Inter-hemispheric prefrontal mechanisms of within- and across-trial working memory

Melanie Tschiersch (mel.tschiersch@gmail.com)
IDIBAPS, Barcelona, Spain

Akash Umakantha (aumakant@andrew.cmu.edu)
Ryan C. Williamson (rwilliamson88@gmail.com)
Matthew A. Smith (mattsmith@cmu.edu)
Carnegie Mellon University, Pittsburgh, PA, USA

Joao Barbosa (palerma@gmail.com)*
École Normale Supérieure, Paris, France

Albert Compte (acompte@recerca.clinic.cat)*
IDIBAPS, Barcelona, Spain

* indicates co-Senior Author

Abstract:

Working memory (WM) refers to short-term information maintenance and its processing. Recent evidence suggests an interplay between persistent activity and activity-silent mechanisms in the prefrontal cortex (PFC) for WM maintenance and history effects (serial dependence) (Barbosa, Stein et al., 2020). However, the PFC has so far been viewed as one entity, even though neurons in the PFC display a contralateral preference (Funahashi et al., 1989). It is therefore still unclear how neural representations of WM maintenance and serial dependence relate across hemispheres. Here, we answer this question by analyzing behavior and bilateral PFC neural recordings from three monkeys performing a visuo-spatial delayed response task. We found behavioral and neural evidence for diffusing memories and serial dependence. Interestingly, the neural correlates of memory drift were weakly, but significantly, correlated across hemispheres, suggesting weak hemispheric interactions. When comparing several two-area bump-attractor models with varying degrees of tuned and untuned across-area connections, we found that tuning of across-area connections is necessary for error correlations to emerge. The model further predicted either private or shared serial dependence drift across hemispheres based on increasing connectivity strength. The data confirmed a private history-drift and we therefore conclude that faint, but specific connections underlie continuous working memory across prefrontal hemispheres.

Keywords: working memory; hemisphere; serial dependence; prefrontal cortex

We analyzed behavioral and neural data (139.2±23.8 multi-units/session from bilateral area 8Ar with implanted Utah arrays; 41 sessions) of three monkeys in a visuospatial delayed response task (Fig. 1a). All three monkeys showed delay-dependent precision errors (Sa: 1.02°/s (p=1.8e-15), Wa: 4.5°/s (p=5.1e-14), Pe: 0.9°/s (p=1.4e-16) from OLS model of absolute error with delay) and varying degrees of serial dependence: An attraction to the previously shown item (Sa: 2.7° (p=6.7e-65), Wa: 0.85° (n.s.,p=0.08), Pe: 0.06° (n.s.,p=0.8), from OLS model of error with Derivative of Gaussian (DoG) fit).
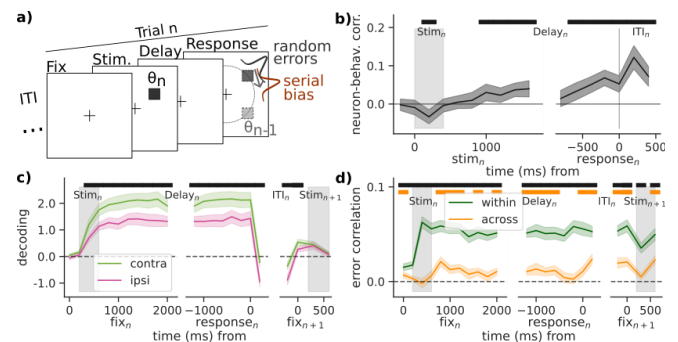

Figure 1: Memory drift is correlated across prefrontal hemispheres. a) Task schematic showing a random response error distribution and a serial dependence distribution shifted towards the previous item. b) The correlation of neural decoder errors with the behavioral response errors increases during the delay suggesting memory drift. Error bars are CI. c) Decoder of stimulus location shows lateralized PFC tuning

during delay but not during reactivations (fix$_{n+1}$). d) Neural decoder errors correlated between the left and right hemisphere (across) or within a hemisphere. Error bars are S.E.M. unless otherwise specified. Black horizontal bars are significance (p<0.05).

Memory errors are weakly correlated across hemispheres

To estimate the evolving encoding of the stimulus location through the delay period, we trained a 5-fold cross-validated linear decoder to predict the stimulus location from the bilaterally recorded neural activity. When correlating the error of the neural predictions in held-out trials with the behavioral response errors, we saw a progressive increase in correlation through the delay period (Fig. 1b). This population analysis generalizes a previous finding of slow memory drift in single neurons, predicted by bump-attractor models (Wimmer et al., 2014) and explains the delay-dependent precision errors seen in behavior. We then trained hemisphere-specific linear decoders on delay-averaged activity and found stronger memory representations for contralateral visual stimuli during the delay period, but not during reactivations (which predict history effects (Barbosa, Stein et al., 2020)) (fix$_{n+1}$ in Fig. 1c). Interestingly, the prediction errors of the hemisphere-specific neural decoders were significantly correlated across the left and right hemisphere, but were weaker than correlations within each area (cross-validated and balanced by number of neurons, Fig. 1d), suggesting weak connections between the two hemispheres.

A biophysical network model suggests tuned connections across hemispheres

To understand the inter-hemispheric neural mechanisms underlying WM, we implemented several two-area bump attractor spiking networks with different connectivity profiles between the two areas (Fig. 2a, weak=0.2%, faint=0.06%, untuned=0.2% of within connectivity), and short-term synaptic plasticity (STP) in local connections (as in Barbosa, Stein et al. (2020)). We found that tuning in the connections between areas was necessary to produce correlated memory drift across areas (Fig 2b, *untuned* at 0) and that tuning strength between areas determined correlation strength (Fig. 2b, *faint* vs *weak*). Therefore, only tuned connections can replicate the finding of correlated errors in the neural data (compare to Fig. 1d, across). Additionally, the model through local STP

displays serial dependence, which develops through the delay (Fig. 2c). Interestingly, we found that based on the connectivity strength of the model, the memory drift towards the reactivations in the same or opposite hemisphere could be either private ("across" reactivation bias at 0 for *faint* model, Fig. 2d) or shared across hemispheres ("across" reactivation bias increases for *weak* model, not shown).

Neural drift to previous memories is private within hemispheres suggesting faint inter-hemispheric connections

To constrain the proposed connectivity strengths, we tested if reactivations had a shared or private effect across hemispheres in the neural data. To this effect, we trained a leave-one-out cross validated decoder for each hemisphere, leading to trial-by-trial predictions in reactivations and delay periods for each hemisphere. We first confirmed that the neural activity during the delay drifted towards the previous stimulus (Fig. 2e). We further tested if the neural code in the delay period in each hemisphere drifted more towards the reactivated location in the same or the opposite hemisphere. We only found a drift towards reactivations within the same hemisphere, but not across hemispheres (Fig. 2f), as predicted by the *faint* model. These findings suggest faint, but tuned inter-area communication.
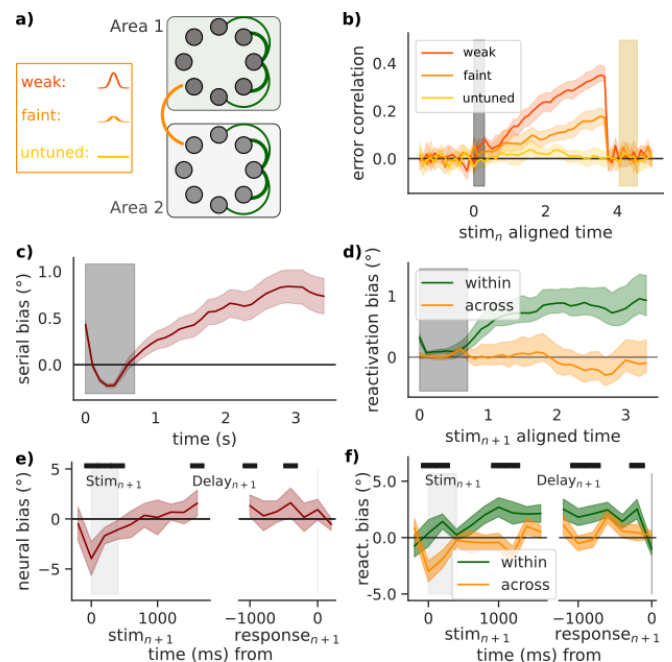

Figure 2: Two faintly tuned connected bump attractors replicate weak error correlations and predict private

reactivations. a) Schematic of two connected bump attractor networks and their ring connectivity profiles with different across-area connectivities. b) Error correlations for models with different inter-area connectivities. c) Serial dependence of the model grows from repulsion to attraction. d) *Faint* model: Memory traces drift exclusively to the reactivations within the same area, but not across areas. e) Data: Decoded locations in the current trial are first repelled (adaptation) and later attracted (serial dependence) to the previous memory item. Neural bias is the DoG fit in each time point. f) Data: Same es e), but viewing attraction to the previous reactivation location in either the same (within) or the other hemisphere (across), shows a similar pattern as the *faint* model.

Conclusions

In conclusion, we found neural evidence for slowly drifting memory traces that were correlated across hemispheres in the monkey prefrontal cortex within a single trial, but not across trials (serial dependence). We propose a model of two bump attractors with tuned, but faint inter-area connections and local short-term plasticity, which can replicate the behavioral and neural effects, as a model for interhemispheric WM.

Acknowledgments

References

Barbosa, J., Stein, H., Martinez, R. L., Galan-Gadea, A., Li, S., Dalmau, J., Adam, K. C. S., Valls-Solé, J., Constantinidis, C., & Compte, A. (2020). Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. *Nature Neuroscience*, 23(8), Article 8.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, 61(2), 331–349.

Wimmer, K., Nykamp, D. Q., Constantinidis, C., & Compte, A. (2014). Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nature Neuroscience*, 17(3), Article 3.