Beyond working memory: the role of persistent activity in decision making

Clayton E. Curtis¹ and Daeyeol Lee²

¹ Department of Psychology and Center for Neural Science, New York University, 6 Washington Place, New York, NY 10003, USA
² Department of Neurobiology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06610, USA

Since its first discovery in the prefrontal cortex, persistent activity during the interval between a transient sensory stimulus and a subsequent behavioral response has been identified in many cortical and subcortical areas. Such persistent activity is thought to reflect the maintenance of working memory representations that bridge past events with future contingent plans. Indeed, the term persistent activity is sometimes used interchangeably with working memory. In this review, we argue that persistent activity observed broadly across many cortical and subcortical areas reflects not only working memory maintenance, but also a variety of other cognitive processes, including perceptual and reward-based decision making.

Properties of persistent activity

For decades now we have known that neuronal activity persists in the prefrontal cortex when the subject is in an active state of remembering a recent event [1–3], a process termed working memory [4,5]. Working memory allows animals to use information that is not currently available in the environment but is crucial for adaptive behavior. In other words, an internal representation of relevant information must be created and maintained until it can be used later to guide behavior.

Typical working memory experiments involve the brief presentation of a stimulus to be remembered, which is then followed by a short delay, or retention interval. Performance is probed by having the subject make a response that is contingent on successful maintenance of the memoranda. Neural activity in the prefrontal cortex persists during the delay period without being contaminated by sensory and motor processes, so it is thought to reflect the maintenance of a working memory representation [1–3].

Several key features of persistent activity suggest that it is a neural mechanism critical for working memory maintenance and presumably other high-level cognitive processes such as decision making. First, neural activity, measured by a variety of methodologies, such as single-neuron recording and neuroimaging, persists during the time epoch when a representation is thought to be maintained in an active state (e.g. when a stimulus needs to be remembered) [1–3,5–7]. Second, sustained neural activity dissipates when the representation is no longer needed (e.g. when a memory-guided response has been generated) [3,8,9]. Third, when activity does not persist throughout a retention interval, memory performance is compromised [3,10]. Fourth, persistent activity scales with working memory load. For example, as the number of items that need to be maintained increases, the level of delay period activity also increases until it plateaus near the limits of the individual’s short-term memory capacity [11–13]. Fifth, persistent activity is selective, such that specific information about the contents of working memory can be decoded from the population activity [14–16]. For example, persistent activity is selective for the location of the memoranda during spatial working memory tasks [3,9] (Figure 1). Finally, although much of the relevant research has focused on the prefrontal cortex, persistent activity has been reported in many other cortical and subcortical areas [17,18], suggesting that its generation and maintenance might depend on a broadly distributed circuit mechanism (Box 1). In this review, we argue that the persistence of neural activity in the absence of external stimulation is the general mechanism by which internal representations are maintained in an active state. Moreover, these representations go beyond traditional notions of working memory and include variables useful for decision making [19,20].

Persistent activity and working memory

The development of neurobiological theories of working memory critically depends on the measurement of neural activity during the delay period. As a result, persistent activity has essentially become a proxy for working memory and the two are often used interchangeably. This is not surprising because persistent neural activity during delay periods seems to carry useful information about the features of the memoranda. Persistent activity arises while the subject maintains information about the spatial location [3,9,10,17,18,21,22], object identity [23–25], word [26], sound [27,28] or haptic sensation [29] of the memorandum. Theoretically, the short-term maintenance of such features is necessary for a variety of high-level cognitive functions because it facilitates temporal integration of past sensory events with future actions [5,30]. Importantly, working memory paradigms have been the test bed for studies addressing the neural mechanisms of persistent activity (Box 1).

Despite its logical fallibility, the reverse inference that the presence of persistent activity indicates the maintenance of a past stimulus continues to tempt researchers. The problem, however, is that persistent activity has been linked to several other cognitive processes in addition to...
Furthermore, neural activity in the prefrontal cortex still persists and encodes the properties of visual stimuli even in the absence of working memory demands \[33\]. Persistent activity can also reflect not the maintenance of past events, but the anticipation of future ones. During an interval prior to or in preparation for an action, persistent activity is thought to represent the metrics or goal of a forthcoming planned action \[34,35\]. Similarly, persistent activity has been linked to stimuli that are anticipated but not yet encountered \[25,36\]. Therefore, persistent activity underlies the representation of retrospective, current and prospective information. Moreover, persistent activity can be observed in the prefrontal cortex during the active representation of more abstract information, such as rules, associations, categories and strategies \[37–39\]. Therefore, in addition to traditional notions of working memory representations, persistent activity can reflect a wider variety of cognitive processes, some of which are used for making decisions and to which we now turn our attention.

**Persistent activity reflecting accumulation of sensory evidence**

As described above, activity of individual neurons and blood-oxygen-level-dependent (BOLD) signals in multiple brain areas can be sustained during the delay period of a working memory task, and therefore might provide the substrate for working memory and storage of other behaviorally relevant information. It is often thought that such persistent activity might arise from a network of recurrently connected neurons \[19\]. However, persistent activity related to a past sensory event might also arise indirectly from the short-term changes in synaptic efficacy induced by transient neural activity \[40,41\] (Box 1). Regardless of its
Precisely how persistent activity is generated and controlled in the mammalian brain is not known, but several possible, and not mutually exclusive, mechanisms have been proposed. The ubiquity of persistent activity, it remains an important area of future research to determine the exact relationship between each of these mechanisms and a specific form of persistent activity observed during a particular behavioral task.

- Persistent activity from reverberations in a recurrent network. Persistent activity might be initiated and sustained by reciprocal positive feedback within a population of neurons. Persistent activity observed in a given cortical area (e.g., prefrontal cortex) might arise locally within a population of neighboring neurons [19], within a population of neurons distributed across multiple cortical regions [21] or from a network of areas including both cortical and subcortical systems, such as the corticobasal ganglia-thalamus-cortical loops [17,18]. At each of these network levels, activation of separate networks could specify different stimuli, different actions or stimulus–response mappings, different behavioral strategies and reward expectancies.

- Persistent activity relying on intracellular signals. Individual cortical neurons might display persistent activity due to specific membrane currents or cumulative changes in the concentration of intracellular calcium. For example, neurons in the entorhinal cortex can display a step-like depolarization following brief stimulation. Increased spiking can lead to calcium influx through voltage-gated calcium channels at dendrites, causing subsequent depolarization and thereby closing a positive feedback loop. This form of persistent activity seems to rely on nonspecific calcium-sensitive cationic currents [67].

- Persistent activity and short-term synaptic plasticity. When a particular synapse is repeatedly stimulated, the amount of neurotransmitter released in response to a given presynaptic action potential can increase in some synapses and decrease in others [68]. The pattern of activity sustained in a network is a function of synaptic weights among its constituent neurons, so short-term synaptic changes induced by a particular pattern of activity might make it possible for the network to resume specific patterns of persistent activity even after a quiescent period [40,41].

In most neurophysiological experiments conducted in behaving animals, neural activity is recorded while the animal performs a particular behavioral task repeatedly in a large number of trials separated by short inter-trial intervals. As a result, properties of persistent activity have been largely studied in relation to behaviorally relevant events within a single trial. For example, persistent activity during the delay period of a working memory task has been viewed as the substrate of working memory for the items presented at the beginning of a trial [3,8–10]. Similarly, gradual changes in neural activity observed during a perceptual decision-making task have been linked to the process of evidence accumulation within a trial [19,42,45,46]. However, the network of neurons capable of temporal integration might be exploited in tasks in which the animal is required to integrate information about the events in multiple trials. Recently, this possibility has been tested during several behavioral tasks in which animals need to change their decision-making strategies based on their previous experience [20,47–51].

The process of decision making, namely, selecting a particular action among a number of alternative actions, is trivial if the outcomes of all the alternative actions are fully known and can be evaluated along a single dimension. Often, however, the outcomes of actions vary along multiple dimensions, such as the magnitude, probability and immediacy of reward expected from each action, and decision makers need to resolve potential conflicts to maximize the magnitude and probability of their reward and to obtain it as soon as possible. Nevertheless, as long as all the information about the expected outcomes is available, such conflicts can be resolved without any additional inputs from the environment. However, it is a much more challenging task to discover an optimal decision-making strategy in an unfamiliar and stochastic environment for which the probabilities of different outcomes from alternative actions are not known. In such cases, the desirability of each action needs to be estimated through experience. The framework of reinforcement learning theory [52] parsimoniously accounts for dynamic changes in the behaviors of humans and animals as they experience unpredictable outcomes [53–55]. In addition, the same framework has played an important role in elucidating cortical and subcortical mechanisms involved in evaluating the action outcomes and exploiting this information to optimize decision-making strategies [47–49,54–58].

In reinforcement learning theory, the entire history of previous actions and their outcomes is abstracted and represented by a set of quantities, referred to as value functions. The value function (Q in Figure 2c) is an estimate of future rewards expected from a particular state and action, and is continuously adjusted according to the animal’s experience. Accordingly, decision makers can make choices at any moment simply by selecting an action with the maximum value function. Value functions are increased (decreased) when the outcome from a given action is better (worse) than expected. The discrepancy between the expected and actual outcomes is referred to as the reward prediction error and might be encoded by the activity of midbrain dopamine neurons [56]. In addition, neurons with
a change in activity according to the value functions of specific actions have been identified in a number of brain areas, including the prefrontal cortex, posterior parietal cortex and basal ganglia \cite{47,49,54,59,61}.

As in working-memory and perceptual decision-making tasks, persistent activity has been observed in multiple brain areas during reward-based decision-making tasks. Activity related to value functions is often manifest as persistent activity \cite{43,47,49,51,61} and therefore can be used to guide the selection of subsequent actions. Moreover, many of the key computations in reinforcement learning correspond to temporal integration and might be implemented in the form of persistent activity. For example, the outcomes from a particular action are often revealed to the animal after substantial delays. In some cases, the animal might execute multiple actions before receiving a reward or penalty. This makes it difficult to determine how the value functions should be adjusted for multiple actions, creating a problem of temporal credit assignment \cite{52,62}. In the so-called temporal-difference (TD) learning algorithms, this problem is addressed by comparing the value functions of two successive states. Therefore, a particular action can produce a positive reward prediction error and increase its value function, even without receiving any reward immediately, if it allows the decision maker to move to a state with a higher value function than that for the current state. Although TD learning provides a solution to the problem of temporal credit assignment, it is important to note that these computations might be implemented in the form of persistent activity.

Figure 2. Persistent activity resulting from temporal integration can subserve multiple types of cognitive processes, including working memory, accumulation of noisy sensory evidence and creation of eligibility trace and other computations in reinforcement learning. (a,b) Schematic illustration of (a) the time course of inputs leading to persistent activity during a working memory task and (b) buildup activity related to the accumulation of noisy sensory inputs. (c) Hypothetical signals related to the animal’s actions (green and blue vertical bars for leftward and rightward actions, respectively) can be temporally integrated to generate eligibility traces (upper and middle left panels), whereas temporal integration of reward (red disks) can lead to signals related to average reward rate (bottom left panel). The value function for each action is incremented by the reward prediction error, namely, the discrepancy between the reward earned and the reward predicted from the current value function, weighted by its eligibility trace. As a result, reward delivered in a given trial increases the value functions according to the recency of each action (upper and middle right panels), even when it was not chosen in the last trial. Dark red disks and colored vertical bars indicate the trials in which the leftward or rightward action led to immediate reward. The difference in value functions for the two alternative actions is also shown (bottom right panel; dark red circles indicate the trials in which reward was obtained after choosing the leftward (top) and rightward (bottom) actions, respectively).
credit assignment, the time course of reward prediction errors encoded by dopamine neurons does not match the pattern predicted by simple TD learning algorithms, suggesting that additional mechanisms for temporal credit assignment exist [63]. Indeed, some reinforcement learning algorithms rely on the memory of previous actions and states visited by the decision makers. These memory signals are referred to as eligibility traces [52] (Figure 2c).
Persistent activity in the prefrontal cortex [20,47,49], posterior parietal cortex [48] and basal ganglia [51,64,65] can be modulated by previous actions of the animal across multiple trials and might correspond to the eligibility traces (Figure 3a,c).

Persistent activity related to the outcomes of previous choices has also been observed in the same regions of the cortex and basal ganglia that showed persistent changes in neural activity according to the animal’s previous choices [20,43,47–50,64,65] (Figures 2c and 3b,c). These results suggest that the signals related to previous actions and their outcomes might influence the ways in which sensory stimuli are encoded and transformed to behavioral responses in multiple regions of the brain. Similar to the activity related to previous choices, persistent activity related to previous outcomes often continues during inter-trial intervals. In addition, in contrast to persistent activity observed during the delay period of a working memory task, persistent activity related to the animal’s previous choices and outcomes tends to decay gradually [43,48,51]. Reward-related persistent activity might provide the substrate for temporal integration of rewards in multiple trials, and could therefore be used to encode the average reward rate [43]. Information about the average reward rate plays an important role in some reinforcement learning algorithms. For example, it might be used to calculate the reward prediction error and influence how the value functions would be updated, as in so-called R-learning [52]. Indeed, the primate anterior cingulate cortex contains neurons encoding reward prediction errors in addition to neurons with persistent activity related to previous rewards [43,66]. A sudden change in the average rate is likely to indicate unexpected changes in the animal’s environment that require appropriate adjustment in the animal’s behavioral strategies [55]. Therefore, persistent activity related to previous reward might provide a substrate for monitoring the productivity of the animal’s behavior and therefore facilitate flexible behavior.

Concluding remarks
The brain uses the information obtained from previous experience to make predictions about the future and to select behaviors that are likely to result in the most beneficial outcomes. The length of the retention interval during which the brain has to store information about past events varies greatly according to the properties of the environment and the animal’s behavioral needs. Presumably, structural changes in the nervous system are necessary for long-term information storage. By contrast, persistent neural activity might underlie short-term storage of behaviorally relevant information. Such persistent activity can support diverse cognitive processes, ranging from working memory maintenance to computation of value functions. An attractive hypothesis is that persistent activity arises from a neural circuitry capable of temporal integration of its inputs, and might flexibly subserve a multitude of functions. Therefore, characterizing persistent activity in more detail under a diverse set of behavioral conditions and understanding its underlying biophysical mechanisms represent important goals if we are to uncover the biological basis of cognition (Box 2).

Box 2. Questions for future research
- What are the precise functions of persistent activity during working memory? Does the persistent activity observed in different brain areas perform different functions, such as selection and maintenance of remembered items? Does persistent activity reflect active control processes as well as maintenance, and how can we characterize these functions separately?
- How is persistent activity modified to store information about multiple sensory items that are presented simultaneously or sequentially [69]? How do the inputs to the neural circuits generate persistent activity gated so that it stores only behaviorally relevant information?
- What are the mechanisms of persistent activity? Does persistent activity depend on intracellular signals, short-term synaptic plasticity or feedback loops formed between multiple brain regions, including the basal ganglia and thalamus?
- What is the relationship between persistent activity and synaptic plasticity? Does persistent activity play a critical role in inducing synaptic plasticity and hence long-term memory? Conversely, what is the contribution of short-term synaptic plasticity to persistent activity?
- How is the time course of persistent activity controlled? Does persistent activity related to the animal’s previous choices and their outcomes depend on the statistics of the task and environment? How is persistent activity terminated?
- What are the functions of neuromodulators, such as dopamine and norepinephrine, in controlling the strength and time course of persistent activity [70]?

Acknowledgements
We are grateful to Hyojung Seo, Min Whan Jung and Xiao-Jing Wang for helpful discussions. This research was supported by National Institutes of Health Grants R01-EY016407 (C.E.C.), R01-DA024853, R01-MH05916 and P05-NS048328 (D. L.).

References
Hikosaka, O. et al. (2000) Role of the basal ganglia in the control of purposive saccadic eye movements. Physiol. Rev. 80, 953–978
Tark, K.J. and Curtis, C.E. (2009) Persistent neural activity in the human frontal cortex when maintaining space that is off the map. Nat. Neurosci. 12, 1463–1468