Rotations of working memory representations protect from interference in a dual task paradigm

A. Mahrach[1], X. Zhang[2,3], D. Li[2,3], C.T. Li[2,3], A. Compte[1]

[1] IDIBAPS, Barcelona, Spain

[2] Institute of Neuroscience, State Key Laboratory of Neuroscience, Chinese Academy of Sciences, CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai Center for Brain Science and Brain-Inspired Technology, Shanghai, China

[3] School of Future Technology, University of Chinese Academy of Sciences, Beijing, China

Working memory (WM) is a cognitive function that allows for the short-term maintenance and manipulation of information when it is no longer accessible to the senses. It relies on the temporary storage of stimulus features in the neuronal activity. However, the mechanisms underlying its maintenance when subject to task-irrelevant influences, are unknown. Recent studies involving WM tasks with distractors have suggested that stimulus representations before and after distraction are orthogonal, thus protecting WM from interferences. However, whether "*orthogonalization*" is a general mechanism for WM preservation remains an open question, and the network mechanisms supporting it are unclear. Here, we investigated WM protection mechanisms in calcium imaging data from the prelimbic cortex (PrL) in mice performing a recently developed olfactory *dual task*. The dual task consists of an outer delayed paired-association task (DPA) combined with an inner Go-NoGo task inserted in the DPA's delay period. To perform the dual-task correctly, mice need to maintain the DPA sample's memory while performing the Go/NoGo task. Mice performed well after a few training days even in distractor trials. We studied how the memory representation of the DPA sample odors was related to the Go/NoGo distractor odor representation as animals learned the dual task. Using a population decoding approach, we inferred the low-dimensional encoding axes of the sample and distractor odors in the neural response space. In the first few days of training, the population activity significantly overlapped with both samples' early memory axis and the distractors' axis. This was no longer the case in the last few training days where sample representation evolved dynamically to become nearly orthogonal to the distractors' axis. We show that this rotation coincides with a more stable sample representation across the DPA delay period. Altogether, our results suggest that rotations of WM representations in PrL play a role in preserving WM from interfering tasks. Finally, we give a mechanistic account of PrL's WM rotations in a network model of strongly recurrent neurons with low-rank connectivity and short-term facilitation.