Thalamocortical contribution to credit assignment in neural systems

Mien Brabeeba Wang

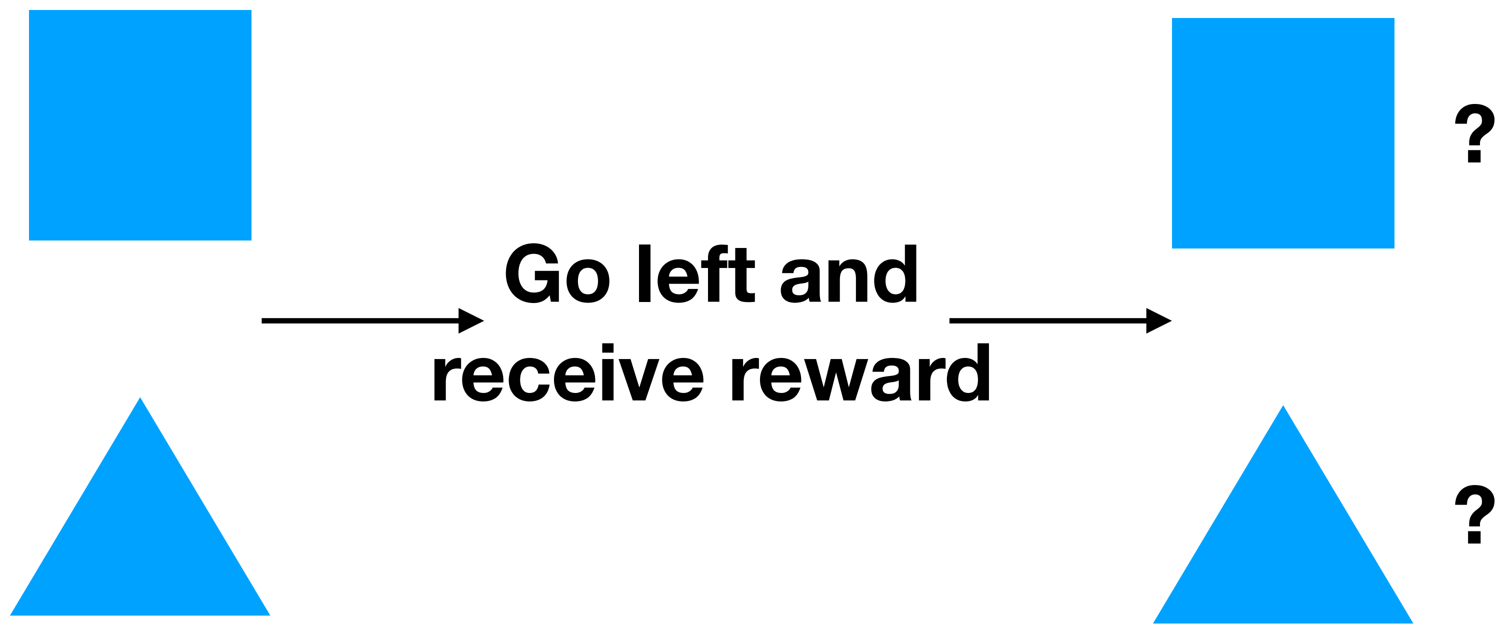
3/26/2021

Learning in brain

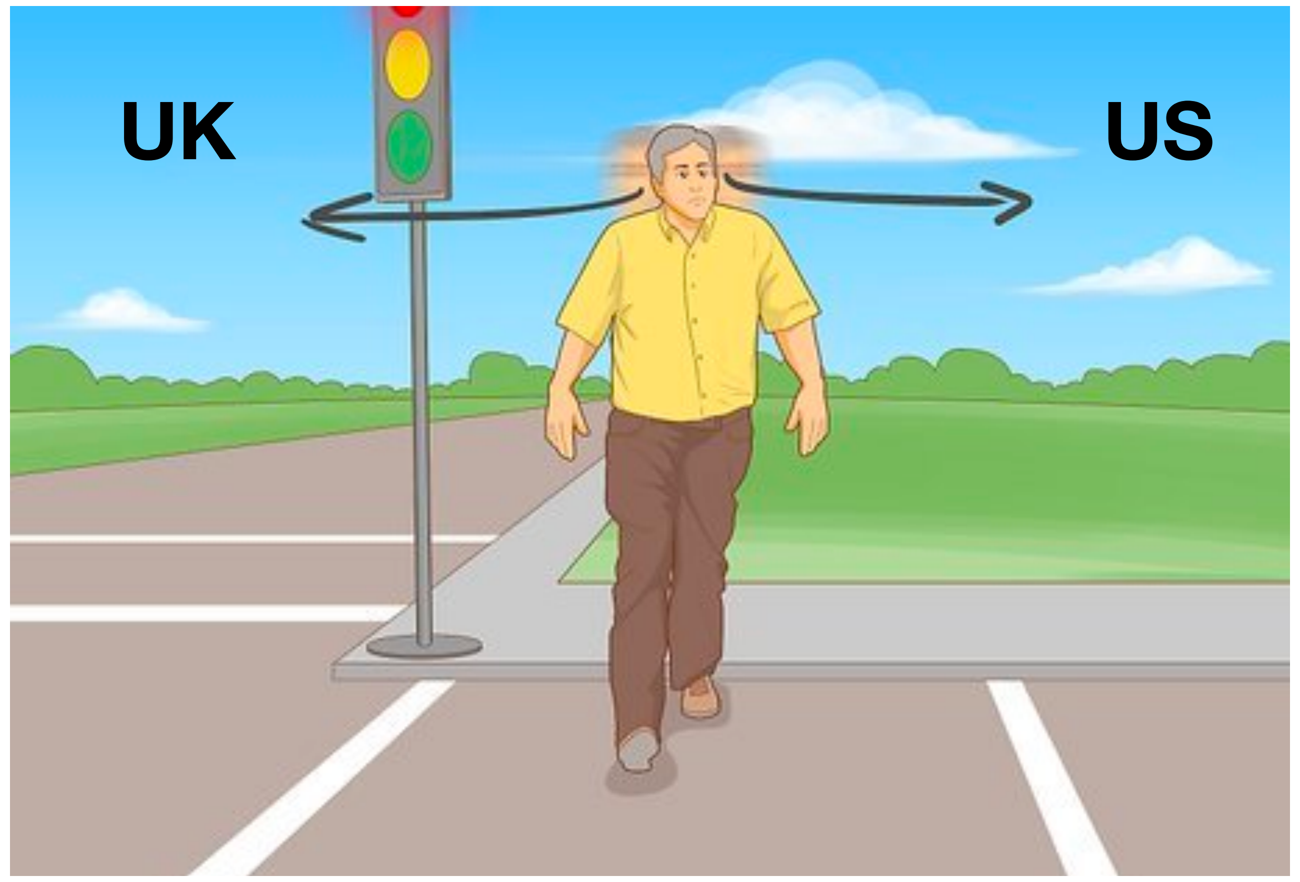
- Dopamine represents reward prediction error.
- Want to reinforce synapses that lead to positive reward prediction error and weaken synapses that lead to a negative ones.
- How can one do it when multiple cues in multiple contexts and multiple actions taken before rewards arrive?

Credit assignment

Structural



Contextual



Temporal

Backpropagation

- Great empirical results and match internal representation in brains
- Weight transport: $\delta W_i \propto \frac{\partial E}{\partial W_i} = e_i f(a_{i-1})^\top,\ e_i = W_{i+1}^\top e_{i+1} \circ f'(a_i)$
- Separation of error v.s. activity

Continual learning

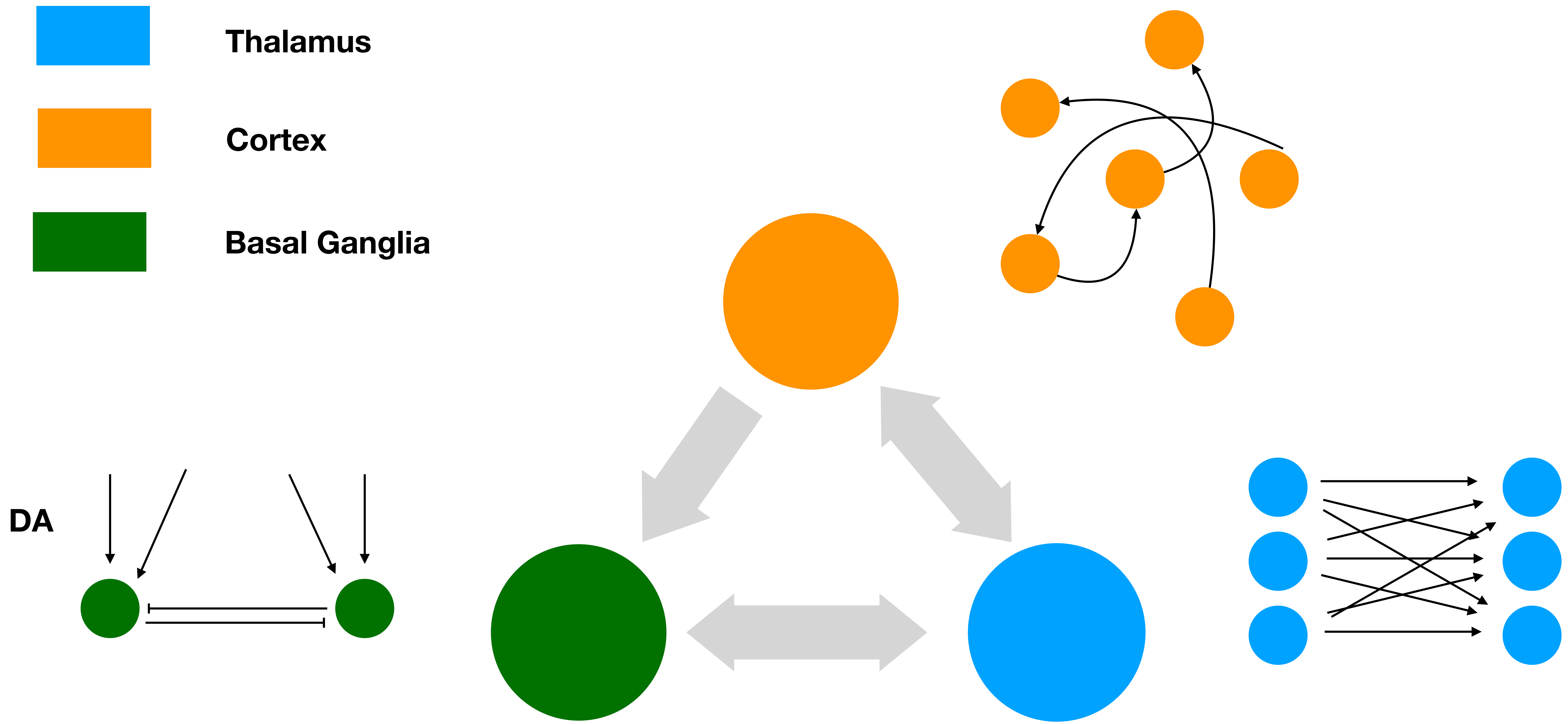
- Forward transfer and backward transfer
- Complementary memory system and replay
- Regularization to protect past task
- Dynamic architecture

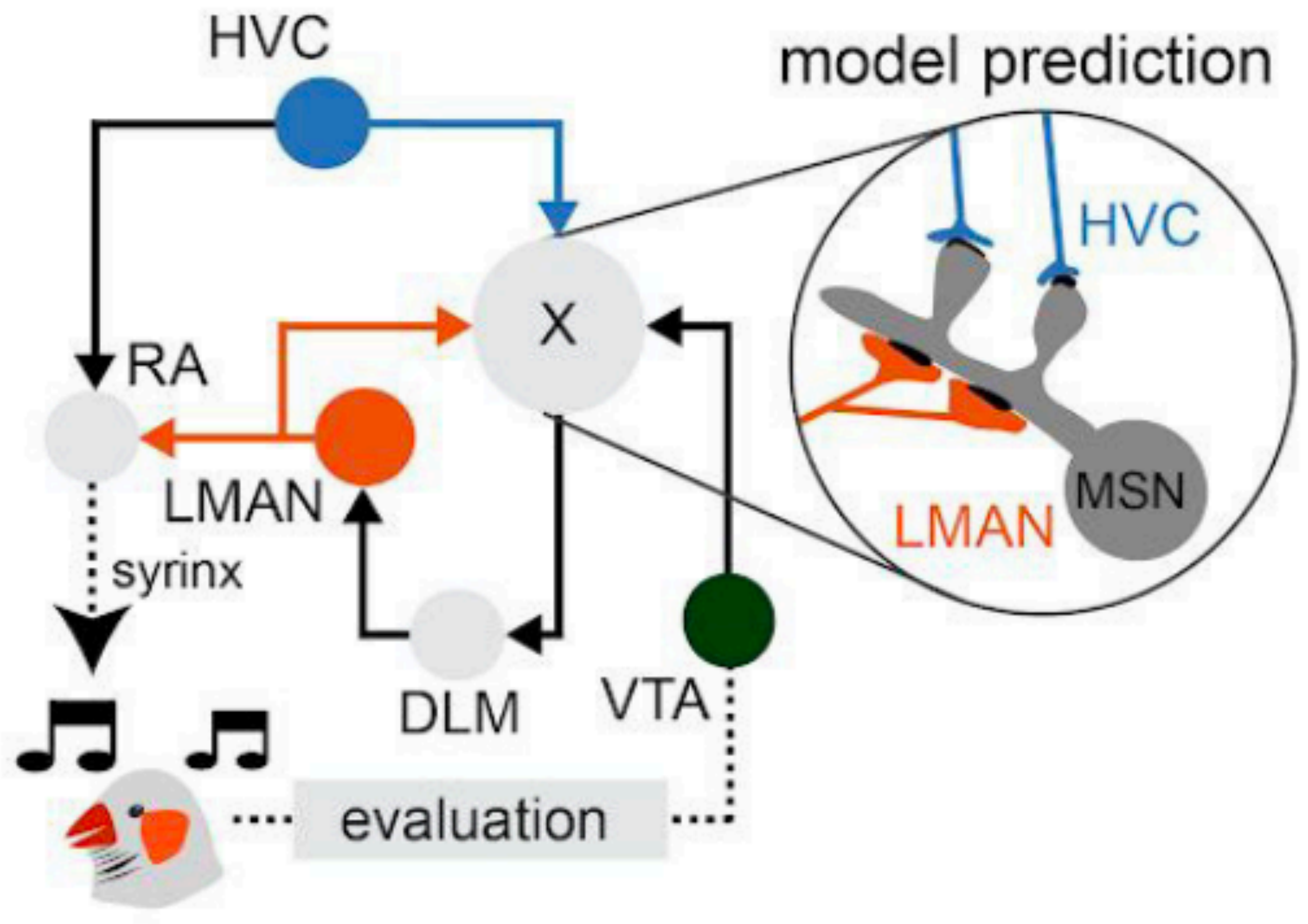
Temporal credit assignment

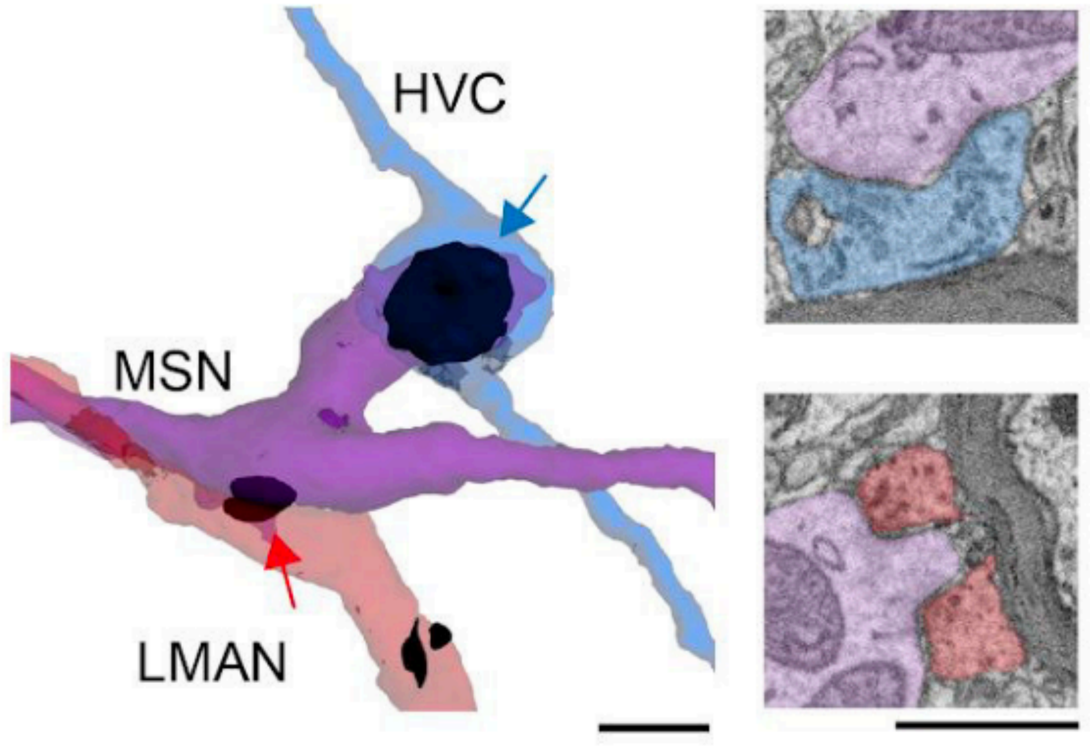
- TD learning: $e_t = r_t + V(s_{t+1}) - V(s_t),\ V(s_t) = \mathbb{E}\left[\sum_{i=0}^{\infty} r_{t+i} \middle| s_t\right].$
- LSTM to make non-Markov environment Markov
- Backpropagation through time

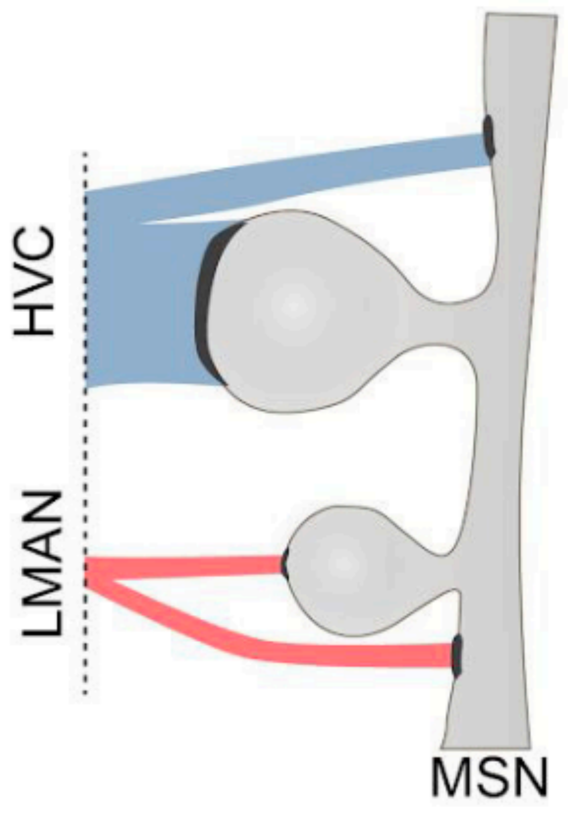
Backpropagation => General

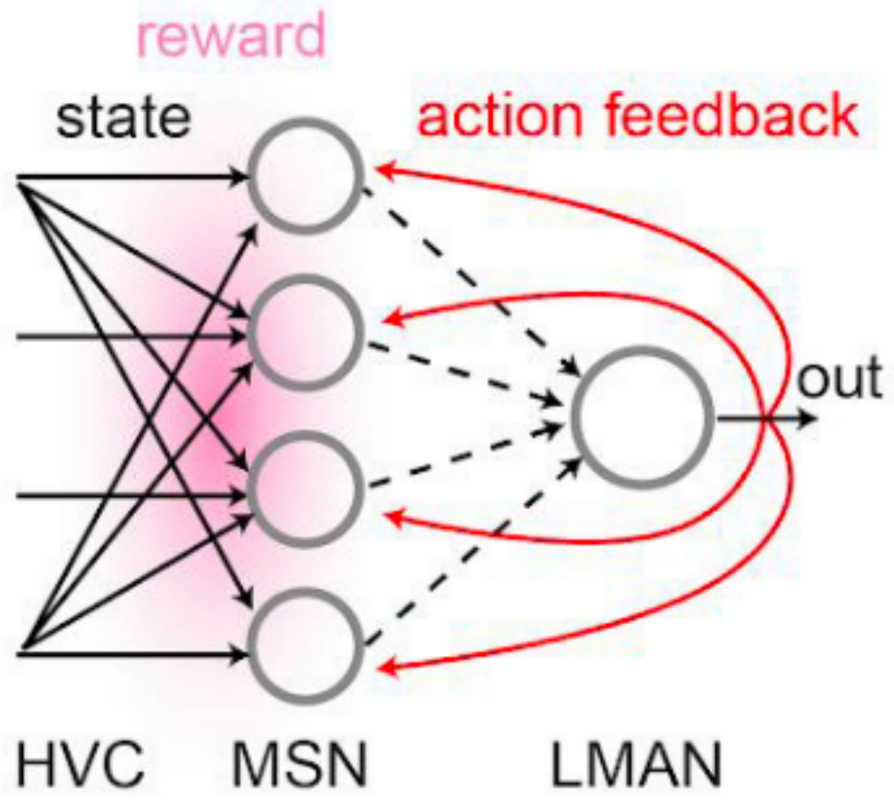
Brains => Specialized hardwares

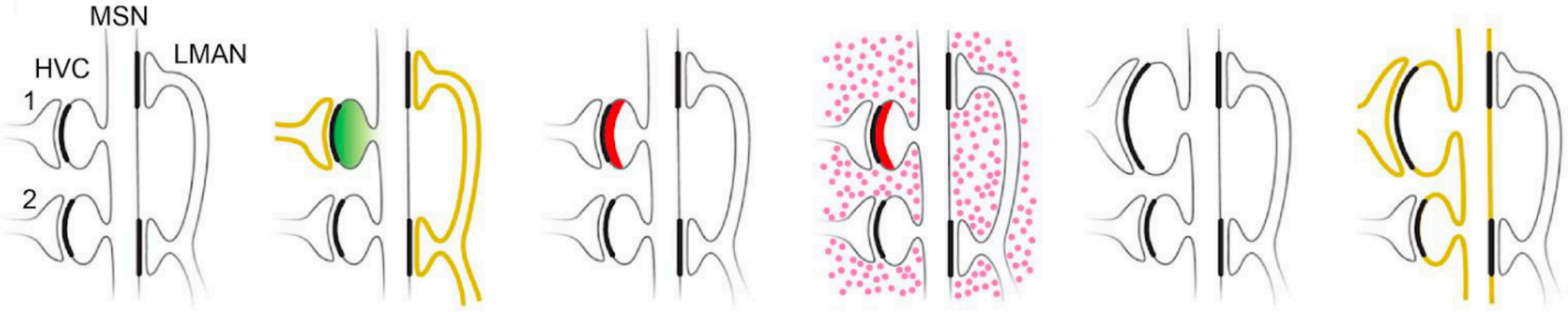

Thalamus
Cortex
Basal Ganglia
DA


HVC
model prediction
HVC
X
RA
LMAN
syrinx
LMAN
MSN
DLM
VTA
evaluation


HVC
MSN
LMAN


HVC
LMAN
MSN


reward
state
action feedback
out
HVC
MSN
LMAN

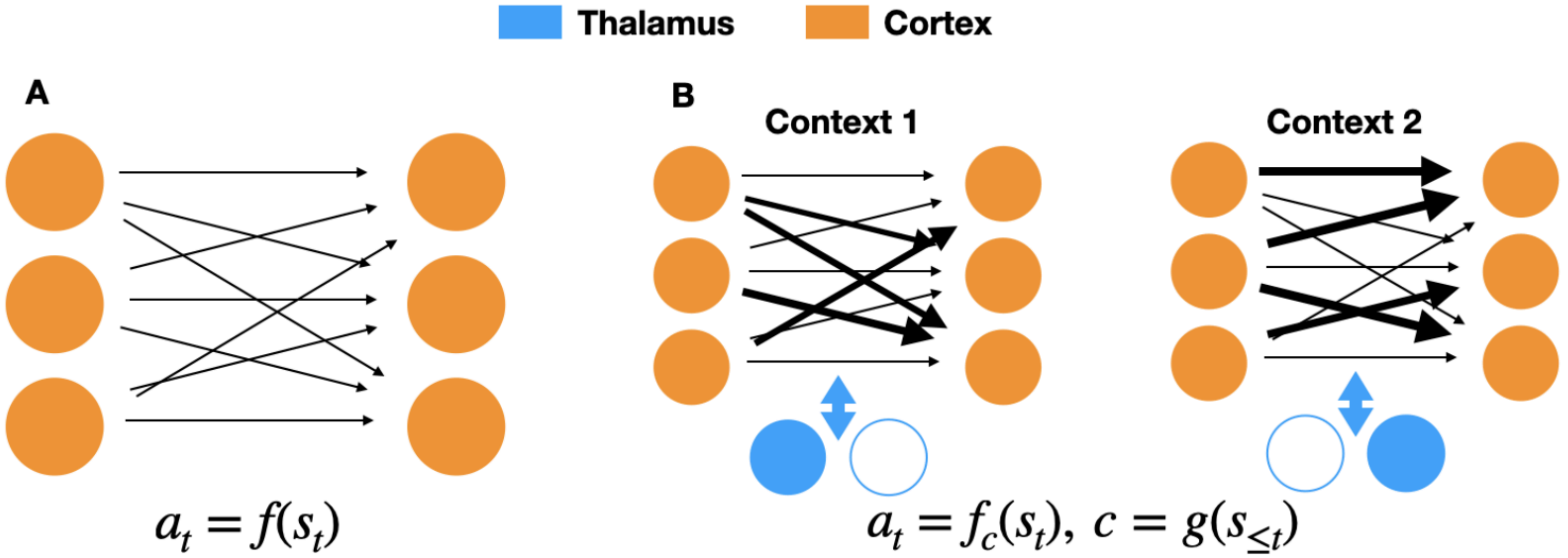

MSN
HVC
LMAN
1
2

Learning in Cortex?

- There are high concentration of dopamine active transporters (DAT) in striatum to make sure the circuit can do precise credit assignment
- Since the cortex has low DAT concentration, dopamine behaves on a much slower timescale (30 minutes-1 hour).
- So DA is unlikely to support RL in a similar manner as in BG
- And there are lots of evidences that cortex is doing more Hebbian learning types of plasticity (unsupervised learning)
- However, one can observe that cortex develop task-relevant representation

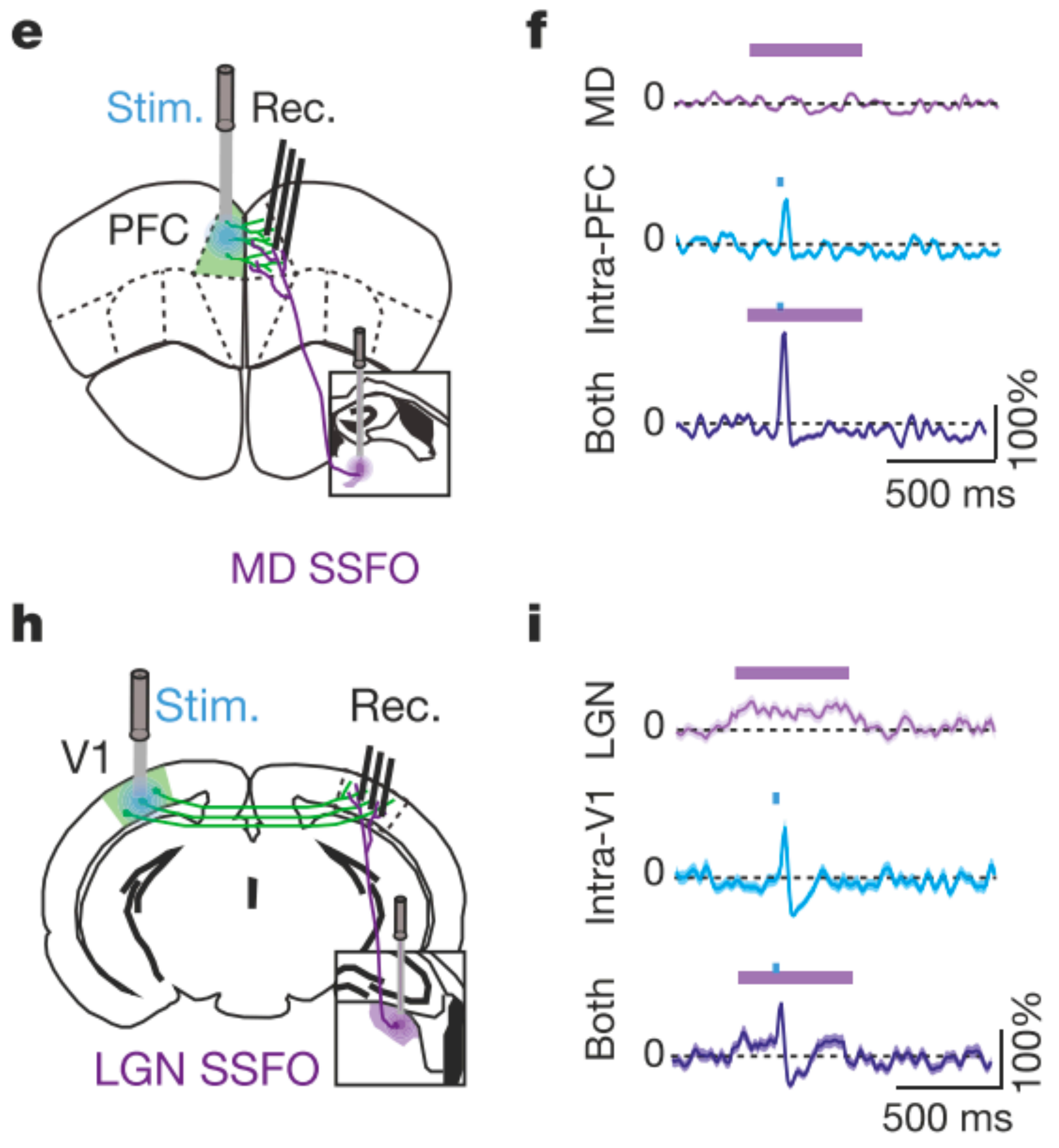
Learning in Cortex?

- Consolidate learning signal from BG? Ans: duplicate and slow
- Cortex => slow plasticity but fast at generalization and flexible behaviors
- Protein synthesis is slow (5-10 minutes), how does one learn on the fly?
- Meta learning: learning to learn


Thalamus
Cortex
A
B
Context 1
Context 2
$a_t = f(s_t)$
$a_t = f_c(s_t),\ c = g(s_{\le t})$

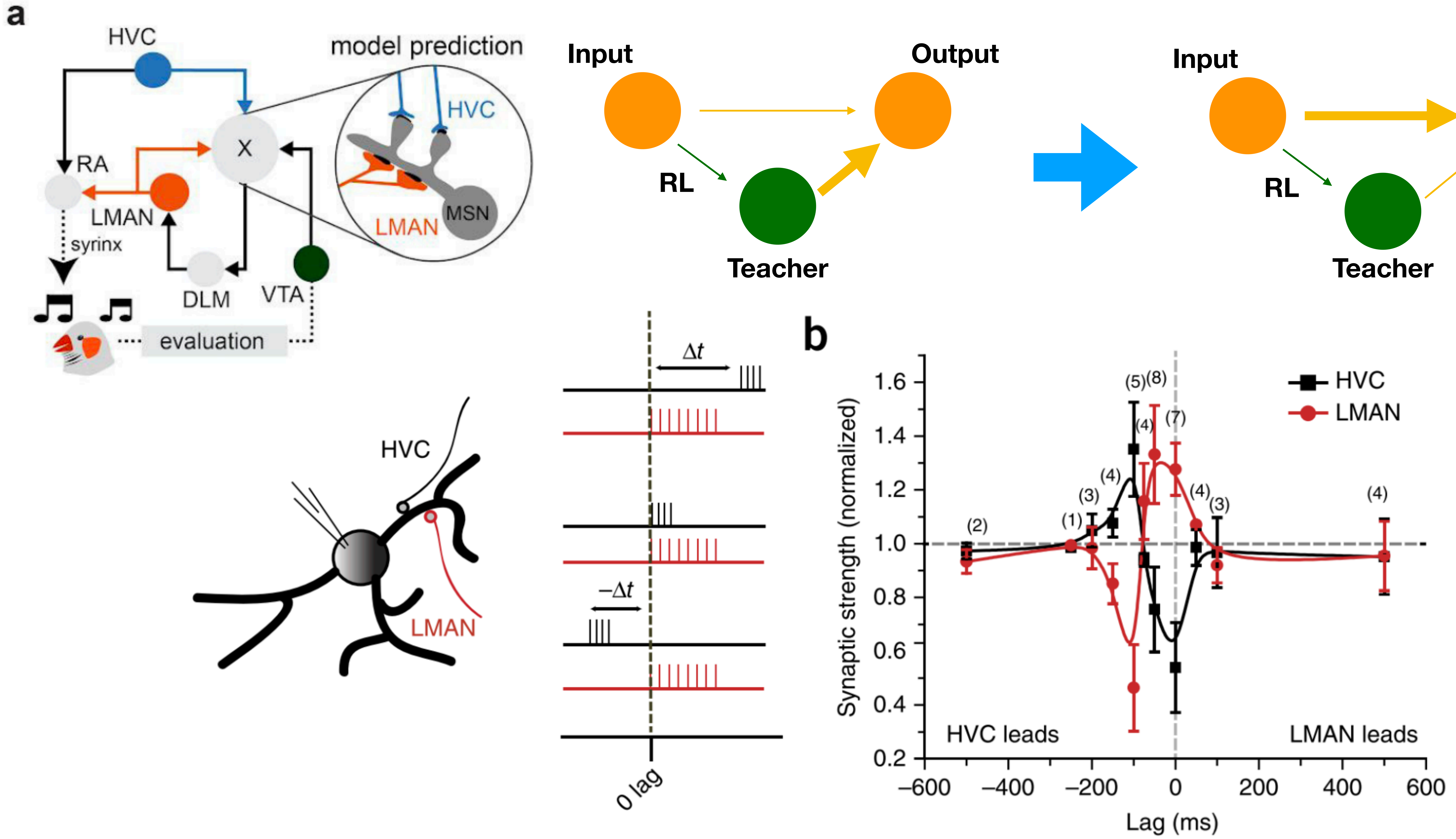
Thalamus as control functions

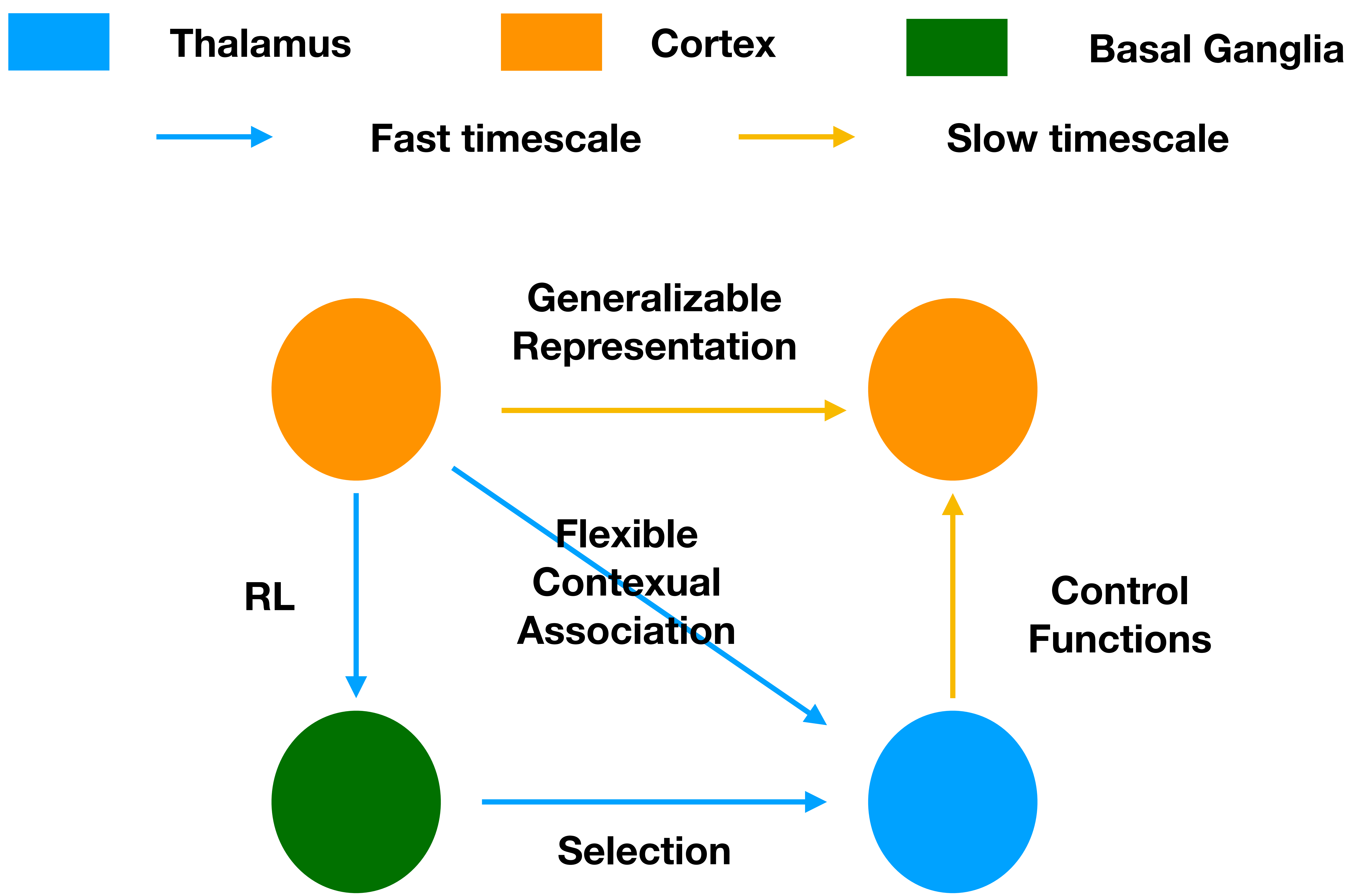
- Traditionally, thalamus is thought to gate and relay information to the cortex
- However, there are works that show some thalamic populations do not change the baseline firing rate of cortical neurons
- Instead, they amplify the functional connectivity

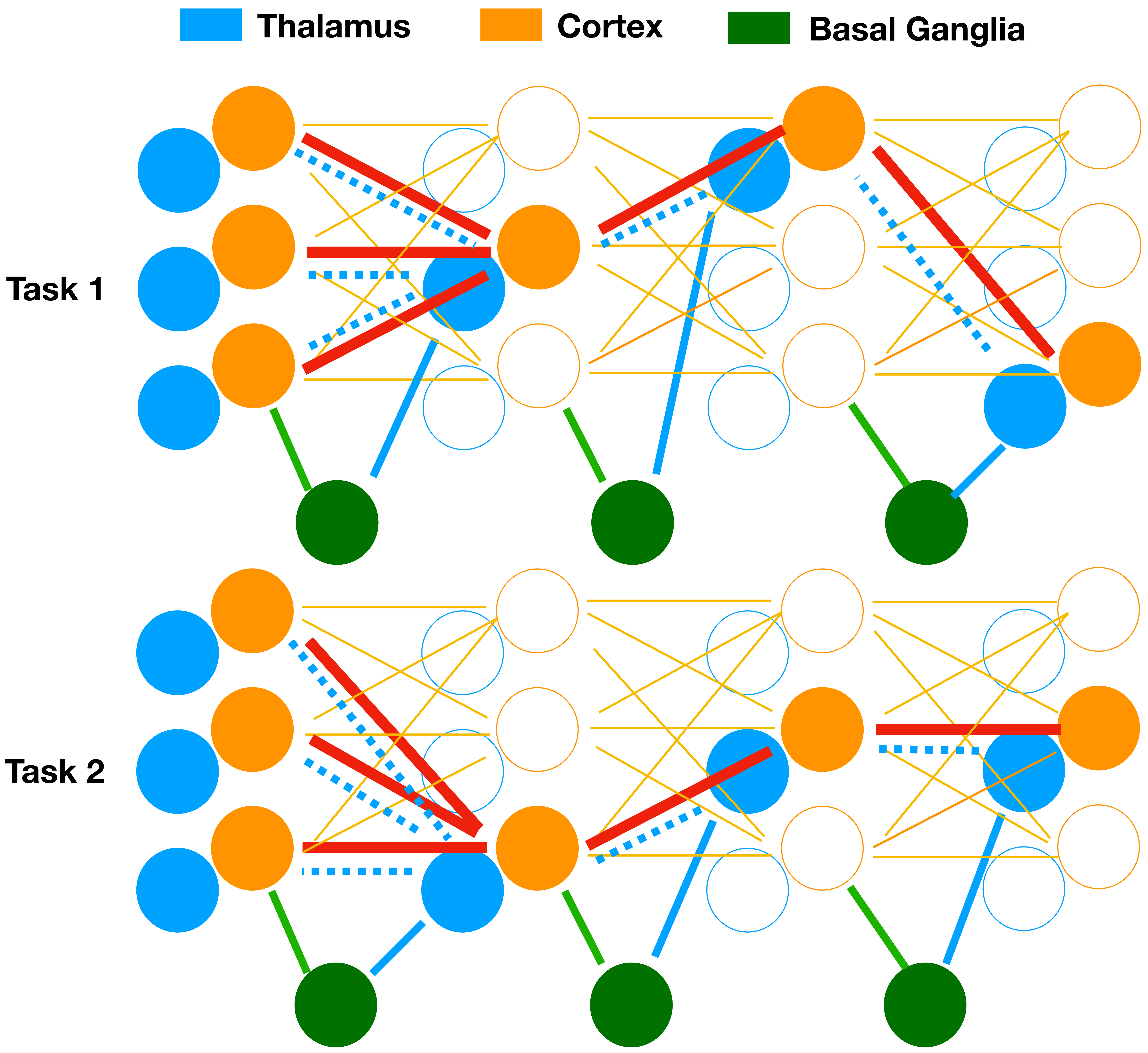


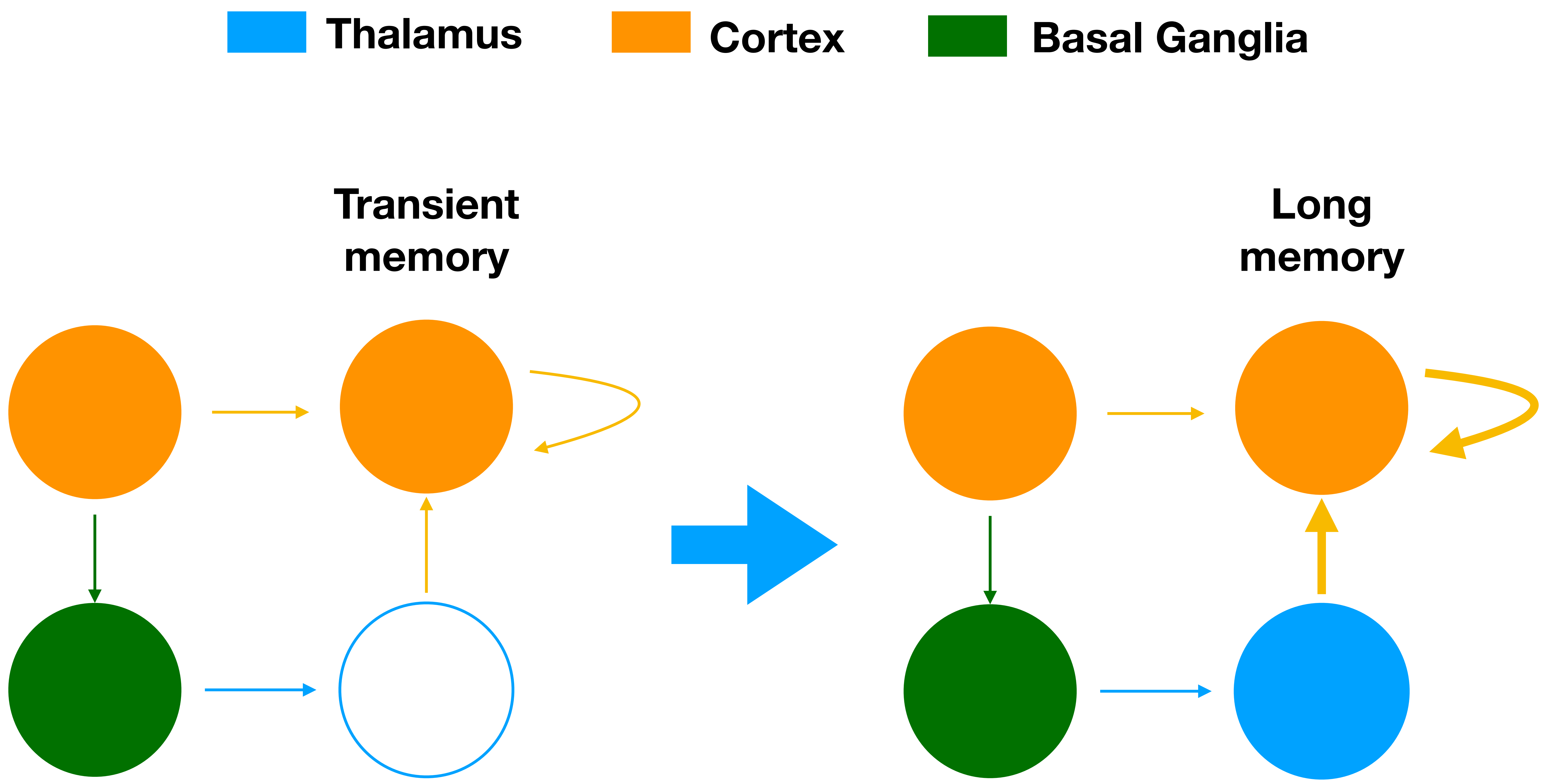
Basal ganglia

- Traditionally, it is thought that basal ganglia is in charge of action/strategy selection based on reinforcement learning
- However, recent studies also found both input/output pathways from all cortex
- We propose that basal ganglia selects these thalamic control functions
- Furthermore, it serves as a trainer for meta learning


a
HVC
model prediction
HVC
X
RA
LMAN
syrinx
MSN
LMAN
DLM
VTA
evaluation
Input
Output
RL
Teacher
Input
Output
RL
Teacher
HVC
LMAN
Δt
−Δt
0 lag
b
Synaptic strength (normalized)
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
HVC
LMAN
(2)
(1)
(3)
(4)
(5)
(8)
(4)
(7)
(4)
(3)
(4)
HVC leads
LMAN leads
−600
−400
−200
0
200
400
600
Lag (ms)


Thalamus
Cortex
Basal Ganglia
Fast timescale
Slow timescale
Generalizable Representation
Flexible Contexual Association
RL
Control Functions
Selection


Thalamus
Cortex
Basal Ganglia
Task 1
Task 2


Thalamus
Cortex
Basal Ganglia
Transient memory
Long memory