popularized by Marr (1971) and later expanded by several detailed neurocomputational models proposing that interactions among distinct subregions within the hippocampus serve to accomplish these two functions (e.g., O'Reilly and McClelland, 1994). In these models, the dentate gyrus (DG) supports sparse representations with fierce inhibitory competition that encourages pattern separation, and biases downstream representations in area CA3 involved in memory storage to be separated as well. However, in the absence of strong DG input, recurrent collaterals within area CA3 support pattern completion based on partial cues coming from entorhinal cortex. Thus, pattern separation and pattern completion are balanced in CA3 via these circuit interactions, such that as inputs become more similar to previously stored representations, CA3 transitions from pattern separation to pattern completion.

Modern neuroscience and neurogenetic methods have tested these specific roles and provided striking evidence for them (Yassa & Stark, 2011). Transgenic mice with particular hippocampal pathways spared or destroyed have shown that keeping the DG in the loop is required for rapid, one-shot learning of conjunctive representations (requiring pattern separation) but not for incremental spatial learning (Nakashiba, Young, McHugh, Buhl, & Tonegawa, 2008). Slight changes in a spatial environment (i.e., greater input pattern overlap) give rise to highly distinct populations of correlated activity in the DG (i.e., greater pattern separation; Leutgeb, Leutgeb, Moser, & Moser, 2007). By contrast, neural activity is highly correlated in CA3 based on the same changed input until these changes are sufficiently large, when a shift to novel neural firing occurs (Leutgeb et al., 2007; also see Vazdarjanova & Guszowski, 2004). This shift from correlated to uncorrelated activity is reflective of a transition from pattern completion to separation within CA3. Genetic knockout of NMDA receptors in the dentate impairs behavioral evidence of pattern separation and also renders the rest of the hippocampus to be more biased toward pattern completion, as predicted (McHugh et al., 2007). Similar distinctions across in hippocampal subregions has been observed in humans using ultra high resolution fMRI (Bakker, Kirwan, Miller, & Stark, 2008; Lacy, Yassa, Stark, Muftuler, & Stark, 2010). To summarize, the basic computational tradeoff between pattern separation and pattern completion has led to an effort to understand how the hippocampus confronts this trade off. This, in turn, has led to a wealth of new neuroscience data that make concrete the relationship between the lower level observations and the high level function.

What these examples from the domain of memory and hippocampal function demonstrate is that despite the

advance of sophisticated neuroscientific and genetic tools independently of cognitive science, without cognitive theory one would not have known where to begin to apply these tools to answer any of these questions. Similarly, without development of cognitive theory, including the use of formal and computational methods, the functional significance of individual neuroscientific observations about the hippocampus would not have been fully realized.

### 3. From motor to cognitive actions: Prefrontal cortex and basal ganglia interactions

The basal ganglia (BG) are a collection of subcortical nuclei traditionally studied for their well-established role in motor control, mostly because of the obvious deficits produced by diseases that affect the BG, including Parkinson's, Huntington's, dystonia, and Tourette's. However, it has become increasingly clear that cognitive changes accompany these disorders and many others that affect BG function.

The BG are also notorious for their convoluted circuit diagrams, with several interacting pathways that mostly involve neurons in one area that inhibit neurons in another area that normally inhibit neurons in yet another area. Anatomically, the circuits that link motor cortex with BG comprise only a small subset of frontal-BG circuits; indeed, virtually every part of the frontal cortex forms a closed loop with corresponding part of the BG-thalamic loop (Alexander, DeLong, & Strick, 1986). For example, the dorsolateral prefrontal cortex (DLPFC) – involved in working memory and rule representations – is interconnected with its own with BG circuit that modulates the thalamic inputs back to DLPFC. Moreover, patients with BG dysfunction exhibit frontal-like cognitive impairments in executive function, impulsivity, and working memory, leading neuropsychologists to propose a scaffolding of principles of motor behavior onto cognitive function (Kozlak & Budding, 2009).

How can cognitive theory provide insight into these neuroanatomical circuits and the somewhat surprising role of classical motor structures in higher cognitive function? Analogous to the computational trade-offs solved by the hippocampus in episodic memory, the incorporation of BG into domains traditionally thought accomplished by PFC may provide a solution to a fundamental computational cognitive tradeoff within working memory. The PFC has been long associated with robust active maintenance of information over delays in the face of distraction. Detailed biophysical models, validated by experimental data, have shown that this robust active maintenance involves recurrent connectivity and persistent NMDA currents that allow a pattern of activity to sustain itself in the absence of external input (Compte, Brunel, Goldman-Rakic, & Wang, 2000; Wang, 2013). However, studying the cognitive demands of working memory (WM) have lent insight into a complementary function: the need to rapidly update the contents of WM when needed by task demands (i.e. to attend to new information rather than continually maintain previous information). These two functions are at odds with one another: stronger recurrent

activity increases maintenance (i.e., stability) in the face of noise but prevents external relevant inputs from disrupting existing memoranda (i.e., flexibility).

The tradeoff between flexibility and stability has led to the proposal that working memory requires a gate or a dissociable mechanism that, when “opened”, can transiently increase the sensitivity of some memory buffers to external events, allowing them to be updated. But when the gate is “closed”, WM representations can then be robustly maintained until no longer required (Braver & Cohen, 2000; Frank, Loughry, & O'Reilly, 2001; Hochreiter & Schmidhuber, 1997). Notably, in the motor domain, the BG is thought to accomplish exactly this gating function: to facilitate desired motor programs while suppressing competing motor actions from interfering with them, and to do so in proportion to the learned reward values of such actions. Indeed, motor deficits associated with BG dysfunction are related to either impoverished ability to select (gate) actions (Parkinson's disease), too much gating of multiple agonist and antagonist actions (dystonia), or too little suppression of unwanted movements (Huntington's and Tourette's).

Thus, considerations of the cognitive demands of WM, the established BG gating role of in motor function, and the analogous circuits linking BG with PFC, led to the development of computational models that specify how the BG may provide an analogous gating function for WM (Frank, Loughry & O'Reilly, 2001) which could also explain counterintuitive findings of the effects of Parkinson's disease on WM function (Frank, 2005). Various subsequent neuroscientific investigations have corroborated these dissociable roles of BG and PFC in WM gating versus maintenance using neuroimaging (McNab & Klingberg, 2008; Chatham, Frank, & Badre, 2014; also reviewed in Chatham & Badre, *in press-b*), patients with circumscribed lesions (Baier et al., 2010), pharmacological manipulations, and genetics (for review, Frank & Fossella, 2011).

The investigation of the neuroscience underlying these processes has reciprocally influenced cognitive models of WM. For example, the implication of BG dopamine function in WM suggested that we might learn from the well-studied role of BG dopamine in motor circuits, in which it supports reinforcement learning. Thus, analogously, neural models of PFC-BG interactions have suggested that through dopamine signals, the BG learn to gate cognitive actions (such as what to maintain in WM) in terms of their reliability in maximizing likelihood of preferred outcomes. Indeed, in tasks that require storing some information in WM and ignoring irrelevant distractors, subjects learn to attend to the relevant information in such a way that accords with reinforcement learning principles. Moreover, pharmacological manipulations of dopamine (DA) that increase reward relative to punishment learning analogously increase learning to update task-relevant information at the cost of making participants more susceptible to updating distracting information (Cools, Miyakawa, Sheridan, & D'Esposito, 2010; Frank & O'Reilly, 2006; Moustafa, Sherman, & Frank, 2008). Indeed, though not linked to DA directly, evidence from human fMRI suggests that the BG may track changes in the utility of items maintained in WM based on their history (Chatham & Badre,

2013). This tight coupling between basic RL and motor functions on the one hand, and executive WM processes on the other, may not have been considered without the influence of cognitive theory on neuroscience and the reciprocal influence of implementational considerations on discovering new cognitive phenomena.

Expanding on the above notions, studies have leveraged the hierarchical nature of cognitive control during action selection and suggested that these may be implemented by nested frontostriatal gating loops, with more anterior frontal regions representing abstract goals and rules which then contextualize gating by more posterior frontostriatal loops (Botvinick, Niv, & Barto, 2009; Collins & Frank, 2013; Frank & Badre, 2012). Neuroscientific evidence for such hierarchical processing has been validated by functional imaging (Badre & Frank, 2012; Jeon, Anwander, & Friederici, 2014; Mestres-Missé, Turner, & Friederici, 2012), EEG (Collins, Cavanagh, & Frank, 2014) and white matter tractography (Jeon et al., 2014; Verstynen, Badre, Jarbo, & Schneider, 2012). Moreover, engagement of this circuitry is related to performance enhancements in task environments that demand hierarchical organization of behavior (Badre, Doll, Long, & Frank, 2012; Badre & Frank, 2012; Badre, Kayser, & D'Esposito, 2010; Collins et al., 2014).

#### 4. Simple mechanisms of decision making

For several decades, cognitive models of simple decision making have posited that decision makers employ a noisy evidence accumulation process, sequentially sampling information from the environment (or from their memory), and committing to a choice once the relative evidence for one option over the other reaches a critical decision threshold (Ratcliff, 1978). Such models provide quantitative fits to a host of empirical patterns – not only of choice proportions, but of their full response time distributions (for review, Ratcliff & McKoon, 2008). Such models have had enormous influence on the neuroscientific enterprise examining the neural basis of simple decisions, and have further elaborated how choice process amounts to Bayesian decision theory, where the likelihood ratio for one option over another is integrated optimally with prior knowledge and uncertainty to reach a choice (Dayan & Daw, 2008; Deneve, 2012; Huang, Friesen, Hanks, Shadlen, & Rao, 2012). Researchers have designed careful experiments based on this notion and found evidence for neuronal populations that accumulate evidence corresponding to the likelihood ratio for one choice option or another, until this evidence reaches a critical threshold (e.g., Gold & Shadlen, 2007). This finding has been highly influential in neuroscience and has led to intense debates about the precise nature of these neural signals, their sensitivity to reward value, and their contributions to choice (Platt & Glimcher, 1999).

Human imaging studies have also corroborated sequential sampling models of choice and how they deal with influence of prior knowledge and choice incentives, which involve common modulation of a single mechanism (Mulder, Wagenmakers, Ratcliff, Boekel, & Forstmann,

(2012) and, separately, decision threshold modulation by speed-accuracy tradeoffs (Forstmann et al., 2008). These model-based cognitive neuroscience studies show that decision threshold is itself modulated by pathways connecting frontal cortex to basal ganglia, and that disruption of these mechanisms leads to impulsive behaviors (Cavanagh et al., 2011; Forstmann et al., 2008; Green et al., 2013). Reciprocally, neural implementational models have provided motivation to refine some aspects of the cognitive model. For example, the neural mechanisms that relate to threshold adjustment are also sensitive to decision conflict, and exhibit a profile such that threshold rises when faced with conflict but then declines across time, in effect corresponding to a ‘collapsing bound’ model (Ratcliff & Frank, 2012). Similarly, nonlinear neural attractor models of decision making have accounted for patterns of choice that are not accommodated by previous cognitive models, for example evidence arriving early in the decision process has greater effect on choice than the same evidence arriving later in the process (Wang, 2008, 2012).

## 5. Conclusions

We have reviewed three examples of domains of neuroscience that have been directly influenced by cognitive theory. Of course, there are many other examples that we have not elaborated here. For example, the cross-species study of response inhibition in the frontal lobes and basal ganglia (Aron & Poldrack, 2006; Aron, Robbins, & Poldrack, 2004; Logan, Schall, Palmeri, in press; Schmidt, Leventhal, Mallet, Chen, & Berke, 2013) has been grounded in basic work from cognitive psychology that conceptualizes inhibition as a horse race between response and inhibitory signals that balances their opposing functions (Logan & Cowan, 1984). Likewise, the cognitive theory that confronts the basic trade off in decision making between immediate versus delayed rewards, in which choices tend to follow a hyperbolic discounting function, have led to intense debates about whether distinct neural signals reflect contributions of separable or unified valuation systems (Kable & Glimcher, 2007; McClure, Laibson, Loewenstein, & Cohen, 2004). The classical exploration-exploitation dilemma, in particular the role of outcome uncertainty in driving exploration highlighted by cognitive theory, has similarly refined neuroscientists’ views of the role of the frontal polar cortex (Badre et al., 2012; Cavanagh, Figueiroa, Cohen, & Frank, 2012; Daw, O’Doherty, Dayan, Seymour, & Dolan, 2006; Frank, Doll, Oas-Terpstra, & Moreno 2009). Thus, across the cases we have reviewed and many more examples, the theoretical distinctions and fundamental trade-offs identified at the cognitive level have provided functional explanation and generated new science at the neural level.

These examples also highlight some potential routes by which cognitive theory is most easily elaborated in order to have an influence on neural data. We emphasize that cognitive science makes many valuable contributions without a demand to constrain or influence neuroscience. Nevertheless, neuroscience is an area where cognitive science can have a broad impact if the cognitive theory

can be elaborated in such a way to make predictions for neural data (Chatham & Badre, in press-a,b). This process is helped greatly when this theory is elaborated in the context of explicit computational models. However, this does not mean that modeling must be at an implementational level to be influential for neuroscience. Cognitive models also provide means with which to translate neuroscientific measurements into latent processes (Frank, in press). It follows, as well, that cognitive models that have been elaborated for neuroscience data also may themselves benefit from new discoveries in neuroscience, and so can enjoy a greater source of data and a broader domain in which to explain and generate new science.

## References

- Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, 9, 357–381.
- Aron, A. R., & Poldrack, R. A. (2006). Cortical and subcortical contributions to stop signal response inhibition: Role of the subthalamic nucleus. *The Journal of Neuroscience*, 26, 2424–2433.
- Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, 8, 170–177.
- Badre, D., Doll, B. B., Long, N. M., & Frank, M. J. (2012). Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron*, 73, 595–607.
- Badre, D., & Frank, M. J. (2012). Mechanisms of hierarchical reinforcement learning in cortico-striatal circuits 2: Evidence from fMRI. *Cerebral Cortex*, 22(3), 527–536.
- Badre, D., Kayser, A., & D’Esposito, M. (2010). Frontal cortex and the discovery of abstract action rules. *Neuron*, 66, 315–326.
- Baier, B., Karnath, H. O., Dieterich, M., Birklein, F., Heinze, C., & Müller, N. G. (2010). Keeping memory clear and stable – the contribution of human basal ganglia and prefrontal cortex to working memory. *Journal of Neuroscience*, 30(29), 9788–9792.
- Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. (2008). Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science*, 319(5870), 1640–1642.
- Bliss, T. V. P., & Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anesthetized rabbit following stimulation of the perforant path. *Journal of Physiology (London)*, 232, 331–356.
- Botvinick, M. M., Niv, Y., & Barto, A. C. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition*, 113(3), 262–280.
- Braver, T. S., & Cohen, J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 713–737). Cambridge, MA: MIT Press.
- Cavanagh, J. F., Figueiroa, C. M., Cohen, M. X., & Frank, M. J. (2012). Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. *Cerebral Cortex*, 22, 2575–2586.
- Cavanagh, J. F., Wiecki, T. V., Cohen, M. X., Figueiroa, C. M., Samanta, J., Sherman, S. J., et al. (2011). Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. *Nature Neuroscience*, 14, 1462–1467.
- Chatham, C. H., & Badre, D. (2015). How to test cognitive theory with fMRI. In D. Spieler & E. Schumacher (Eds.), *New methods in cognitive psychology*. New York, NY: Routledge (in press).
- Chatham, C. H., & Badre, D. (2015). Multiple gates on working memory. *Current Opinion in Behavioral Sciences* (in press).
- Chatham, C. H., & Badre, D. (2013). Working memory management and predicted utility. *Frontiers in Behavioral Neuroscience*, 7(83), 1–12.
- Chatham, C. H., Frank, M. J., & Badre, D. (2014). Corticostriatal output gating during selection from working memory. *Neuron*, 81, 930–942.
- Cohen, N. J., Poldrack, R. A., & Eichenbaum, H. (1997). Memory for items and memory for relations in the procedural/declarative memory framework. *Memory*, 5, 131–178.
- Collins, A. G. E., Cavanagh, J. F., & Frank, M. J. (2014). Human EEG uncovers latent generalizable task-set structure during learning. *Journal of Neuroscience*, 34, 4677–4685.
- Collins, A. G. E., & Frank, M. J. (2013). Cognitive control over learning: Creating, clustering and generalizing task-set structure. *Psychological Review*, 120, 190–229.

- Compte, A., Brunel, N., Goldman-Rakic, P. S., & Wang, X. J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral Cortex*, 10, 910–923.
- Cools, R., Miyakawa, A., Sheridan, M., & D'Esposito, M. (2010). Enhanced frontal function in Parkinson's disease. *Brain*, 133(Pt 1), 225–233.
- Corkin, S. (1968). Acquisition of a motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia*, 6, 255–265.
- Curran, T., DeBuse, C., Woroch, B., & Hishman, E. (2006). Combined pharmacological and electrophysiological dissociation of familiarity and recollection. *Journal of Neuroscience*, 26, 1979–1985.
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, 69, 1204–1215.
- Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, 441, 876–879.
- Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, and Behavioral Neuroscience*, 8, 429–453. <http://dx.doi.org/10.3758/CABN.8.4.429>.
- Deneve, S. (2012). Making decisions with unknown sensory reliability. *Frontiers in Neuroscience*, 6, 75.
- Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K. R., et al. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences of the United States of America*, 105(45), 17538–17542.
- Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience*, 17(1), 51–72.
- Frank, M. J. (2015). Linking across levels of computation in model-based cognitive neuroscience. In B. U. Forstmann & E. Wagenmakers (Eds.), *An introduction to model-based cognitive neuroscience*. New York: Springer (in press).
- Frank, M. J., & Badre, D. (2012). Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: Computational analysis. *Cerebral Cortex*, 22(3), 509–526.
- Frank, M. J., Doll, B. B., Oas-Terpstra, J., & Moreno, F. (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nature Neuroscience*, 12, 1062–1068.
- Frank, M. J., & Fossella, J. A. (2011). Neurogenetics and pharmacology of learning, motivation, and cognition. *Neuropsychopharmacology*, 36, 133–152.
- Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: A computational model. *Cognitive Affective and Behavioral Neuroscience*, 1(2), 137–160.
- Frank, M. J., & O'Reilly, R. C. (2006). A mechanistic account of striatal dopamine function in human cognition: Psychopharmacological studies with cabergoline and haloperidol. *Behavioral Neuroscience*, 120(3), 497–517.
- Frank, M. J., O'Reilly, R. C., & Curran, T. (2006). When memory fails, intuition reigns: Midazolam enhances implicit inference in humans. *Psychological Science*, 17, 700–707.
- Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, 30, 535–574.
- Green, N., Bogacz, R., Huebl, J., Beyer, A.-K., Kuhn, A. A., & Heekeren, H. R. (2013). Reduction of influence of task difficulty on perceptual decision making by STN deep brain stimulation. *Current Biology*, 23(17), 1681–1684.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9, 1735–1780.
- Huang, Y., Friesen, A., Hanks, T., Shadlen, M., & Rao, R. (2012). How prior probability influences decision making: A unifying probabilistic model. In *Advances in neural information processing systems* (Vol. 25). Lake Tahoe, NV.
- Huffman, D. J., & Stark, C. E. L. (2014). Multivariate pattern analysis of the human medial temporal lobe revealed representationally categorical cortex and representationally agnostic hippocampus. *Hippocampus*, 24, 1394–1403.
- Jeon, H. A., Anwander, A., & Friederici, A. D. (2014). Functional network mirrored in the prefrontal cortex, caudate nucleus, and thalamus: High-resolution functional imaging and structural connectivity. *Journal of Neuroscience*, 34, 9202–9212.
- Johnson, A., & Redish, A. D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *Journal of Neuroscience*, 27(45), 12176–12189.
- Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, 10, 1625–1633.
- Kozlak, L. F., & Budding, D. E. (2009). *Subcortical structures and cognition: Implications for neuropsychological assessment*. New York, NY: Springer.
- Lacy, J. W., Yassa, M. A., Stark, S. M., Muftuler, L. T., & Stark, C. E. (2010). Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. *Learning and Memory*, 18, 15–18.
- Leutgeb, J. K., Leutgeb, S., Moser, M. B., & Moser, E. I. (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science*, 315(5814), 961–966.
- Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, 91, 295–327.
- Logan, G. D., Schall, J. D., Palmeri, T. J. (2015). Inhibitory control in mind and brain: The mathematics and neurophysiology of the underlying computation. In B. U. Forstmann & E. Wagenmakers (Eds.), *An introduction to model-based cognitive neuroscience*. New York: Springer (in press).
- Marr, D. (1971). Simple memory: A theory for archicortex. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences*, B262, 23–81.
- McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102, 419–457.
- McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, 306(5695), 503–507.
- McGeoch, J. A., & Melton, A. W. (1929). The comparative retention values of maze habits and of nonsense syllables. *Journal of Experimental Psychology*, 12, 392–414.
- McHugh, T. J., Jones, M. W., Quinn, J. J., Balthasar, N., Coppari, R., Elmquist, J. K., et al. (2007). Dentate gyrus NMDA receptors mediate rapid pattern separation in the hippocampal network. *Science*, 317(5834), 94–99.
- McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*, 11(1), 103–107.
- Mestres-Missé, A., Turner, R., & Friederici, A. D. (2012). An anterior-posterior gradient of cognitive control within the dorsomedial striatum. *Neuroimage*, 62, 41–47.
- Moustafa, A. A., Sherman, S. J., & Frank, M. J. (2008). A dopaminergic basis for working memory, learning and attentional shifting in Parkinsonism. *Neuropsychologia*, 46, 3144–3156.
- Mulder, M. J., Wagenmakers, E. J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: A diffusion model analysis of prior probability and potential payoff. *The Journal of Neuroscience*, 32(7), 2335–2343.
- Myers, C. E., Shohamy, D., Gluck, M. A., Grossman, S., Kluger, A., Ferris, S., et al. (2003). Dissociating hippocampal versus basal ganglia contributions to learning and transfer. *Journal of Cognitive Neuroscience*, 15(2), 185–193.
- Nakahashi, T., Young, J. Z., McHugh, T. J., Buhl, D. L., & Tonegawa, S. (2008). Transgenic inhibition of synaptic transmission reveals role of CA3 output in hippocampal learning. *Science*, 319(5867), 1260–1264.
- Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary learning systems approach. *Psychological Review*, 110, 611–646.
- O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Clarendon Press.
- O'Reilly, R. C., & McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: Avoiding a tradeoff. *Hippocampus*, 4, 661–682.
- O'Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends in Cognitive Science*, 6, 505–510.
- O'Reilly, R. C., & Rudy, J. W. (2001). Conjunctive representations in learning and memory: Principles of cortical and hippocampal function. *Psychological Review*, 108, 311–345.
- Packard, M. G., & McGaugh, J. L. (1996). Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiology of Learning and Memory*, 65(1), 65–72.
- Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, 400, 233–238.
- Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia*, 41(3), 245–251.

- Posner, M. I. (1966). Components of skilled performance. *Science*, 152, 1712–1718.
- Posner, M. I. (1967). Characteristics of visual and kinesthetic memory codes. *Journal of Experimental Psychology*, 75, 103–107.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Ratcliff, R., & Frank, M. J. (2012). Reinforcement-based decision making in corticostriatal circuits: Mutual constraints by neurocomputational and diffusion models. *Neural Computation*, 24, 1186–1229.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922.
- Sadeh, T., Shohamy, D., Levy, D. R., Reggev, N., & Maril, A. (2011). Cooperation between the hippocampus and the striatum during episodic encoding. *Journal of Cognitive Neuroscience*, 23(7), 1597–1608.
- Schmidt, R., Leventhal, D. K., Mallet, N., Chen, F., & Berke, J. D. (2013). Canceling actions involves a race between basal ganglia pathways. *Nature Neuroscience*, 16, 1118–1124.
- Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, 20, 11–21.
- Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, 99, 195–231.
- Squire, L. R., & Knowlton, B. J. (1995). Learning about categories in the absence of memory. *Proceedings of the National Academy of Sciences of the United States of America*, 92(26), 12470–12474.
- van der Meer, M. A., Johnson, A., Schmitzer-Torbert, N. C., & Redish, A. D. (2010). Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. *Neuron*, 67(1), 25–32. 15.
- Vazdarjanova, A., & Guzowski, J. F. (2004). Differences in hippocampal neuronal population responses to modifications of an environmental context: Evidence for distinct, yet complementary, functions of CA3 and CA1 ensembles. *Journal of Neuroscience*, 24, 6489–6496.
- Verstynen, T. D., Badre, D., Jarbo, K., & Schneider, W. (2012). Microstructural organizational patterns in the human corticostriatal system. *Journal of Neurophysiology*, 107(11), 2984–2995.
- Wang, X. J. (2008). Decision making in recurrent neuronal circuits. *Neuron*, 60(2), 215–234.
- Wang, X. J. (2012). Neural dynamics and circuit mechanisms of decision-making. *Current Opinion in Neurobiology*, 22, 1039–1046.
- Wang, X. J. (2013). The prefrontal cortex as a quintessential ‘cognitive-type’ neural circuit: Working memory and decision making. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (2nd ed., pp. 226–248). New York: Cambridge University Press.
- Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, 34, 515–525.