

NMDA-Receptor Dysfunction Disrupts Serial Biases in Spatial Working Memory

Heike Stein¹ (heike.c.stein@gmail.com)
Joao Barbosa¹ (palerma@gmail.com)
Josep Dalmau² (jdalmau@clinic.cat)
Albert Compte¹ (acompte@clinic.cat)

Theoretical Neurobiology (1) and Neuroimmunology (2), IDIBAPS
Carrer Rosselló 149-151, Barcelona, Spain

Abstract

In working memory (WM) tasks, attractive biases to previous items are evidence for continuous temporal integration of memories. These serial biases have been modeled as a product of synaptic short-term plasticity, allowing WM representations to endure in a synaptic trace and interfere with the next trial even when neural activity returns to baseline values. We hypothesized that the NMDAR, a key component of both short-term potentiation (STP) and stable WM delay activity, would be of central importance to serial biases in a visuospatial WM task. Confirming this hypothesis, we found drastically reduced biases in patients with anti-NMDAR encephalitis and schizophrenia, both diseases that have been related to NMDAR hypofunction. We simulated serial biases in a spiking neural network supported by a Hebbian STP mechanism that builds up during persistent delay-activity. We found a close correspondence between patient and model behavior when gradually lowering levels of STP, suggesting a disruption of short-term plasticity in associative cortices of schizophrenic and anti-NMDAR encephalitis patients. Further, we explored the capability of the model to explain reduced biases in light of the disinhibition theory of schizophrenia.

Keywords: working memory; NMDAR; schizophrenia; spiking neural network; E-I balance

Introduction

Continuity of mnemonic contents in time contributes to forming coherent memory representations. Recently, attractive response biases towards previously memorized features in delayed-response tasks have been reported as evidence for the continuous integration of working memory (WM) contents between trials (Fischer & Whitney, 2014). In

turn, brain disorders with reported executive and memory dysfunction may be characterized by reduced WM serial bias (Lieder et al., 2019), revealing reduced temporal coherence of memory representations. To gain mechanistic insight into this effect, we tested a unique population of patients recovering from anti-NMDAR encephalitis, an immune-mediated brain disease causing a drastic reduction of NMDAR, and a population of patients suffering from schizophrenia, another disorder associated with hypofunctional NMDARs (e.g., Javitt, 2010). NMDAR hypofunction is a candidate mechanism for a disturbed balance between cortical excitation and inhibition (E-I balance, Belforte et al., 2010), supporting one major glutamatergic theory of schizophrenia, the cortical disinhibition hypothesis (Lewis, Hashimoto, & Volk, 2005). On the other hand, there is evidence for reduced excitation of cortical networks under the influence of NMDAR antagonists in the presence of cognitive demand (Wang et al., 2013; Driesen et al., 2013). Further, NMDARs are an essential component of synaptic plasticity, including short-term potentiation (STP), leading to reduced STP when NMDAR are inhibited (Volianskis et al., 2013).

We hypothesized that NMDAR hypofunction in anti-NMDAR encephalitis and schizophrenic patients would change the temporal dynamics of working memory, influencing how information is maintained and integrated between trials. We collected behavioral data from anti-NMDAR encephalitis patients, schizophrenic patients, and healthy controls performing a visuospatial delayed-response task, and analyzed serial biases as a behavioral read-out of information integration across time.

Results

Serial biases are reduced in anti-NMDAR encephalitis and schizophrenia

Serial biases in visuospatial WM are characterized by delay-dependent attraction of current-trial responses towards previously remembered stimuli. This



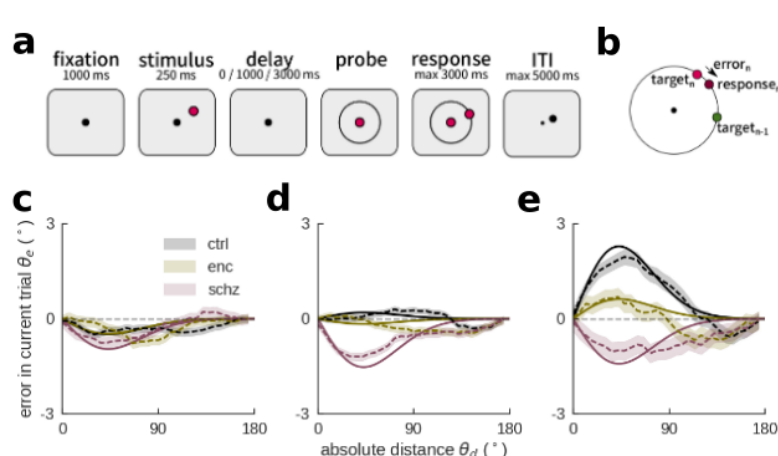


Figure 1: WM task and serial biases. **a)** Task protocol. After 1 s of fixation, a stimulus appeared at any of 360 angular positions. After a delay of 0, 1, or 3 s, subjects should make a mouse click to the remembered location. **b)** Serial bias in trial n can be measured as the angular displacement of responses towards stimuli in trial $n-1$. **c-e)** Serial biases for 0, 1 and 3 s delay (left to right). Dashed lines show biases as a function of absolute previous-current distances θ^d . For negative θ^d , θ^e were multiplied by -1 . Error shading indicates 1 SEM. Solid lines show linear model fits (Methods).

attraction results in a slight, but systematic shift of the error distribution towards the previous stimulus. The functional form in which response errors depend on the previous-current stimulus distance follows approximately the form of the first derivative of Gaussian (DoG; Fischer & Whitney, 2014). To statistically assess differences in delay-dependent biases, we thus modeled single-trial errors θ^e as a linear model of delay length, group, and the DoG of previous-current distance θ^d (DoG(θ^d), Methods). Interaction terms of delay and group with DoG(θ^d) allow to determine systematic influences of these factors on the amplitude of serial biases.

In accordance with previous results (Bliss, Sun, & D'Esposito, 2017; Papadimitriou, White, & Snyder, 2015), we found that the strength of attractive bias depended on WM delay (delay \times DoG(θ^d), $F(2,43) = 11.47$, $p = .0001$). Moreover, biases differed between groups of participants (group \times DoG(θ^d), $F(2,46) = 8.83$, $p = .0005$), especially when comparing each group for different WM delays (group \times delay \times DoG(θ^d), $F(4,43) = 8.04$, $p = 6.2e-5$). Figure 1c-e shows model fits and

averaged errors θ^e as a function of θ^d for 0, 1 and 3 s delay lengths, left to right). Groupwise models allowed to assess the time course of biases in each population: initially slightly repulsive biases became gradually attractive with delay length for healthy controls ($F(2,16) = 24.39$, $p = 1.4e-5$) and for encephalitis patients ($F(2,22.8) = 5.01$, $p = 0.016$). In contrast, delay length did not affect serial biases in schizophrenic patients ($F(2,20.5) = 1.14$, $p = .34$). Rather, a significant repulsive bias ($F(1,14.3) = 7.19$, $p = .018$) dominated all delay lengths. Mean biases for each group and delay are shown in Figure 1c.

Reduced NMDAR-dependent STP causes reduced serial bias in spiking neural networks

To explore potential network mechanisms of reduced serial biases in NMDAR-related diseases, we ran simulations of consecutive WM trials in a spiking neural network model of prefrontal cortex (Compte, Brunel, Goldman-Rakic, & Wang, 2000). In this model, a pool of inhibitory and a pool of excitatory leaky integrate-and-fire (LIF) neurons are connected all-to-

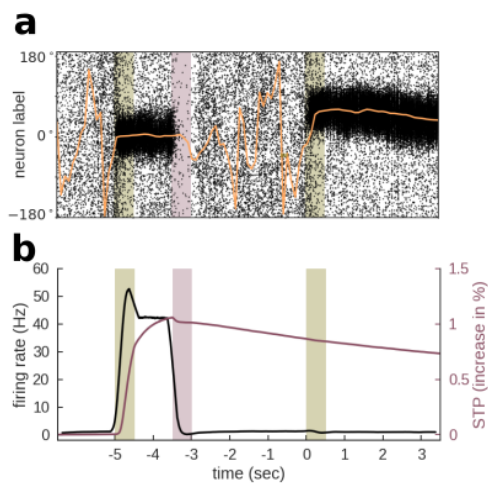


Figure 2: Simulation of two consecutive trials in a network model of WM. **a)** Raster plot of spike times in the excitatory pool of neurons. Neurons are ordered along the y-axis by their preferred angular location. Stimulation times are marked by green bars, and the red bar marks the response time in trial $n-1$. The orange line shows the population vector decoded from firing rates of all neurons in the network in sliding windows of 250 ms. **b)** Trial-averaged firing rates (black) and STP traces (maroon) of neurons selective to stimulus $n-1$. During trial $n-1$, firing rates and STP are maximal and decrease after the response. At the time of onset of stimulus n , information about the previous trial is available only in the STP trace.

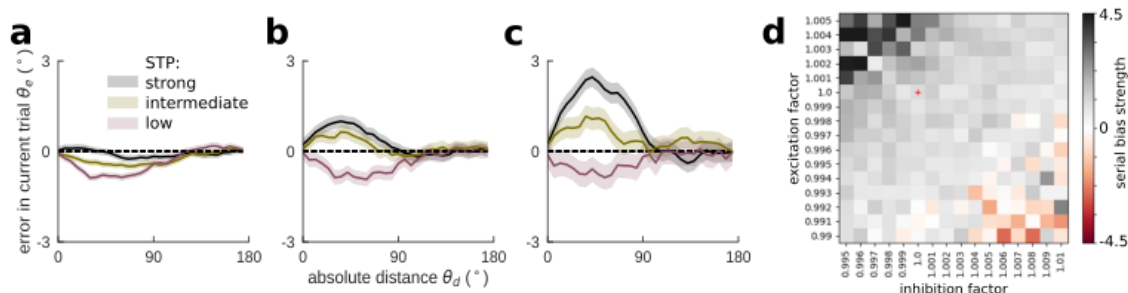


Figure 3: Simulating serial bias in schizophrenia and anti-NMDAR encephalitis by reducing synaptic short-term potentiation (STP) and by disturbing E-I balance. **a-c)** Reducing STP decreases attractive serial bias gradually. Left to right, bias plots after 0, 1, and 3 s delay. Error shading depicts 1 SEM. **d)** A reduction of excitation to inhibitory neurons (disinhibition) increases serial bias, while reducing excitation to excitatory neurons decreases serial bias. Note that decreasing or increasing NMDA-mediated excitation to both excitatory and inhibitory neurons concurrently does not affect biases (diagonal pattern), while inducing a strong E-I imbalance causes a loss of network bistability (loss of the memory attractor for $I > E$, lower right orange triangle, and loss of stable low baseline firing when $E > I$, upper left black triangle).

all. Crucially, the strength of connectivity within the excitatory pool is determined by the similarity of any two neurons' preferred stimuli, so that neurons with similar preferred locations excite each other more strongly. During the delay period, attractor dynamics cause persistent neural firing even in the absence of external input (Figure 2a). In the response period, neural activity is suppressed with an external input and the network returns to a stable baseline state of low firing rates.

We extended this network model by adding an associative STP mechanism in excitatory synapses. Synapses between neurons that participated in previous-trial delay activity are potentiated by $\sim 1\%$ (Figure 2b) and this potentiation decays between trials as a result of sparse, spontaneous presynaptic spikes (Volianskis et al., 2015). At the onset of the next stimulus, a location-specific synaptic trace persists from the previous trial and attracts the current trial's bump of activity progressively over the course of the delay (orange line in Figure 2a).

To investigate WM behavior, we simulated 5,000 pairs of consecutive trials with randomized stimulus locations. Network inputs were slightly transformed to mimic the repulsive bias away from previous stimulus locations after 0 s delay, reflecting adaptation effects taking place in sensory cortices (Bliss & D'Esposito, 2017). To obtain behavioral readouts from the network, the population vector was calculated from firing rates measured just after stimulus offset (0 s), after 1 s, and after 3 s. Errors θ^e were calculated as the distance between the decoded angle and the stimulus angle for each trial. Figure 3a-c shows serial biases after 0, 1, and 3 s for different levels of STP

(strong, intermediate, and low STP). Similarly to patients' data, serial biases were initially repulsive, and increased over the course of WM delay for intact STP. In contrast, conditions of reduced STP led to lower serial biases, even after a 3 s delay (low STP).

We further investigated whether disturbing the E-I balance of our network could account for reduced serial biases. We gradually reduced/increased the efficiency of NMDAR in excitatory and inhibitory synapses (Figure 3d) and measured peak serial bias amplitudes for all conditions. Our analyses show that a disinhibition of the circuit, by reducing excitatory NMDA conductances onto inhibitory neurons, increased serial bias, rather than decreasing it. In contrast, a decrease in excitatory conductances onto excitatory neurons reduced bias.

Together, these simulations are inconsistent with the cortical disinhibition hypothesis of schizophrenia, at least during the mnemonic period of a WM task. Rather, they support two alternative hypotheses: (1) a reduction in NMDA-dependent STP or (2) reduced NMDA-mediated recurrent excitation between delay-active neurons.

Conclusions

We found a drastic reduction of attractive WM serial biases in patients with anti-NMDAR encephalitis and schizophrenia, as compared to healthy controls. Using a biologically-constrained model, we propose that synaptic dysfunction underlies these reductions in bias, either as a result of reduced STP or diminished recurrent cortical excitation.

Methods

Experimental Procedures

We tested 16 patients with anti-NMDAR encephalitis (enc), 16 schizophrenic patients (schz), and 18 healthy controls (ctrl) in a visuospatial WM task shown in Figure 1a. In each trial, subjects were to memorize a stimulus that appeared at a random angular location with fixed eccentricity. After a delay of 0, 1 or 3 seconds, they were prompted to respond by making a mouse click at the remembered location. Response errors θ^e were measured as the angular displacement of responses from the target stimulus in each trial.

Data analysis

A linear mixed model of errors θ_{ij}^e in trial i and subject j included the factors group (ctrl, enc or schz), delay (0, 1, or 3 s) and first derivative of Gaussian (DoG) basis functions of previous-current stimulus distance θ_{ij}^d :

$$\begin{aligned} \theta_{ij}^e = & \beta_0 + \beta_1 \text{group}_j + \beta_2 \text{delay}_{ij} - \beta_3 \text{DoG}(\theta_{ij}^d) \\ & + \beta_4 \text{group}_j \text{delay}_{ij} - \beta_5 \text{group}_j \text{DoG}(\theta_{ij}^d) \\ & - \beta_6 \text{delay}_{ij} \text{DoG}(\theta_{ij}^d) - \beta_7 \text{group}_j \text{delay}_{ij} \text{DoG}(\theta_{ij}^d) \\ & + \gamma_{0j} - \gamma_{1j} \text{delay}_{ij} \text{DoG}(\theta_{ij}^d) + \varepsilon_{ij} \end{aligned}$$

The DoG location hyperparameter was fixed at 0° and the variance hyperparameter was determined using cross-validation.

Neural network simulations

We simulated 5,000 consecutive pairs of trials using a LIF bump-attractor network model of spatial WM as previously described by Compte, Brunel, Goldman-Rakic and Wang (2000). Synaptic conductance dynamics modeled NMDARs, AMPARs, and GABA_ARs. We extended the model by including Hebbian STP at recurrent excitatory synapses.

Acknowledgments

Funding provided by Institute Carlos III, Spain (grant PIE 16/00014), Cellex Foundation, the Spanish Ministry of Science, Innovation and Universities (grants BFU 2015-65318-R, RTI2018-094190-B-100), the European Regional Development Fund, the Generalitat de Catalunya (grant AGAUR 2017 SGR 1565), "la Caixa" (LCF/BQ/IN17/11620008, H.S.), and the European Union's Horizon 2020 Marie Skłodowska-Curie grant (713673, H.S.).

References

Belforte, J. E., Zsiros, V., Sklar, E. R., Jiang, Z., Yu, G., Li, Y., ... Nakazawa, K. (2010). Postnatal NMDA receptor ablation in corticolimbic interneurons confers schizophrenia-like phenotypes. *Nature Neuroscience*, 13, 76–83.

Bliss, D. P., & Esposito, M. D. (2017). Synaptic augmentation in a cortical circuit model reproduces serial dependence in visual working memory. *PLoS ONE*, 12, 1–19.

Bliss, D. P., Sun, J. J., & D'Esposito, M. D. (2017). Serial dependence is absent at the time of perception but increases in visual working memory. *Scientific Reports*, 7, 1–13.

Compte, A., Brunel, N., Goldman-Rakic, P. S., & Wang, X.-J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral Cortex*, 10, 910–923.

Dalmau, J., Tüzün, E., Wu, H. Y., Masjuan, J., Rossi, J. E., Voloschin, A., ... Lynch, D. R. (2007). Paraneoplastic anti-N-methyl-D-aspartate receptor encephalitis associated with ovarian teratoma. *Annals of Neurology*, 61, 25–36.

Dalmau, J., Lancaster, E., Martinez-Hernandez, E., Rosenfeld, M. R., & Balice-Gordon, R. (2011). Clinical experience and laboratory investigations in patients with anti-NMDAR encephalitis. *The Lancet Neurology*, 10, 63–74.

Driesen, N. R., McCarthy, G., Bhagwagar, Z., Bloch, M. H., Calhoun, V. D., D'Souza, C., ... Krystal, J. H. (2013). The impact of NMDA receptor blockade on human working memory-related prefrontal function and connectivity. *Neuropsychopharmacology*, 38, 2613–22.

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, 17, 738–743.

Javitt, D. C. (2010). Glutamatergic theories of schizophrenia. *Israel Journal of Psychiatry and Related Sciences*, 47, 4–16.

Lewis, D. A., Hashimoto, T., & Volk, D. W. (2005). Cortical inhibitory neurons and schizophrenia. *Nature Reviews Neuroscience*, 6, 312–324.

Lieder, I., Adam, V., Frenkel, O., Jaffe-Dax, S., Sahani, M., & Ahissar, M. (2019). Perceptual bias reveals slow-updating in autism and fast-forgetting in dyslexia. *Nature Neuroscience*, 22, 256–264.

Papadimitriou, C., White, R. L., & Snyder, L. H. (2015). Ghosts in the machine: memory interference from the previous trial. *Journal of Neurophysiology*, 113, 567–577.

Volienskis, A., France, G., Jensen, M. S., Bortolotto, Z. A., Jane, D. E., & Collingridge, G. L. (2015). Long-term potentiation and the role of N-methyl-D-aspartate receptors. *Brain Research*, 1621, 5–16.

Wang, M., Yang, Y., Wang, C. J., Gamo, N. J., Jin, L. E., Mazer, J. A., ... Arnsten, A. F. T. (2013). NMDA receptors subserve persistent neuronal firing during working memory in dorsolateral prefrontal cortex. *Neuron*, 77, 736–749.