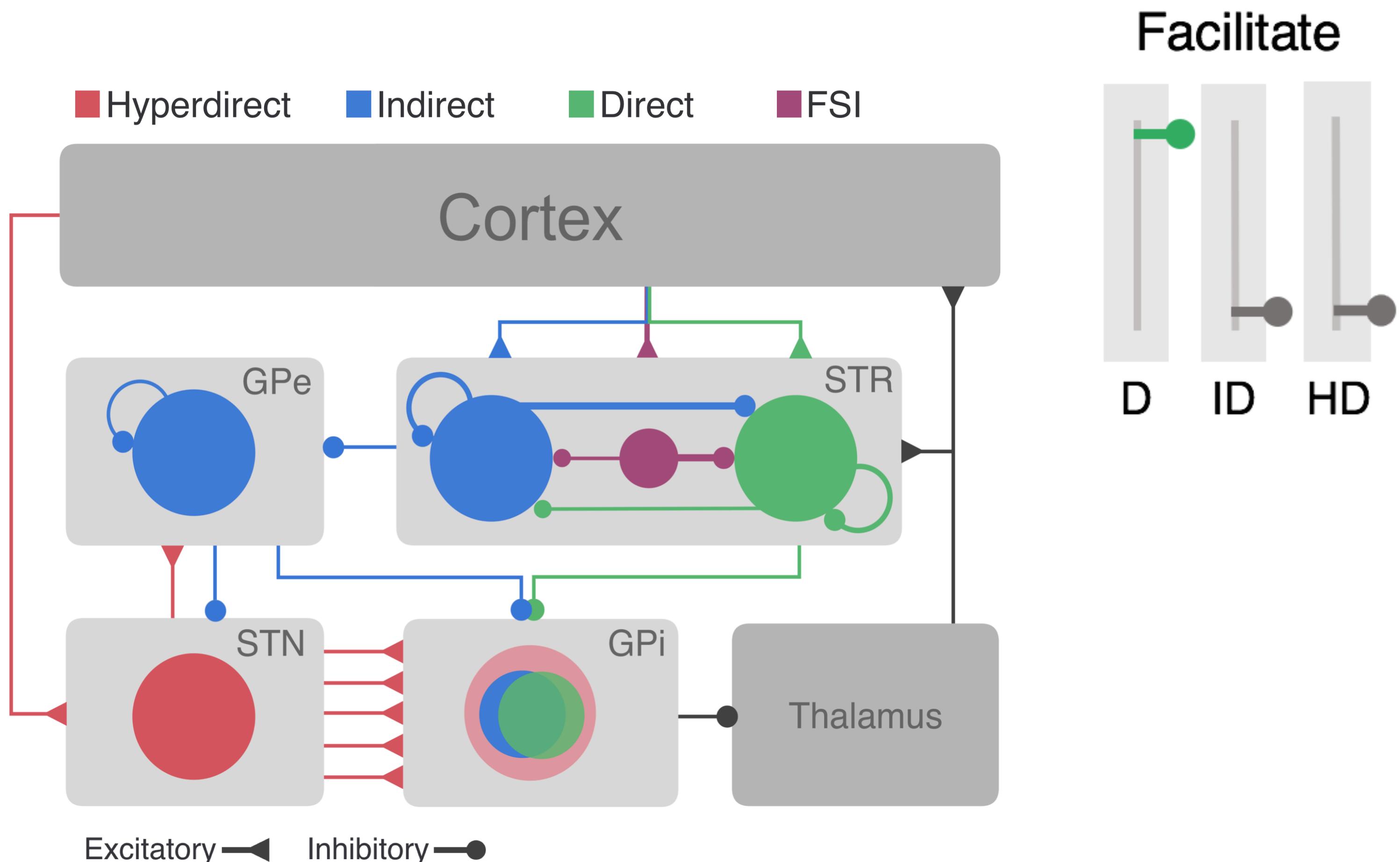


# **How does the brain learn what is the best decision?**

# Readings for today

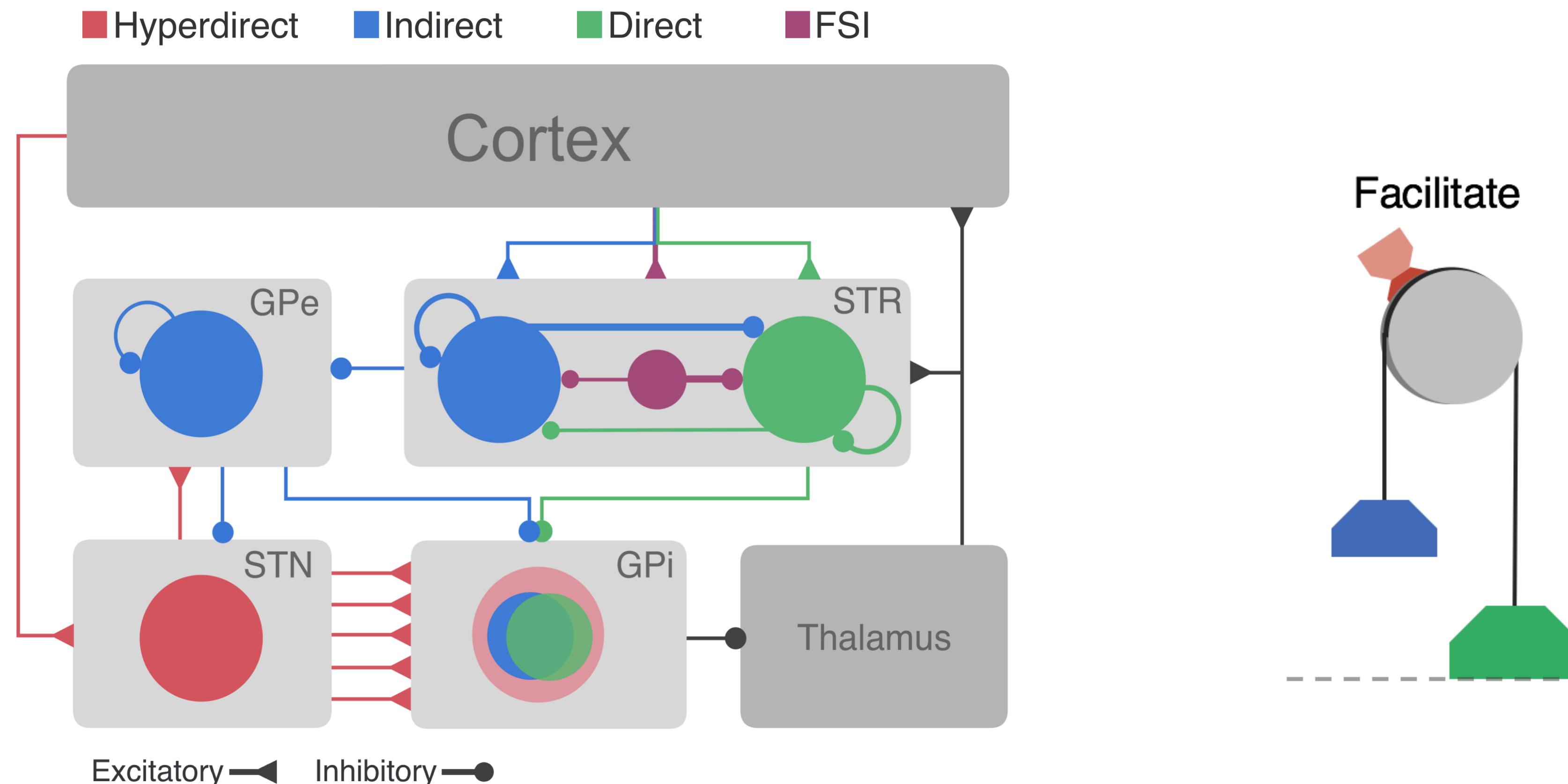
- Dunovan, K., & Verstynen, T. (2016). Believer-skeptic meets actor-critic: rethinking the role of basal ganglia pathways during decision-making and reinforcement learning. *Frontiers in neuroscience*, 10, 106.
- Dunovan, K., Vich, C., Clapp, M., Verstynen, T., & Rubin, J. (2019). Reward-driven changes in striatal pathway competition shape evidence evaluation in decision-making. *PLoS computational biology*, 15(5), e1006998.

# The independent levers model



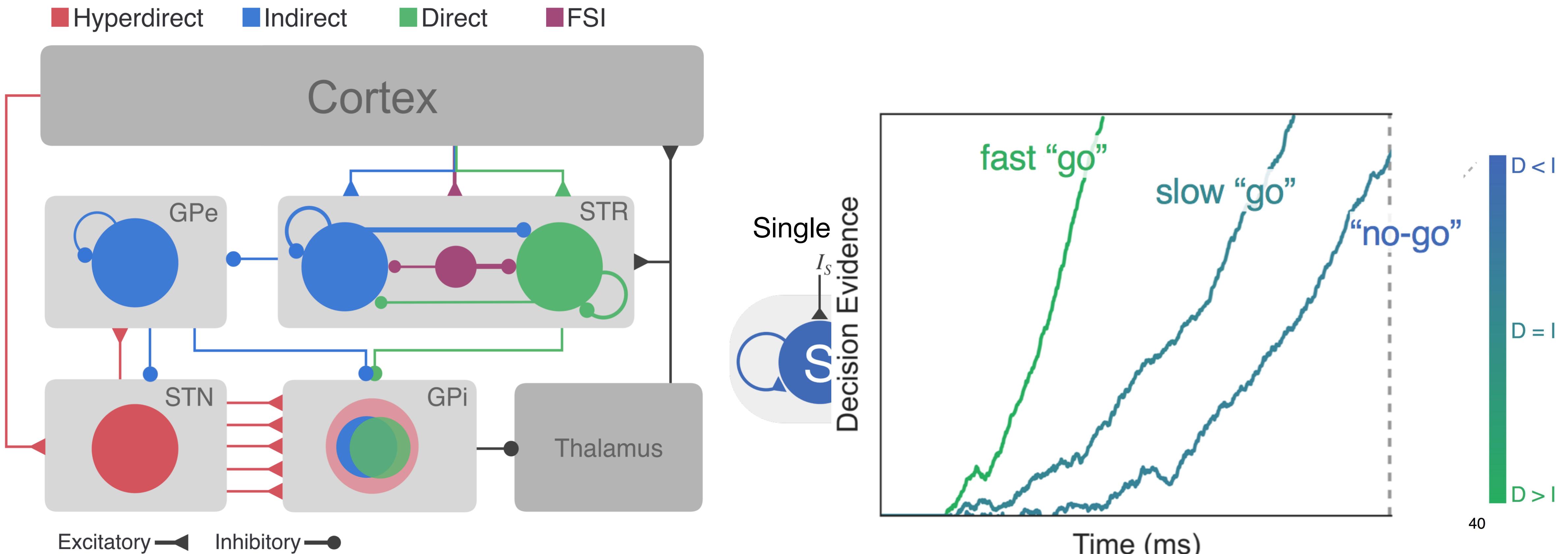
Str: striatum; GPe: external globus pallidus;  
GPi: internal globus pallidus; STN: Subthalamic nucleus

# The pulley model



Str: striatum; GPe: external globus pallidus;  
GPi: internal globus pallidus; STN: Subthalamic nucleus

# The Believer-Skeptic model



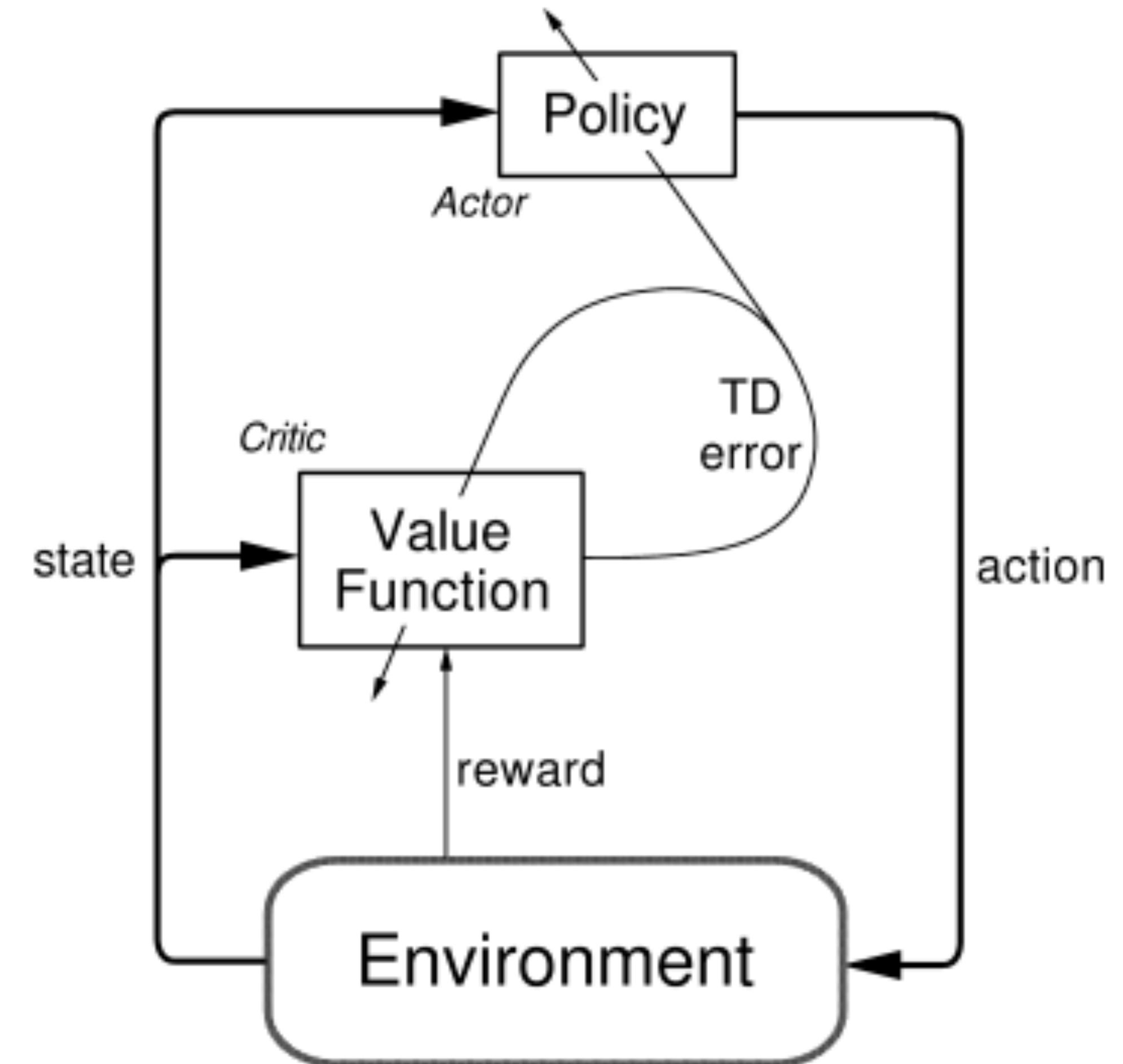
Str: striatum; GPe: external globus pallidus;  
GPi: internal globus pallidus; STN: Subthalamic nucleus

# Actor-Critic Learning

Two interacting processes (networks):

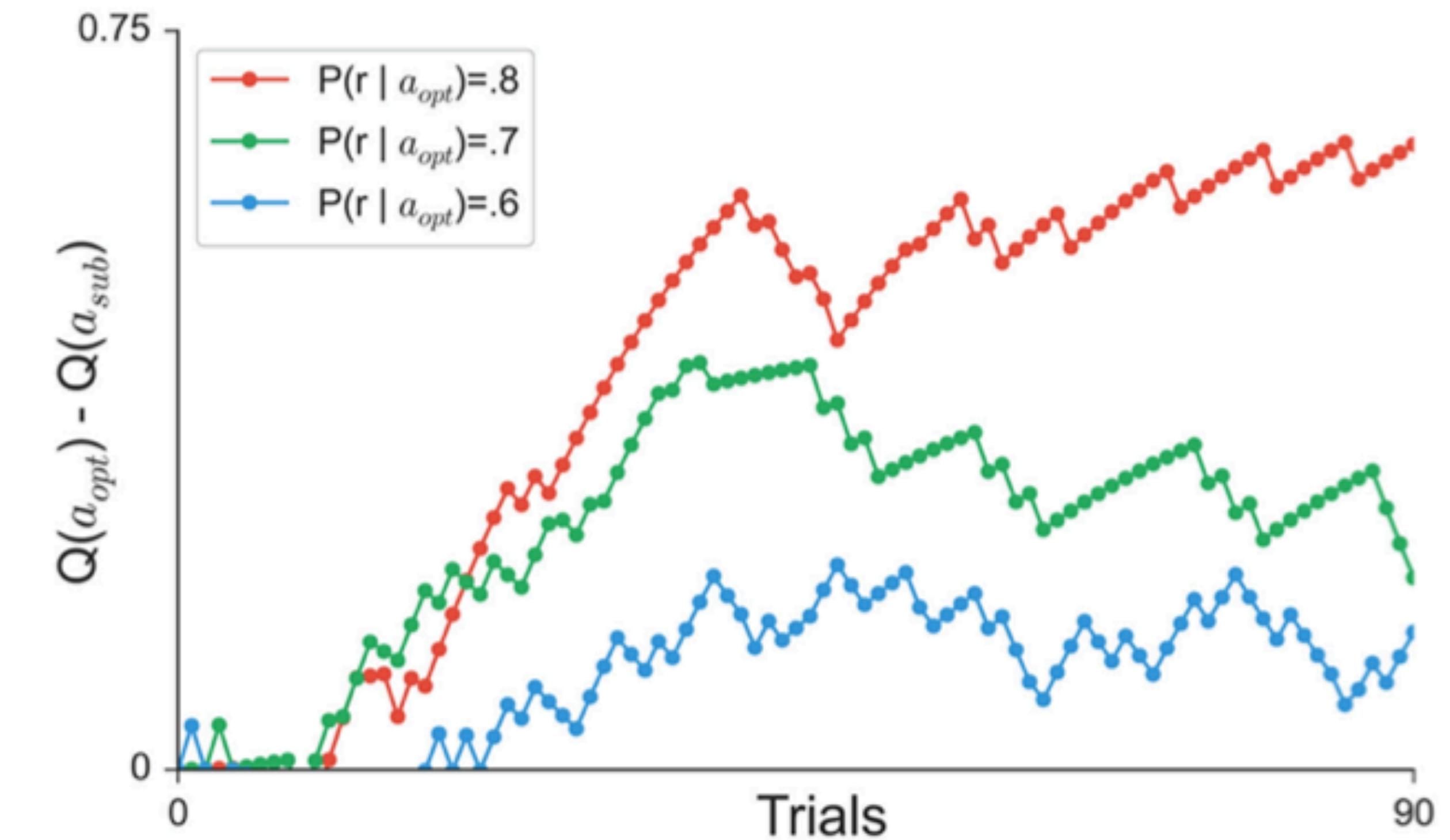
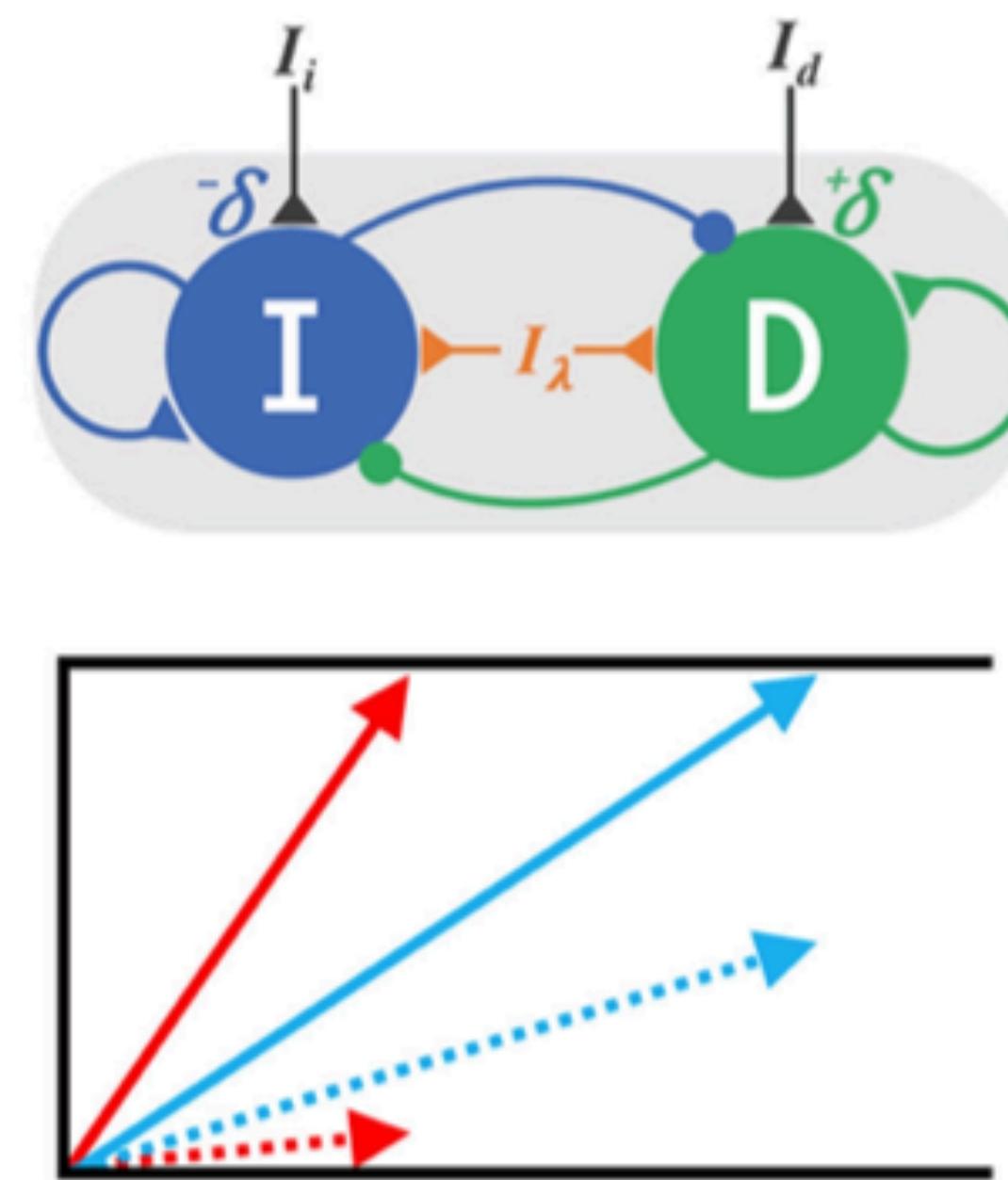
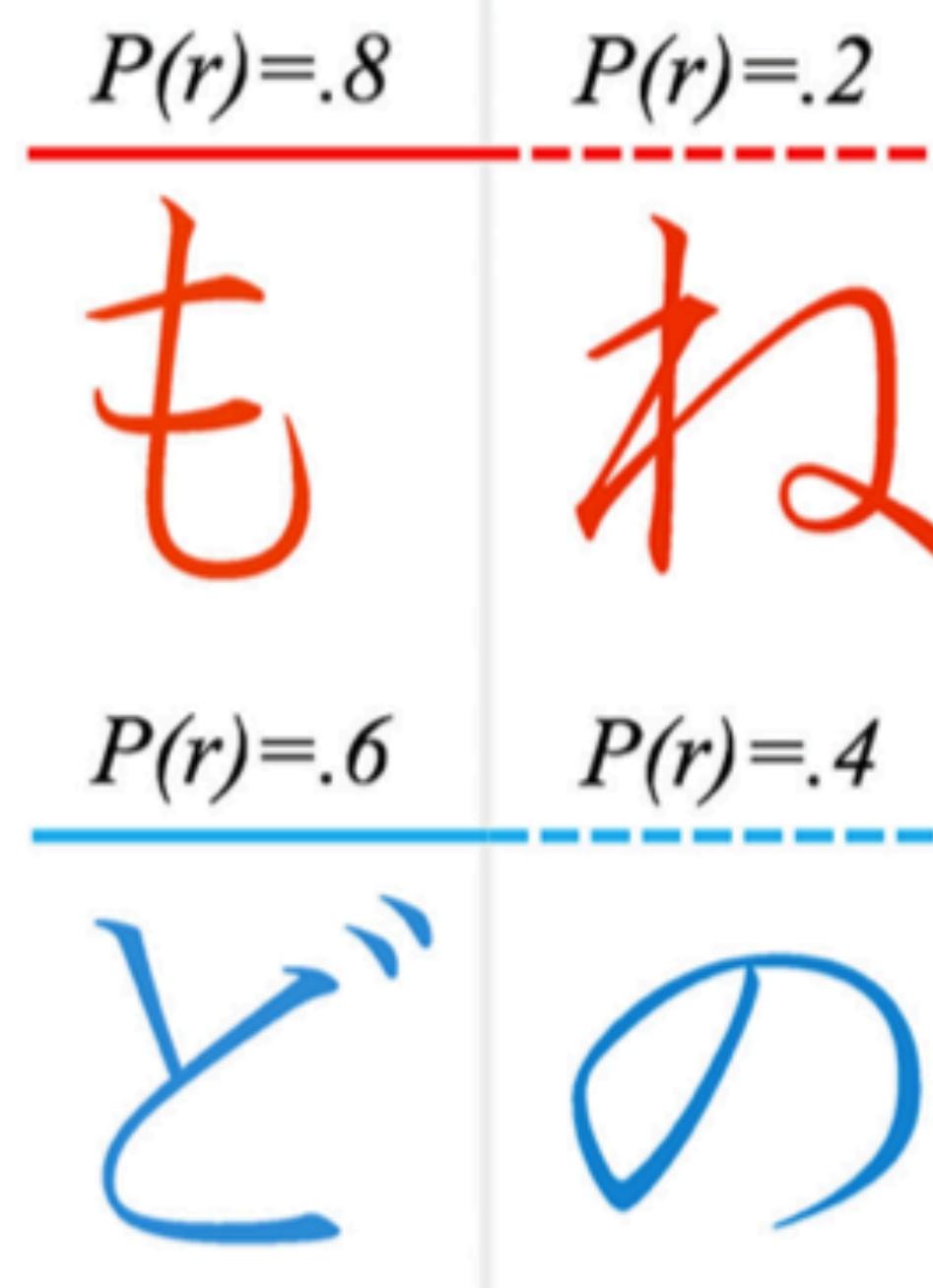
- The actor: decides which action should be taken.
- The critic: informs the actor how good the action was and to adjust adjust.

The learning of the actor is based on policy gradient approach (i.e., RL)

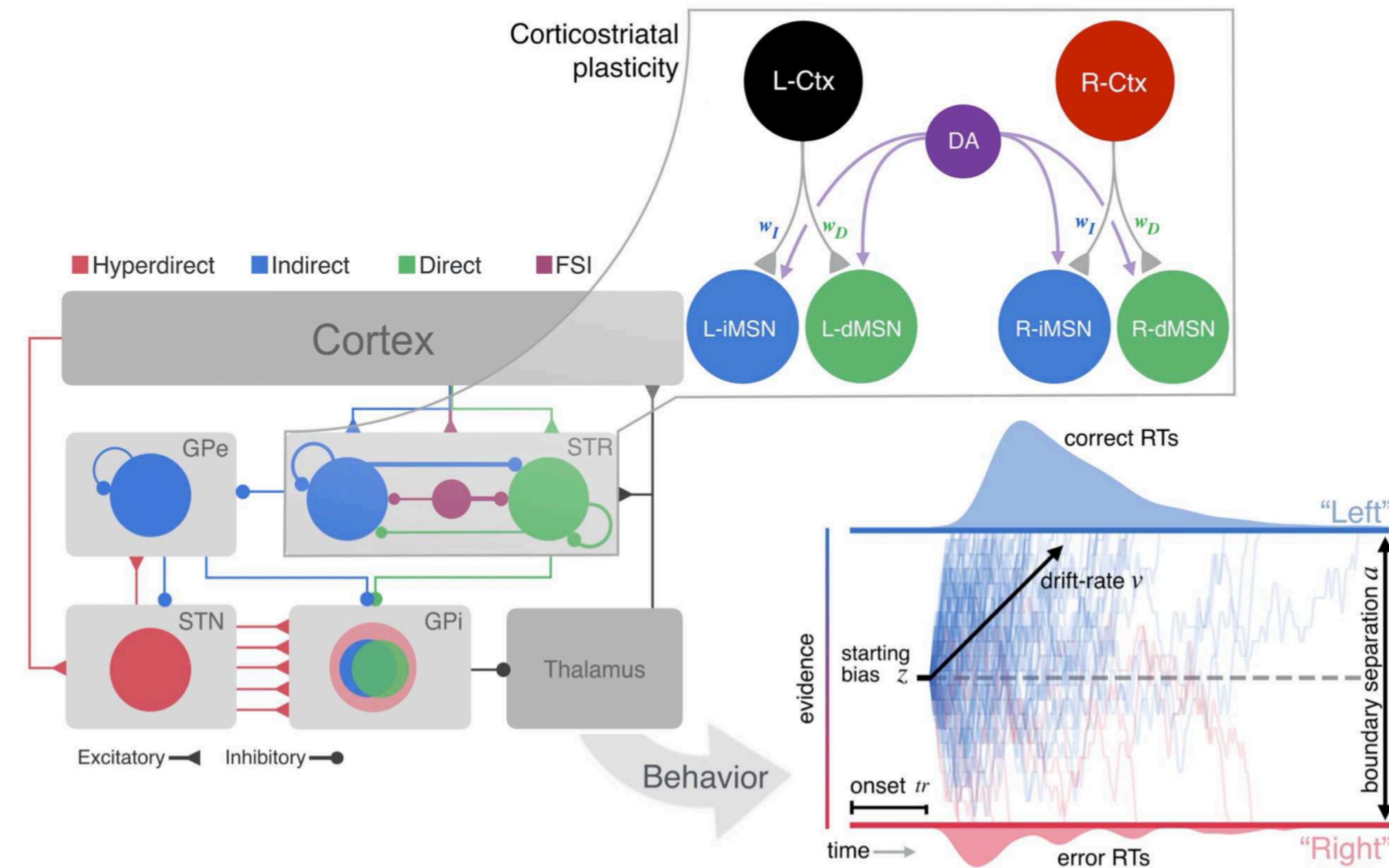


# Just add (reinforcement) learning

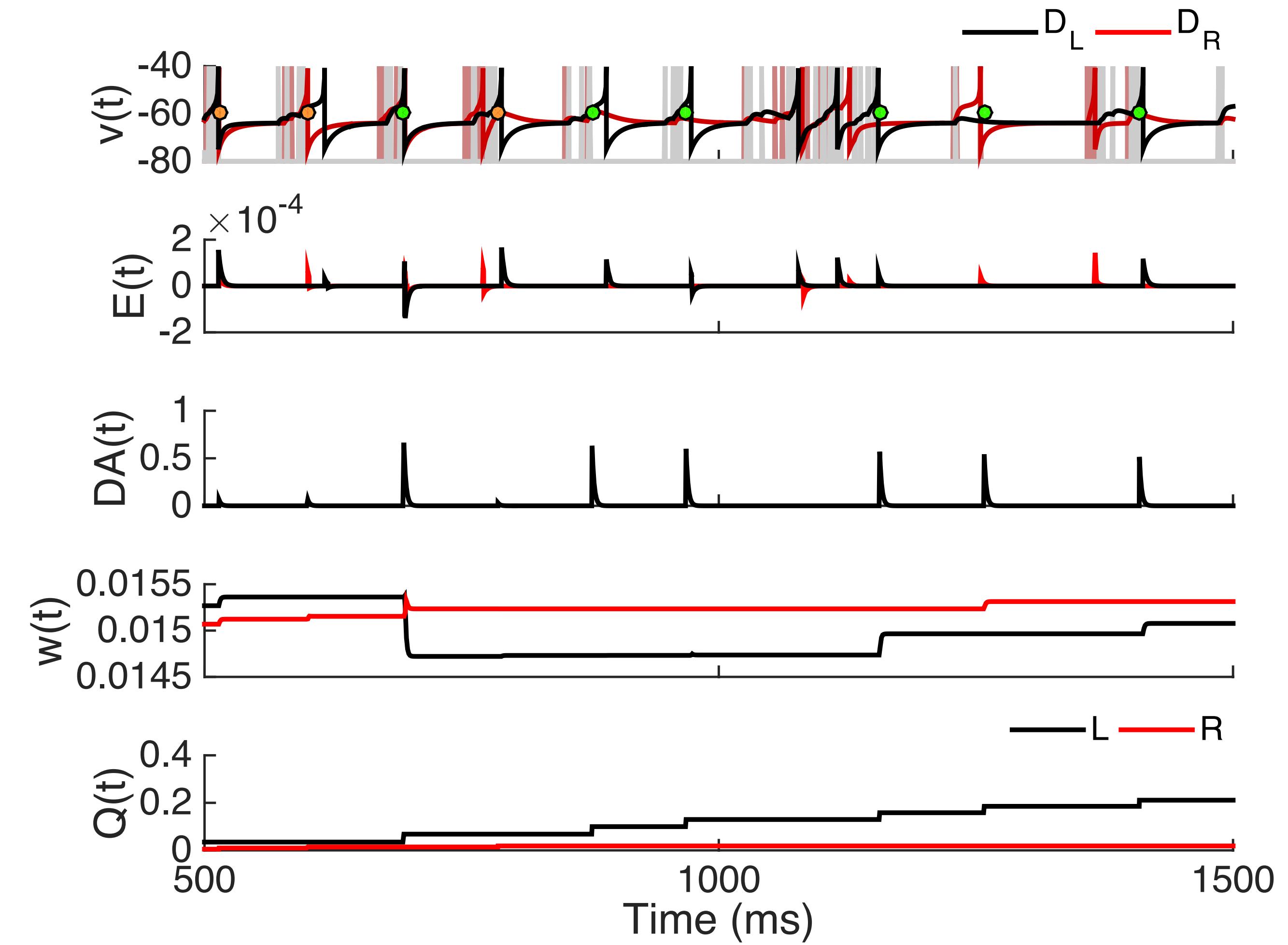
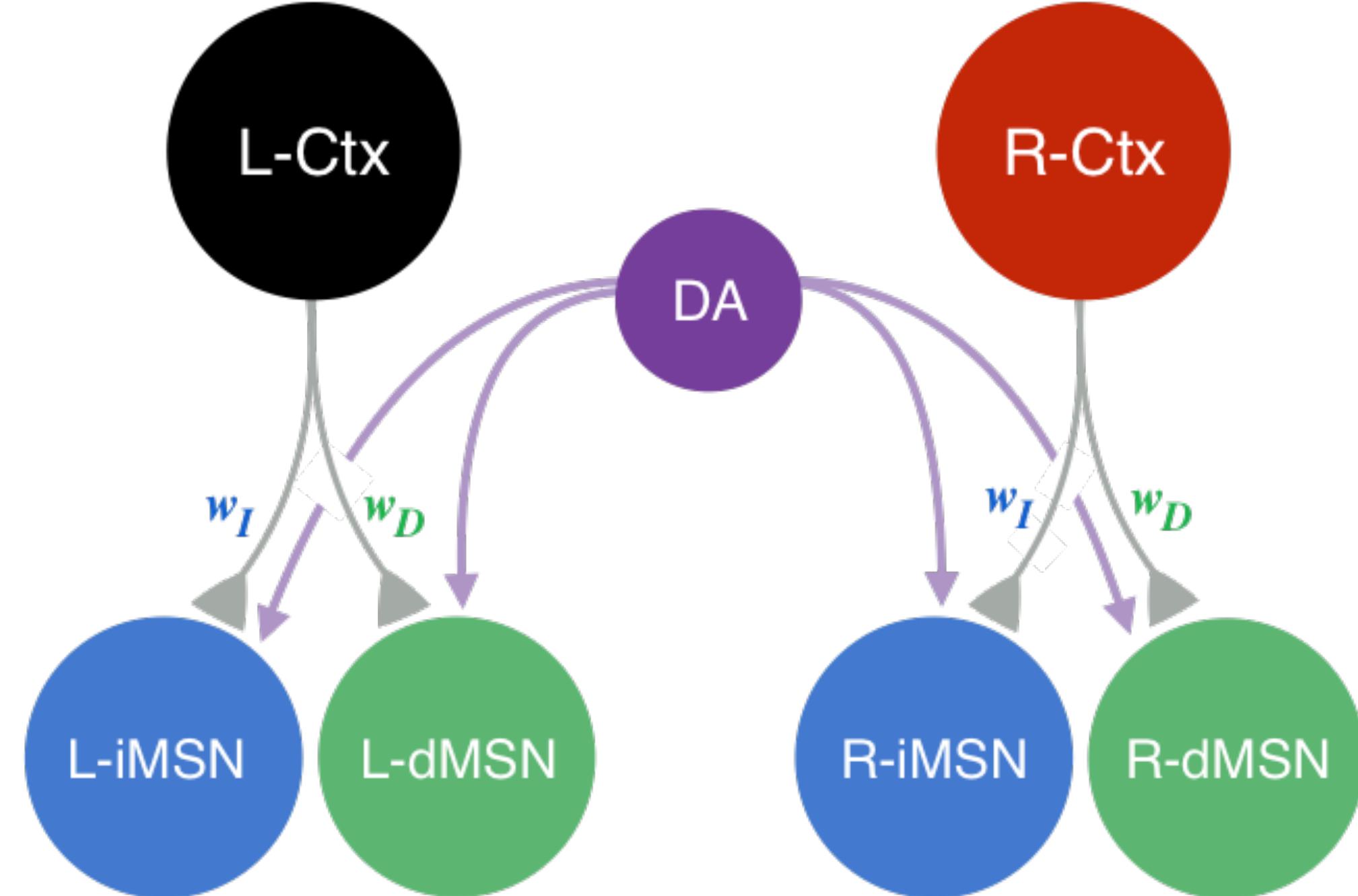
Phasic Dopamine ( $\pm \delta$ ) and RL



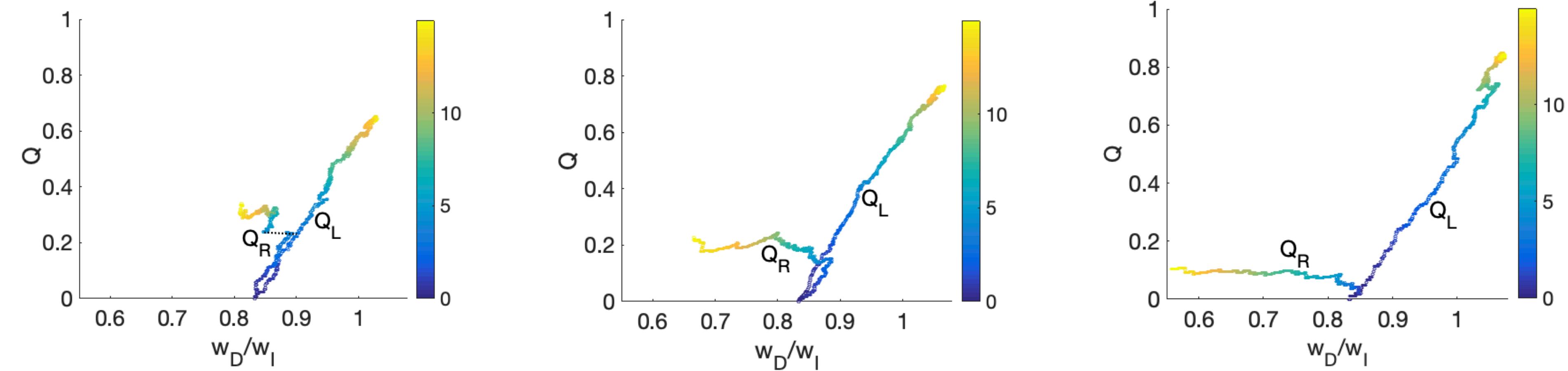
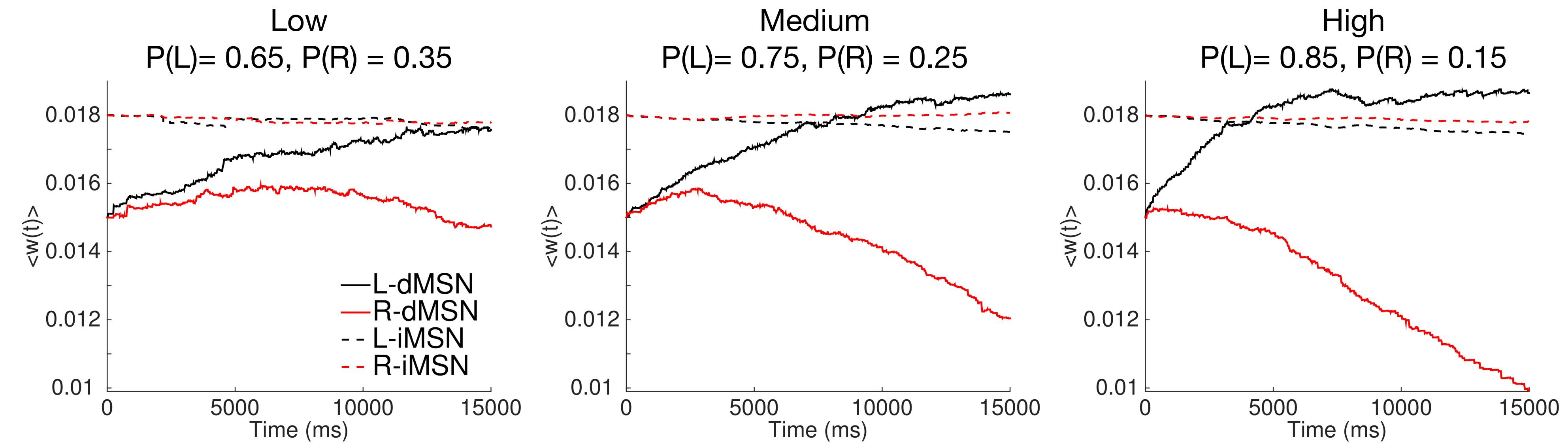
# Upwards mapping



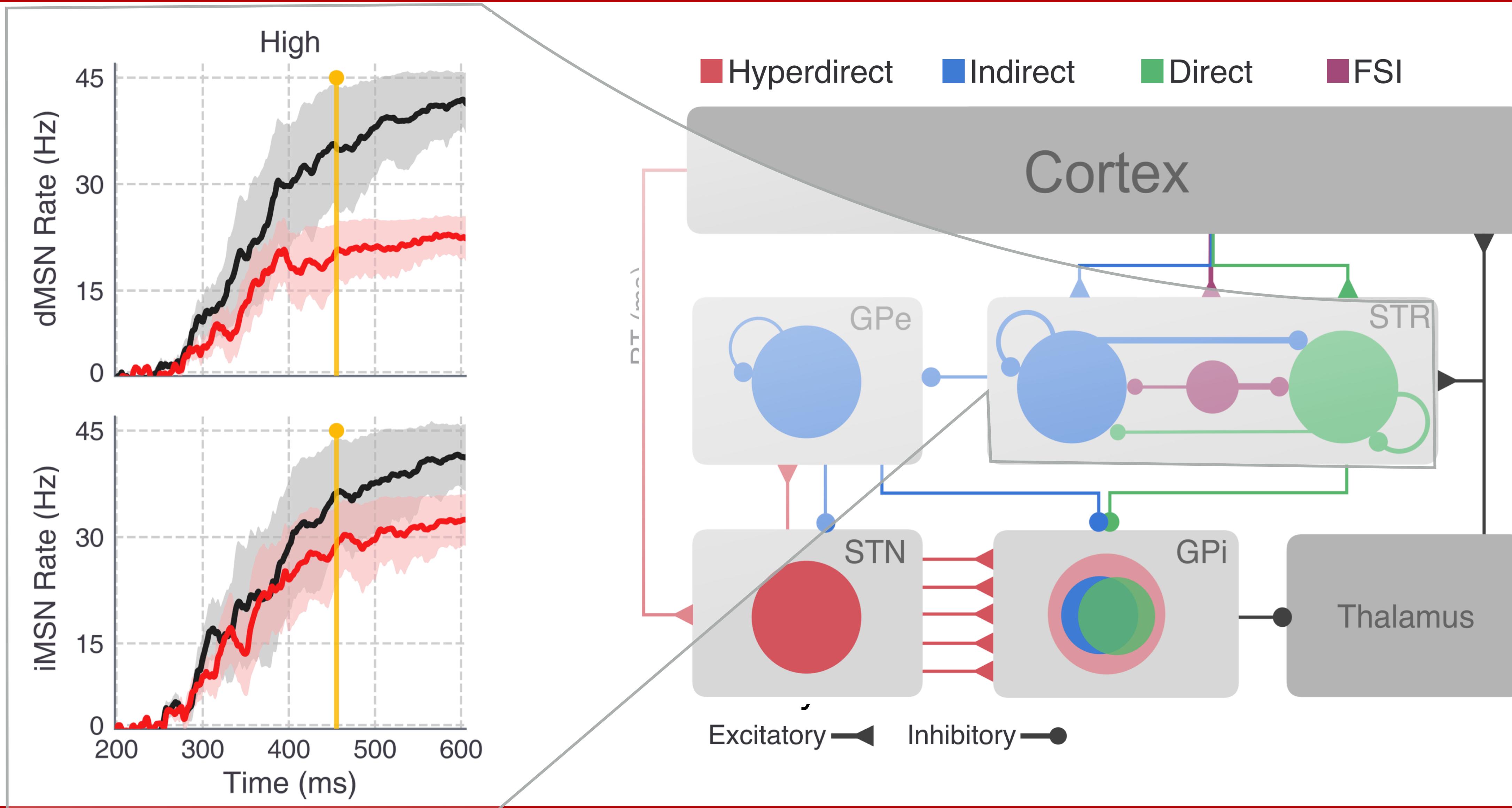
# Spike-timing dependent plasticity



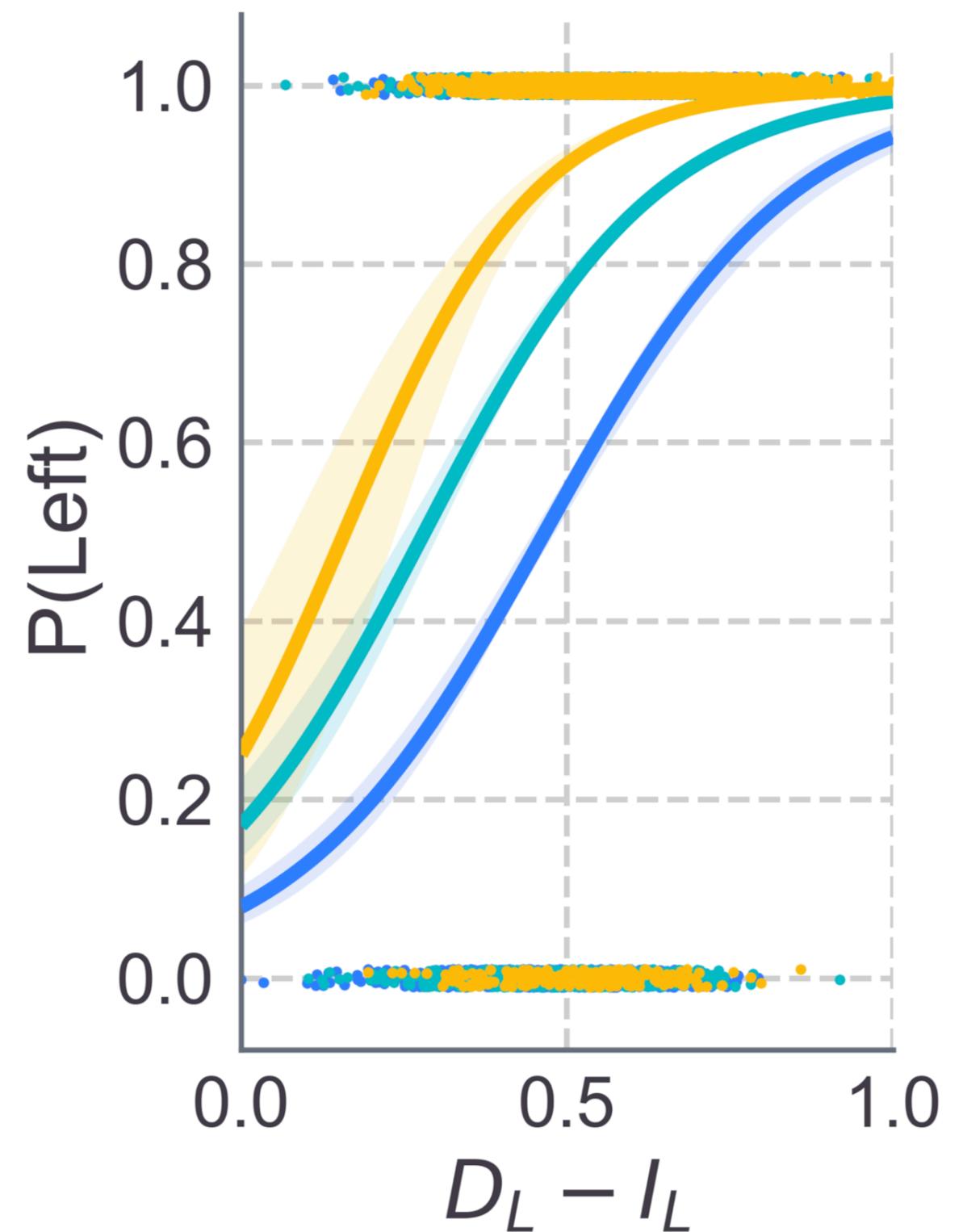
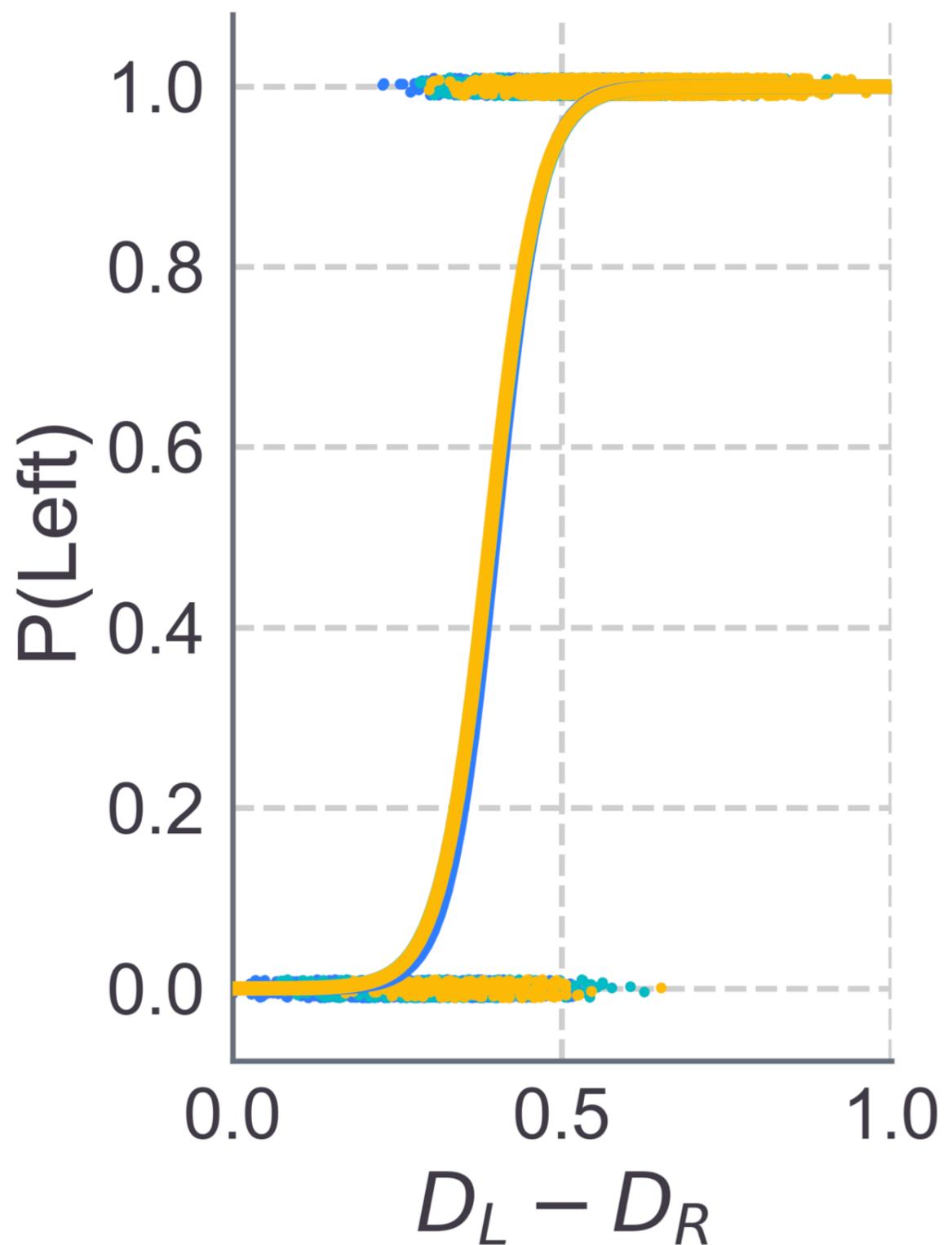
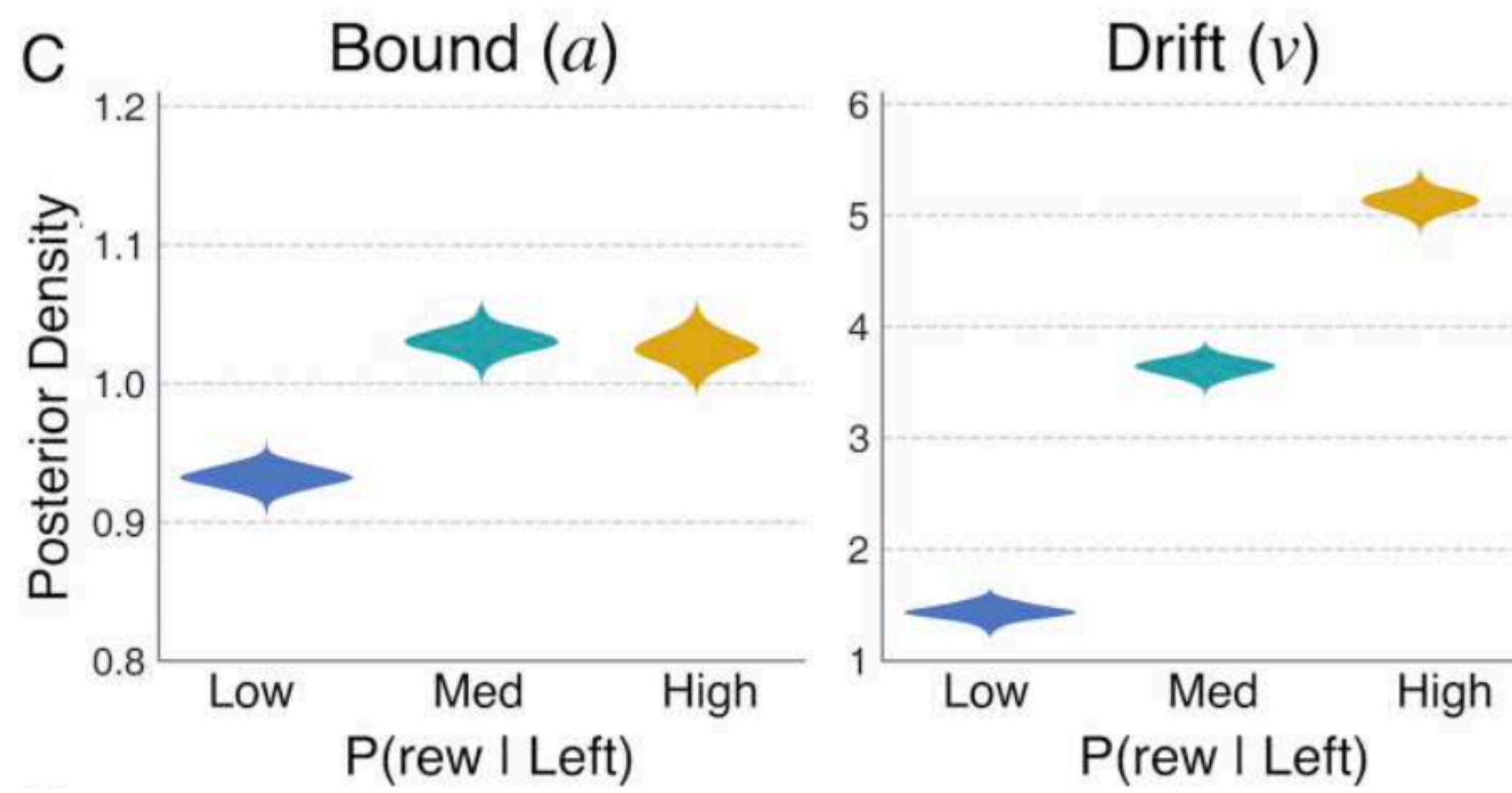
# Temporal difference update



# Dopamine learning changes B-S balance

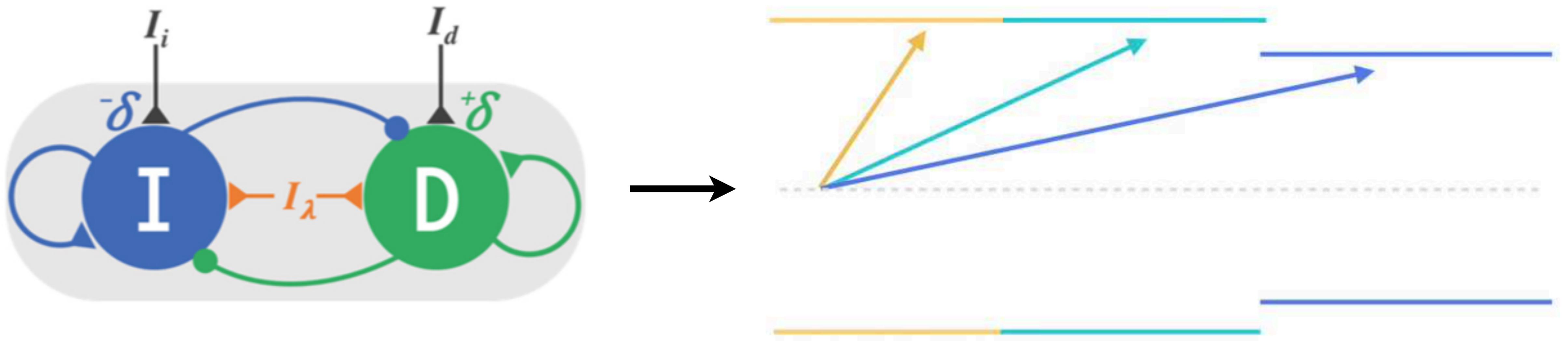


# The $k$ -armed bandit problem



Increased separation of direct & indirect pathways (driven largely by changes in direct pathway gain) track with changes in drift rate ( $v$ ) and action choices.

# Believer-Skeptic meets Actor-Critic



# Food for thought

## Group discussions

- Given that the decision policy *\*and\** learning mechanisms are integrated into the same neural circuits, is it reasonable to think of these as 2 separate processes or a single, unified process?