Understanding transformations between “decision-potent” and “decision-null” representations in visual working memory using demixed PCA

Quan Wan[1], Jorge A. Menendez[2], Bradley R. Postle[1,3]

[1]Department of Psychology, University of Wisconsin–Madison
[2]Gatsy Computational Neuroscience Unit, University College London, UK
[3]Department of Psychiatry, University of Wisconsin–Madison

Abstract: In the 2-back working memory task the representation of a stimulus item (*n*) transitions from memory probe (for comparison vs. *n - 2*) to unprioritized memory item (UMI; while *n + 1* is compared to *n - 1*), and then to prioritized memory item (PMI; for comparison to *n + 2*). Inverted encoding modeling (IEM) of EEG voltage data (*N* = 30) revealed “opposite” reconstruction of stimulus representations as UMI relative to their reconstructions as PMI, suggesting that stimulus representations might be rotationally “remapped” with changes of behavioral priority status (Wan et al., 2020). Such rotational remapping has also been observed in the dynamics of a recurrent neural network (RNN; with an LSTM architecture) trained to perform this task. Moreover, only the PMI coding dimensions seemed to align with the dimensions relevant to the RNN's output, presumably facilitating its influence on the decision and response. In the current work, we use demixed Principal Component Analysis (dPCA; Kobak et al., 2016), a dimensionality reduction method, to better understand the representational transformations between UMI and PMI, both *in vitro* and *in silico*. Combining the advantages of traditional approaches such as LDA and PCA, dPCA can parse neural representation into a few components related to task parameters (e.g., stimulus, decision, and time), while capturing most of the variance of the data. In both EEG and RNN data, we are using dPCA to identify stimulus- and decision-dependent axes, to characterize how they interact with each other, and to track the representational dynamics along these dimensions as a function of behavioral priority. These analyses will enable us to distill essential mechanisms of how brains and RNNs dynamically transform working memory representations to flexibly guide decisions.