Striatal prediction errors in a decision-making task are modulated by action execution failures

Samuel D. McDougle[1], Peter A. Butcher[2], Darius E. Parvin[1], Faisal Mushtaq[3], Yael Niv[2], Richard B. Ivry[1], & Jordan A. Taylor[2]

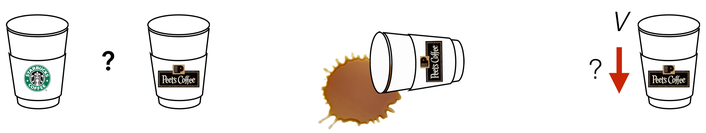
[1]University of California, Berkeley; [2]Princeton University; [3]University of Leeds

mcdougle@berkeley.edu

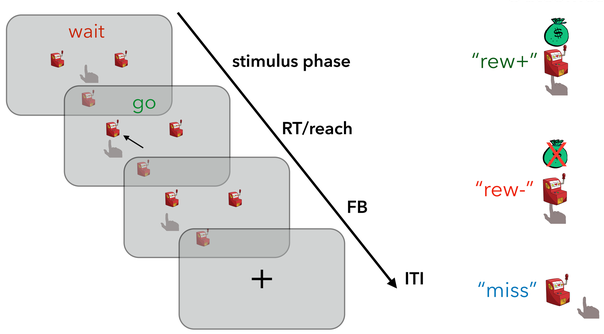
This research was supported by NIH grant #NS092079

Introduction

- Reward prediction error (RPE) signals in the striatum are shaped by a number of cognitive factors, including attention (Leong, Radalescu, et al., 2017), episodic memory (Bornstein et al., 2017), working memory (Collins et al., 2017), and model-based computations (Daw et al., 2011)
- We previously showed that in a simple decision-making task, when reward was withheld due to salient movement errors, human subjects were less likely to "punish" the stimulus that they had chosen (McDougle et al., 2016)
- We hypothesized that this resulted from attenuated negative RPEs in the striatum when credit for a negative outcome was assigned to action execution rather than action selection



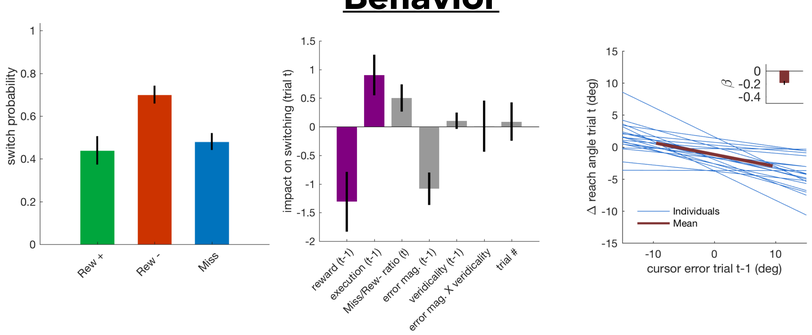
Method



fMRI Acquisition & Pre-Processing

Movements recorded on MR-compatible tablet
2s TR; 28ms TE; 3mm^3 voxel size; 36 interleaved axial slices
Smoothing: Gaussian kernel (8mm FWHM)
ICA-AROMA used for aggressive ICA-based motion correction (Pruim et al., 2015)
N = 24; 4 subjects excluded for movements > 3 mm in any direction; final N = **20**

Behavior



- Subjects were more likely to persist after Miss trials versus Rew- trials ($p < 0.001$)
- Subjects did not appear to be sensitive to the veridicality of feedback

3-α model provides best fit to behavior

Standard Model (2 α)

$V_{t+1}(a|rew+_t) = V_t(a) + \alpha_{rew+}\,\delta_t$

$V_{t+1}(a|rew-_t \text{ or } miss_t) = V_t(a) + \alpha_{rew-/miss}\,\delta_t$

Gating Model (3 α)

$V_{t+1}(a|rew+_t) = V_t(a) + \alpha_{rew+}\,\delta_t$

$V_{t+1}(a|rew-_t) = V_t(a) + \alpha_{rew-}\,\delta_t$

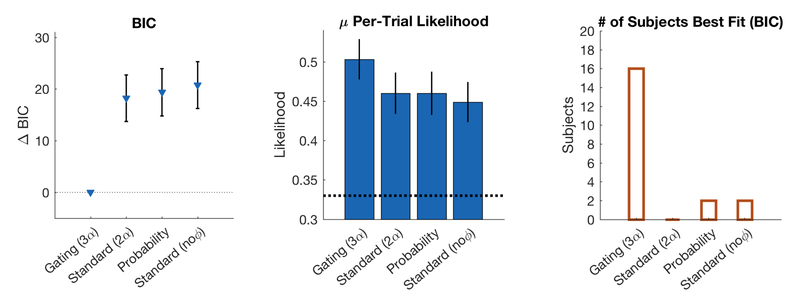
$V_{t+1}(a|miss_t) = V_t(a) + \alpha_{miss}\,\delta_t$

Probability Model

$P_{t+1}(a) = P_t(a) + \alpha_{prob}\,\delta_{t,prob}$

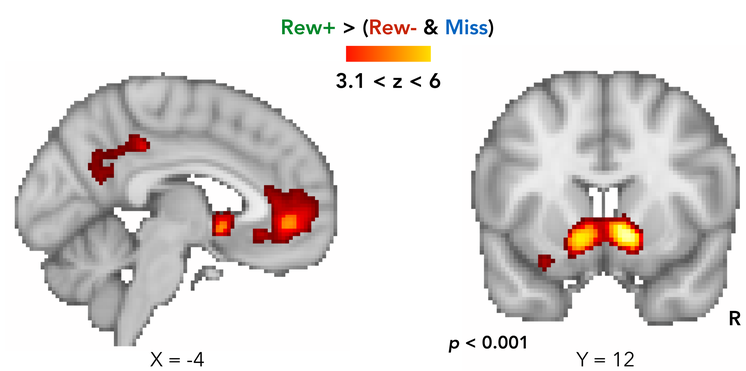
$E_{t+1}(a|hit_t) = E_t(a) + \alpha_{payoff}\,\delta_{t,payoff}$

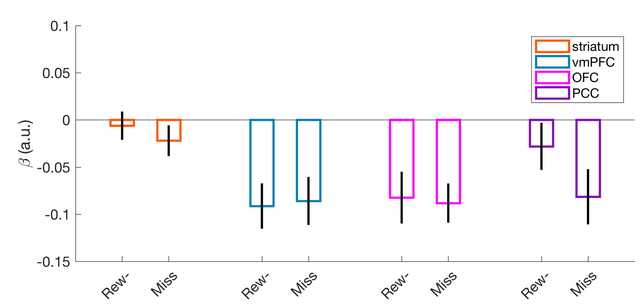
$V_{t+1}(a) = P_t(a)\,E_t(a)$



- Of the models tested, the "Gating" model — where Miss trials are associated with their own learning rate parameter — provided the best fit to subjects' choices, consistent with previous computational modeling results (McDougle et al., 2016)

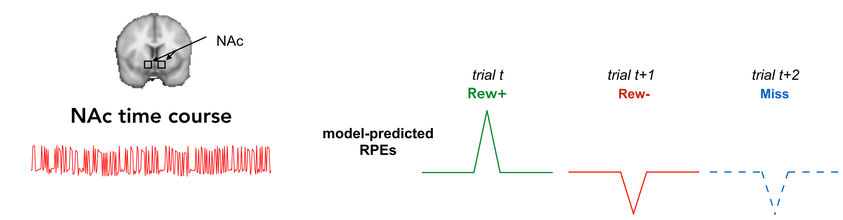
No clear difference between unrewarded conditions in simple contrast analysis

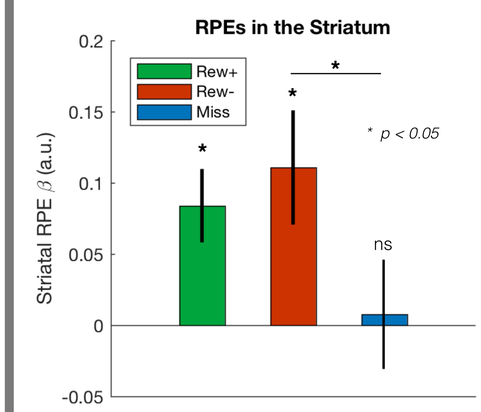




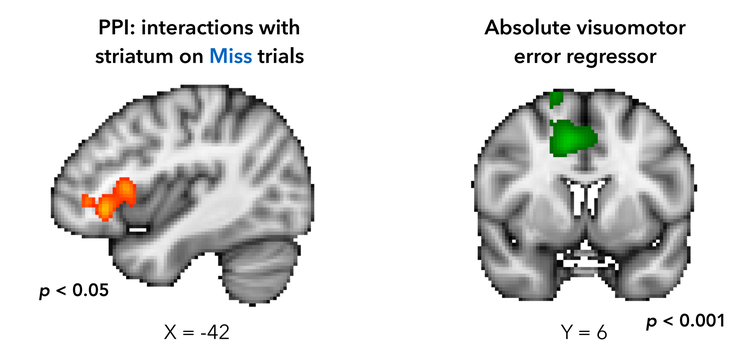
- Reward clusters found in ventral striatum, ventromedial prefrontal cortex (vmPFC), left orbitofrontal cortex (OFC), and posterior cingulate cortex (PCC)
- No clear differentiation of Miss and Rew- trials using simple contrast approach

Model-driven analysis reveals differences in RPE encoding





- Significant RPE encoding in the NAc for Rew+ & Rew- trials
- No significant encoding of Miss trial RPEs



Conclusions

- Decision-making behavior in tasks that dissociate action selection errors from action execution errors shows that these errors are treated differently during learning, with credit being assigned to either the object of a decision or the implementation of a decision
- This "gating" of negative RPEs on execution error trials is reflected in diminished RPE encoding in the ventral striatum
- These results suggest that striatal RPEs are sensitive to high-level information related to the implementation of decisions, perhaps mediated by a sense of agency over motor execution aspects of a task (Parvin et al., 2018)

References: McDougle et al., PNAS (2016); Leong, Radalescu, et al., Neuron (2017); Bornstein et al., Nat Comm (2017); Collins et al., J Neuro (2017); Daw et al., Neuron (2011); Pruim et al., Neuroimage (2015); Parvin et al., J Neuro (2018)