Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Neurobiology

Working models of working memory

Omri Barak[1] and Misha Tsodyks[2]

Working memory is a system that maintains and manipulates information for several seconds during the planning and execution of many cognitive tasks. Traditionally, it was believed that the neuronal underpinning of working memory is stationary persistent firing of selective neuronal populations. Recent advances introduced new ideas regarding possible mechanisms of working memory, such as short-term synaptic facilitation, precise tuning of recurrent excitation and inhibition, and intrinsic network dynamics. These ideas are motivated by computational considerations and careful analysis of experimental data. Taken together, they may indicate the plethora of different processes underlying working memory in the brain.

Addresses
[1] Faculty of Medicine, Technion – Israel Institute of Technology, 1 Efron St., Haifa 31096, Israel
[2] Department of Neurobiology, Weizmann Institute of Science, Herzl St., Rehovot 76100, Israel

Corresponding authors: Tsodyks, Misha (misha@weizmann.ac.il)

Current Opinion in Neurobiology 2014, **25**:20–24

This review comes from a themed issue on **Theoretical and computational neuroscience**

Edited by **Adrienne Fairhall** and **Haim Sompolinsky**

For a complete overview see the Issue and the Editorial

Available online 4th December 2013

0959-4388/$ – see front matter, © 2013 Elsevier Ltd. All rights reserved.

http://dx.doi.org/10.1016/j.conb.2013.10.008

Introduction

Working memory is a crucial component in the execution of many cognitive tasks that require holding and manipulating information for short periods of time (see e.g., [1]). In this review, we will focus on the holding of information for a time period of several seconds. From the mechanistic point of view, working memory differs from long term memory in that no structural changes are hypothesized to be involved – it is a transient phenomenon. Models of working memory are presented with two types of challenges: data-driven and computational-driven (Figure 1, middle). The data-driven challenges arise from the analysis of behavior and neuronal recordings in animals performing working memory tasks. Animals were shown to be able to maintain several items simultaneously in memory, remember their order, and manipulate them (see e.g., [2] for a recent account). Among the common physiological observations, it was reported that neurons typically exhibit irregular firing activity at a low rate, the activity related to storing a fixed item is not stationary, and there is a large heterogeneity in the firing profiles of different neurons [3,4,5•,6]. From the computational side, the network activity representing a memorized item should exhibit a sufficient degree of stability to ensure memory retainment. This requirement is especially challenging for storing continuous variables, such as orientation or spatial position of a visual cue, because of an inevitable drift along the variable's representation. Furthermore, integrating the various data-driven challenges in a self-consistent manner is often a non-trivial computational problem.

To cope with these challenges, various models incorporate different amounts of biophysical detail – highlighting the contribution of model elements to the various challenges (Figure 1, right). In the current review, we will briefly present the classic models of working memory, and proceed to highlight several recent attempts at addressing the different challenges. The focus of this review is on network mechanisms of working memory. For alternative mechanisms based on single cells persistent activity see [7,8].

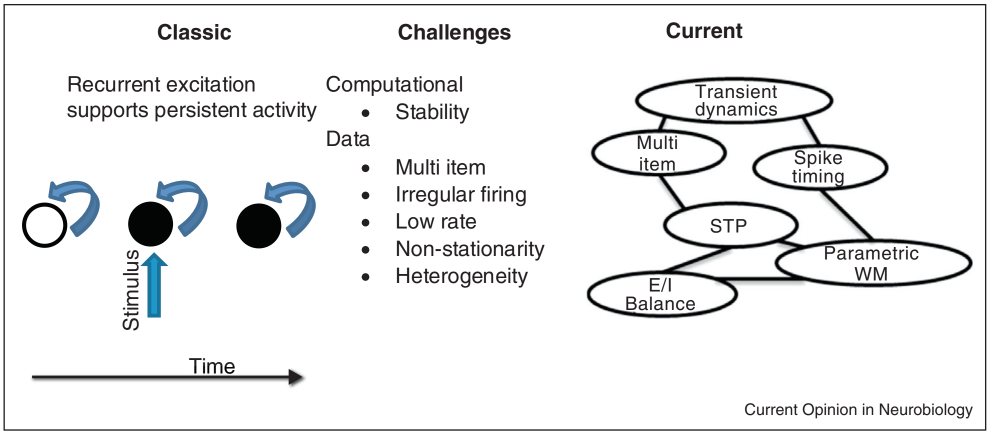
The classic models

The classic view is that items are embedded in long term memory via specific synaptic modifications, and presentation of these items leads to activation of stable activity patterns in the network ('attractors') [9,10]. Thus the information on which item is currently in working memory is stored in the persistent firing of these attractors. Supporting this theory, neurons exhibiting persistent activity after the removal of a stimulus were observed in the inferior temporal and prefrontal cortices of monkeys [11,12] (Figure 1, left).

Multi-item memory

The majority of foundational work on models of working memory were motivated by delayed memory experiments where only one item had to be retained in memory [11,13,14]. Working memory, of course, is not limited to a single item [15], and accordingly electrophysiological recordings were done on monkeys performing tasks requiring the maintenance of several items in working memory [12]. The mechanistic challenge of maintaining more than one item arises due to interference between the activations of the different items. Amit *et al.* [16] proposed that such interference is reduced when items are encoded by sparse patterns – every item is represented by a small fraction of the neuronal population. This approach was extended by [17] to account for the storage of both learned and novel items.

Figure 1



Current Opinion in Neurobiology

Concepts in working memory models. **Left**: the classic account of working memory is that a strong recurrent excitation enables the network to sustain persistent activity after removal of a transient stimulus. **Middle**: models of working memory face challenges on computational and data-driven fronts. **Right**: current models of working memory introduce various biophysical considerations to cope with the challenges, while attempting to remain simple enough to understand.

Both inhibition [18] and excitation [19] were shown to influence the capacity of multi item working memory. In both of these works, the authors showed how a network storing a continuous value can be dynamically partitioned to maintain several localized bumps of activity, each representing one memorized value of this variable. The balance of excitation and inhibition determines both the number of items that can be held, and their mode of failure (fade out or merge). Continuous attractors required tuned connectivity, but this tuning can be relaxed by incorporating more biological detail into the model. Specifically, Rolls *et al.* [20] showed that synaptic facilitation (detailed in the next section) increases the capacity of working memory. Moving beyond capacity considerations, Dempere-Marco *et al.* [21] showed that salient items (those presented with higher intensity) can be guaranteed a higher chance of maintenance at the expense of less salient items.

A conceptually different method of holding multiple items in working memory is to multiplex them in time rather than in space [22,23]. In this approach, the activated items are all oscillating at some frequency in different phases, and capacity is determined by the ratio of this frequency to the temporal width of each activation. In principle, this method can store information about the order of the items as well as their identity.

Effects of NMDA receptors on persistent activity

Early network models of persistent activity used a highly simplified description of neuronal and synaptic dynamics, resulting in certain difficulties in reproducing a realistic range of firing rates during working memory [24]. As first pointed out by [25], this issue can be resolved by considering networks with slow recurrent excitatory currents that are reminiscent of NMDA currents. Indeed, it was recently reported that blocking NMDA, but not AMPA, receptors during a working memory task abolishes persistent activity in prefrontal neurons [26•]. Moreover, the relative efficacy of NMDA currents is sensitive to Dopamine modulation, thus providing a possible mechanism of regulating working memory [27]. In particular, strengthening NMDA currents during the delay period of memory tasks can enhance the robustness of persistent activity to intervening stimuli. More intriguingly, NMDA currents can also affect the temporal aspects of neuronal activity, for example, by enhancing the burstiness of firing, thus potentially mediating the more complex forms of persistent activity compared to simple steady-state asynchronous states [27].

Short term synaptic plasticity

The model of [20] mentioned above relied on the slow timescale of synaptic facilitation to stabilize the persistent firing state (see also [28]). Synaptic facilitation, and other forms of short term synaptic plasticity, enable synapses to temporarily modify their efficacy in response to stimuli [29,30]. Recently, Itskov *et al.* [31] examined the effect of synaptic facilitation on a network storing a continuous variable via the 'line attractor' mechanism, that is, a continuous one-dimensional set of marginally stable activity states, and showed that facilitation reduces the inherent drift of the system, thus prolonging memory lifetime significantly.

A more dominant role for synaptic facilitation was suggested by Mongillo *et al.* [23], who proposed that a stimulus-selective pattern of synaptic facilitation can itself maintain working memory in the absence of