Endogenous Midbrain Dopaminergic Fluctuations Selectively Modulate Late-Stage Reward Learning

Kadri Kenza[1], Horsch Belana[1], Dome Lenard[1,2] Kakaei Ehsan[1], T.U.Hauser[1,2,3,4]

[1]Department of Psychiatry and Psychotherapy, Faculty of Medicine, University Tuebingen (Germany),
[2]German Center for Mental Health (DZPG), Partner Site Tübingen (Germany),
[3]Functional Imaging Laboratory (FIL), Department of Imaging Neuroscience, University College London, London (United Kingdom),
[4]Max Planck UCL Centre for Computational Psychiatry and Ageing Research, University College London (United Kingdom)

Main findings

- **Endogenous dopaminergic fluctuations** in the Ventral Tegmental Area (VTA) and the Substantia Nigra (SN) **selectively enhance discrimination** between high- and low-reward stimuli in the **later stages** of a probabilistic learning task
- This modulation has a **persistent effect on stimulus valuation**, as revealed by stronger reward probability-based **confidence** in a post-scan preference task
- Dopaminergic fluctuations in the SN/VTA elicit **sustained BOLD responses** in the midbrain that may reflect the neural basis of these learning effects

Methods

n = 28 healthy participants (15 females) age = 28.28, SD = 5.24

fMRI probabistic reward learning task task with **5 blocks** of **24 trials**
4 objects in each condition with **low** [0.2 0.4] and **high** [0.6 0.8] reward probability

Introduction

How do endogenous midbrain fluctuations drive variability in human learning and decision-making?

Humans do not always make the same decision, even when facing identical options, a pervasive source of behavioural variability whose origins remain poorly understood.

Endogeneous fluctuations in dopaminergic midbrain activity have been shown to contribute to this variability, modulating risky decision-making (Chew et al., 2019).

Our aim was to extend these findings to reward learning, probing whether dopaminergic state in the SN/VTA shapes how well individuals acquire and maintain stimulus-reward associations.

Experimental design

Condition are triggered using the real-time BOLD average in the SN/VTA

Baseline calculated over 2min moving windows

Trough ≤ 15th percentile
Peak ≥ 85th percentile

ITI between 20-50s

Day 1

Structural scans (MPM)

Extraction of the SN/VTA

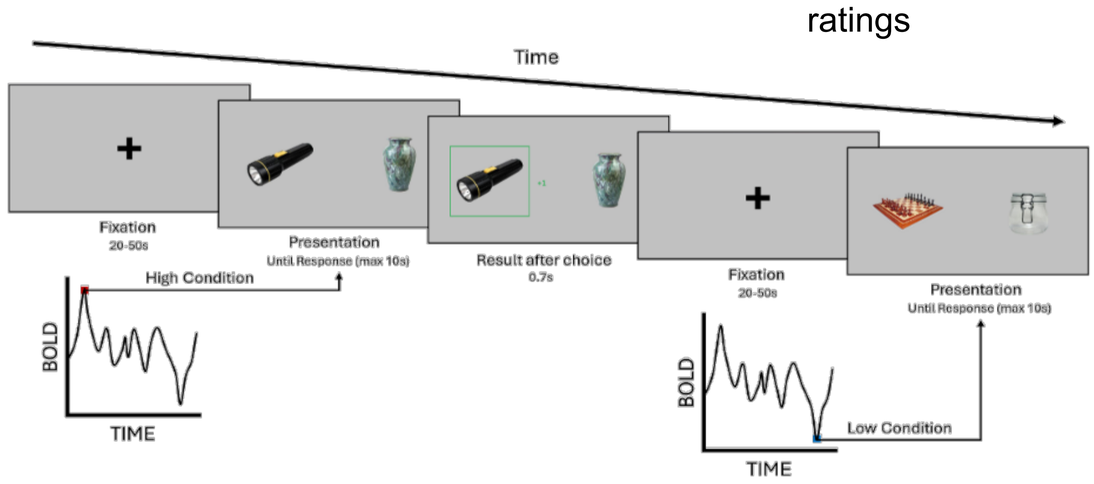
Day 2

Structural scan (T1)

Coregistration

Task-based real-time fMRI

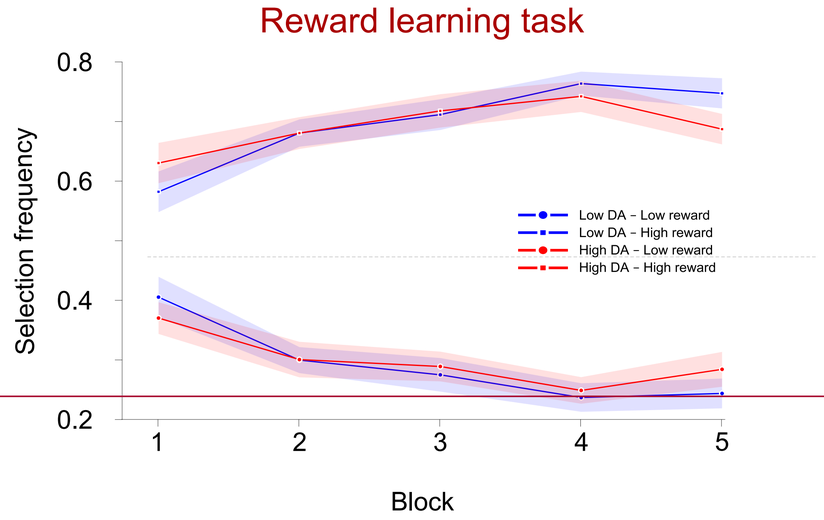
Offline preference task with confidence ratings



Behavioural results

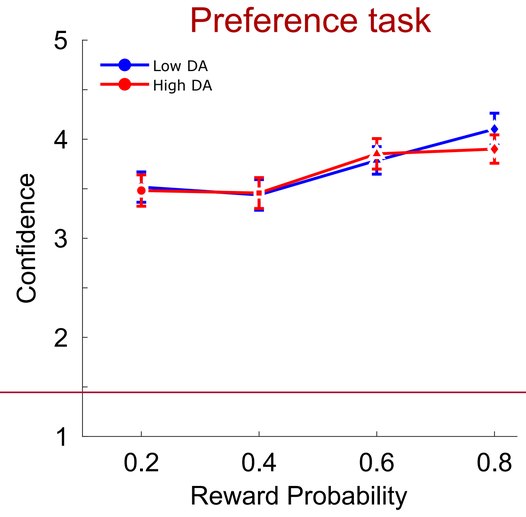
By the final blocks, low dopaminergic activity in the SN/VTA selectively enhanced discrimination between high- and low-reward objects, driven by a significant Probability × Condition × Block interaction ($F(1,552) = 5.49$, p **= .019**).

Reward learning task



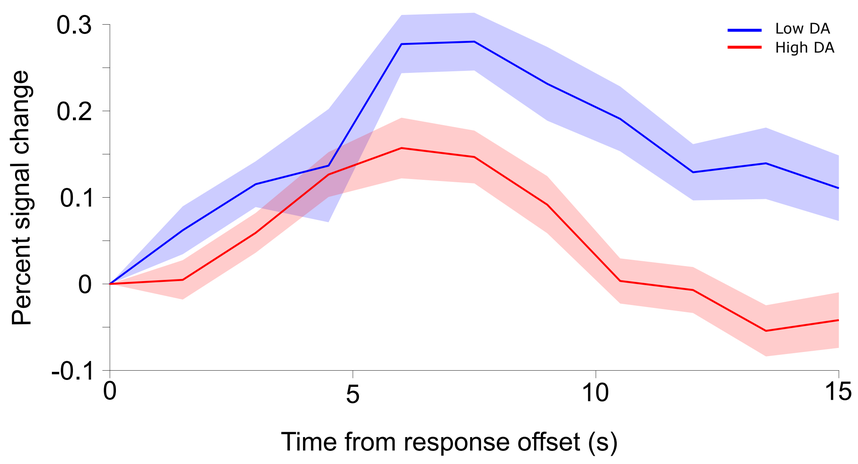
Confidence ratings increased with stimulus reward probability in both conditions, but this relationship was stronger under low midbrain dopaminergic activity, as indicated by a significant Condition × Reward Probability interaction ($F(1,1564) = 4.63$, p **= .032**).

Preference task



Post-feedback BOLD response

Following feedback, midbrain BOLD activity peaked then rapidly declined under high dopaminergic states, whereas low dopaminergic states showed a sustained elevation, consistent with a post-trough rebound in dopaminergic tone.



Outlook

- Three candidate **models** showed the best fit to behavioural data will be compared:
  - Decaying value for non-presented stimuli
  - Decaying learning rate
  - Anticorrelated updating (non-selected stimuli value tend toward the opposite feedback)
- Sustained midbrain BOLD will be correlated with behaviour to link neural dynamics to learning
- Analyses will be extended to other regions of interest, including the vmPFC

References Chew, B., Hauser, T. U., Papoutsi, M., Magerkurth, J., Dolan, R. J., & Rutledge, R. B. (2019). Endogenous fluctuations in the dopaminergic midbrain drive behavioral choice variability. Proceedings of the National Academy of Sciences of the United States of America, 116(37), 18732–18737. https://doi.org/10.1073/pnas.1900872116; Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature, 442(7106), 1042–1045. https://doi.org/10.1038/nature05051 ; Weiskopf, N.; Suckling, J.; Williams, G.; Correia, M. M.; Inkster, B.; Tait, R.; Ooi, C.; Bullmore, E. T.; Lutti, A.: Quantitative multi-parameter mapping of R1, PD(*), MT, and R2(*) at 3T: A multi-center validation. Frontiers in Neuroscience 7, 95 (2013)

Acknowledgment This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 946055). TUH is also supported by the Carl-Zeiss-Stiftung. TUH consults for limbic ltd and holds shares in the company, which is unrelated to the current project.