PRINCIPLES OF
FRONTAL LOBE FUNCTION

SECOND EDITION

EDITED BY

Donald T. Stuss, PhD
PRESIDENT AND SCIENTIFIC DIRECTOR, ONTARIO BRAIN INSTITUTE
ROTMAN RESEARCH INSTITUTE OF BAYCREST
DEPARTMENTS OF PSYCHOLOGY AND MEDICINE (NEUROLOGY, REHABILITATION SCIENCES)
UNIVERSITY OF TORONTO
TORONTO, ONTARIO
CANADA

Robert T. Knight, MD
DEPARTMENT OF PSYCHOLOGY
HELEN WILLS NEUROSCIENCE INSTITUTE
UNIVERSITY OF CALIFORNIA AT BERKELEY
BERKELEY, CA

OXFORD UNIVERSITY PRESS
The well-established role of the prefrontal cortex (PFC) in a wide range of executive functions (Fuster, 2008; Miller & Cohen, 2001) begs the question: what are the key properties that enable the PFC to subserve cognitive processes, in contrast to primary sensory or motor systems? The answer, in part, lies in its privileged position in the brain network (Averbeck & Seo, 2008; Nauta, 1971; Petrides & Pandya, 2002): at the top of the cortical hierarchy, the PFC is well situated for representational processing of the highest order; and its extensive input-output connections with the rest of the brain allow the PFC to combine information from various sensory, motor, and limbic areas, provide a top-down attention signal to modulate sensory processing, exert inhibitory control of motor action, and so on.

Equally important is the local circuit, the microcircuit operation within each of two dozen subregions of the PFC. Whereas sensory neurons are characterized by rapid and transient responses to external stimuli, PFC neurons commonly exhibit mnemonic persistent neural activity in the absence of direct stimulation (Fuster & Alexander, 1971; Funahashi, Bruce, & Goldman-Rakic, 1989; Kubota & Niki, 1971; Miller, Erickson, & Desimone, 1996; Romo, Brody, Hernandez, & Lemus, 1999). Such sustained activity is a neural correlate of working memory, the brain’s ability to internally maintain and manipulate information (e.g., across a short time interval between sensory stimulation and motor response). The persistence time of sustained firing activity during working memory is orders of magnitude longer than the biophysical time constants (tens of milliseconds) of fast electrical signals in neurons and synapses. For this reason, persistent activity is believed to be generated by feedback dynamics, or reverberation, in a local circuit (Amir, 1995; Arnsten, Paspalas, Gamo, Yang, & Wang, 2010; Goldman-Rakic, 1995; Hebb, 1949; Lorente de Nó, 1933; Wang, 2001). As Hebb wrote more than sixty years ago: “To the extent that anatomical and physiological observations establish the possibility of reverberatory after-effects of a sensory event, it is established that such a process would be the physiological basis of a transient ‘memory’ of the stimulus” (Hebb, 1949, p. 61). The characteristic horizontal connections found in the superficial layers II–III of the monkey dorsolateral PFC may provide the anatomical substrate for such a recurrent circuit (Kritzer & Goldman-Rakic, 1995; Levitt, Lewis, Yoshioka, & Lund, 1993), and a recent monkey study showed that persistent activity in the PFC decreases with age and that this decline may be associated with reduced recurrent excitation (Wang et al., 2011a). The idea of reverberation is made precise in theoretical work where persistent activity is described as “dynamical attractors” (Amari, 1977; Amit, 1995; Wang, 2001). The mathematical term “attractor” simply means any self-sustained and stable state of a dynamic system, such as a neural network. According to this picture, in a working memory system, the spontaneous state and stimulus-selective memory states are assumed to represent multiple attractors, such that a memory state can be switched on and off by transient inputs.

P. Milner (a former colleague of Hebb) expressed doubts about the attractor theory on the ground that a reverberatory neural assembly is “liable to fire out of control” (Milner, 1996). Positive feedback could drive neurons to fire at higher and higher rates until saturation is reached. This would be inconsistent with the observation that mnemonic persistent activity in the cortex occurs at moderate rates (typically 10 to 50 Hz), significantly above the spontaneous firing rate (a few hertz) but well below the maximum firing capability of cortical neurons. When inhibition is incorporated to compensate for feedback excitation, inhibitory damping might prevent reverberation altogether (Milner, 1999).

Indeed, quantitative circuit modeling has shown that this problem of instability poses a serious challenge. However, in the same year that Milner’s words were written, it was realized that cortical circuits with multiple attractor states are dynamically stable if the recurrent excitation is slow compared to inhibition—for example, partially
mediated by the N-methyl-d-aspartate (NMDA) receptors (Wang, 1999). Conceptually, this means that working memory local circuits do not operate as fast switches with millisecond-scale transition times. Instead, neural computation is more like an integration in the sense of calculus, at least up to a point, converting a brief pulsatile input into a persistent output pattern (Figure 15–1, left panel). Surprisingly, this turns out to be precisely what is needed for gradual accumulation over time, in the form of ramping activity, of information about choice options in a decision process (Figure 15–1, right panel; Wang, 2002). Synaptic excitation is balanced by inhibition, which gives rise to selectivity and winner-take-all competition—important for both working memory and decision making. Therefore, a common local circuit mechanism may be capable of generating neural signals necessary for decision computations as well as working memory (Wang, 2002, 2008). This computational finding is supported by the observation that delay period persistent activity in working memory tasks, and decision-related neural signals in perceptual and reward-based decision tasks, are often observed in the same brain regions, especially the PFC but also the posterior parietal cortex and other areas (Gold & Shadlen, 2007; Heekeren, Marrett, & Ungerleider, 2008; Wang, 2008). Therefore, a common “cognitive-type” local circuit mechanism may underlie both working memory and decision making.

In this chapter, I will expound on the notion of cognitive-type local circuits. First, I will summarize the basic features of a working memory/decision circuit and its neuromodulation. Then I will show how such a circuit model applies to inhibitory control of behavioral responses and how this framework can be extended to coding of behavioral rules in the PFC, which has led to the idea of a “reservoir” of randomly connected neurons endowed with mixed selectivity. Implications for understanding cognitive deficits in schizophrenia will be briefly discussed.

A BIOPHYSICALLY BASED MODEL OF WORKING MEMORY

A well-known working memory paradigm is the delayed oculomotor response task, in which a subject is required to remember a visual cue (a directional angle) across a delay period in order to perform a memory-guided saccade (Chafee & Goldman-Rakic, 1998; Constantinidis & Goldman-Rakic, 2002; Constantinidis & Wang, 2004; Funahashi et al., 1989). We have developed a network model for this spatial working memory experiment (Figure 16–2A; Compte, Brunel, Goldman-Rakic, & Wang, 2000; Renart, Brunel, & Wang, 2003; Tegnér, Compte, & Wang, 2002). The key feature is the pre-eminence of recurrent connections (“loops”) between neurons, so that a cell receives not only external stimulation (via afferents from upstream neurons) but also inputs from other cells within the same microcircuit (via “horizontal” connections). A commonly assumed network architecture is the so called “Mexican hat”: localized recurrent excitation between pyramidal cells with a similar preference for spatial cues and broader inhibition mediated by interneurons. Models of synapses and single cells are calibrated quantitatively by cortical electrophysiological studies. This is important: even though network function is determined by the collective dynamics of many thousands of neurons, the emergent population behavior depends critically on the properties of single cells and synapses.

Figure 15–2B shows a model simulation of the delayed oculomotor task. Initially, the network is in a resting state in which all cells fire spontaneously at low rates. A transient input (in this case at 180 degrees) drives a subpopulation of cells to fire at high rates. As a result, they send recruited excitation to each other via horizontal connections. This internal excitation is large enough to sustain elevated
activity, so that the firing pattern persists after the stimulus is withdrawn. Synaptic inhibition ensures that the activity does not spread to the rest of the network, and persistent activity has a bell shape ("bump attractor"). At the end of a mnemonic delay period the cue information can be retrieved by reading out the peak location of the persistent activity pattern, and the network is reset back to the resting state.

In different trials, a cue can be presented at different locations, and the firing activity of a single cell can be compared with the single-unit recording data from monkey’s prefrontal cortex (Funahashi et al., 1989). At the network level, each cue triggers a persistent firing pattern of the same bell shape but peaked at a different location. A spatial working memory network thus requires a continuous family of bump attractors, each encoding a potential location (Ben-Yishai, Lev Bar-Or, & Sompolinsky, 1995; Camperi & Wang, 1998; Compte et al., 2000).

This instantiation of such a continuous attractor can be rendered robust by regulatory homeostatic mechanisms in a biophysically realistic cortical network in spite of cellular heterogeneities (Renart et al., 2003). Thus, this biologically constrained model captures salient experimental observations from behaving monkeys. What lessons have we learned from such a model?
SLOW EXCITATORY REVERBERATION: ROLE OF NMDA RECEPTORS

A system with fast positive and slow negative feedbacks, both powerful, is prone to dynamic instability (Douglas, Koch, Mahowald, Martin, & Suarez, 1995; Wang, 1999). In biologically realistic models, persistent activity is often disrupted in the middle of a delay period, so the memory is lost (Compte et al., 2000; Renart et al., 2003; Tegnér et al., 2002; Wang, 1999). The same destabilization problem is present if negative feedback is instantated by spike-frequency adaptation (McCormick, Connors, Lighthall, & Prince, 1985) or short-term synaptic depression (Abbott & Regehr, 2004; Markram, Wang, & Tsodyks, 1998). Such instability does not occur if the excitation is sufficiently slow compared to negative feedback, when recurrent synapses are primarily mediated by NMDA receptors (time constant 50–100 ms) (Compte et al., 2000; Wang, 1999). Moreover, the slow NMDA receptor (NMDAR) unbinding to glutamate gives rise to saturation of the NMDA synaptic current with repetitive stimulation at high frequencies. As a result, a further increase in neural firing rates does not lead to a larger excitatory drive, and the explosive positive feedback is curtailed. Therefore, it helps to control the firing rate in a persistent activity state (Wang, 1999).

A specific suggestion from modeling work, then, is that in a working memory microcircuit, if persistent activity is sustained primarily by synaptic reverberation, local excitatory synapses should have a sufficiently high NMDA/alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (NMDA/AMPA) ratio. How high is high enough? The answer depends on the details of network biophysics and connectivity. For instance, the time constant of a synaptic current depends on the subunit composition of its receptors. If gamma-aminobutyric acid A (GABA_A) receptor (GABA_A,R)-mediated inhibition is unusually fast in a working memory circuit, instability due to the time constant mismatch with AMPA receptor (AMPA)-mediated excitation would be less severe and the required NMDA/AMPA ratio would be lower (Tegnér et al., 2002). Recently, the contribution of NMDARs to synaptic transmission locally between prefrontal pyramidal cells has been measured using intracellular recording from connected pairs of neighboring cells in the rat frontal cortex in vitro. It was found that NMDAR-mediated currents at prefrontal synapses in the adult, but not young, rats exhibit a twofold longer decay time constant and temporally summate a train of stimuli more effectively, compared to those in the primary visual cortex, in support of our working memory circuit model (Wang, Stradtman, Wang, & Gao, 2008). Moreover, importantly, the NMDA model prediction has been directly tested in vivo during working memory. Preliminary data indicate that iontophoresis of a drug that selectively blocks NMDARs in recorded cells effectively suppresses stimulus-selective persistent activity of PFC neurons in behaving monkeys (Wang et al., 2011b), confirming the model’s prediction.

Quantitative differences breed qualitatively different behaviors. That a cortical area exhibits a new type of behavior does not necessarily mean that the circuit must possess unique biological machineries completely different from those of other areas. Hence, persistent activity may be generated in the PFC when the strength of recurrent excitation (mediated by AMPARs + NMDARs combined) exceeds a critical threshold, whereas this may not be the case for a sensory area such as the primary visual cortex. Based on our modeling results, we can extend this idea and propose that, for stable function of a working memory circuit, the NMDA/AMPA ratio at recurrent synapses should also be above a certain threshold. It is important to emphasize that what matters for persistent activity is not the unitary amplitude of excitatory postsynaptic currents (EPSCs) at resting potential, but the ratio of the average NMDA and AMPA synaptic currents during repetitive neural discharges. This ratio depends on multiple factors, such as presynaptic short-term plasticity, postsynaptic summation and saturation, and the voltage dependence of the NMDA channel conductance. Further, a relatively high NMDA/AMPA ratio at local synapses can be compatible with a low total NMDA/AMPA ratio in a neuron—for instance, if feedforward inputs from outside of the network are predominantly mediated by AMPARs. Last but not least, this ratio can be enhanced by neuromodulators, such as dopamine (Seamans & Yang, 2004).

MULTIPLE FEEDBACK PROCESSES

In addition to the NMDAR-mediated recurrent excitation, other cellular and synaptic processes may contribute to the generation of persistent activity. We have examined in computational models three such processes, all dependent on intracellular calcium signaling. First, single neurons exhibit a large repertoire of active ion channels (Llinás, 1988; Magee, Hoffman, Colbert, & Johnston, 1998); some could provide intrinsic positive feedback leading to persistent activity (Camperi & Wang, 1998; Egorov, Hamam, Fransen, Hasselmo, & Alonso, 2002; Koukalov, Raghavachari, Kepecs, & Lisman, 2002; Loewenstein & Sompolinsky, 2003). In rat frontal neurons, accumulation of intracellular calcium during spiking activity activates a slow inward current, which in turn further increases the neuronal excitability, providing a positive feedback loop that can lead to spike firing outlasting a transient stimulus (Haj-Dahmane & Andrade, 1998). When such an intrinsic current is added in single pyramidal neurons of our network model for spatial working memory, less slow NMDAR-mediated synaptic transmission is needed to ensure stable persistent activity, as shown in Figure 15–2C.
(Tegnér et al., 2002). Therefore, the principle that a positive feedback process should not be too fast compared to a negative feedback process holds true for intrinsic membrane mechanisms of single neurons as well.

Second, synapses commonly display slow-term plasticity on a time scale of hundreds of milliseconds to seconds (Abbott & Regehr, 2004; Markram et al., 1998). Interestingly, it was found that, in rodents, whereas synapses between two neighboring pyramidal cells typically show short-term depression in the primary visual cortex, local synaptic connections display pronounced short-term facilitation in the frontal cortex (Hempel, Hartmann, Wang, Turrigiano, & Nelson, 2000; Wang et al., 2006). Modeling work showed that prominent synaptic facilitation can contribute to the maintenance of a short-term memory trace (Hempel et al., 2000; Mongillo, Barak, & Tsodyks, 2008; Szatmary & Izhikevich, 2010). It should be emphasized that both the NMDAR-mediated glutamate transmission and facilitation are parts of a synaptic mechanism; they could work together to provide sufficient recurrent excitation underlying persistent activity.

Third, depolarization-induced suppression of inhibition (DSI) refers to the phenomenon in which depolarization of pyramidal neurons induces a reduced inhibition of the same neurons by a subclass of GABAergic cells through a process mediated by endogenous cannabinoids (Freund, Katona, & Piomelli, 2003; Wilson & Nicoll, 2001). As a result, the more active pyramidal cells are, the more they are disinhibited, leading to enhanced reverberation. A modeling study (Fig. 15–3A) that incorporates DSI further revealed the balancing acts of positive and negative feedbacks in the complex dynamics of a working memory system (Carter & Wang, 2007). Pyramidal neurons commonly display spike-frequency adaptation: spiking activity triggers a negative feedback process (with a time constant of ~100 ms) leading to a reduced firing rate (McCormick et al., 1985; Wang, 1998). In a working memory model where pyramidal cells show spike-frequency adaptation, a “bump” activity pattern (like that in Figure 15–2B) often becomes unstable and persistently active cells become less excitable over time (Hansel & Sompolinsky, 1998; Laing & Chow, 2001). This instability might still be present in a working memory circuit dominated by the NMDA mechanism, which has a time constant comparable to that of the adaptation process, but the addition of DSI with a slower time constant (seconds) could restore the stability and ensure robust working memory function (Carter & Wang, 2007). Furthermore, DSI reduces noise-induced random drifts of a persistent activity pattern during a delay period, so the readout of a remembered cue at the end of a delay period is reliable (Figure 15–3B,C). Interestingly, a simulated agonist for cannabinoid receptors leads to the opposite effect: random drifts are larger, and the information decoded from persistent population activity deteriorates over time (Figure 15–3D). This is because the action of a cannabinoid agonist is not activity dependent. Hence, disinhibition is not selective only for those pyramidal cells that are active (for their preferred cue) but globally in the entire network, which is detrimental to stable, persistent activity. This model thus provides an explanation for the observation that the use of marijuana reduces the accuracy of readout from working memory in human subjects (Ploner et al., 2002).

Taken together, our modeling work has led to the “slow reverberation hypothesis” about strongly positive feedback mechanisms required for the generation of persistent activity in a working memory circuit. In addition to ensuring circuit stability, slow reverberating neural activity also provides a time integration mechanism of critical importance for decision computations (see below). It is worth noting that very slow cellular processes with a time scale of seconds should not be predominant; otherwise, it would be difficult to switch on and off persistent activity with relatively brief (hundreds of milliseconds) external inputs. Instead, these processes could aid the NMDA-mediated synaptic mechanism, which we hypothesize is the workinghorse of reverberating neural dynamics in a working memory circuit.

**EXCITATION-INHIBITION BALANCE**

A general principle of cortical circuit organization is a dynamic balance between synaptic excitation and inhibition. We found that such a balance is important for normal functions of a PFC network model for several reasons. Some are in common with sensory systems; others are more specially relevant to working memory.

**DYNAMIC STABILITY**

A conspicuous feature of our network model is multistability: a resting state coexists with a number of stimulus-selective memory states, so that transient inputs lead to switching between self-sustained network firing patterns, or “attractors.” The resting state should be stable to small perturbations due to noisy spontaneous neural firing, in spite of strong excitatory recurrence. This is achieved by a tight balance between excitation and inhibition (E-I balance). In fact, in the resting state, feedback inhibition is slightly greater than excitation; hence, the overall recurrent input to a neuron is inhibitory and spontaneous spike firing is driven by random background external inputs (Amit & Brunel, 1997). Interestingly, in a memory state in which stronger reverberatory excitation is recruited to sustain an elevated firing rate, synaptic inhibition increases proportionally with excitation; this dynamically maintained E-I balance contributes to the control of the firing rates and prevent runaway excitation (Brunel & Wang, 2001). Other experimental and theoretical work suggests that a fixed E-I balance, regardless of changing neuronal firing rates, may
be a general characteristic of cortical network dynamics (Compte et al., 2003b; Shadlen & Newsome, 1994; Shu, Hasenstaub, & McCormick, 2003; van Vreeswijk and Sompolinsky 1996). The balancing act of recurrent excitation and inhibition may contribute to an explanation for the highly irregular spike discharges in prefrontal cells (Compte et al., 2003a).

**STIMULUS SELECTIVITY**

Synaptic inhibition plays a critical role in sculpting the stimulus selectivity of mnemonic persistent firing patterns, in consonance with the observation that GABA_\text{A} antagonists result in the loss of spatial tuning of prefrontal neurons during a delayed oculomotor task (Rao, Williams, & Goldman-Rakic, et al., 2000).
RESISTANCE AGAINST DISTRACTORS

A key aspect of memory maintenance, in which inhibition plays an important role, is resistance against distractors: while behaviorally relevant information is actively held in mind, irrelevant sensory stimuli should be denied entrance to the working memory system. In delayed response experiments using intervening stimuli (distractors), mnemonic activity has been shown to be easily disrupted by distractors in inferotemporal neurons but not in prefrontal neurons (Miller et al., 1996). Similarly, delay period activity in posterior parietal cortex appears to be sensitive to distractors (Constantinidis & Wang, 2004; Powell & Goldberg, 2000). Therefore, the evidence suggests that although multiple cortical areas exhibit delay period activity, mnemonic neural signals in PFC may persist when those in the temporal lobe and parietal lobe are lost, so that behaviorally relevant information is maintained in the brain in spite of distractors. This observation at the single-cell level suggests that PFC is a pivotal part of the attention network that focuses brain resources on selective information (Corbetta & Shulman, 2002; Desimone & Duncan, 1995; Mesulam, 2000; Zanto, Rubens, Thangavel, Gazzaley, 2011).

What enables PFC to resist distracting stimuli? A gating mechanism may be involved in deciding which stimulus is behaviorally relevant and thus should be held in working memory (Cohen, Braver, & Brown, 2002; Cohen, Braver, & O’Reilly, 1996). On the other hand, it is desirable that a working memory circuit be endowed with mechanisms to filter out, “by default,” external inputs that constantly bombard the senses. We found that synaptic inhibition naturally gives rise to this capability (Brunel & Wang, 2001; Compte et al., 2000). This is because, in a memory delay period, active neurons recruit inhibition that projects to the rest of the network. Consequently, those cells not encoding the initial cue are less excitatory than when they are in the resting state (see Figure 15–2B) and, hence, are less responsive to distracting stimuli presented during the delay. For working memory, the impact of a distractor depends on its strength (saliency) and the distance to the memorized cue (Compte et al., 2000), and the network’s ability to ignore distractors is sensitive to modulation of recurrent excitation and inhibition (Brunel & Wang, 2001; Compte et al., 2000; Durstewitz, Seamans, & Sejnowski, 2000a).

SYNCHRONIZED NETWORK FAST OSCILLATIONS

The E-I balance often manifests itself in the form of coherent network oscillations, typically in the gamma (40 Hz) frequency range (Compte et al., 2000; Tegnér et al., 2002; Wang, 1999). This is because fast excitation followed by slower inhibition is a common recipe for rhythmmogenesis in neural networks (Wang, 2010; Wilson & Cowan, 1972). Synaptic inhibition mediated by GABA, Rs is typically about three to five times slower than fast synaptic excitation mediated by AMPARs, the latter having a decay time constant of a few milliseconds (Hestrin, Sah, & Nicoll, 1990; Xiang, Huguenard, & Prince, 1998). Modeling studies showed that coherent oscillations resulting from an interplay between AMPAR-mediated excitation and GABA, R-mediated inhibition have a preferred frequency range around 40 Hz (Brunel & Wang, 2003; Geisler, Brunel, & Wang, 2005). This theoretical result suggests that synchronous 40 Hz oscillations may be observed in mnemonic persistent activity, a notion that has found some experimental support (Pesaran, Pezaris, Sähni, Mitra, & Andersen, 2002).

According to this view, fast γ rhythms may be a characteristic sign of the engagement of strongly reverberatory cortical circuits (Wang, 2010). In particular, selective attention involves top-down signals to sensory neurons that originate from the parieto-prefrontal circuit largely overlapping with the working memory system. Hence, enhanced γ oscillations and synchrony associated with selective attention (Fries, Reynolds, Rorie, & Desimone, 2001; Gregoriou, Gotts, Zhou, & Desimone, 2009; Womelsdorf, Fries, Mitra, & Desimone 2006) could be explained by this mechanism, as demonstrated by a two-module model of spiking neurons that consists of a reciprocal loop between a sensory network and a working memory (e.g., PFC) network (Ardid, Wang, & Compte, 2007; Ardid, Wang, Gomez-Cabrero, & Compte, 2010).

CONTRIBUTIONS OF DIFFERENT GABAERGIC CELL SUBTYPES

Traditionally, fast-spiking, perisomatic targeting basket cells have been the focus of studies of synaptic inhibition. However, in the cortex, there is a wide range of GABAergic interneurons with regard to their morphology, electrophysiology, chemical markers, synaptic connections, short-term plasticity, and molecular characteristics (Buzsáki, Geisler, Henze, & Wang, 2004; Cauli et al., 1997; DeFelipe, 1997; Freund & Buzsáki, 1996; Kawaguchi, 1997; Markram et al., 2004; Somogyi, Tamas, Lujan, & Buhl, 1998). Three largely nonoverlapping subclasses of inhibitory cells can be identified according to the expression of the calcium-binding proteins parvalbumin (PV), calbindin (CB), and calretinin (CR). Interestingly, in the macaque monkey, the distributions of PV, CB, and CR interneurons appear to be quite different in PFC compared to V1. In primary visual cortex, PV-containing interneurons (including fast-spiking basket cells) are prevalent (~75%), whereas the other two (CB- and CR-containing) interneuron types constitute about 10% each of the total GABAergic neural population (Brederode, Mulligan, & Hendrickson, 1990; Meskenaite, 1997). By contrast, in the PFC, the proportions are about 24% (PV), 24% (CB), and 45% (CR), respectively (Condé,
activity pattern is shaped by synaptic inhibition from PV
tained persistent activity in these neurons, and the network
interneuron-mediated disinhibition could generate self-sus-
tion (disinhibition) of the dendrites of the same pyramidal
rons within the same column, leading to reduced inhibi-
time, activated CR interneurons suppress CB interneu-
excite each other through interconnections. At the same
in a column are excited by a transient extrinsic input, they
geting inhibitory neurons in the same layer (Gonchar &
Gonzalez-Albo, Del Rio, & Elston, 1999). It is also possible
and preferentially target CB interneurons (DeFelipe,
1996), at least in the same cortical layer (Meskenaite, 1997).
CR cells avoid pyramidal cells (Gulyías, Hajós, & Freund,
Parkinson, Baimbridge, & Lewis, 1994; Gabbott & Bacon, 1996). Other studies (Dombrowski, Hilgetag, & Barbas, 2001; Elston & Gonzalez-Albo, 2003; Kondo, Tanaka, Hashikawa, & Jones, 1999) also found that CB and CR interneurons are predominant in PFC areas, espe-
cially in the superficial layer 2/3, but the precise percentage
differs among studies presumably due to method-
ological differences.
A circuit model (Wang et al., 2004) suggests how these
different interneuron types may work together in the PFC (Figure 15–4). This model incorporates three sub-
types of interneurons classified according to their synap-
tic targets and their prevalent interconnections. First, PV
interneurons project widely and preferentially target the
persomatic region of pyramidal neurons, thereby con-
trolling the spike output of principal cells and sculpturing
the tuning of the network activity pattern. Second, CB
interneurons act locally within a cortical column. They
predominantly target dendritic sites of pyramidal neurons,
ence controlling the inputs onto principal cells. Third, CR
interneurons also act locally and project preferentially
to CB interneurons. Note that the three interneuron types
in the model should be more appropriately interpreted
according to their synaptic targets rather than their cal-
cium-binding protein expressions. For example, PV cells
display a variety of axonal arbors, among which the large
basket cells are likely candidates for the widely projecting
perisoma-targeting cells (Kisvarday et al., 2003; Krimer &
Goldman-Rakic, 2001). Similarly, CB interneurons show
a high degree of heterogeneity, but some of them (such as
double bouquet cells or Martinotti cells) are known to act
locally and preferentially target dendritic spines and shafts
of pyramidal cells (DeFelipe, 1997; Somogyi et al., 1998).
Finally, although many CR interneurons do project to
pyramidal cells, anatomical studies show that a subset of
CR cells avoid pyramidal cells (Gulyás, Hajós, & Freund,
1996), at least in the same cortical layer (Meskenaite, 1997)
and preferentially target CB interneurons (DeFelipe,
Gonzalez-Albo, Del Rio, & Elston, 1999). It is also possible
that axonal innervations of a CR cell project onto pyra-
midal cells in a different cortical layer while selectively
targeting inhibitory neurons in the same layer (Gonchar &
Burkhalter, 1999; Meskenaite, 1997).
Figure 15–5A-B shows a computer simulation of a bio-
physically detailed implementation of this circuit model
for a spatial working memory task. When pyramidal cells
in a column are excited by a transient extrinsic input, they
excite each other through interconnections. At the same
time, activated CR interneurons suppress CB interneu-
rons within the same column, leading to reduced inhibi-
tion (disinhibition) of the dendrites of the same pyramidal
cells. The concerted action of recurrent excitation and CR
interneuron-mediated disinhibition could generate self-sus-
tained persistent activity in these neurons, and the network
activity pattern is shaped by synaptic inhibition from PV
interneurons. Moreover, CB interneurons in other columns
might be driven to enhance their firing activity. Therefore,
pyramidal cells in the rest of the network would become
less sensitive to external inputs, ensuring that working
memory storage is not vulnerable to behaviorally irrelevant
distracters. In the model, fast-spiking PV interneurons have
broader spatial tuning curves than pyramidal cells, consist-
ent with physiological observations from monkey experi-
ments (Constantinidis & Goldman-Rakic, 2002). Another
prediction of this model is that a small fraction of (putative
CB) PFC neurons recorded from a behaving monkey should
show a reduced firing rate during the delay relative to sponta-
aneous activity selectively for some sensory cues (inverted
tuning of mnemonic delay period activity). This prediction
was confirmed in data analysis from a monkey spatial work-
ning memory task (Figure 15–5C). Roughly 5% of recorded
neurons in that experiment showed behavior that was pre-
picted by the model for dendrite-targeting CB interneu-
rons, consistent with the crude estimate of CB-containing
interneurons (~24% of GABAergic cells, which, in turn,
represent ~20% of all neurons). Future physiological work
is needed to test the hypothesized disinhibition mechanism
and assess whether it may be especially prominent in work-
ing memory circuits.

DEcision MAKING

Unexpectedly, the same model originally developed for
working memory turned out to be well suited to account for
decision-making processes as well (Wang, 2002). In ret-
spect, this is because both working memory and decision
making rely on slow reverberatory excitation for time integra-
tion of noisy inputs and persistent activity, as well as inhibi-
tion to sculpt the selectivity and winner-take-all competition
underlying categorical choice formation. Consider percep-
tual decision making (Newsome, Britten, & Movshon, 1989;
Parker & Newsome, 1998). In a two-alternative forced-choice
task, subjects are trained to make a judgment about the direc-
tion of motion (say, left or right) in a near-threshold stochas-
tic random dot display and to report the perceived direction
with a saccadic eye movement. Neurons in posterior parietal
cortex (Roitman & Shadlen, 2002; Shadlen & Newsome,
2001) and PFC (Kim & Shadlen, 1999) were found to exhibit
firing activity correlated with the animal’s perceptual choice.
In particular, in a reaction time version of the task, neurons in
the posterior parietal cortex are correlated with the subject’s
choice, and display quasi-linear ramping activity over time
that is slower in more difficult trials when motion coherence
is lower and the subject’s reaction times are longer. This neural
activity pattern is reminiscent of the noisy integrator (or drift
diffusion) model known in cognitive psychology (Bogacz,
Brown, Moehlis, Holmes, & Cohen, 2006; Gold & Shadlen,
2007; Usher & McClelland, 2001; Smith & Ratcliff, 2004;
Wang, 2008).
We used the same model designed for working memory to simulate this decision experiment. The only difference is that for the delayed response task only one stimulus is presented, whereas for the perceptual discrimination task conflicting sensory inputs are fed into competing neural subpopulations in a decision circuit (Furman & Wang, 2008; Liu & Wang, 2008; Wang, 2002; Wong, Huk, Shadlen, & Wang, et al., 2007; Wong & Wang, 2006). Specifically, in a two-pool version of the model, subpopulations of spiking neurons are selective for two choice options.
alternatives (e.g., A = left motion, B = right motion). Within each pyramidal neural group, strong recurrent excitatory connections can sustain persistent activity triggered by a transient preferred stimulus. The two neural groups compete through feedback inhibition from interneurons. Conflicting sensory inputs are fed into both neural pools in the circuit, with the motion strength c implemented as the relative difference in the inputs. Our model accounts not only for salient characteristics of the observed decision-correlated neural activity, but also quantitatively for the animal’s behavioral performance (psychometric function and reaction times; Wang 2002; Wong et al., 2007; Wong & Wang 2006). Figure 15–6A shows an extended circuit model (Lo & Wang 2006) with a cortical decision circuit and a downstream motor circuit (superior colliculus, SC). Also included is the direct pathway in the basal ganglia, with an input layer (caudate, CD) and an output layer (substantia nigra pars reticulata, SNr), which is known to play a major role in controlling voluntary movements (Hikosaka, Takikawa, & Kawagoe, 2000). As a neural pool in the cortex ramps up in time, so does its synaptic inputs to the corresponding pool of SC movement neurons as well as CD neurons. When this input exceeds a well-defined threshold level, an all-or-none burst of spikes is triggered in the SC movement cells, signaling a particular (A or B) motor output. In this scenario, a decision threshold (as a bound of the firing rate of decision neurons) is instantiated by a hard threshold of synaptic input for triggering a special event in downstream motor neurons. Figure 15–6B shows a sample trial of such a model simulation for
the visual motion direction discrimination experiment. The rate of ramping activity fluctuates from trial to trial, as a result of stochastic firing dynamics in the cortex, and is inversely related to the decision time on a trial-by-trial basis (Figure 15–6C). Moreover, when the task is more difficult (with lower motion coherence), ramping activity is slower, leading to longer reaction times. However, the threshold of cortical firing activity that is read out by the downstream motion system has the same narrow distribution (inserts in Figure 15–6C), regardless of the ramping speed or reaction times.

Therefore, the variability of reaction times is mostly attributed to the irregular ramping of neural activity itself rather than to the trial-to-trial variability of the decision bound. This model reproduced the monkey’s behavioral performance and reaction times quantitatively (Figure 15–6D). Interestingly, it was found that, in this model, the decision threshold can be most effectively adjusted (e.g., in a speed-accuracy trade-off) by tuning the strength of projections from the cortex to the striatum, suggesting a specific role of reward-dependent plasticity at the corticostriatal synapses in flexible decision making (Lo & Wang, 2006). This attractor network model has also been applied to a number of other decision experiments, especially a two-interval discrimination task with vibrotactile stimuli (Deco, Pérez-Sanagustín, de Lafuente, & Romo, et al., 2007a; Deco, Scarano, & Soto-Faracos, 2007b; Machens, Romo, & Brody, 2005).

Furthermore, our model endowed with learning can describe adaptive choice behavior in reward-seeking tasks. Consider the neural network shown in Figure 15–6A. Recall that the network’s behavior is described by a softmax decision criterion: in a single trial, the probability of choosing A versus B is a sigmoid function of the difference in the inputs to the two competing neural pools (Figure 15–6D, upper panel). Suppose that the strengths of the two synaptic connections are plastic; then synaptic modifications will alter the network’s decision behavior over time. Specifically, we used binary synapses that undergo Hebbian learning, namely, that synaptic
In addition, it is assumed that synaptic learning depends on reward signals, based on the observation that the dopamine signal could gate synaptic plasticity in the striatum (Shen, Flajolet, Greengard, & Surmeier, 2008; Wickens, Reynolds, & Hyland, 2003) and PFC (Matsuda, Marzo, & Otani, 2006; Otani, Daniel, Roisin, & Crépel, 2003; Xu & Yao, 2010). This is a synaptic implementation of reinforcement learning (Glimcher, 2011; Montague, Dayan, & Sejnowski, 1996; Schultz, 1998; Sutton & Barto, 1998), which is a driving force for valuation of choice options through experienced choice-outcome associations (Glimcher, 2003; Rushworth & Behrens, 2008; Sugrue, Corrado, & Newsome, 2005). For instance, synapses for inputs to decision neurons are potentiated only if the choice is rewarded; otherwise, they are depressed. Therefore, in a learning process, synapses acquire information about reward outcomes of chosen responses (action-specific values). As a result of synaptic modifications, the input strengths for the competing neural groups of the decision network vary from trial to trial, leading to adaptive dynamics of choice behavior (Figure 15–7A). Such a model has been shown to account for behavioral data and single-neuron physiological data in several experimental paradigms, such as foraging with probabilistic reward delivery on choice options (Lau & Glimcher, 2005; Soltani & Wang, 2006; Sugrue, Corrado, & Newsome, 2004); competitive games (Barraclough, Conroy, & Lee, 2004; Dorris & Glimcher, 2004; Glimcher, 2003; Soltani, Lee, & Wang, 2006), arbitrary sensorimotor mapping (Asaad, Rainer, & Miller, 1998; Fusi et al., 2007; Wise & Murray, 2000), and probabilistic inference in a weather prediction task (Soltani & Wang, 2010; Yang & Shadlen, 2007).

Figure 15–7 shows the performance of such a model for probabilistic inference in the weather prediction task. In this task, several (say, four) sensory cues are shown; each is associated with a weight of evidence (WOE), defined by log likelihood odds, that one of the two outcomes A (rain) or B (shine) is true. The subject is required to make a decision (rain or shine) based on the combined evidence, the sum of the WOEs of the four cues presented in a single trial (Gluck, Shohamy, & Myers, 2002). We found (Soltani & Wang, 2010) that summing log posterior odds, a seemingly complicated calculation, can be readily achieved, through approximations, by a plausible plasticity mechanism with bounded synapses in a decision circuit. A biophysically based neural network model implementation of the monkey weather-prediction task (Yang & Shadlen, 2007) quantitatively accounted for many behavioral and single-unit neurophysiological observations. Furthermore, when the choice alternatives have unequal priors, the model predicts deviations from the Bayes decision rule that are akin to an effect called “base-rate neglect” commonly observed in human studies, namely, there is an overestimate of the predictive power of each cue for the less probable outcome (Soltani & Wang, 2010). Therefore, the core mechanisms in our model might be sufficiently general to describe not only simple reward-seeking tasks, but also more complex probabilistic problem solving.

**Figure 15–7** Probabilistic inference in a decision-making circuit endowed with reinforcement learning. (A) Schematic of the three-layer model for a weather prediction task. The first layer consists of cue-selective neural populations; each is activated upon the presentation of a cue. The sensory cue-selective neurons provide, through synapses that undergo reward-dependent Hebbian plasticity, inputs to two neural populations in an intermediate layer that encode reward values of two choice alternatives (action values). Combination of cues is accomplished through convergence of cue-selective neurons onto action value-encoding neurons. The latter project to a decision-making circuit (gray box, same as the cortical circuit in Figure 15–6A). The choice (A or B) is determined by which of the two decision neural populations wins the competition on a trial. Depending on the reward schedule, a chosen action may be rewarded or not. The presence (or absence) of a modulatory reward signal leads to potentiation (or depression) of plastic synapses. (B) Choice behavior of the model in the weather prediction task. This is the probability of choosing alternative A as a function of the sensory evidence favoring this option, defined by the sum of log likelihood odds of four cues. Source: Adapted from Soltani and Wang (2010) with permission.
PROACTIVE CONTROL OF MOTOR RESPONSES

More recently, we extended our recurrent network to inhibitory control of behavioral responses, an important executive function that depends on the PFC (Aron, 2007) and is impaired in psychiatric illnesses such as attention-deficit hyperactivity disorder (ADHD; Armstrong & Munoz, 2003; Schachar, Tannock, & Logan, 1993). In a countermanding (or stop-signal) task, a subject has to withhold the response to a go signal when an infrequent and delayed stop signal appears (Boucher, Palmeri, Logan, & Schall, 2007; Logan & Cowan, 1984). Neurophysiological recordings from the frontal eye field and the superior colliculus of rhesus monkeys suggested that the GO and STOP processes may be instantiated by movement and fixation neurons, respectively, and that the antagonistic interplay between them may be responsible for the interruption of the GO process (Hanes, Patterson, & Schall, 1998). We used the recurrent (attractor) network approach to build a spiking network model for countermanding, in which a population of movement (GO) neurons interacts with a population of fixation (STOP) neurons through a shared population of inhibitory GABAergic cells (Figure 15–8A; Lo, Boucher, Pare, Schall, & Wang, 2009). We found that this model can capture quantitatively the monkey’s performance in a stop-signal task, such as the inhibition function (the probability that a planned response is not canceled by a stop signal as a function of the delay between the go and stop signals) and reaction times (Figure 15–8B).

Figure 15–8C shows the neural dynamics of two sample stop-signal trials, where an external stop input (to the STOP neural population) is presented 160 ms after a go input (to the GO neural population). Importantly, fixation neurons display high baseline activity from the start, as observed in physiological studies (Hanes et al., 1998), unlike in the conventional race model for the stop-signal task, where the STOP process is zero until a stop signal is presented (Boucher et al., 2007; Logan & Cowan, 1984). After the go signal onset, while firing activity of movement neurons ramps up, firing activity of fixation neurons decreases (due to effectively inhibitory interactions between the two and the withdrawal of external input) but remains at a relatively high level. This internally maintained tonic activity of fixation neurons, observed experimentally (Hanes et al., 1998), tends to be higher or longer-lasting in a trial when the planned response is successfully canceled by a stop signal (top) than otherwise (bottom). Critically, this proactive component of an inhibitory control signal, and the reciprocal interactions between the GO and STOP processes, are present even in go trials when no stop signal is presented. This idea is further illustrated in Figure 15–8D, where the firing rates of movement neurons (r_mov) and fixation neurons (r_fix) are plotted against each other in a two-dimensional “phase-plane.” After the go signal onset (Figure 15–8D, top panel), the system moves from its initial state toward an attractor state corresponding to the generation of a motor response (indicated by the black filled circle, with high r_mov and low r_fix). When a stop signal is presented (Figure 15–8D, bottom panel), the phase plane landscape is changed, with the creation of another attractor state corresponding to response cancellation (with high r_fix and low r_mov). The yellow and gray regions define the “basins of attraction” of these two attractors: in the absence of noise, the system would converge to one of the attractors, depending on whether its internal state at the stop-signal onset is within one or the other basins of attraction (green and gray filled circles). In this sense, inhibitory control is proactive: to a large extent (depending on the amount of noise present), whether a planned response is ultimately canceled in a stop-signal trial is determined by the internal state and network dynamics before any external stop input is shown.

This model suggests that a possible way to adjust the amount of inhibitory control depending on behavioral demands is to tune the baseline activity level of fixation neurons (STOP process) by top-down signaling presumably from the PFC (Aron, 2007). Generally, this work demonstrates that the framework of strongly recurrent circuit models is applicable to inhibitory control of behavior.

REPRESENTATION OF TASK RULES: NEURAL HETEROGENEITY AND MIXED SELECTIVITY

So far, we have discussed persistent activity underlying working memory of sensory stimuli. However, internal representation is not limited to sensory information but can be more abstract. In particular, behavioral rules, prescribed guides for actions and problem solving, must be actively maintained in order to carry out tasks in our daily lives, and the PFC is known to play a major role in rule representation and rule learning (Buckley et al., 2009; Bunge, 2004; Goldman-Rakic, 1987; Miller & Cohen, 2001; Milner, 1963; Wallis, Anderson, & Miller, 2001). Can our attractor network models be generalized to represent task rules?

We found that our existing models are inadequate for describing rule-based behavior. The difficulty arises from the fact that flexible context or rule-based behaviors involve neural computations akin to logic operations like exclusive Or (XOR). For example, consider a simplified version of the Wisconsin Card Sorting Test. Given a sensory cue (a colored shape, e.g., a red circle), a subject selects one of two test stimuli that matches the cue either in color or shape, depending on the task rule (color or shape; Mansouri, Matsumoto, & Tanaka, 2006). Presumably, the rule currently in play, say color, is represented internally by persistent activity of “color rule cells,” which must be maintained across trials but switched off when the rule has changed—for example, from color to shape. As explained in
Figure 15–8 A spiking neural network model of inhibitory control for the countermanding task. (A) The schematic model architecture. The premovement module consists of two populations of movement (GO) neurons, a population of fixation (STOP) neurons, and a population of inhibitory interneurons. The fixation neurons receive a top-down signal for proactive inhibitory control. (B) The model quantitatively fits to the behavioral data from a rhesus monkey (Boucher et al., 2007). Top panel: the inhibition function shows the probability of a noncanceled response in stop-signal trials as a function of stop-signal delay (SSD, the difference in the go and stop signal onset times). Bottom panel: cumulative reaction time distributions of nostop-signal trials (rightmost) and noncanceled stop-signal trials for different SSDs. (C) Simulated population firing rates of MOV (black lines) and FIX (gray lines) neurons on top of corresponding spike rasters (color-matched dots) from two trials of the model simulation (top, a canceled trial; bottom, a noncanceled trial). Each row in the rastergrams represents the spike train from one neuron. The vertical dashed line indicates the onset of the stop signal (SSD of 160 ms). Arrows indicate the offset of the top-down control signal. The model suggests that a quicker offset of the top-down control results in faster ramping-down activity of fixation neurons and faster ramping-up activity of movement neurons, thus increases the probability of making a saccade to the target. (D) Phase plane plots demonstrate the attractor dynamics of the model network. Black and red curves represent the nullclines for the MOV and FIX neurons, respectively, and the intersections between the nullclines determine the equilibrium points (black circles: stable; gray circles: unstable) of the network. The brown curves (canceled trials) and green curves (noncanceled trials) depict the trajectories of eight trials. Depending on the state of the system at the moment when the stop-signal input starts (indicated by the color-matched circles in the top panel), the network may continue converging into the GO attractor and trigger a response or may turn back into the STOP attractor and cancel the response (bottom panel). Dashed lines mark the threshold for the saccade response. Insets: schematic plots illustrating how the one-dimensional “action landscape” changes with stimulus inputs. Source: Adapted from Lo et al. (2009) with permission.
Figure 15–9A, the required state transitions of a color rule cell as a function of its recurrent input and feedback input amounts to an XOR operation. This problem is equivalent to the known problem of nonlinear separability of the Boolean operation of XOR, and it plagues most neural networks implementing context-dependent tasks.

We proposed a solution to this problem (Rigotti, Rubin, Wang, & Fusi, 2010) by adding to a decision-making circuit a large “reservoir” of randomly connected neurons (RCNs; Figure 15–9B). The basic idea is that by virtue of random connections, RCNs are naturally activated by a combination of synaptic inputs from external stimuli as well as rule-coding neurons (e.g., the color rule is currently in play and the network receives a negative feedback signal), and such mixed selectivity is exactly what is needed to perform the task (see also Dayan, 2007). This model provides a general framework for describing context- or rule-dependent tasks (Rigotti et al., 2010). Figure 15–9C–D shows an implementation of such a network for the simplified Wisconsin Card Sort Test. Notable is the high degree of variability of firing activity, across cells and for a single neuron across task epochs. Heterogeneity and mixed selectivity are salient yet puzzling characteristics of frontal cortical neurons recorded from behaving animals (Asaad et al., 1998; Lapish, Durstewitz, Chandler, & Seamans, 2008; Miller & Cohen, 2001; Sigala, Kusunoki, Nimmo-Smith, Gaffan, & Duncan, 2008). Our model suggests that mixed selectivity is computationally desirable, as it allows the network to encode a large number of facts, memories, events, and, importantly, their combinations that are critically important for enabling the PFC to subserve context- and rule-dependent flexible behavior.

**INSIGHTS INTO PREFRONTAL DYSFUNCTION IN MENTAL ILLNESS**

Inasmuch as the PFC is central to multiple facets of cognition and executive control, it is a major focus of basic and clinic research on the brain mechanisms of cognitive deficits associated with neuropsychiatric disorders, including schizophrenia (Goldman-Rakic, 1999), autism (Amaral, Schumann, & Nordahl, 2008; Shalom, 2009), ADHD (Casey, Nigg, & Durston, 2007), and obsessive-compulsive behavior (Sachdev & Malhi, 2005).

Reciprocal network modeling has given rise to a number of specific candidate explanations for the frontal lobe dysfunction associated with mental disorders, especially schizophrenia (Durstewitz & Seamans, 2008; Rolls, Deco, & Winterer, 2008; Wang, 2006b). Traditionally, the function of NMDA conductance is almost exclusively emphasized in terms of its role in long-term synaptic potentiation and depression. Thus, an abundance of NMDARs (Scherzer et al., 1998) could reflect a high degree of plasticity of prefrontal microcircuits, which could subserve learning of flexible and adaptive behaviors (Fusi et al., 2007; Miller & Cohen, 2001). That may be, but we propose that NMDARs also directly mediate the slow excitatory synaptic transmission critically important to working memory and decision-making processes. If so, effects on cognitive behavior due to NMDA signaling alternations may also be partly accounted for by impaired working memory and decision-making functions in addition to long-term memory.

We showed that hypofunction of NMDARs at intrinsic prefrontal synapses is detrimental to the persistent activity underlying working memory. These results provide a mechanistic explanation for why working memory dysfunction similar to that observed in schizophrenic patients can be induced in healthy subjects by subanesthetic doses of ketamine, a noncompetitive NMDAR antagonist; and such impairment presumably does not involve learning and long-term plasticity (Krystal et al., 1994). Postmortem studies showed significant alterations in NMDAR mRNA expression (Akbarian et al., 1996), but revealed either no abnormality (Healy et al., 1998) or a slight increase (Dracheva, McGurk, & Haroutunian, 2005) in the AMPAR level. Available information does not yet permit a more precise explanation for why and how impairment of the NMDAR system causes the cognitive deficits associated with schizophrenia. It was previously suggested that impairment can occur outside of PFC, such as in hippocampus (Grune et al., 1996; Jodo et al., 2005; Rowland et al., 2005) or in the dopamine system (Carlsson et al., 2001). Again, functional implications tend to be discussed in the realms of learning and synaptic modification. By contrast, our modeling work suggests a novel scenario focused on the role of NMDARs in persistent activity. Of course, this scenario is compatible with other proposals, given that impairment of NMDARs may not be restricted to a single pathway and that NMDARs play a major role in long-term synaptic plasticity (Stephan, Friston, & Frith, 2009). These different facets of NMDAR function are also under the influence of dopamine modulation (Chen, Greengard, & Yan, 2004; Huang et al., 2004).

On the other hand, there is mounting evidence that the dorsolateral PFC of schizophrenic patients shows an abnormality of selective interneuron subtypes, especially fast-spiking basket and chandelier cells (Lewis, Hashimoto, & Volk, 2005). Our model suggests that this may be the case for two reasons. Modeling work (Brunel & Wang, 2001; Compte et al., 2000; Wang, Tegnér, Constantinidis, & Goldman-Rakic, 2004), in concordance with physiological experiments (Constantinidis & Goldman-Rakic, 2002; Rao et al., 2000), demonstrates that inhibition mediated by fast-spiking and broadly projecting interneurons is critical to the stimulus selectivity, and hence information specificity, of mnemonic persistent activity. Moreover, fast-spiking GABAergic cells are critical to the generation of coherent gamma (40 Hz) oscillations (Traub, Bibbig,
LeBeau, Buhl, & Whittington, 2004; Traub, Whittington, Collins, Buzsáki, & Jefferys, 1996; Wang, 2010; Wang & Buzsáki, 1996), which may contribute to a variety of cognitive processes (Wang, 2010). Revealingly, gamma oscillations appear to be decreased in schizophrenic brains compared to those of control subjects (Lee, 2003; Spencer et al., 2004; Uhlhaas, Haenschel, Nikolić, & Singer, 2008). Thus, deficits in synaptic inhibition could impair the quality of information stored in working memory as well as neural communication across brain areas.
We found that inhibition is also crucial for robust working memory despite ongoing sensory flow. This finding provides another insight into how dopamine may affect prefrontal functions (Brunel & Wang, 2001; Durstewitz, Seamans, & Sejnowski, 2000b). It is known that dopamine acts on PFC partly through modulation of glutamatergic and GABAergic synaptic transmissions (Seamans & Yang, 2004).

Modeling work suggests several ways in which dopamine modulation might affect a cognitive circuit function. First, a relatively small increase by dopamine of recurrent connections (while preserving the E-I balance) can lead to significant enhancement of the network’s resistance against distractors (Brunel & Wang, 2001; Durstewitz et al., 2000a). Conversely, mild impairment of dopamine signaling in the PFC can result in the behavioral distractibility associated with mental disorders such as schizophrenia. Second, if dopamine D1 receptor activation modulates NMDARs with a higher sensitivity in pyramidal cells than in inhibitory cells (Muly, Szigeti, & Goldman-Rakic, 1998), then too little dopamine might imply insufficient excitation and too much dopamine might lead to excessive inhibition, resulting an inverted U-shaped dopamine dependence (Figure 15–10), as observed experimentally (see Arnst et al., 2010, for a review). Caution is warranted here, as the actual biological mechanism underlying the inverted-U dopamine action is presently unknown. Certain subtypes of inhibitory neurons have significant NMDARs, whereas others do not (Wang & Gao, 2009), and dopamine D1 and D2 receptors modulate a number of synaptic and ion channel targets (Arnst et al., 2010; Seamans & Yang, 2004).

Furthermore, according to the disinhibition mechanism (Figures 15–4 and 15–5), dendritic inhibition is reduced locally in activated pyramidal cells but increased in those pyramidal cells not engaged in encoding the shown stimulus. This mechanism, mediated by CB interneurons, could serve to filter out distracting stimuli, and it is enhanced with a larger dendritic/somatic inhibition ratio (Wang et al., 2004). A high dendritic/somatic inhibition ratio can be achieved in a working memory circuit be hard-wired, for example with a large proportion of CB cells in PFC. Alternatively, it can also be dynamically controlled by neuromodulators such as dopamine. Interestingly, an in vitro work suggests that dopamine D1 receptor activation precisely increases the ratio of dendritic/somatic inhibition onto pyramidal cells in PFC (Gao, Wang, & Goldman-Rakic, 2003). Using double intracellular recording in PFC slices and morphological reconstruction, it was found that bath application of dopamine had a dual effect on inhibitory synaptic transmission in a pyramidal cell of the PFC. Dopamine was found to reduce the efficacy of inhibitory synapses onto the perisomatic domains of a pyramidal cell, mediated by fast-spiking interneurons, whereas it enhanced inhibition at synapses from accommodating or low-threshold spiking interneurons that target the dendritic domains of a pyramidal cell (Gao et al., 2003). Our model predicts a specific function for such a dual dopamine action: it could boost the ability of a working memory network to filter out behaviorally irrelevant distracting stimuli. Our modeling work (Brunel & Wang, 2001), as well as brain imaging (Sakai, Rowe, & Passingham, 2002), points to a possible physiological basis of the clinical literature documenting distractibility as a common symptom of frontal lobe damage (Fuster, 2008; Goldman-Rakic, 1987; Mesulam, 2000).

**CONCLUDING REMARKS**

In this chapter, I discussed biophysically based neural modeling that, in concert with experiments, provides a powerful tool for investigating the cellular and circuit mechanisms of mnemonic persistent activity in delayed response tasks. This approach has been used to assess whether the attractor model for working memory and decision making can be instantiated by biologically plausible mechanisms. Theoretical work suggests that slow excitatory reverberation underlies persistent activity in working memory and time integration in decision making. A candidate cellular substrate is the NMDARs at local recurrent synapses; an alternative/complementary scenario involves intrinsic

![Differential D1 modulation of NMDA conductances](Image)

*Figure 15–10 Inverted-U dopamine action implemented by differential D1 modulation of NMDA conductances in pyramidal neurons and interneurons. The state diagram shows that persistent activity is the highest in an intermediate range of D1 modulation. Three simulations at different levels of D1 modulation, indicated by filled circles, are shown in the upper panels, demonstrating that too high or too low D1 activation would be suboptimal or detrimental to working memory behavior. Source: Adapted from Brunel and Wang (2001) with permission.
channels and calcium dynamics in single cells. Recurrent excitation must be balanced by feedback inhibition, which is mediated by several types of GABAergic interneurons. We found that inhibitory circuitry plays a key role in stimulus selectivity (similar to that in sensory areas) and network resistance against distracting stimuli (a cardinal requirement for robust working memory), as well as in winner-take-all competition in decision making.

We have confined ourselves to models in which working memory storage is maintained by roughly tonic (constant) spike discharges in a neural assembly across a delay period. However, many cortical cells exhibit delay activity that is not stationary but ramps up or down over time (Barak, Tsodyks, & Romo, 2010; Brody, Hernández, Zainos, & Romo, 2003; Chafee & Goldman-Rakic, 1998; Fuster, 2008). Such ramping activity can conceivably be achieved in a two-layer network, in which first-layer neurons show tonic delay activity and second-layer neurons slowly integrate inputs from the first-layer neurons in the form of ramping activity (Miller, Brody, Romo, & Wang, 2003), or it can result from a very slow biophysical process such as DSI (Carter & Wang, 2007). Functionally, such ramping activity may represent elapsed time during the delay period (Brody et al., 2003; Machens, Romo, & Brody, 2010). Moreover, persistent activity patterns can exhibit chaotic dynamics, or occur as a firing pattern that moves from one neural group to another in a circuit (Baeg et al., 2003). A challenge in the field is to understand the heterogeneity and temporal variations of mnemonic activity and how the specificity of stored information can be preserved in dynamically moving neural activities. It is worth stressing that rich temporal behavior does not necessarily contradict the attractor network paradigm, since an attractor state does not have to be a steady state and can be a complex spatiotemporal pattern. The more basic question is whether there exist multiple stable persistent states (each with potentially very complex dynamics) or, alternatively, whether there is no multistability and memory traces are transient events that can be decoded long after stimulus presentation (Ganguli, Huh, & Sompolinsky, 2008; Goldman, 2009).

Our emphasis on internal representations by no means underestimates the importance of processes such as action selection. Rather, we propose that PFC does not simply send out nonspecific “control signals” and that representational information is indispensable to processes. As it turns out, our model is capable of both working memory maintenance and decision-making computations. These results suggest that it may not be a coincidence that decision-related neural activity has been found in the same cortical areas that also exhibit persistent activity during working memory (Gold & Shadlen, 2007; Romo & Salinas, 2001; Schall, 2001; Wang, 2008). In our model, both working memory and decision making rely on slow reverberatory dynamics that gives rise to persistent activity and time integration, and inhibitory circuitry that leads to selectivity and winner-take-all competition. Thus, we are beginning to unravel the microcircuit properties of a “cognitive” cortical area (such as PFC in contrast to, say, primary visual cortex) that enable it to serve multiple cognitive functions. At a fundamental level, these studies point to a unified view of why and how “cognitive” cortical area can serve both internal representation (active working memory) and processing (decision, action selection, etc.).

Microcircuitry is at a level of complexity ideally suited for bridging the gap between cognitive network functions and the underlying biophysical mechanisms. The delicate balancing act of recurrent excitation and feedback inhibition is at the heart of the strongly nonlinear dynamics that underlies cognitive processes in PFC. In this sense, microcircuit neurodynamics provides the critical link from molecules to behavior, and ultimately holds the key to a theoretical foundation for neuropharmacology and molecular psychiatry (Harrison & Weinberger 2005). One of the major challenges for future research is to elucidate the differential dynamics, computations, and functions of distinct frontal subregions, and to understand how they work together in an interconnected circuit as well as with the rest of the brain. Progress in this direction at the system level of large-scale brain circuitry with multiple interacting modules, together with the tremendous advances in genome science, will prove to be especially promising in the next decade.

ACKNOWLEDGMENTS
I am grateful to Albert Compte, Nicolas Brunel, Alfonso Renart, Jesper Tegnér, Paul Miller, Alireza Soltani, Gene Carter, Chung-Chuan Lo, Stefano Fusi, Feng Liu, Marcelo Camperi, Mattia Rigotti and Daniel Ben Dayan Rubin for their contributions. I would also like to thank Chung-Chuan Lo, John Murray, Mattia Rigotti, and Rishidev Chaudhuri for their help with figures.

DISCLOSURE STATEMENT
This work was supported by NIH grant MH062349 and by the Kavli Foundation. This chapter has drawn material from Wang (2006a). X.-J. W. has no conflicts of interest to disclose.

REFERENCES


15. WORKING MEMORY AND DECISION MAKING


15. WORKING MEMORY AND DECISION MAKING

15. WORKING MEMORY AND DECISION MAKING


Corrigendum


Figure caption 15-4 should be for figure 15-3, caption 15-5 for figure 15-4, and caption 15-3 for figure 15-5.