51 Ignatowski, T.A. *et al.* (1999) Brain-derived TNFα mediates neuropathic pain. *Brain Res.* 841, 70–77
52 Solari, R. *et al.* (1994) Receptor mediated endocytosis and intracellular fate of interleukin 1. *Biochem. Pharmacol.* 47, 93–101
53 Ramer, M.S. *et al.* (1998) Spinal nerve lesion-induced mechanoallodynia and adrenergic sprouting in sensory ganglia are attenuated in interleukin-6 knockout mice. *Pain* 78, 115–121
54 Covey, W.C. *et al.* (2000) Brain-derived TNFalpha: involvement in neuroplastic changes implicated in the conscious perception of persistent pain. *Brain Res.* 859, 113–122
55 DeLeo, J.A. *et al.* (2000) Transgenic expression of TNF by astrocytes increases mechanical allodynia in a mouse neuropathy model. *NeuroReport* 11, 599–602
56 Sweitzer, S.M. *et al.* IL1ra and sTNFr reduce mechanical allodynia and spinal cytokine expression in a model of neuropathic pain. *Neuroscience* (in press)
57 Laughlin, T.M. *et al.* (2000) Cytokine involvement in dynorphin-induced allodynia. *Pain* 84, 159–167
58 Tadano, T. *et al.* (1999) Induction of nociceptive responses by intrathecal injection of interleukin-1 in mice. *Life Sci.* 65, 255–261
59 Arruda, J.L. *et al.* (2000) Antibody and IgG attenuates mechanical allodynia in a mononeuropathy model in the rat: potential role of immune modulation in neuropathic pain. *Brain Res.* 879, 216–225
60 Murphy, P.G. *et al.* (1999) Endogenous interleukin-6 contributes to hypersensitivity to cutaneous stimuli and changes in neuropeptides associated with chronic nerve constriction in mice. *Eur. J. Neurosci.* 11, 2243–2253
61 Watkins, L.R. *et al.* (1999) Dynamic regulation of the proinflammatory cytokine, interleukin-1 beta: molecular biology for non-molecular biologists. *Life Sci.* 65, 449–481
62 Watkins, L.R. *et al.* Glial proinflammatory cytokines mediate exaggerated pain states: implications for clinical pain. In *Immune Mechanisms of Pain & Analgesia* (Machelska, H. and Stein, C., eds), Landes Biosciences (in press)
63 DeLeo, J.A. *et al.* (1996) Interleukin (IL)-6 mediated hyperalgesia/allodynia and increased spinal IL-6 in a rat mononeuropathy model. *J. Interferon Cytokine Res.* 16, 695–700
64 Tyor, W.R. *et al.* (1992) Cytokine expression in the brain during the acquired immunodeficiency syndrome. *Ann. Neurol.* 31, 349–360
65 VanLeeuwen, R. *et al.* (1996) Evaluation of safety and efficacy of 3TC (lamivudine) in patients with asymptomatic or mildly symptomatic human immunodeficiency virus infection: a phase I/II study. *J. Infect. Dis.* 171, 1166–1171
66 Watkins, L.R. and Maier, S.F. (1997) The case of the missing brain: arguments for a role of brain-to-spinal cord pathways in pain facilitation. *Behav. Brain Sci.* 20, 469
67 Chacur, M. *et al.* A new model of sciatic inflammatory neuritis (SIN): induction of unilateral and bilateral mechanical allodynia following acute unilateral peri-sciatic immune activation in rats. *Pain* (in press)
68 Watkins, L. *et al.* Spinal cord glia: new players in pain. *Pain* (in press)

Synaptic reverberation underlying mnemonic persistent activity

Xiao-Jing Wang

Stimulus-specific persistent neural activity is the neural process underlying active (working) memory. Since its discovery 30 years ago, mnemonic activity has been hypothesized to be sustained by synaptic reverberation in a recurrent circuit. Recently, experimental and modeling work has begun to test the reverberation hypothesis at the cellular level. Moreover, theory has been developed to describe memory storage of an analog stimulus (such as spatial location or eye position), in terms of continuous 'bump attractors' and 'line attractors'. This review summarizes new studies, and discusses insights and predictions from biophysically based models. The stability of a working memory network is recognized as a serious problem; stability can be achieved if reverberation is largely mediated by NMDA receptors at recurrent synapses.

Xiao-Jing Wang
Volen Center for Complex Systems, Brandeis University, 415 South Street, Waltham, MA 02254-9110, USA.
e-mail: xjwang@brandeis.edu

A fundamental ability of the brain is to actively hold an item of information 'on-line' in short-term memory. The stored information can be a sensory stimulus that guides a prospective action, such as a delayed perceptual decision or a delayed behavioral response. It can also be an item retrieved from long-term memory, for example when the memory of a face is activated and used in the visual search of a friend in a crowd. The obligatory physical process underlying active (working) memory is persistent neural activity that is sustained internally in the brain, rather than driven by inputs from the external world. Persistent activity provides the cellular basis of 'a central neural mechanism', as postulated by Hebb, 'to account for the delay, between stimulation and response, that seems so characteristic of thought'[1]. In order for a neural persistent activity to subserve working memory, it must be stimulus-selective, and therefore information-specific. Moreover, it must be able to be turned on and switched off rapidly (≈100 ms) by transient inputs.

For 30 years, persistent activity in the cortex has been documented by numerous unit recordings from behaving monkeys during working memory tasks (Box 1). How does stimulus-selective persistent activity arise in a neural network? Can we explain persistent activity in terms of the biophysics of neurons and synapses, and circuit connectivity? Recent experiments and computational modeling have been devoted to these mechanistic questions; these studies have led to new insights into the cellular basis of mnemonic neural activity in behaving animals.

Reverberatory excitation: how localized can it be?

The central idea is that recurrent excitatory loops within a neural network can sustain a persistent activity in the absence of external inputs[1,2]. To test this hypothesis, a key issue is to identify the crucial and minimum anatomical substrate for reverberation. In the neocortex, several scenarios are conceivable.

Thalamocortical loop

Persistent activity can arise from a large neural network that involves subcortical systems, through

 0166-2236/01/$ – see front matter PII: S0166-2236(00)01868-3

Box 1. Persistent activity as neural correlate of working memory

Cortical 'memory neurons' that show persistent activity are typically recorded during a delayed response task, in which a monkey is required to retain the information of a sensory cue across a delay period between the stimulus and behavioral response. Memory cells were first found, and seem to be especially abundant, in the prefrontal cortex (PFC)[a–e]. The crucial role of the PFC in working memory is also supported by lesion[f,g] and brain imaging studies[h]. However, neural persistent activity is a widespread phenonemon in association cortices, including the posterior parietal cortex[i–k], and the inferotemporal cortex[l,m]. According to the type of sensory stimulus that is encoded for storage, one can distinguish three kinds of working memory.

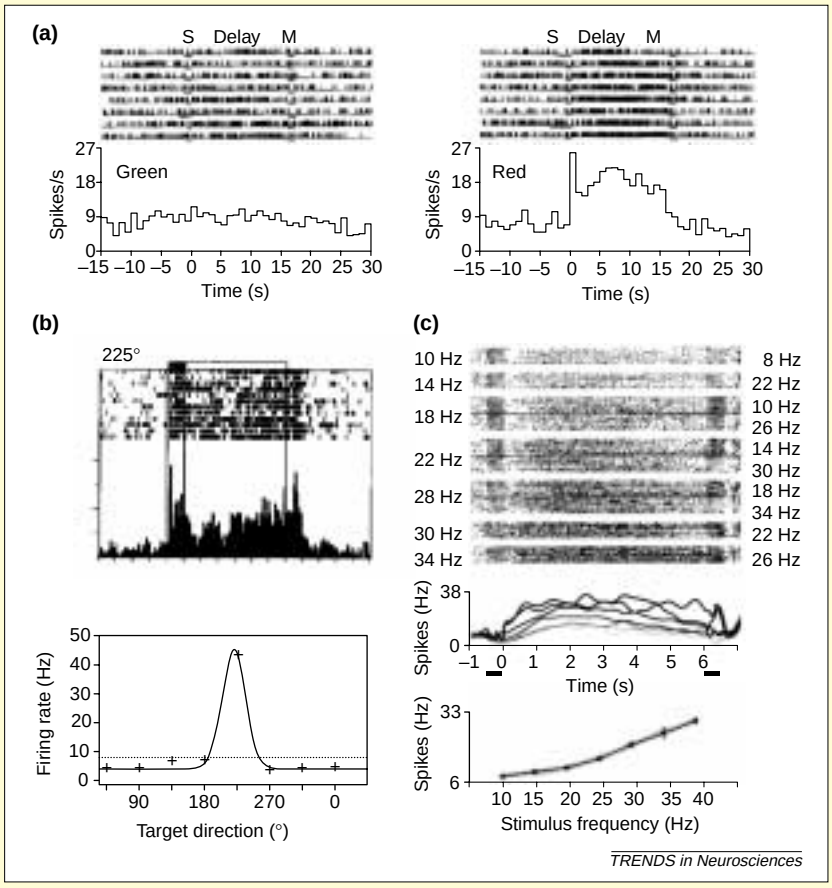

Fig. I. Three types of working memory encoding. (a) Discrete working memory. In a delayed match (M)-to-sample (S) experiment, an inferotemporal neuron shows sustained high activity for the color red (but not green) of a visual cue, during a delay period of 16 s. Redrawn, with permission, from Fuster and Jervey[l]. (b) Spatial working memory. In a delayed saccade experiment, a prefrontal neuron shows persistent activity that is tuned to a preferred location of a visual cue. Upper panel: rasters and cumulative spike histogram for a preferred cue; lower panel: spatial tuning curve of delay period activity. Redrawn, with permission, from Funahashi *et al.*[c] (c) Parametric working memory. In a delayed somatosensory discrimination task, a neuron in the inferior convexity shows persistent activity with a firing rate proportional to the cue frequency. Upper panel: rasters. Cue stimulus frequency indicated on the left, comparison stimulus frequency indicated on the right. Middle panel: trial-averaged firing rates as a function of time. Lower panel: mean firing rates, averaged across the entire delay period, as a function of the cue frequency. Redrawn, with permission, from Romo *et al.*[e]

Discrete working memory

Figure Ia shows a delayed match-to-sample experiment, in which the behavioral response depends on the memory of one of the two items (the stimulus color, red or green). An inferotemporal neuron displays elevated activity through the entire delay period (16 s), which is selective to the color red. Such tasks engage a working memory circuit in which the stored information is a categorical feature of the stimulus or an object (a face, color or word). Arguably, the items form a discrete collection, and a given neuron or neural assembly is selective to one or a few items.

Spatial working memory

Figure Ib illustrates a delayed oculomotor experiment, in which a saccadic eye movement is guided by the memory of a spatial stimulus. In this case, the stored information is spatial location, which is an analog quantity. Neurons in the dorsolateral PFC display persistent delay activity that is spatially selective. The 'memory field' of a cell is characterized by a smooth tuning curve, peaked at a preferred spatial cue, which is different from cell to cell. The memory of a given spatial location is stored by the neural population in the form of a spatially localized persistent firing pattern, or a 'bump attractor'. Bump states are common to

reciprocal excitation between the cortex and thalamus, or via the cortico–striato–thalamo–cortical loop (Fig. 1a). Consistent with this scenario, during the delay period of a memory task, thalamic[3] and caudate[4] neurons show an elevated persistent activity; whereas the output neurons from basal ganglia show a sustained inhibition[5]. It is unknown whether the subcortical mnemonic activities are a mere reflection of persistent cortical inputs. To answer this question, it would be interesting to see whether cortical persistent activity survives the disruption of the closed loop, for example, by lesioning the thalamocortical pathway.

Reciprocal loop between cortical areas

It is also possible that persistent activity is maintained by reciprocal loops between cortical areas, such as between the prefrontal cortex (PFC) and the posterior parietal cortex, or between the PFC and the inferotemporal cortex (Fig. 1b). Persistent activity is present in all these areas (Box 1). However, to date there is no convincing evidence that such a loop is necessary for the generation of persistent activity. In an object-working-memory experiment using intervening stimuli (distractors), stimulus-specific mnemonic activity has been shown to be easily disrupted by distractors in inferotemporal neurons,

neural circuits that encode directional or spatial information, such as head-direction cells[n] and place cells[o].

Parametric working memory

Figure Ic shows a delayed somatosensory discrimination task, in which the monkey was trained to compare and discriminate the frequencies of two vibrotactile stimuli separated by a delay period. Here, the encoded information is the cue frequency, again an analog quantity. PFC cells in the inferior convexity show persistent activity that varies monotonically with the cue frequency. Presumably, there is a continuous range of persistent activity rates that vary in proportion to the stimulus parameter. Therefore, the stimulus is encoded by a rate code continuum: not by 'what' neurons show delay period activity, but by the firing rates at which they fire. A similar monotonic encoding scheme is used by oculomotor neurons for memory of the current eye position[p].

References

a Fuster, J. and Alexander, G. (1971) Neuron activity related to short-term memory. *Science* 173, 652–654
b Kubota, K. and Niki, H. (1971) Prefrontal cortical unit activity and delayed alternation performance in monkeys. *J. Neurophysiol.* 34, 337–347
c Funahashi, S. *et al.* (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61, 331–349
d Ranier, G. *et al.* (1998) Memory fields of neurons in the primate prefrontal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 95, 15008–15013
e Romo, R. *et al.* (1999) Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* 399, 470–473
f Fuster, J. (1988) *The Prefrontal Cortex.* (2nd edn), Raven Press
g Goldman-Rakic, P. (1987) Circuitry of the prefrontal cortex and the regulation of behavior by representational knowledge. In *Handbook of Physiology* (Plum, F. and Mountcastle, V., eds), pp. 373–417, American Physiological Society
h D'Esposito, M. *et al.*(1998) Functional MRI studies of spatial and non-spatial working memory. *Cognit. Brain Res.* 7, 1–13
i Gnadt, J.W. and Andersen, R.A. (1988) Memory related motor planning activity in posterior parietal cortex of macaque. *Exp. Brain Res.* 70, 216–220
j Colby, C.L. *et al.* (1996) Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *J. Neurophysiol.* 76, 2841–2852
k Chafee, M.V. and Goldman-Rakic, P.S. (1998) Neuronal activity in macaque prefrontal area 8a and posterior parietal area 7ip related to memory guided saccades. *J. Neurophysiol.* 79, 2919–2940
l Fuster, J.M. and Jervey, J.P. (1981) Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science* 212, 952–955
m Miyashita, Y. and Chang, H.S. (1988) Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* 331, 68–70
n Taube, J.S. (1998) Head direction cells and the neurophysiological basis for a sense of direction. *Prog. Neurobiol.* 55, 1–32
o Samsonovich, A. and McNaughton, B.L. (1997) Path integration and cognitive mapping in a continuous attractor neural network model. *J. Neurosci.* 17, 5900–5920
p Robinson, D.A. (1989) Integrating with neurons. *Annu. Rev. Neurosci.* 12, 33–45

but not in PFC neurons[6]. This result suggests that a neural circuit within PFC is by itself capable of sustaining a stimulus-selective persistent activity.

Recurrent local cortical network

The third view is that persistent activity can be produced locally within a cortical area (Fig. 1c). In particular, Goldman-Rakic (1995) proposed a columnarly organized cortical network model for the PFC, in which persistent activity arises from reverberatory excitation, and stimulus selectivity is formed by recurrent inhibition[7]. This proposal is consistent with the anatomical and electrophysiological data that show extensive horizontal excitatory connections in the PFC, especially in layer II/III (Refs 8–10). Moreover, prefrontal synapses display short-term plasticity on a time scale of 0.1–10 s (Refs 11,12), which might contribute to increase the synaptic efficacy transiently during a reverberatory activity.

Intrinsic cellular bistability

Finally, feedback excitation could originate from regenerative membrane dynamics of single neurons (Fig. 1d). Voltage- and Ca^{2+}-gated ion channels could in principle generate bistability between a resting state and an active state sustained by a 'plateau potential'[13,14]. Activation of relevant ion currents (such as a Ca^{2+}-activated cation current I_{Can}) could require neuromodulatory signals such as acetylcholine[15]. Even if single cells are not bistable, the behavior of a cortical network is always the result of the interplay between intrinsic cellular properties and synaptic mechanisms[16].

Experiments and biophysical modeling on the neural basis of persistent activity have so far been focused on the scenario of reverberation within a brain area. The present article will be confined to synaptic mechanisms in a local recurrent network.

Attractor paradigm for persistent activity

Since the 1970s, it has been proposed that delay activity patterns can be theoretically described by 'dynamical attractors'[17–21]. In this framework, a working memory network displays multiple attractor states, each of which stores the memory of a specific stimulus by a unique persistent firing pattern. The emergence of persistent attractor states requires that excitatory connections in a recurrent network are sufficiently strong. It is only recently, beginning with the work by Amit and colleagues, that attractor network models have been implemented with realistic models of cortical neurons and synapses[22–27].

Figure 2 illustrates the biophysics of an attractor network. In an object-working-memory model by Amit and Brunel[22], subpopulations of neurons are selective to different object stimuli. When the strength of excitatory connections between neurons within each subpopulation is increased beyond a critical threshold, persistent activity appears as an all-or-none phenomenon (Fig. 2a,b). Below the critical threshold, only the spontaneous state exists. Above the threshold, the spontaneous activity state is still dynamically stable to small perturbations, because at low firing rates excitation is effectively counteracted by feedback inhibition (Fig. 2c). However, if a stimulus generates a transient high activity in a neural subpopulation, now recurrent reverberation is sufficiently powerful to drive this group of cells to 'escape' from the spontaneous state. A higher firing activity leads to an even larger recurrent synaptic excitation, which becomes sufficient to sustain a persistent active state after the stimulus is

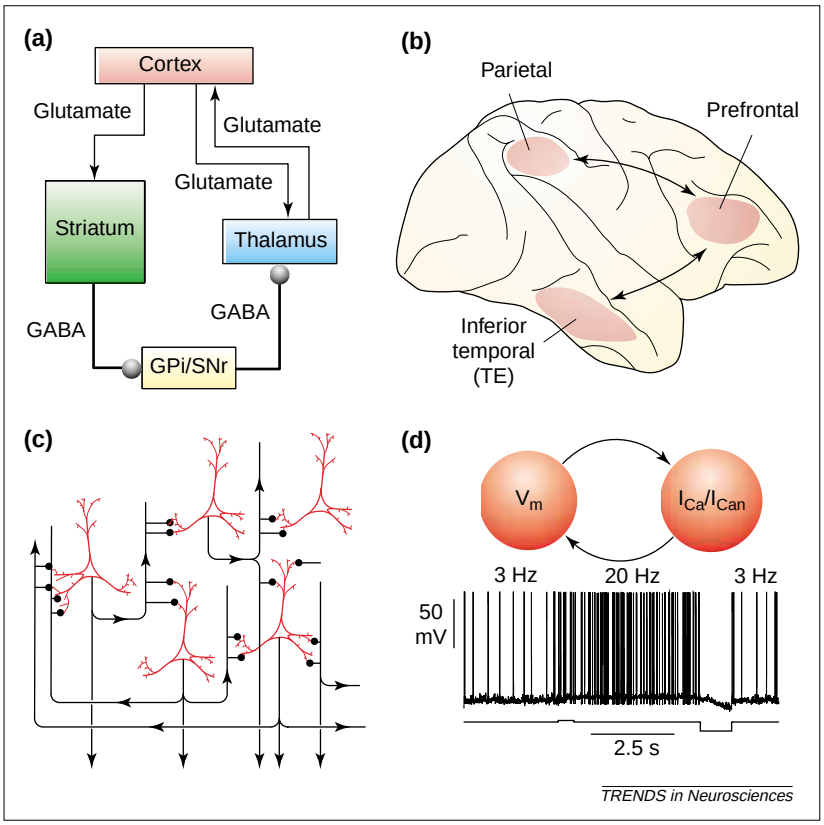

Fig. 1. Various scenarios for the anatomical substrate of excitatory reverberation in the cortex. (a) A closed thalamo–cortical loop and/or cortico–striato–thalamic–cortical circuit. In the latter case, cortical excitation of the caudate nucleus leads to an inhibition of the output from the cells of the basal ganglia and thus a disinhibition of thalamic neurons, which in turn send increased excitation back to the cortex. Abbreviations: GPi, internal segment of the globus pallidus; SNr, substantia nigra pars reticulata. (b) Reciprocal interactions between two cortical areas, prefrontal and posterior cortices (respectively prefrontal and inferotemporal cortices) for spatial (resp. object) visual memory. (c) Excitatory recurrent collaterals within a local circuit. (d) Intrinsic regenerative dynamics of single neurons. Positive feedback between membrane depolarization/spike discharges and active inward currents (voltage-gated I_{Ca}/Ca^{2+}-gated I_{Can}) can produce persistent activity that outlasts a transient input current pulse (J. Tegnér and X-J. Wang, unpublished).

withdrawn. The firing rate is eventually stabilized by negative feedback (Fig. 2c). As a result, a stable attractor of persistent activity with an elevated firing rate is realized, that coexists with the stable spontaneous state (Fig. 2c). Biophysical mechanisms that control the firing rates in a working memory network remain to be identified. Among possible contributors are outward ion currents in the cell, feedback inhibition, short-term synaptic depression, and saturation of the synaptic drive at high frequencies[24].

These simulations of biologically based models clearly show that quantitative differences in intrinsic connections between cortical networks (e.g. association areas versus primary sensory areas) can lead to qualitatively different behaviors (with or without persistent activity). A prediction from attractor models is that persistent activity depends on the strength of recurrent excitation in an abrupt manner, so that mnemonic activity could disappear suddenly when excitatory synaptic transmission is gradually reduced by pharmacological means.

Spatial working memory and bump attractors

Another major issue concerns the circuit architecture that gives rise to stimulus selectivity of persistent neural activity. Structured excitatory connectivity could arise from a columnar organization[7,28] or through Hebbian long-term plasticity[21]. For spatial working memory, such as in the delayed oculomotor experiment of Funahashi *et al.*[29] (Box 1), the network encodes a cue location and stores its memory in the form of a 'bump attractor', which is a spatially localized persistent activity pattern. Such 'bump attractors' naturally arise from a network connectivity where the strength of synaptic coupling between two pyramidal neurons decreases with the difference in their preferred cues[14,18,30–34]. Figure 3a illustrates a bump attractor network model of spiking neurons for spatial working memory[25]. In the model, the persistent bump state is sustained by recurrent synaptic excitation within a local group of pyramidal cells. The spatial tuning is sculptured by synaptic inhibition from interneurons, in agreement with physiological data[35]. Stable bump attractors typically require that lateral inhibition is spatially more widespread than excitation, with interneurons showing a broader tuning curve compared with pyramidal cells (Fig. 3b) and/or projecting widely to their targets.

In a spatial working memory model, a localized persistent activity pattern tends to drift randomly as a diffusion process during the delay period[14,25]. This is because there is a continuous family of 'bump attractors' (each encoding a potential location), and noise induces drifts between them. Interestingly, psychophysical studies have shown that the accuracy of the memory-guided saccade decreases with the delay duration in a way similar to diffusion[36,37]. It would be interesting to see whether random drifts do occur in PFC neurons during working memory, and whether they correlate with quantitative errors in the memory-guided saccades, and thus short-term memory decay in behaving animals.

Bump attractors are of general interest to diverse neural systems (Box 1). Interestingly, persistent activity in head-direction cells is crucially dependent on subcortical structures, where excitatory collaterals seem to be scarce[38]. The circuit mechanisms underlying persistent activity of head-direction cells remain to be elucidated in future experiments and model studies[33,39].

Parametric working memory and line attractors

Persistent activity with monotonic stimulus tuning (Box 1) requires a different type of network architecture, which is still not understood. At a phenomenological level, there is certain similarity between prefrontal delay activity with monotonic tuning[40] and 'neural integrators' that store the short-term memory of eye position. During eye fixation, oculomotor neurons display persistent activity, with the firing rate varying in proportion to the current eye

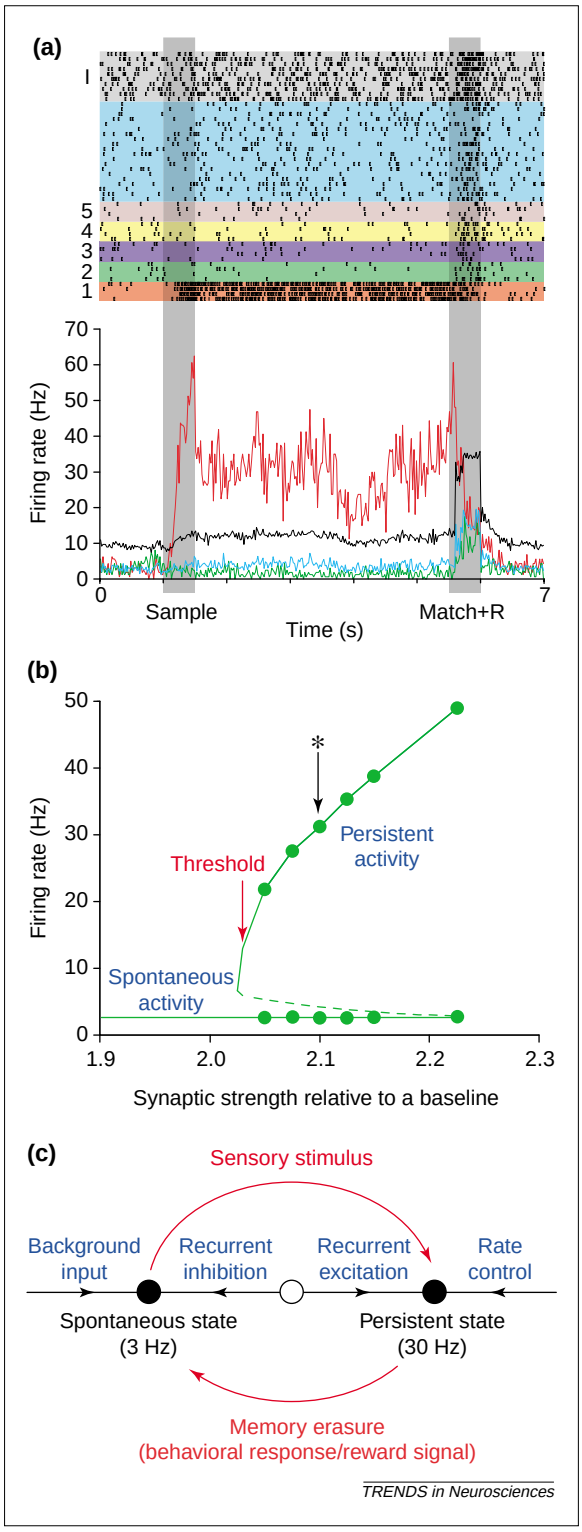

Fig. 2. Attractor paradigm for persistent activity. (a) Delayed match-to-sample simulation of an object working memory model[22]. Neural subpopulations (labeled 1 to 5) are selective to different stimuli. Average activity from each neural group is plotted below the rastergram. Inhibitory (I) cell population is shown in black. In the simulation, the shown stimulus triggers persistent activity in a pyramidal cell group (red) at about 30 Hz, but not in other pyramidal groups. Delay period activity is switched off by a transient excitatory input generating a brief surge of activity in all neurons. Adapted, with permission, from Ref. 65. (b) Dependence of network activity on the strength of recurrent synaptic connectivity. Solid lines: spontaneous state and persistent memory state; dashed line: unstable states. There is a critical threshold of synaptic strength, above which persistent activity appears abruptly as an all-or-none phenomenon. The example in (a) corresponds to the parameter indicated by the asterisk. (c) Illustration of the biophysics underlying an attractor dynamics. An attractor is a neural firing state that is stable to perturbations: when a small input perturbs the network to a lower or higher activity level, there is a 'restoring force' to bring the network back to the attractor state. In this case, the spontaneous state is stabilized from below by background inputs, and from above by feedback synaptic inhibition. The persistent activity state is stabilized from below by excitatory reverberation, and from above by various negative feedback 'rate control' mechanisms. A transient sensory stimulus can switch the network from the spontaneous state to the memory state, and a behavioral response or reward signal can turn the network off and erase the memory.

position[41]. Based on the idea of synaptic reverberation, Seung and colleagues[42] constructed a Hodgkin–Huxley-type conductance model of neural integrators (Fig. 4). Because persistent firing rates of individual cells are proportional to the eye position, when they are plotted against each other in a 'firing rate space', they always fall on a straight line. Each eye position is encoded by a unique set of persistent firing rates of neurons, corresponding to a particular point on this 'line attractor'[43]. Tank and colleagues investigated neural integrators experimentally in a goldfish preparation. Using intracellular recording from the behaving animal, it was shown that an increase in the persistent firing rate is correlated with a step change in the membrane potential as a result of network inputs, consistent with a synaptic reverberation mechanism[44].

Line attractors require fine tuning of network parameters. When parameters such as the synaptic strengths deviate from the precisely tuned values, persistency deteriorates, and neural activity decays or increases during fixation. It is unclear how such fine-tuning of parameters could be realized by biological mechanisms. The problem of robustness also exists for attractor networks of spatial working memory. A continuum of bump attractors assumes that the network is homogeneous, so that any activity pattern that is displaced spatially leads to another activity pattern of the same shape but peaked at a different location[14,25,30,33]. A major challenge is to understand how such a recurrent network could still function in the presence of network heterogeneity.

NMDA receptors and the stability of a memory network

An emergent notion from recent model studies is that working memory behavior is usually easier to realize and is more stable, if the network's recurrent synapses are primarily mediated by the slow and voltage-gated NMDA receptors (NMDARs). The reasons are several fold. First, the slow dynamics of I_{NMDA} could be crucial to stabilize a working memory network. Negative feedback processes, such as inhibition and synaptic depression, are typically slower than the AMPA receptor (AMPAR)-mediated synaptic transmission. For example, the AMPAR-mediated EPSCs (Refs 45,46) are about three times

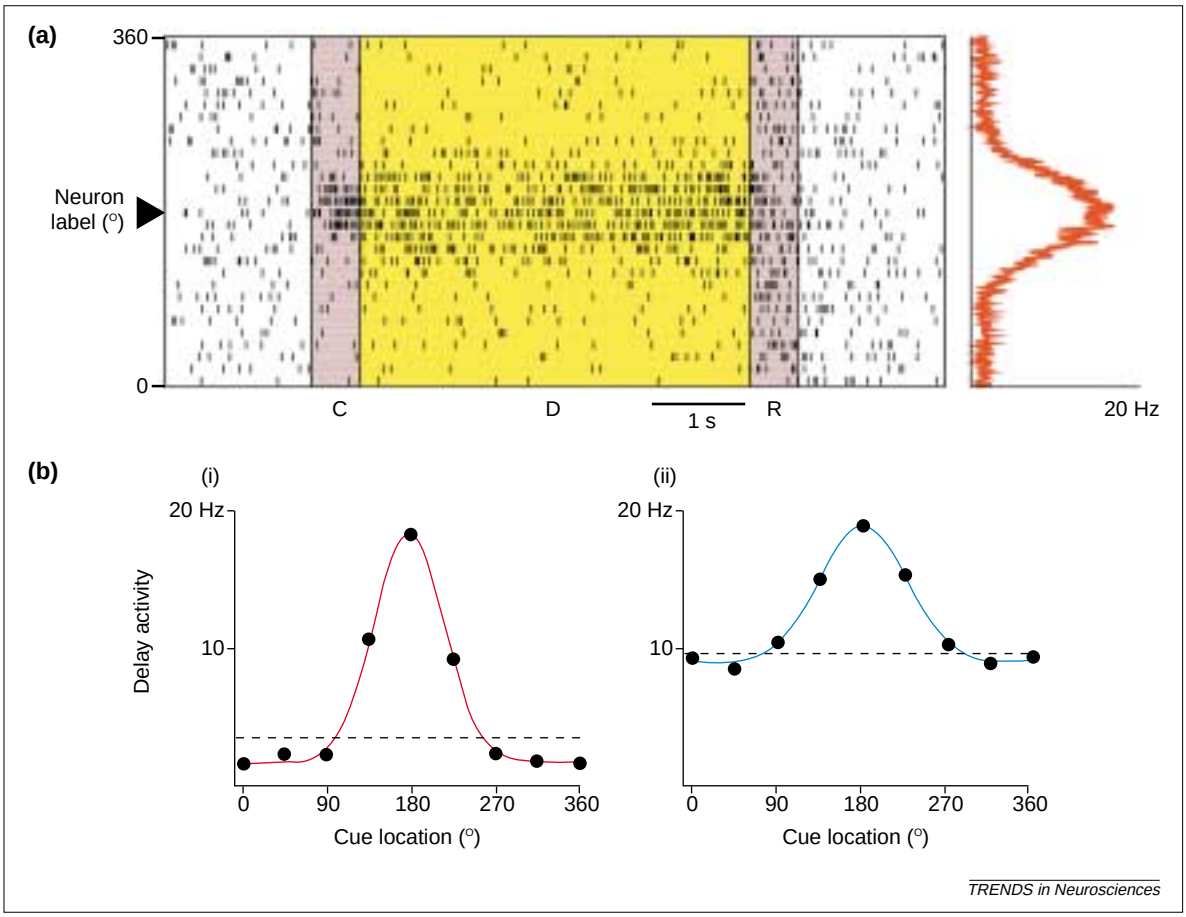

Fig. 3. Bump attractors for spatial working memory, as shown in a cortical network model of the Funahashi experiment[29]. (a) Spatio-temporal raster of the pyramidal cell population in a delayed saccade simulation. Abscissa: time, ordinate: neurons (labeled by their preferred cues). C: cue period, D: delay period, R: response period. The elevated and localized neural activity is triggered by the cue stimulus at 180°, and persists during the delay period. On the right is shown the spatial profile, where the average firing rate during the delay period is plotted versus the preferred cue of the neuron. In another trial, a different cue would elicit another bump attractor of the same spatial shape, peaked at a different location (see movie at http://www.wanglab.brandeis.edu/movie/spatial_wm.html). (b) Tuning curves of delay period activity for a pyramidal neuron (i) and an interneuron (ii), as a function of the cue location. Adapted, with permission, from Ref. 25.

faster than the $GABA_A$ receptor-mediated IPSCs (Refs 47,48). Synaptic depression has a decay time constant of 200–500 ms or longer[49,50]. A system with fast-positive and slow-negative feedback is dynamically unstable. Persistent activity is often disrupted in the middle of a delay period, and thereby the memory is lost[24,25]. Such instability does not occur, if excitation is sufficiently slow compared with negative feedbacks, that is, when recurrent synapses are primarily mediated by NMDA receptors[24,25]. In this case, persistent activity is stable even in the presence of network oscillations, typically in the gamma frequency range[24,25].

Second, the slow decay time of I_{NMDA} could be required for robust behavior of a memory network with continuous attractors. In the case of line attractor model for the oculomotor integrators, Seung[43] pointed out that if the strength of feedback excitation w is mistuned from the desired value w* by a small amount, $\Delta w = (w - w^*)/w^*$, then mnemonic activity can no longer persist indefinitely. Instead, it will decay with a time constant of $\tau \approx \tau_{syn}/\Delta w$, where τ_{syn} is the decay time constant of the recurrent synaptic current. For example, if memory storage requires $\tau \approx 5$ s, then with AMPA synapses ($\tau_{syn} < 5$ ms) w would need to be extremely precisely tuned ($\Delta w \leq 0.001 = 0.1\%$), whereas with NMDA synapses ($\tau_{syn} \approx 100$ ms) the tolerance is increased to $\Delta w = 0.02 = 2\%$.

Third, the voltage-dependent gating of NMDAR-mediated currents I_{NMDA} could contribute to trigger mnemonic activity selectively in a subgroup of neurons for a particular memory item[23]. The idea is that, during stimulus presentation, those cells activated by the stimulus are more depolarized than the rest of the network. Therefore the magnesium block of NMDA receptor channels are differentially removed in these cells, resulting in an enhanced synaptic reverberation and persistent activity within this group of neurons.

Fourth, the slow NMDAR unbinding to glutamate could lead to a saturation of the steady-state I_{NMDA} with repetitive stimulation at high frequencies. This saturation curtails the explosive positive feedback between neural firing and synaptic drive in a recurrent network. Therefore it can help to control the firing rate in a persistent activity state[24].

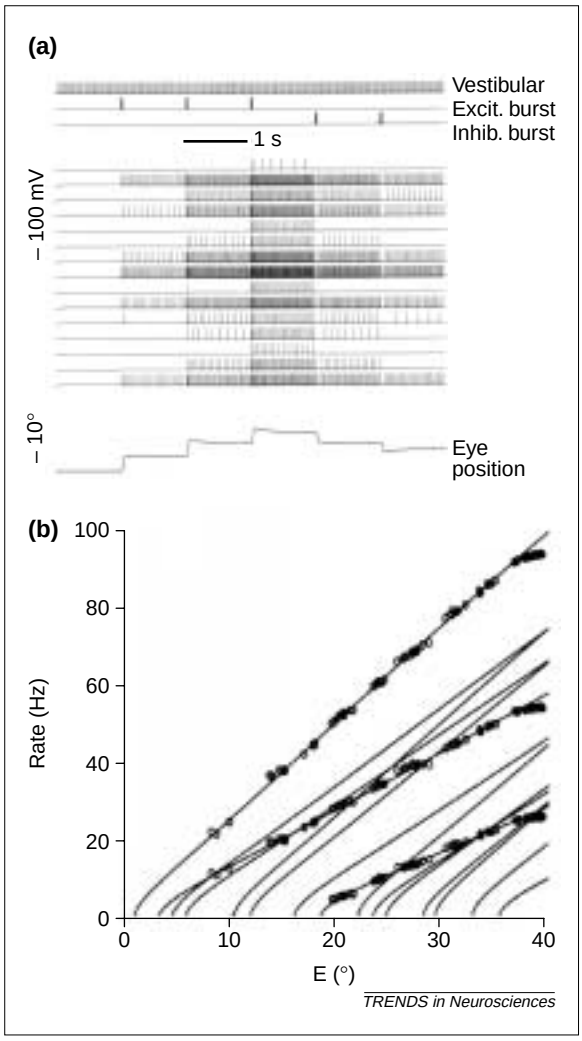

Fig. 4. Line attractor in a conductance-based memory model of oculomotor control. (a) Persistent changes of neural activity and eye position induced by transient inputs from excitatory and inhibitory burst neurons. Each transient burst causes a persistent change in activity of integrator neurons during interburst intervals. (b) Threshold linear encoding of eye position: firing rates of persistent activity of integrator neurons as function of eye position during fixation. Adapted, with permission, from Ref. 42.

Taken together, these results suggest that a reverberatory circuit would be much more stable, and working memory function more robust, if recurrent excitation is primarily mediated by NMDA receptors.

NMDA:AMPA ratio, dopamine modulation and schizophrenia

The role of NMDARs in working memory can be tested in both *in vivo* and *in vitro* experiments. In a working memory experiment, Shima and Tanji[51] found that delay period activity in the motor cortex was more effectively abolished by iontophoresis of NMDAR antagonists compared with AMPAR antagonists, indicating a predominant role of NMDA receptors in delay period activity. To confirm this result it would be worthwhile to carry out further studies using a similar approach, especially on PFC neurons.

The prediction that the NMDA:AMPA ratio at recurrent synapses should be high in a memory network can be tested using *in vitro* experiments. In models, the precise NMDA:AMPA ratio required for stability depends on the details and parameters. The general proposal is that a network with stronger recurrent connectivity should show a higher NMDA:AMPA ratio at its intrinsic synapses. Presently, available data indicate that the NMDA:AMPA ratio is area- and pathway-specific[52–56]. The PFC has the highest mRNA expression for the NMDAR subunits among cortical areas of the human brain[57], but the NMDA:AMPA ratio at intrinsic excitatory synapses of PFC remains to be quantitatively determined.

It should be noted that what is important for persistent activity is not unitary EPSCs at resting potential, but the ratio of the average NMDA and AMPA synaptic currents during repetitive neural discharges. This ratio could be boosted at high firing frequencies in several ways. First, a large depolarization reduces magnesium block of I_{NMDA}, so that the peak I_{NMDA}:I_{AMPA} ratio is increased[58]. Second, there is evidence that during sustained activity, an increase in the intracellular Na^+ concentration enhances NMDAR but not AMPAR currents[59]. Third, NMDARs have a 100-fold higher affinity for glutamate binding compared with AMPARs, thus extrasynaptic NMDARs are more likely to be activated by glutamate spillover during persistent activity[60], leading to a larger effective NMDA:AMPA ratio.

The NMDA:AMPA ratio can also be enhanced by neuromodulation. In particular, dopamine has been shown to differentially enhance NMDAR EPSCs via D1 and D5 receptors, while slightly reducing non-NMDA EPSCs in layer 5 PFC neurons[61]. This result, if confirmed, is especially interesting, given the importance of dopamine regulation of the prefrontal lobe. Experiments on behaving animals showed that working memory performance is modulated by the D1 but not the D2 class of dopamine receptors[62]. Dopamine modulation might be important to enhance the 'signal-to-noise' ratio and the robustness of memory storage in the presence of behaviorally irrelevant intervening stimuli[26,27,63–65].

The notion that reverberation is crucially dependent on NMDA receptors, if proven correct, has potential clinical implications. On the one hand there is strong evidence that working memory deficits in schizophrenia are associated with abnormal function of the frontal lobe[66,67]. On the other, working memory dysfunction can be induced by drugs that block NMDA receptors[68]. Postmortem studies showed significant alterations of NMDAR expression[69], but did not reveal abnormality in the AMPAR level[70], in the PFC of schizophrenics. An intriguing possibility is that working memory dysfunction results from an abnormally low NMDA:AMPA ratio at recurrent synapses which, according to our model studies, would give rise to dynamical instability of the mnemonic cortical circuit. Moreover, because dopamine D1 receptors selectively modulate the NMDAR current, the well-documented link between dysfunction of the dopamine system and working memory deficits in schizophrenia could be explained in part through the NMDAR pathway.

Acknowledgements
I thank P. Miller, J. Tegnér and A. Compte for help with the figures and for comments on the paper. I also thank D. Tank for stimulating discussions; N. Brunel, J. Lisman and G. Turrigiano for a critical reading of the manuscript. This work was supported by the NSF (IBN-9733006) and the Alfred P. Sloan Foundation.

Box 2. Outstanding questions

Recent theoretical models have raised several neurophysiological questions that can be investigated experimentally. Answers to these questions will help to elucidate the mechanisms of neural persistent activity.

- What is the minimum anatomical substrate of a reverberatory circuit capable of persistent neural activity?
- Is persistent activity primarily sustained by synaptic reverberation, or by bistable dynamics of single neurons?
- What is the NMDA:AMPA ratio at recurrent synapses of association cortices, especially in the prefrontal cortex?
- How does this ratio depend on the frequency of repetitive stimulation and on neuromodulation?
- What are the negative feedback mechanisms responsible for the rate control in a working memory network?
- Is delay period activity asynchronous between neurons, or does it display partial network synchrony and coherent oscillations?
- Is delay period activity more sensitive to NMDAR antagonists compared with AMPAR antagonists?
- Does persistent activity disappear in an abrupt fashion, with a graded block of NMDAR and AMPAR channels, as predicted by the attractor model?
- How significant are drifts of persistent activity during working memory? Are drifts random or systematic over trials?
- What are the biological mechanisms underlying the robustness of a memory network with a continuum of persistent activity patterns?

Concluding remarks

Excitatory reverberation is a leading candidate mechanism for mnemonic persistent activity. Recent work on the biophysics of reverberatory networks has led to two significant advances. First, attractor models were constructed with realistic neurons and synapses. Thus, it is possible for the first time to directly compare a model with cellular neurophysiology. These models have led to specific predictions and questions (Box 2) which, albeit difficult, seem to be within the reach of available experimental techniques. Second, the attractor paradigm has now been extended to memory networks that encode an analog stimulus, in the form of 'bump attractors' or 'line attractors'.

It was recognized that reverberatory networks tend to be dynamically unstable, but stability can be achieved if recurrent excitation is primarily mediated by the slow NMDA receptors. This hypothesis remains to be tested experimentally. The idea should also be challenged theoretically, by exploring alternative biological mechanisms (such as heterogeneity) that could stabilize a memory circuit without NMDA receptors. A related issue is the robustness of memory networks with a continuous family of attractors, which in existing models typically require fine-tuning of parameters with unrealistic precision. Ultimately, the problem of robustness needs to be solved by adaptive biological mechanisms[71].

References

1 Hebb, D.O. (1949) *Organization of Behavior*, Wiley
2 Lorente de Nó, R. (1933) Vestibulo-ocular reflex arc. *Arch. Neurol. Psychiatry* 30, 245–291
3 Fuster, J.M. and Alexander, G.E. (1973) Firing changes in cells of the nucleus medialis dorsalis associated with delayed response behavior. *Brain Res.* 61, 79–91
4 Hikosaka, O. *et al.* (1989) Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *J. Neurophysiol.* 61, 814–832
5 Hikosaka, O. and Wurtz, R.H. (1983) Visual and oculomotor functions of monkey substantia nigra pars reticulata. III. Memory-contingent visual and saccade responses. *J. Neurophysiol.* 49, 1268–1284
6 Miller, E.K. *et al.* (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J. Neurosci.* 16, 5154–5167
7 Goldman-Rakic, P.S. (1995) Cellular basis of working memory. *Neuron* 14, 477–485
8 Levitt, B. *et al.* (1993) Topography of pyramidal neuron intrinsic connections in macaque monkey prefrontal cortex (areas 9 and 46). *J. Comp. Neurol.* 338, 360–376
9 Kritzer, M.F. and Goldman-Rakic, P.S. (1995) Intrinsic circuit organization of the major layers and sublayers of the dorsolateral prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.* 359, 131–143
10 González-Burgos, G. *et al.* (2000) Horizontal synaptic connections in monkey prefrontal cortex: an *in vitro* electrophysiological study. *Cereb. Cortex* 10, 82–92
11 Hirsch, J.C. and Crépel, F. (1990) Use-dependent changes in synaptic efficacy in rat prefrontal neurons *in vitro*. *J. Physiol.* 427, 31–49
12 Hempel, C.M. *et al.* (2000) Multiple forms of short-term plasticity at excitatory synapses in rat medial prefrontal cortex. *J. Neurophysiol.* 83, 3031–3041
13 Marder, E. *et al.* (1996) Memory from the dynamics of intrinsic membrane currents. *Proc. Natl. Acad. Sci. U. S. A.* 93, 13482–13486
14 Camperi, M. and Wang, X-J. (1998) A model of visuospatial short-term memory in prefrontal cortex: recurrent network and cellular bistability. *J. Comput. Neurosci.* 5, 383–405
15 Haj-Dahmane, S.and Andrade, R. (1998) Ionic mechanism of the slow afterdepolarization induced by muscarinic receptor activation in rat prefrontal cortex. *J. Neurophysiol.* 80, 1197–1210
16 Sanchez-Vives, M. and McCormick, D.A. (2000) Cellular and network mechanisms of rhythmic recurrent activity in neocortex. *Nat. Neurosci.* 3, 1027–1034
17 Wilson, H. and Cowan, J. (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13, 55-80
18 Amari, S. (1977) Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* 27, 77–87
19 Hopfield, J.J. (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U. S. A.* 79, 2554–2558
20 Zipser, D. *et al.* (1993) A spiking network model of short-term active memory. *J. Neurosci.* 13, 3406–3420
21 Amit, D.J. (1995) The Hebbian paradigm reintegrated: local reverberations as internal representations. *Behav. Brain Sci.* 18, 617–626
22 Amit, D.J. and Brunel, N. (1997) Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb. Cortex* 7, 237–252
23 Lisman, J.E. *et al.* (1998) A role for NMDA-receptor channels in working memory. *Nat. Neurosci.* 1, 273–275
24 Wang, X-J. (1999) Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. *J. Neurosci.* 19, 9587–9603
25 Compte, A. *et al.* (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb. Cortex* 10, 910–923
26 Durstewitz, D. *et al.* (2000) Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *J. Neurophysiol.* 83, 1733–1750
27 Durstewitz, D. (2000) Neurocomputational models of working memory. *Nat. Neurosci.* 3, 1184–1191
28 Gilbert, C.D. (1993) Circuitry, architecture, and functional dynamics of visual cortex. *Cereb. Cortex* 3, 373–386
29 Funahashi, S. *et al.* (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61, 331–349
30 Ben-Yishai, R. *et al.* (1995) Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 92, 3844–3848
31 Somers, D.C. *et al.* (1995) An emergent model of orientation selectivity in cat visual cortical simple cells. *J. Neurosci.* 15, 5448–5465
32 Douglas, R.J. *et al.* (1995) Recurrent excitation in neocortical circuits. *Science* 269, 981–985
33 Zhang, K. (1996) Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensembles: a theory. *J. Neurosci.* 16, 2112–2126

34 Hahnloser, R.H.R. *et al.* (2000) Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature* 405, 947–951
35 Rao, S.G. *et al.* (1999) Isodirectional tuning of adjacent interneurons and pyramidal cells during working memory: evidence for microcolumnar organization in PFC. *J. Neurophysiol.* 81, 1903–1916
36 White, J.M. *et al.* (1994) Saccades to remembered target locations: and analysis of systematic and variable errors. *Vis. Res.* 34, 79–92
37 Ploner, C.J. *et al.* (1998) Temporal limits of spatial working memory in humans. *Eur. J. Neurosci.* 10, 794–797
38 Sharp, P.E. *et al.* (2001) The anatomical and computational basis of the rat head direction cell signal. *Trends Neurosci.* 24, 289–294
39 Rubin, J. *et al.* (2001) Localized bumps of activity sustained by inhibition in a two-layer thalamic network. *J. Comput. Neurosci.* (in press)
40 Romo, R. *et al.* (1999) Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* 399, 470–474
41 Robinson, D.A. (1989) Integrating with neurons. *Annu. Rev. Neurosci.* 12, 33–45
42 Seung, H.S. *et al.* (2000) Stability of the memory of eye position in a recurrent network of conductance-based model neurons. *Neuron* 26, 259–271
43 Seung, H.S. (1996) How the brain keeps the eyes still. *Proc. Natl. Acad. Sci. U. S. A.* 93, 13339–13344
44 Aksay, E. *et al.* (2001) *In vivo* intracellular recording and perturbation of persistent activity in a neural integrator. *Nat. Neurosci.* 4, 184-193
45 Hestrin, S. *et al.* (1990) Mechanisms generating the time course of dual component excitatory synaptic currents recorded in hippocampal slices. *Neuron* 5, 247–253
46 Hestrin, S. (1993) Different glutamate receptor channels mediate fast excitatory synaptic currents in inhibitory and excitatory cortical neurons. *Neuron* 11, 1083–1091
47 Salin, P.A. and Prince, D.A. (1996) Spontaneous $GABA_A$ receptor mediated inhibitory currents in adult rat somatosensory cortex. *J. Neurophysiol.* 75, 1573–1588
48 Xiang, Z. *et al.* (1998) $GABA_A$ receptor mediated currents in interneurons and pyramidal cells of rat visual cortex. *J. Physiol.* 506, 715–730
49 Abbott, L.F. *et al.* (1997) Synaptic depression and cortical gain control. *Science* 275, 220–224
50 Tsodyks, M.V. and Markram, H. (1997) The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. U. S. A.* 94, 719–723
51 Shima, K. and Tanji, J. (1998) Involvement of NMDA and non-NMDA receptors in the neuronal responses of the primary motor cortex to input from the supplementary motor area and somatosensory cortex: studies of task-performing monkeys. *Jpn. J. Physiol.* 48, 275–290
52 Daw, N.W. *et al.* (1993) The role of NMDA receptors in information processing. *Annu. Rev. Neurosci.* 16, 207–222
53 Rivadulla, C. *et al.* (2001) Specific roles of NMDA and AMPA receptors in direction-selective and spatial phase-selective responses in visual cortex. *J. Neurosci.* 21, 1710–1719
54 Markram, H. *et al.* (1997) Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *J. Physiol.* 500, 409–440
55 Feldmeyer, D. *et al.* (1999) Reliable synaptic connections between pairs of excitatory layer 4 neurones within a single 'barrel' of developing rat somatosensory cortex. *J. Physiol.* 521, 169–190
56 Fleidervish, I.A. *et al.* (1998) Functionally distinct NMDA receptors mediate horizontal connectivity within layer 4 of mouse barrel cortex. *Neuron* 21, 1055–1065
57 Scherzer, C.R. *et al.* (1998) Expression of N-methyl-D-aspartate receptor subunit mRNAs in the human brain: hippocampus and cortex. *J. Comp. Neurol.* 390, 75–90
58 Collingridge, G.L. *et al.* (1988) Frequency-dependent N-methyl-D-aspartate receptor-mediated synaptic transmission in rat hippocampus. *J. Physiol.* 399, 301–312
59 Yu, X-M. and Salter, M.W. (1998) Gain control of NMDA-receptor currents by intracellular sodium. *Nature* 396, 469–474
60 Kullmann, D.M. and Asztely, F. (1998) Extrasynaptic glutamate spillover in the hippocampus: evidence and implications. *Trends Neurosci.* 21, 8–14
61 Seamans, J.K. *et al.* (2001) Dopamine D1/D5 receptor modulation of excitatory synaptic inputs to layer V prefrontal cortex neurons. *Proc. Natl. Acad. Sci. U. S. A.* 98, 301–306
62 Arnsten, A.F.T. (1998) Catecholamine modulation of prefrontal cortical cognitive function. *Trends Cognit. Sci.* 2, 436–447
63 Servan-Schreiber, D. *et al.* (1990) A network model of catecholamine effects: gain, signal-to-noise ratio, and behavior. *Science* 249, 892–895
64 Yang, C.R. and Seamans, J.K. (1996) Dopamine D1 receptor actions in layers V-VI rat prefrontal cortex neurons *in vitro*: modulation of dendritic-somatic signal integration. *J. Neurosci.* 16, 1922–1935
65 Brunel, N. and Wang, X-J. (2001) Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J. Comput. Neurosci.* (in press)
66 Goldman-Rakic, P.S. (1994) Working memory dysfunction in schizophrenia. *J. Neuropsychiatry Clin. Neurosci.* 6, 348–357
67 Weinberger, D.R. and Berman, K.F. (1996) Prefrontal function in schizophrenia: confounds and controversies. *Philos. Trans. R. Soc. London Ser. B* 351, 1495–1503
68 Krystal, J.H. *et al.* (1994) Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Arch. Gen. Psychiatry* 51, 199–214
69 Akbarian, S. *et al.* (1996) Selective alterations in gene expression for NMDA receptor subunits in prefrontal cortex of schizophrenics. *J. Neurosci.* 16, 19–30
70 Healy, D.J. *et al.* (1998) AMPA receptor binding and subunit mRNA expression in prefrontal cortex and striatum of elderly schizophrenics. *Neuropsychopharmacology* 19, 278–286
71 Turrigiano, G. (1999) Homeostatic plasticity in neuronal networks: the more things change, the more they stay the same. *Trends Neurosci.* 22, 221–227