Research Proposal

Title: Improving models of visual working memory by assessing representations in early visual cortex

BAA: No. W911NF-17-S-0002-09 ARO Neurophysiology of Cognition

Background: The proposed research investigates the involvement of early visual cortex (EVC) in maintaining highly precise representations of visual information for short periods of time in working memory (WM, sometimes also referred to as short-term memory). Visual working memory flexibly supports a wide range of cognitive tasks by maintaining information that is no longer present in the environment to guide decision making and action planning -- everything from remembering the layout of a scene to operating machinery. One theory, termed the *sensory recruitment hypothesis*, holds that visual WM is supported by the sustained activation of the same early sensory neurons that play a role in encoding visual information[1, 2, 3 ,4, 5]. Early visual cortex may play a particularly important role when highly detailed memories must be maintained, as neurons in EVC might be uniquely capable of supporting high-precision representations of relevant information[5]. For example, previous fMRI research has shown that mnemonic information can be gleaned by using response patterns in visual cortex to generate model-based reconstructions of remembered stimuli, demonstrating that memory representations are maintained even after the stimulus input has ended[6]. However, this hypothesis has been challenged. First, electrophysiology work in non-human primates has sometimes failed to observe activity in early visual cortex during the delay period of visual WM tasks[7]. Moreover, others have argued that using the same neurons to encode new sensory inputs and to remember previously seen information could lead to interference and confusion between seen and remembered stimuli[6, 8, 9]. For example, if a pilot encodes and remembers the color of the warning light, then this memory might be overwritten by new inputs as soon as they make an eye movement to look at their altimeter. Because we typically move our eyes several times a second, it is highly important to understand the neural mechanisms that support our ability to remember visual information while receiving new visual inputs. However, given the lack of consensus regarding how visual codes are affected by irrelevant visual inputs, it is still unclear whether EVC plays a critical role in WM or whether other mechanisms are recruited to prevent interference between seen and remembered stimuli. Here I will address this gap by using artificial recurrent neural networks (RNNs) to generate predictions about the neural mechanisms underlying WM, and then I will test these predictions *in vivo* using human behavioral and fMRI experiments. By combining levels of analysis, this project will provide new traction on how the brain maintains visual information in the presence of disrupting sensory inputs.

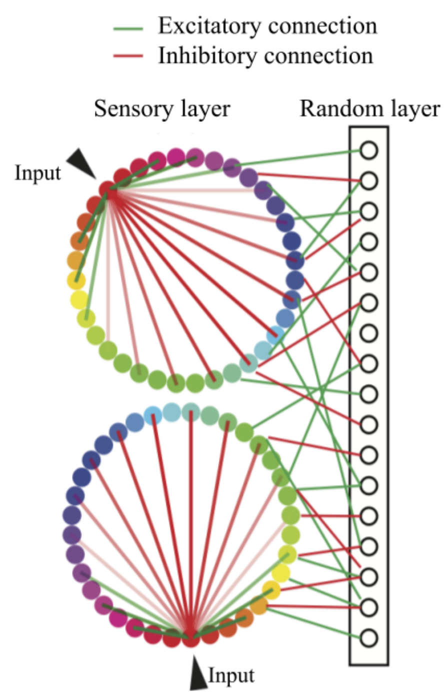

Fig 1. Two sensory rings randomly and reciprocally connected to neurons in the random layer (8 rings in the actual model). Neurons in the sensory layers have excitatory connections with neighboring neurons and inhibitory connections with distant neurons, yielding circular Gaussian tuning for sensory features.

Aim 1: Using a RNN to assess induced memory-related activity in EVC. The proposed project combines computational modeling, behavioral experiments, and neuroimaging to critically test whether or not EVC plays a key role in visual WM. In Aim 1, I use a two-layer, spiking, RNN[10] and demonstrate that stochastic visual stimulation during a memory-retention interval can induce sustained and selective neural responses in the network that resemble neural responses associated with real memories (i.e. stochastic inputs induce ‘false memories’). In turn, I can determine if stochastic stimulation actually induces similarly selective responses in human visual cortex, and if so, whether these responses bias memory representations. If induced activity in sensory cortex does bias the contents of memory, then that would strongly implicate EVC as being one of the key neural mechanisms of visual WM.

The first layer of the RNN contains 8 pools (or “rings”) of feature selective sensory neurons that encode incoming visual stimuli. These sensory neurons are then connected randomly and reciprocally with neurons in a second layer (termed the *random* layer, Figure 1)[10]. The reciprocal connections between the sensory and random layers supports persistent spiking in response to an input stimulus, even after the stimulus is removed, and these representations exhibit many of the key properties of human visual WM such as competition between remembered stimuli and a limited memory capacity. In addition, because of the random projections from the sensory layer to the random layer, the neurons in the random layer receive input from sensory neurons that are tuned to different features and this creates multidimensional tuning profiles -- or mixed selectivity -- much like tuning profiles that are observed in parietal[11] and frontal cortex[12, 13]. In pilot work, I presented one to-be-remembered stimulus to the network and then injected stochastic white noise stimulation into all eight sensory rings to simulate irrelevant sensory inputs that might occur as a human subject moved their eyes and encoded new information while maintaining information in WM. This noise drove activity in the sensory network, which then stimulated a subset of neurons in the random layer. In turn, the activated neurons in the random layer provide feedback to neurons in the sensory layers that are tuned to different features. Preliminary data, shown in Figure 2, shows that this feedback can result in sustained and highly selective activity in sensory rings that were only stimulated with noise, just as if a real stimulus had actually been presented (note the onset of sustained and selective spiking in Figure 2 around 200ms with the onset of noise stimulation). The observation of these “false memory” representations induced by the presentation of noise provides an ideal test case for examining the role of sensory neurons in WM because I can determine if sustained activity in sensory cortex -- induced either by a real memory stimulus or by white noise -- biases the contents of WM.

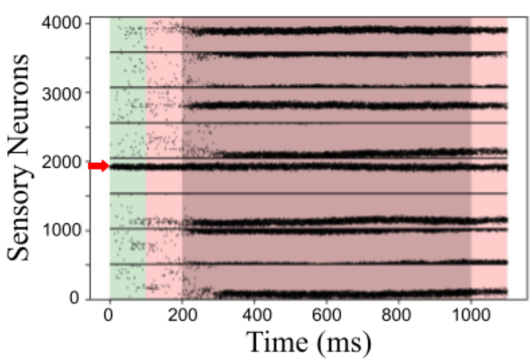

Fig 2. Raster plot of spiking activity in each of the 8 sensory rings (neurons in each ring are separated by a horizontal solid line). At time 0, a memory stimulus was presented to the 4th sensory ring (red arrow). 200ms later, stochastic white noise was presented to all of the sensory rings for 800ms (shaded grey region), leading to sustained and selective neural activity that resembles the activity associated with a real memory.

Aim 2: Using behavior and fMRI to assess representations of ‘false’ memories in EVC. In Experiment 1, participants will be presented with an oriented memory stimulus, selected from a uniform distribution ranging from 0-180 degrees, followed by filtered noise distractors appearing in different quadrants of the visual field (Figure 3). On each trial, subjects will encode the oriented memory stimulus. After a brief delay, a flickering distractor “noise patch” will appear at the location of the memory stimulus. At the end of the trial, participants will perform a recall task where they have to indicate whether a memory stimulus was present -- and if so what orientation it was -- in each of the four possible locations.

The task in Experiment 1 will allow us to test the RNN-based prediction that injecting noise into EVC can create a false memory (as shown in Figure 2). Here the distractor stimuli serve as the noise that is injected into the sensory network, EVC, and the behavioral performance on the visual WM task allows us to assess whether the noise-induced false memories lead to a memory bias. For example, if the participants report perceiving a target stimulus in an incorrect location, this would indicate that false memory representations in visual cortex do influence the contents of memory. In anticipation of response bias from participants learning to expect only one target stimulus, I will also manipulate the set size of the stimulus presentations between one, two, and four. There is prior evidence that changes in set size will have an effect on memory strength[14]. With this in mind, the manipulation of set size will allow us to determine if people are more or less likely to report false memories of perceived targets as set size increases.

Experiment 1 will test the prediction that presenting noise stimuli to visual cortex – in addition to the visual WM target stimuli – will produce false memories as seen in the RNN (Figure 2). This observation would be *consistent* with a key role of EVC in visual WM. However, in Experiment 2, I will more directly examine the role of visual cortex in this process using fMRI and the visual task used in Experiment 1. I will examine multivariate patterns of neural activity across voxels to decode the remembered items using an inverted encoding model[15, 16]. I predict that presenting noise to EVC will result in a decodable representation that will track behavioral indications of false memories (as in Experiment 1). This finding would provide more evidence supporting the hypothesis that EVC is involved in visual WM. However, if there is no decodable representation in EVC and there is still a behavioral false memory bias in Experiment 1, then this would argue against the hypothesis that EVC plays a critical role in visual WM. Irrespective of the outcome, the information gained from the behavioral and the fMRI experiment can be used to further constrain the RNN and to generate further hypotheses about the nature of WM representations in human cortex. For example, I could provide targeted stimulation to neurons in the random layer based on their selectivity during the delay period. This would increase the gain of the subset of neurons in the random layer that are most engaged in a given memory representation, and should render the sensory codes more precise even in the face of new sensory inputs (and render the system less susceptible to false memories).

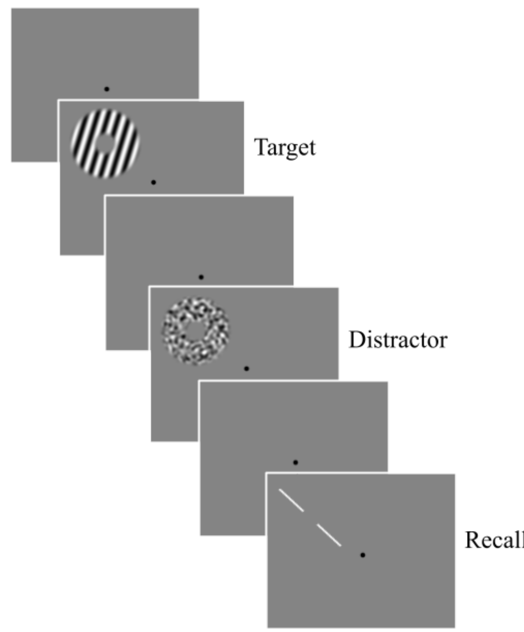

Fig 3. Experiment 1 task paradigm

Impact: An important aspect of understanding human cognition is developing accurate and useful models of perception and memory. Here, I will improve our understanding of WM, which is essential for behaviors such as navigating complex graphics and control displays and operating in novel environments. This approach is therefore aligned with the core NDSEG mission of improving models of neural processes in sensory cognition as described in BAA: W911NF-17-S-0002-09.