Dynamics in the primary sensory, medial temporal and striatal regions during implicit learning of temporal regularity

Wei Tang (wt1@iu.edu)
Department of Psychological & Brain Sciences, Indiana University Bloomington
Bloomington, IN 47404, USA

Zhenghan Qi (z.qi@northeastern.edu)
Departments of Communication Sciences & Disorders and Psychology, Northeastern University
Boston, MA 02115, USA

Abstract

Temporal regularities in sensory inputs can sharpen our perception of the input. How the brain develops sensitivity for temporally structured stimuli through statistical learning (SL) remains unknown. In this study, we investigate brain activities in the early visual, medial temporal and striatal regions during repeated exposure to a temporal structure embedded in sequential visual stimuli. In the late stage of learning, we observed increased sensitivity to stimulus category in V1, and such sensitivity was mirrored by activity in the entorhinal cortex. The entorhinal activity also became correlated with that of the nucleus accumbens towards the end of the task. These results reveal complementary roles of the above regions in the hierarchical processing for learning temporal regularity.

Keywords: Statistical learning; fMRI; V1; Entorhinal cortex; Striatum; Hidden Markov model; Temporal regularity

Introduction

Statistical learning (SL) enhances our sensitivity to repeated patterns in a sensory stream without the need for explicit feedback, even just after short exposure (Fiser & Aslin, 2001; Sherman & Turk-Browne, 2020; Thiessen, Girard, & Erickson, 2016). The enhanced sensitivity facilitates the perception and memory of the input (Batterink, 2017; Isbilen et al., 2022). The neural mechanisms to account for SL are believed to recruit both sensory- and cognitive-level processing (Reber, 2013; Karuza et al., 2013; Frost et al., 2015), and yet, when and where the neural plasticity takes place during SL remains elusive.

In this study we investigate how multi-region brain activity unfolds during the learning of an embedded temporal pattern in sequential visual inputs, using functional magnetic resonance imaging (fMRI). We focused on 6 pre-select regions of interest (ROIs) based on their reported involvement in SL tasks from previous studies (Karuza et al., 2013; McNealy et al., 2006; Schapiro et al., 2012; Richter & de Lange, 2019): the primary visual cortex (V1), hippocampal formation (HF), entorhinal cortex (EC), caudate (Caud), putamen (Put) and nucleus accumbens (NAcc). We aim to understand the multi-level sensory and cognitive processing that leads to facilitated perception by learning the temporal regularity. To overcome the challenge of rapid hemodynamic changes in a fast-paced task, we applied a hidden Markov model (HMM) framework (Baldassano et al., 2017; Bishop, 2006) to dynamically estimate blood-oxygenation-level-dependent (BOLD) responses at each time frame. The observed dynamics reveal complementary roles of V1, EC and NAcc in sensory sharpening through SL.

Methods

Participants. Twenty-two adults (age= mean 20.79 ± s.d. 2.89 years, 7 males) participated in this study. All participants gave written consent.

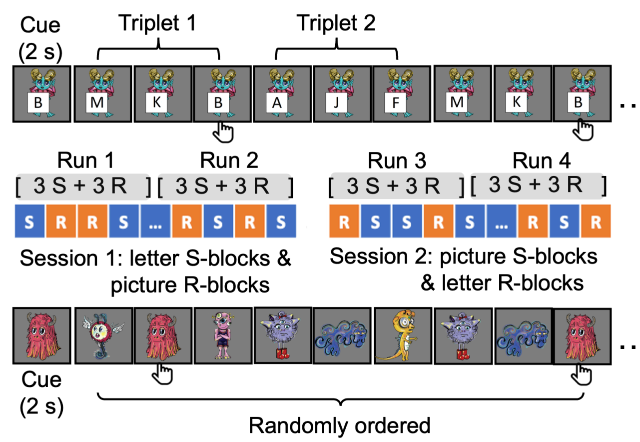

Figure 1: Schematic of the task design.

Stimuli and procedure. Participants viewed sequentially presented images while responding to target images embedded in each sequence (Fig. 1). The stimuli in each sequence were (1) *Letters* or *Pictures* and (2) temporally arranged into triplets ("S-block") or randomly ("R-block"). The target location followed no systematic pattern and the participants were not informed of the embedded structure.

MRI data acquisition & preprocessing. MRI data were acquired on a Siemens 3T Magnetom Prisma scanner with a 64-channel head coil. Functional images used multi-slice T2*-weighted echo-planar scans (TR=800 ms, TE=32 ms, matrix=64 × 64, FOV=21 cm, 61° flip angle, acceleration factor=6, voxel size 2.5 mm isotropic). Preprocessing used the standard fMRIprep pipeline. ROI definition followed FreeSurfer's "aparc" atlas.

HMM design and fitting. We hypothesize that at each time point, the BOLD activity from the 6 ROIs follow a multivariate Gaussian distribution. An *emission probability function* (Bishop, 2006) links the BOLD activity to a finite set of brain states by specifying the mean $\boldsymbol{\mu}_i$ and covariance matrix Σ_i for each state i. Model fitting used the Python toolbox Dynamax (MIT License 2022). One HMM was fit to the same type (S or R) of sequences in each run, treating subjects as independent samples. By leave-one-out cross-validation, the number of states was fixed to 6.

Results

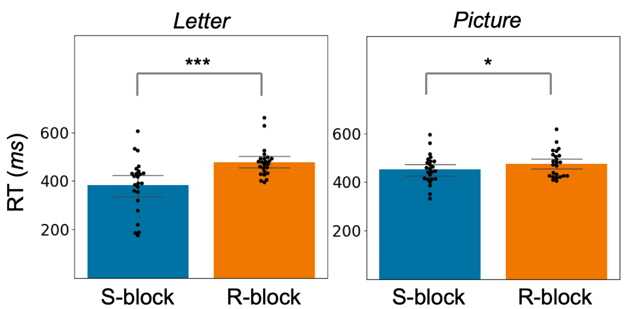

Figure 2: RT shown as median ± s.e. across subjects, superimposed with individual data in black dots. *: $p < 0.05$, ***: $p < 0.001$.

Behavior. Comparisons of RT between S- and R-blocks showed a facilitation effect on target detection. For both types of stimuli, across subjects, RT was significantly lower in S-blocks than in R-blocks (Fig. 2; *Letter* stimuli: Wilcoxon signed-rank W = 23.0, p < 0.00028; *Picture* stimuli: W =71.0, p < 0.024).

BOLD responses in V1 mirrored by EC. We divided the trials into "fast" and "slow" by the median RT, separately for S- and R-blocks. Then, for each RT group and each ROI, we computed the *Letter-Picture* difference of the HMM-estimated BOLD activity in the second run of each learning session (Fig. 3). For V1, this difference was positive in S-blocks and negative in R-blocks. This pattern was mirrored by the EC but not any other ROI. Comparing the difference score between S- and R-blocks, both V1 and EC showed a more significant contrast in fast than slow trials (V1 fast: $t_{163.8} = 3.68$, $p < 0.00032$, slow: $t_{159.6} = 2.04$, $p < 0.042$; EC fast: $t_{157.1} = 3.29$, $p < 0.0012$, slow: $t_{145.4} = 2.71$, $p < 0.0074$). NAcc also showed a significant S-R contrast in fast trials ($t_{130.1} = 2.33$, $p < 0.022$) but not in slow trials ($t_{145.3} = 0.03$, $p < 0.98$).

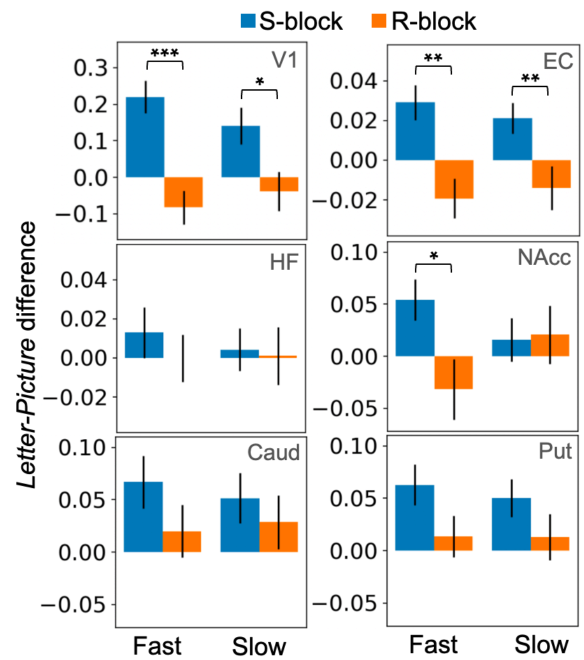

Figure 3: *Letter-Picture* sensitivity in fast and slow trials. Scales of the y-axis were adjusted separately for V1, medial temporal (EC & HF) and striatal (Caud, Put & NAcc) regions to accommodate regional baseline differences for visualization. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

Correlated activity of EC & NAcc. We observed an increase of the HMM-estimated covariance between the EC and NAcc at Run 4, for the S-blocks but not the R-blocks (veridical covariance > 100 out of 100 HMM estimate from permuted sequences).

Discussion

The greater *Letter-Picture* difference of V1 activity in S- than R-blocks suggests sharpened representation of stimulus category for structured sequences. This sharpening effect was greater in fast than slow trials, suggesting its link to the facilitation of target detection. Importantly, the V1 pattern was mirrored by EC, while the latter activity became correlated with NAcc through learning. This suggests that higher-level processing may contribute to the sharpening effect, reflecting a hierarchy in the processes related to SL.

Acknowledgement

This work was supported in part by National Institute on Deafness and Other Communication Disorders (NIDCD) Grant R21DC010576 to Z.Q. The data collection was conducted at the University of Delaware. The main analysis was conducted on the high-performance computing platform Big Red 3 and Carbonate, supported in part by Lilly Endowment, Inc. for the Indiana University Pervasive Technology Institute. We thank the staff at the Center for Biological and Brain Imaging at the University of Delaware for their critical support for our neuroimaging data collection. We thank Violet Kozloff and An Nguyen for their assistance in stimulus construction and programming for the experiment, Julie Schneider, Jennifer Legault, and Yi-Lun Weng for their contribution in data collection and organization, Anqi Hu for her assistance in behavioral data processing, and Pradyumna Shankar for his assistance in fMRI preprocessing.

References

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron*, *95*(3), 709-721.

Batterink, L. J. (2017). Rapid statistical learning supporting word extraction from continuous speech. Psychological Science, 28(7), 921-928.

Bishop, C. M., & Nasrabadi, N. M. (2006). Pattern recognition and machine learning (Vol. 4, No. 4, p. 738). New York: springer.

Conway, C. M. (2020). How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. Neuroscience and Biobehavioral Reviews, 112, 279–299.

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. Psychological science, 12(6), 499-504.

Frost, R., Armstrong, B.C., Siegelman, N. & Christiansen, M.H. (2015). Domain generality vs. modality specificity: The paradox of statistical learning. Trends in Cognitive Sciences, 19, 117-125.

Isbilen, E. S., Frost, R. L., Monaghan, P., & Christiansen, M. H. (2022). Statistically based chunking of nonadjacent dependencies. Journal of Experimental Psychology: General, 151(11), 2623.

Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. Brain and Language, 127(1), 46–54.

McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the language code: neural mechanisms underlying speech parsing. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 26(29), 7629–7639.

Reber, P. J. (2013). The neural basis of implicit learning and memory: a review of neuropsychological and neuroimaging research. Neuropsychologia, 51(10), 2026–2042.

Richter, D., & de Lange, F. P. (2019). Statistical learning attenuates visual activity only for attended stimuli. *elife*, *8*, e47869.

Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. Current Biology: CB, 22(17), 1622–1627.

Sherman, B. E., & Turk-Browne, N. B. (2020). Statistical prediction of the future impairs episodic encoding of the present. Proceedings of the National Academy of Sciences, 117(37), 22760-22770.

Thiessen, E. D., Girard, S., & Erickson, L. C. (2016). Statistical learning and the critical period: how a continuous learning mechanism can give rise to discontinuous learning. Wiley Interdisciplinary Reviews: Cognitive Science, 7(4), 276-288.