A biologically-constrained hybridization of reinforcement learning and accumulator models for adaptive decision-making

Kyle Dunovan[1,2] (ked64@pitt.edu)
1. Psychology Department &
2. Center for the Neural Basis of Cognition
University of Pittsburgh, Pittsburgh, PA

Timothy Verstynen[3,4]
3. Psychology Department
4. Center for the Neural Basis of Cognition
Carnegie Mellon University, Pittsburgh, PA

CNBC Center for the Neural Basis of Cognition
University of Pittsburgh
Carnegie Mellon University

Background & Motivation

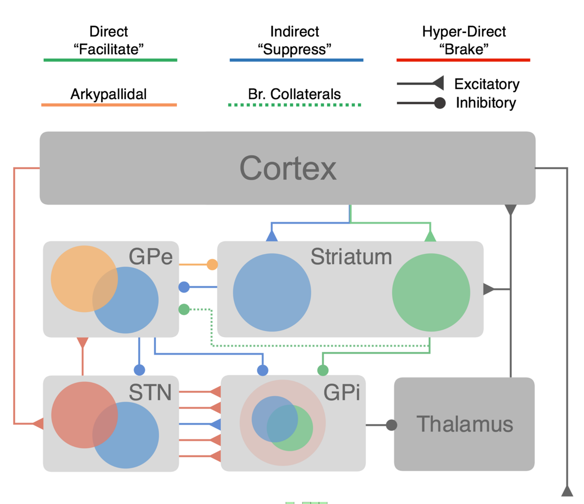
Motivating synthesis of reinforcement learning (RL) and action decisions via overlapping cortico-basal ganglia substrates

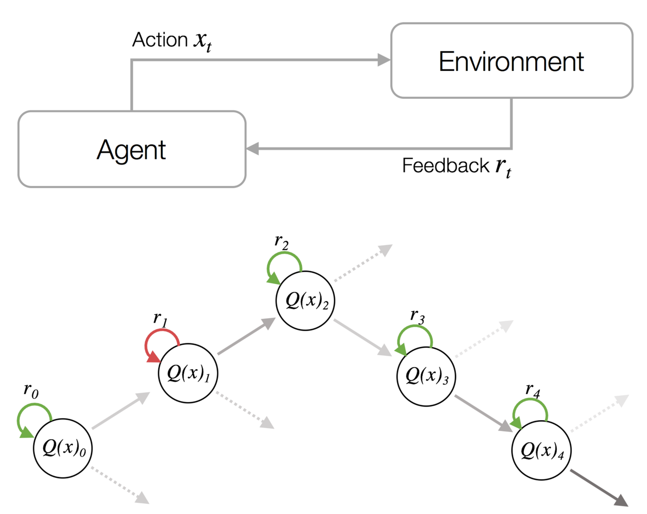
Reinforcement Learning (RL):

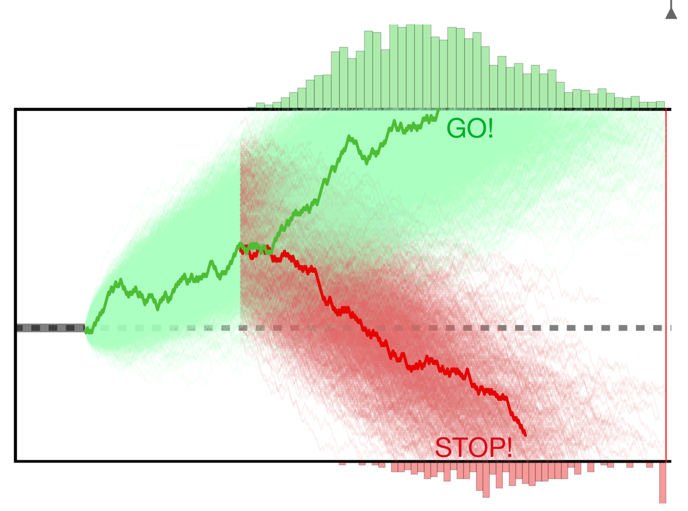
- Dopaminergic modulation of Direct ('Go') and Indirect ('NoGo') pathways tunes action kinematics feedback[1]

Control Decisions (Accumulator Model):

- Cortico-striatal modulation of 'Go' and 'NoGo' pathways mediates proactive control (Direct-Indirect)[2]
- Cortico-subthalamic modulation of 'Braking' pathway mediates reactive control (Hyperdirect)[3]







Questions:

1. Does feedback-dependent plasticity target proactive control mechanisms in accumulator framework?
2. Can this learning mechanism account for temporal dynamics of adaptive control across environments?

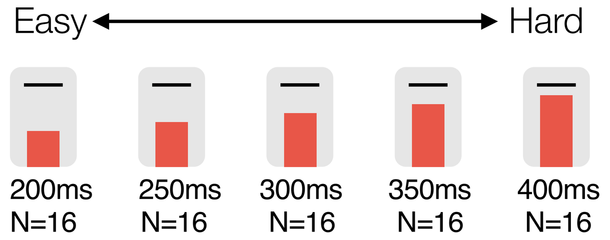
Probabilistic Stop-Signal Task

Go Trials (N_{Go}=600):
Respond when vertically rising bar crosses the Target (T^G = 520 ms)

Stop Trials (N_{Stop}=280):
Bar turns red, stops before T^G prompting subject to withhold their response

Probe SSDs (80 trials)

Easy ← → Hard

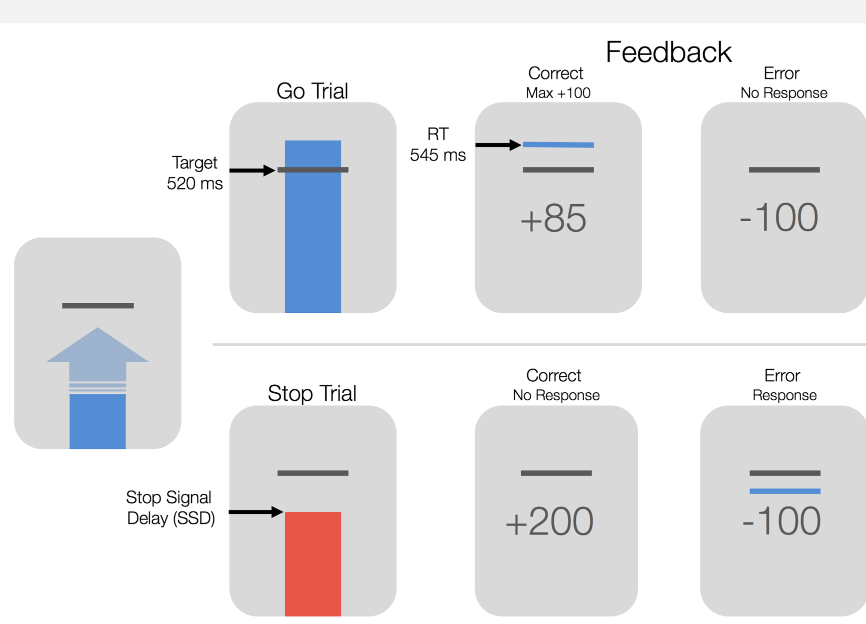


Context SSDs (200 trials)

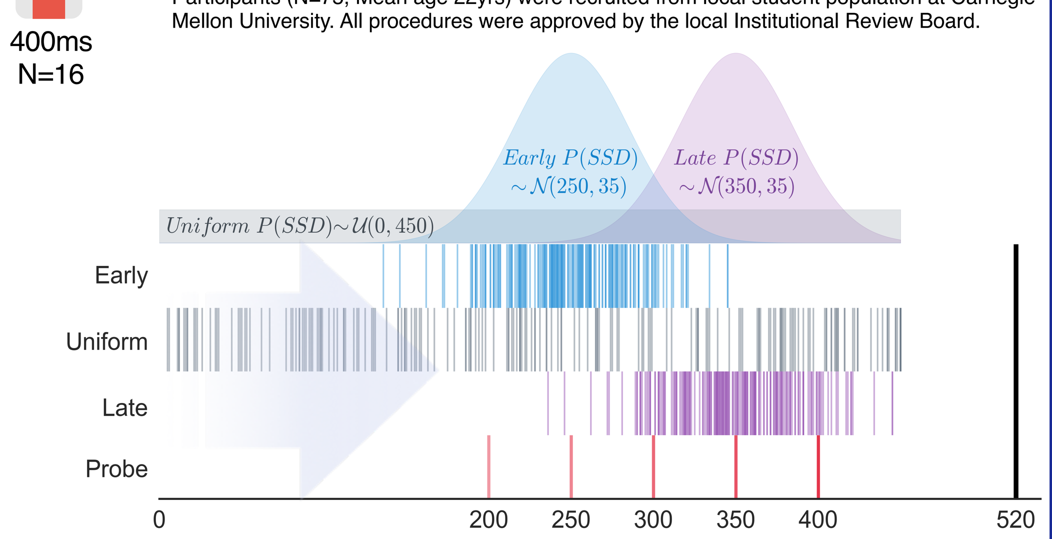
Early Group (n=25):
$P(SSD) \sim \mathcal{N}(250ms, 35)$

Uniform Group (n=25):
$P(SSD) \sim \mathcal{U}(10ms, 450)$

Late Group (n=25):
$P(SSD) \sim \mathcal{N}(350ms, 35)$

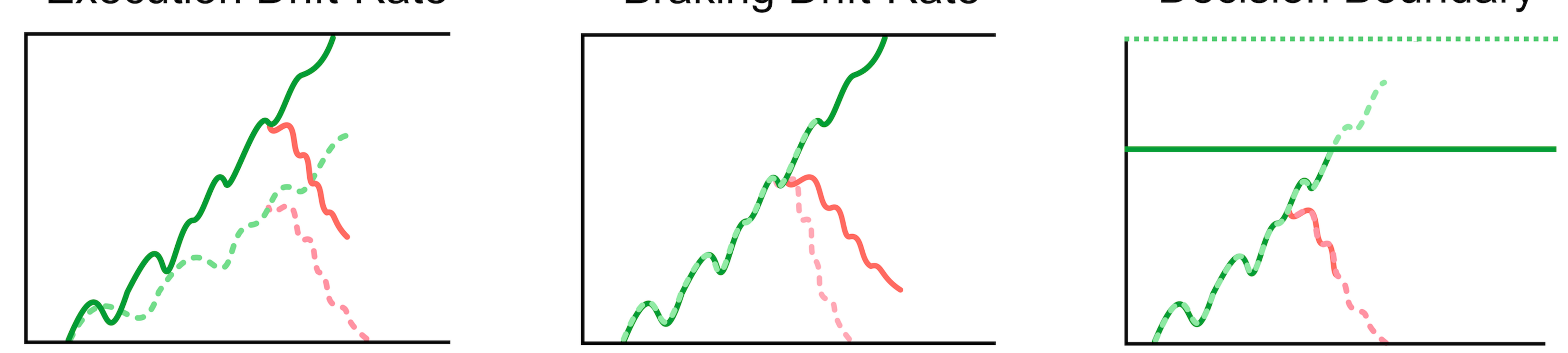


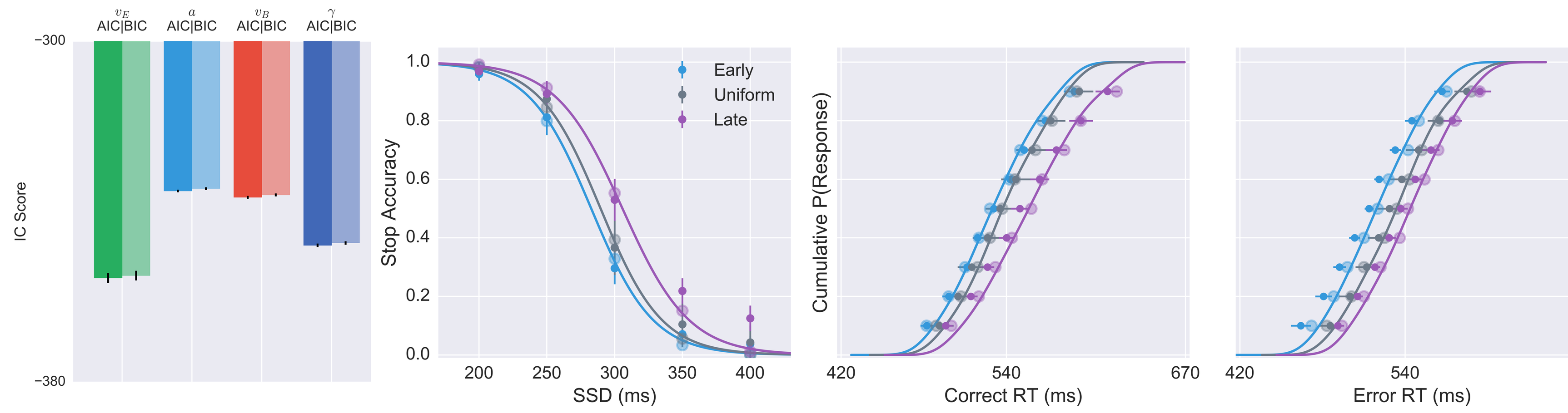
Participants (N=75, Mean age 22yrs) were recruited from local student population at Carnegie Mellon University. All procedures were approved by the local Institutional Review Board.



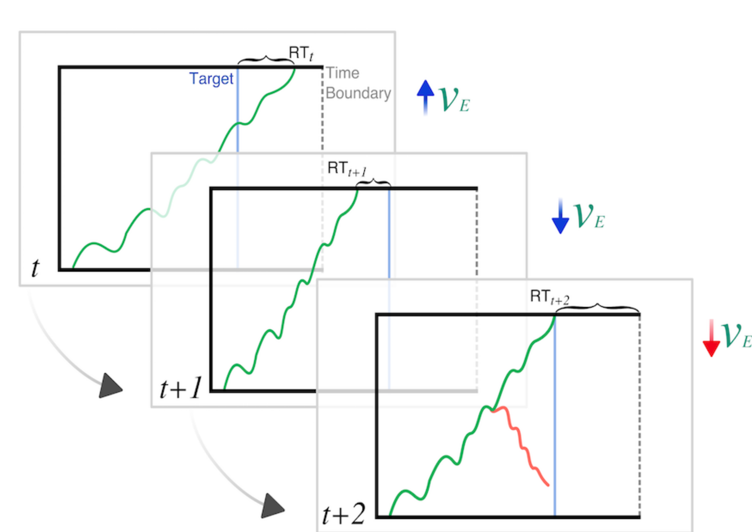
Behavior and Static Model Fits

How does Context effect inhibitory control dynamics *on average*?





Feedback-Dependent Control



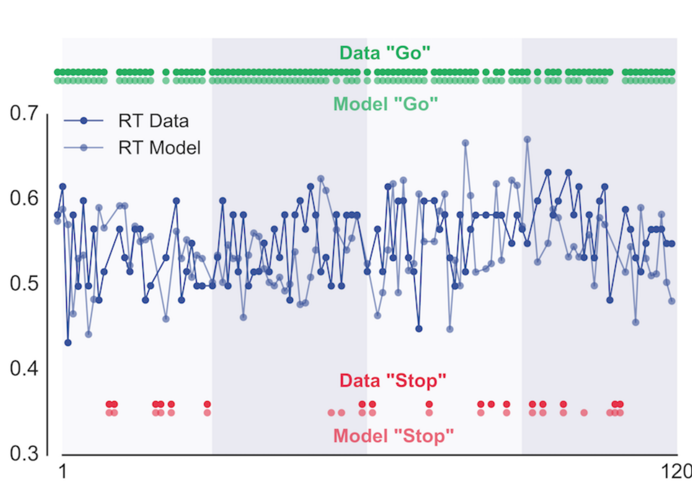
Are Contextual differences in drift-rate a result of feedback-dependent learning?

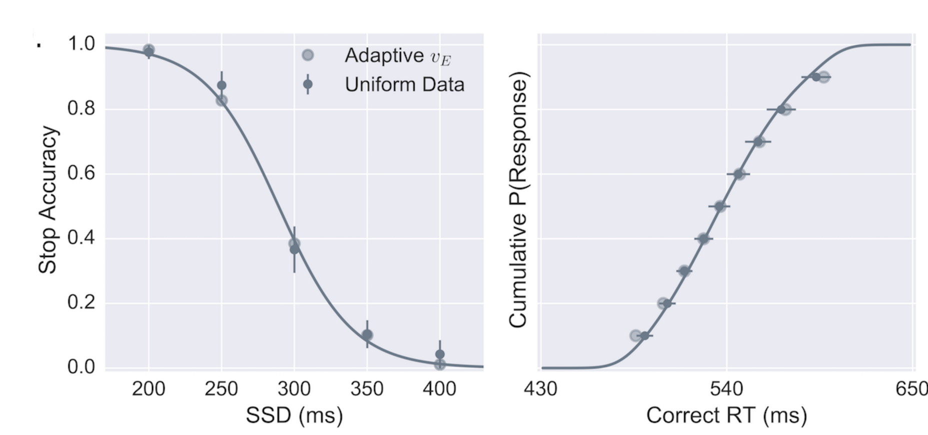
'Go' trials

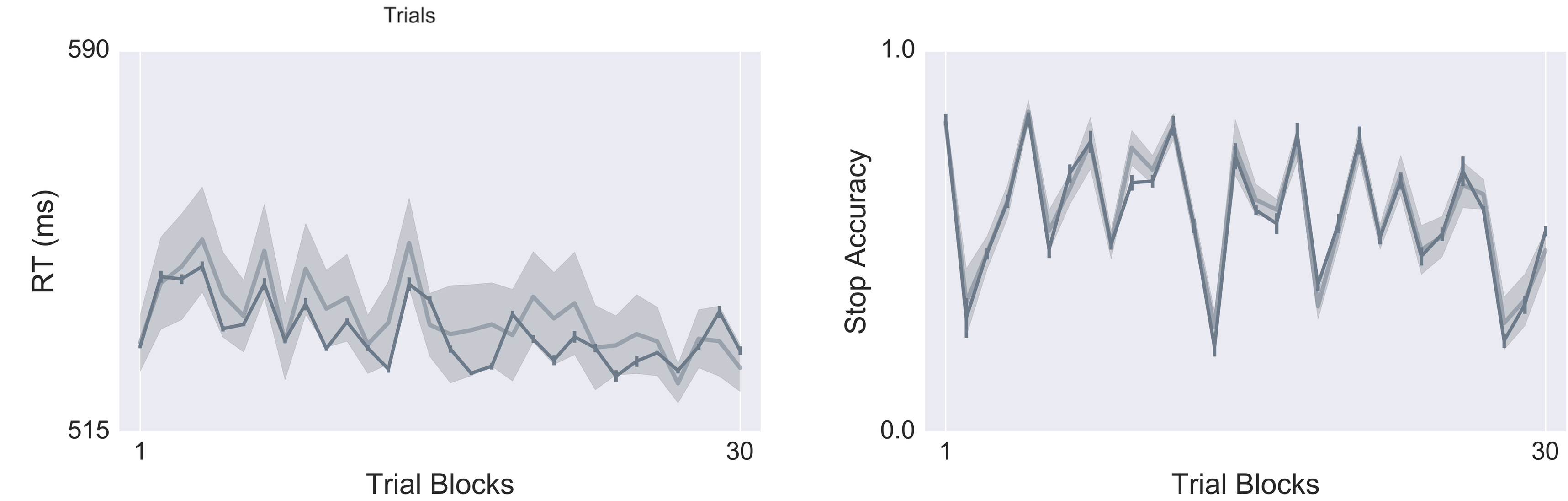
$$v_{t+1} = v_t + \alpha \cdot (v_t - v_t \cdot e^{[T_t^{G/S} - RT_t]})$$

'Stop' trials

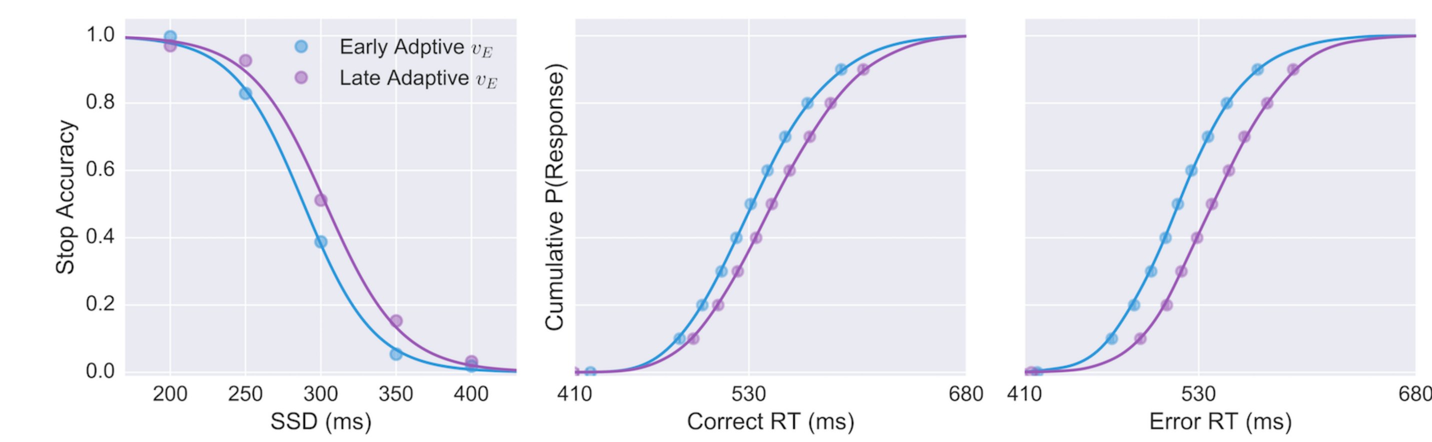
$$T_{t+1}^S = \begin{cases} T_t^S + \beta \cdot T_t^S, & \text{if Stop Failure} \\ T_t^S - \beta \cdot T_t^S, & \text{if Stop Success} \end{cases}$$







Adaptive Control Across Contexts

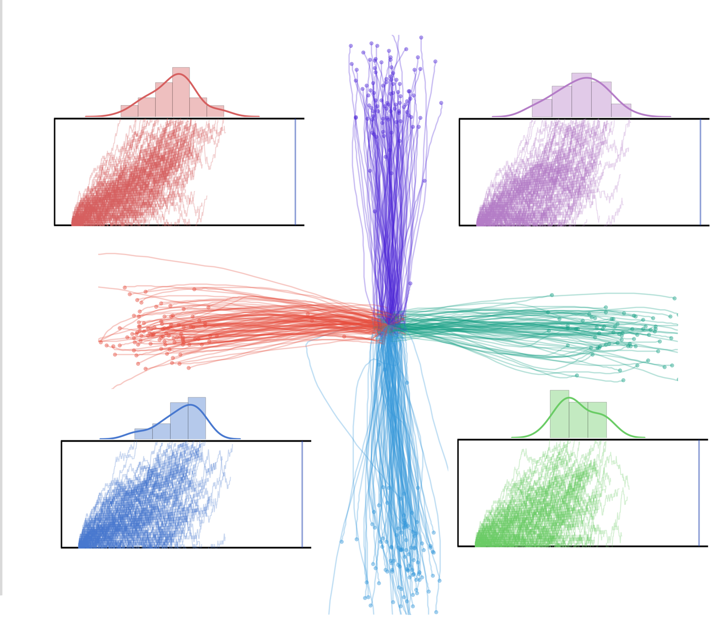


Summary & Future Directions

Summary of Inhibitory Control Findings

- Proactive drift-rate modulation accounts for average effect of Context on inhibitory control
- Feedback-dependent plasticity in drift-rate accounts for trialwise adaptation to timing and control errors
- Synthesis of RL and accumulator models can account for adaptation to errors in action outcome and action timing

Adaptive Multi-Alternative Decisions



References & Acknowledgements

1. Yttri, E. A. & Dudman, J. T. (2016) Opponent and bidirectional control of movement velocity in the basal ganglia. *Nature* 533, 1–16.
2. Dunovan, K., Lynch, B., Molesworth, T. & Verstynen, T. (2015) Competing basal-ganglia pathways determine the difference between stopping and deciding not to go. *Elife* 1–24. doi:10.7554/eLife.087
3. Schmidt, R., Leventhal, D. K., Mallet, N., Chen, F. & Berke, J. D. (2013) Canceling actions involves a race between basal ganglia pathways. *Nature Neuroscience.* 16, 1118–24.

*This work was funded by a National Science Foundation Career Award #1351748 to TV