Project 24: Policy gradient learning in the striatum

1) Implement the striatal policy gradient algorithm described in Chapter 11 (Section 2.1) of *Computational Foundations of Cognitive Neuroscience*. Simulate correlations between dSPNs and iSPNs before and after action selection, comparing to data reported in Lindsey et al. (2025).

2) Show the responses of these neurons following low vs. high probability actions, comparing to data reported in Markowitz et al. (2018).

3) Discuss the possible neural correlate of action prediction errors (Greenstreet et al., 2025), and how they might interact with reward prediction errors, as implied by Eq. 16 in Section 2.1.

References:

Greenstreet, F., Vergara, H. M., Johansson, Y., Pati, S., Schwarz, L., Lenzi, S. C., ... & Stephenson-Jones, M. (2025). Dopaminergic action prediction errors serve as a value-free teaching signal. *Nature*, *643*, 1333–1342.

Lindsey, J. W., Markowitz, J., Gillis, W. F., Datta, S. R., & Litwin-Kumar, A. (2025). Dynamics of striatal action selection and reinforcement learning. *eLife, 13*, RP101747.

Markowitz, J. E., Gillis, W. F., Beron, C. C., Neufeld, S. Q., Robertson, K., Bhagat, N. D., ... & Datta, S. R. (2018). The striatum organizes 3D behavior via moment-to-moment action selection. *Cell, 174*, 44-58.