A potential reset mechanism for the modulation of decision processes under uncertainty

Krista Bond, Alexis Porter, and Timothy Verstynen

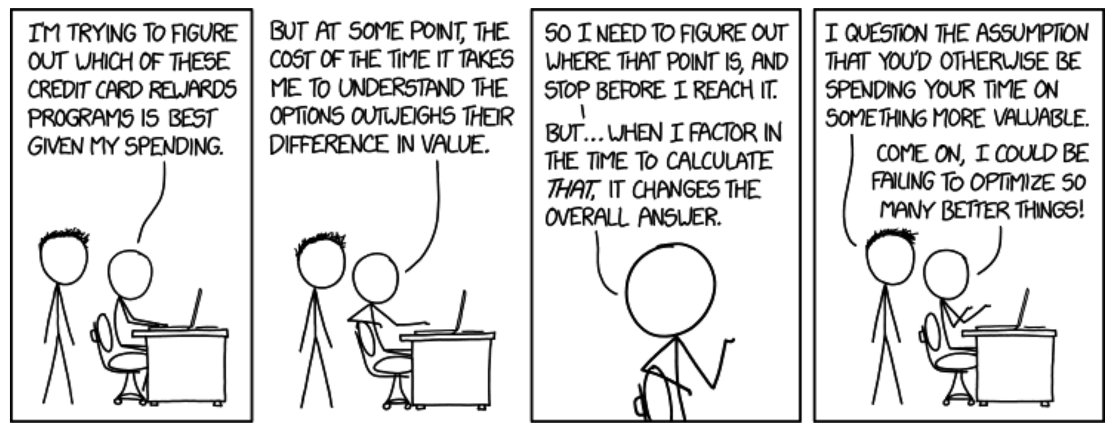
kbond@andrew.cmu.edu; alexisp@andrew.cmu.edu; timothyv@andrew.cmu.edu

Introduction

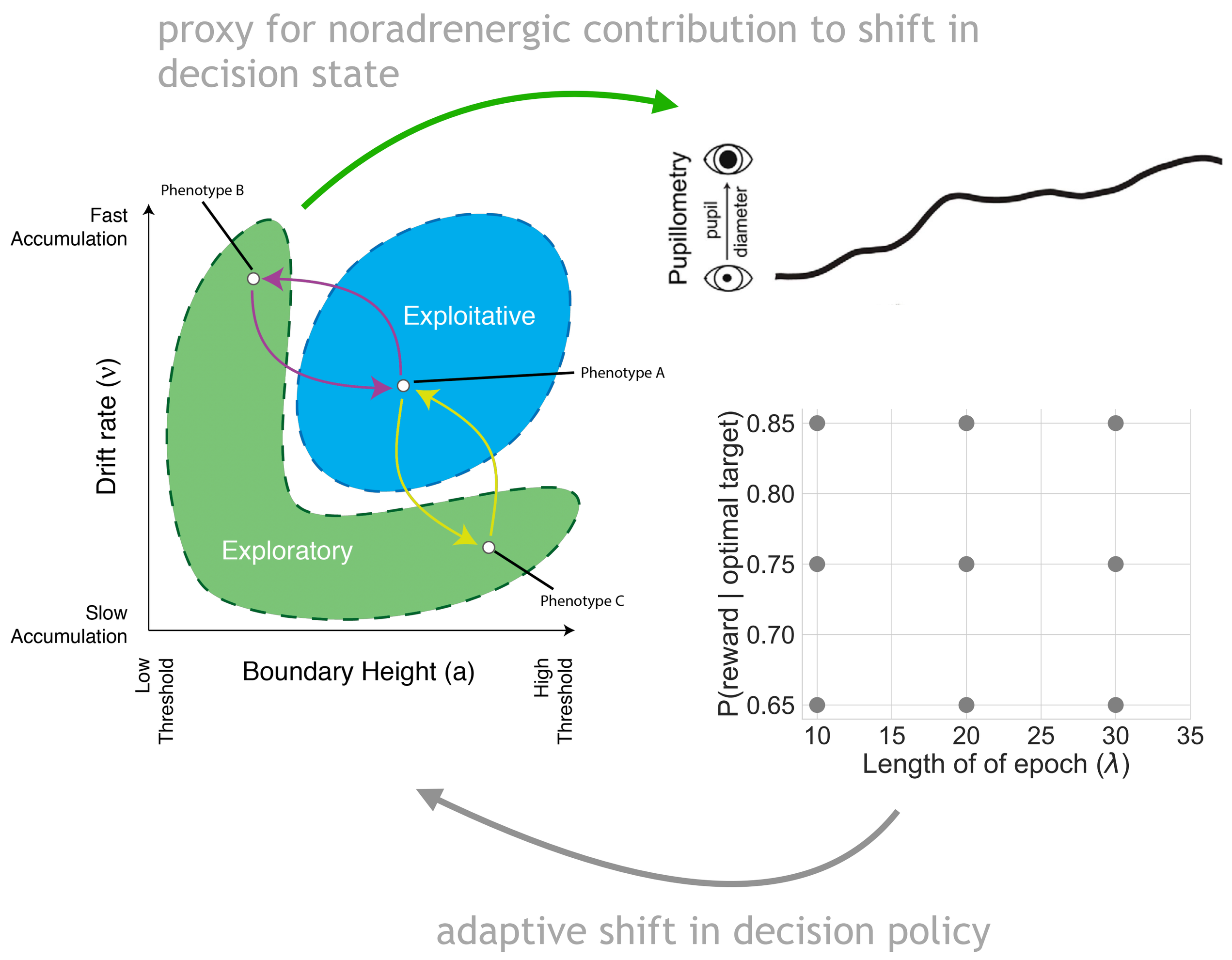
How do we make good decisions under multiple forms of uncertainty?

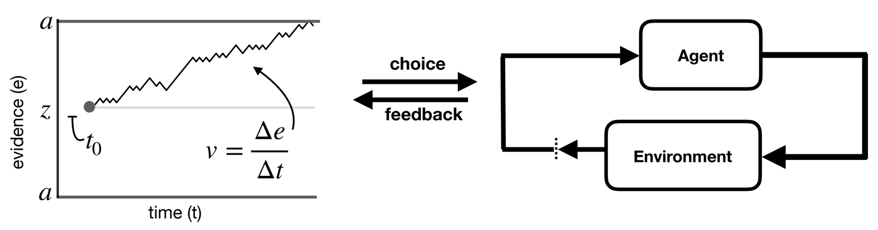
conflict: *a state of uncertainty in the relative estimate of action-value*

volatility: *sudden shifts in action-value associations over time*



Hypotheses





- The amount of evidence needed to make a decision decreases with environmental instability while the rate of evidence accumulation increases with belief in optimal choice value (Bond, Dunovan, & Verstynen, 2018)
- Hypothesized replication using a more expansive manipulation of belief and change point probability
- A phasic increase in noradrenaline will mediate the relationship between environmental instability and the amount of information needed to make a decision, facilitating the exploration process in response to change point detection.

Relative action-value

$$B_{t+1,c} = B_{t,c} + \alpha_t \delta_t$$
$$B_{t+1,u} = B_{t,u}(1-\Omega_t) + \Omega_t E(r)$$
$$E(r) = \frac{\bar{r}_{t_0} + \bar{r}_{t_1}}{2}$$
$$B_{\Delta_{t+1}} = B_{t,1} - B_{t,0}$$

Action-value updates

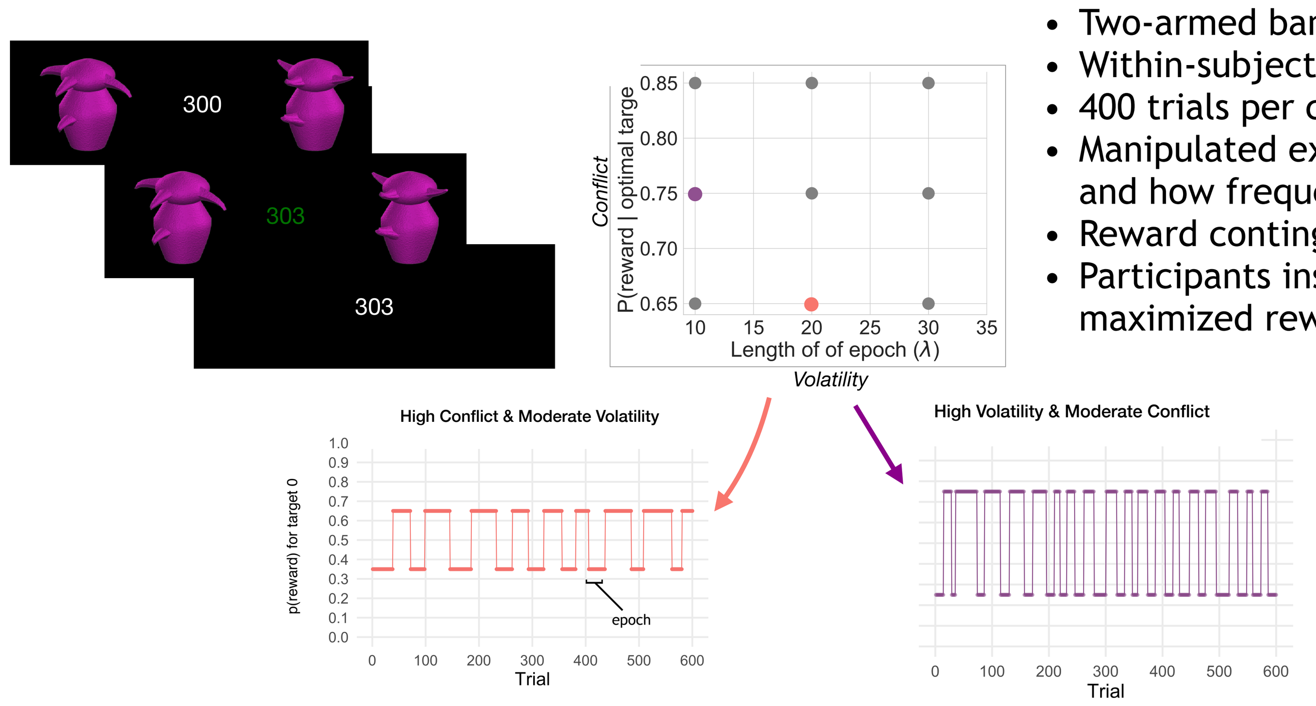
$$\Omega_t = \frac{U(r_{\Delta_t})H}{U(r_{\Delta_t})H + N(r_{\Delta_t}|B_{\Delta_t}, \sigma_t^2)(1-H)}$$
$$RU_t = \frac{\Omega_t\sigma_n^2 + (1-\Omega_t)(1-\phi_t)\sigma_n^2 + \Omega_t(1-\Omega_t)(\delta_t\phi_t)^2}{\Omega_t\sigma_n^2 + (1-\Omega_t)(1-\phi_t)\sigma_n^2 + \Omega_t(1-\Omega_t)(\delta_t\phi_t)^2 + \sigma_n^2}$$
$$\phi_{t+1} = 1 - RU$$

Update rules

$$\alpha_t = \Omega_t + (1-\Omega)(1-\phi_t)$$
$$\delta_t = r_t - B_{t,c} \qquad a_{t+1} = \hat{\beta}_a \cdot \Omega_t + a_0$$
$$\sigma_t^2 = \sigma_n^2 + \frac{(1-\phi_t)\sigma_n^2}{\phi_t} \qquad v_{t+1} = \hat{\beta}_v \cdot B_{\Delta_t} + v_t$$

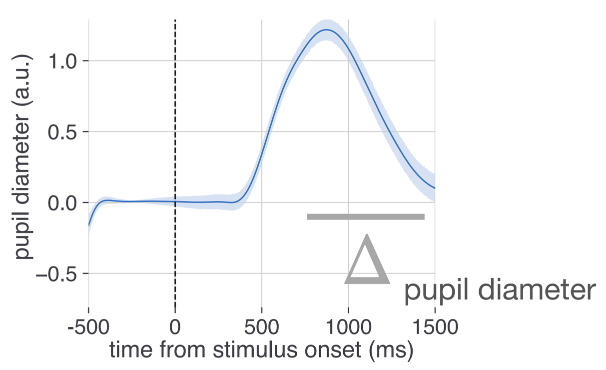
Methods

Task



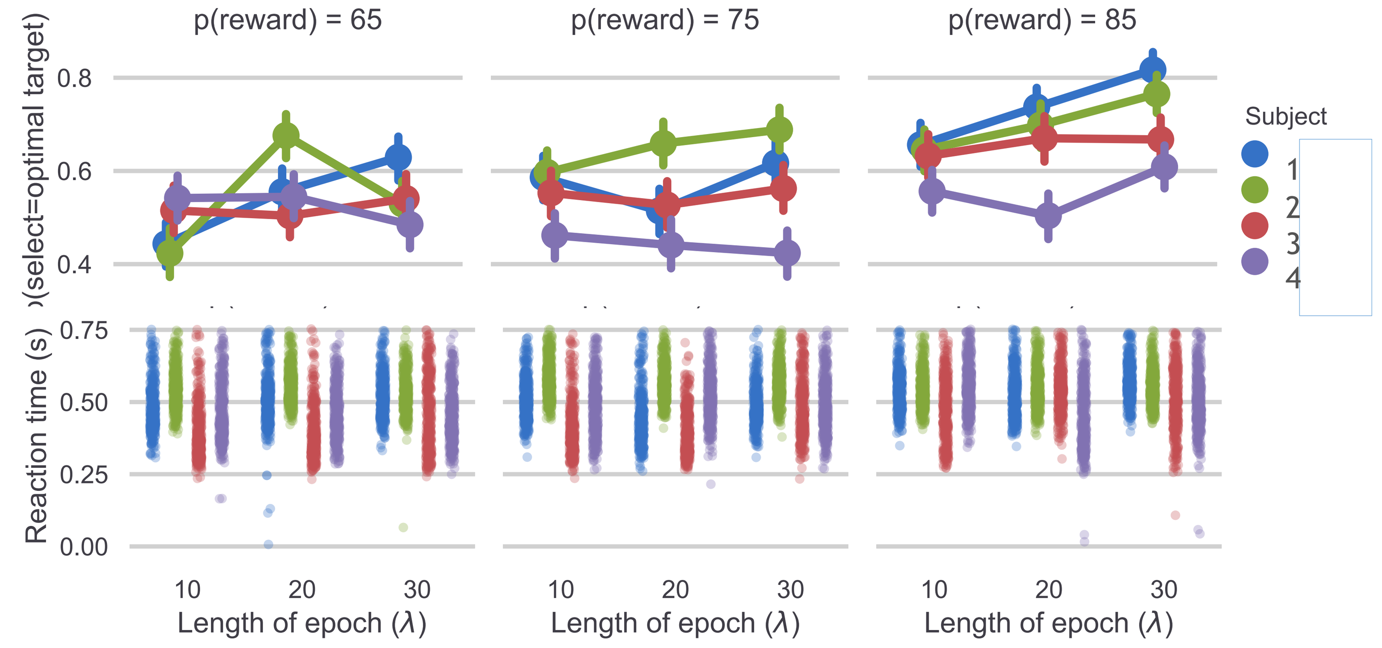
- Two-armed bandit task with constrained reaction time (.1 - .7 s)
- Within-subjects, replication-based design (N = 4)
- 400 trials per condition, 3600 trials per subject
- Manipulated expected value of rewards for each target (conflict) and how frequently the most rewarding target changed (volatility)
- Reward contingencies changed across epochs
- Participants instructed to select the sex identity of the greeble that maximized reward

Pupillometry



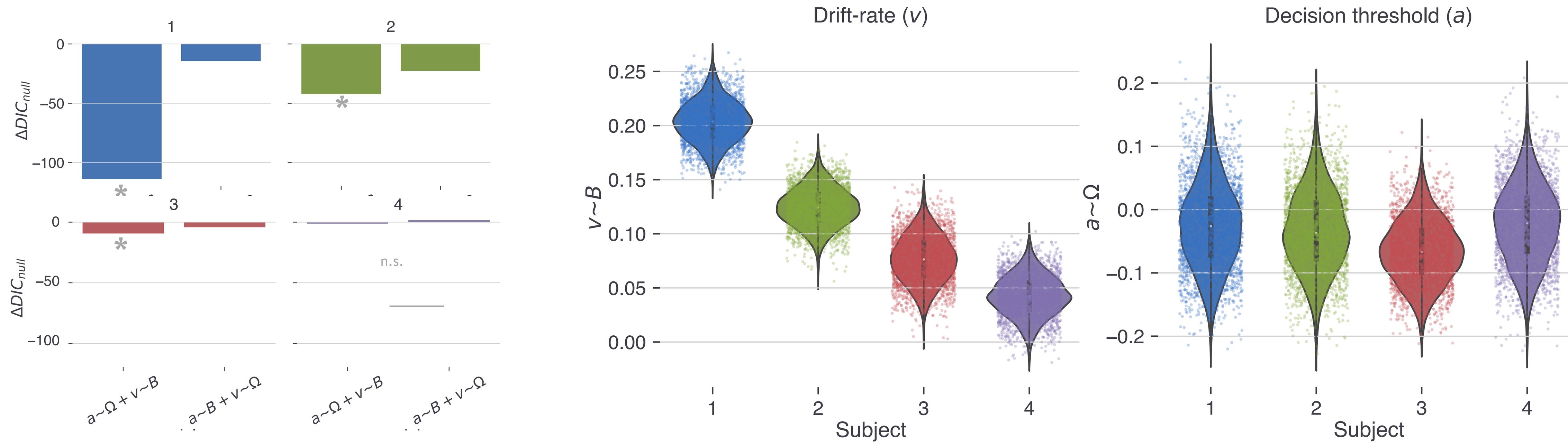
- Pupillometry data were sampled at 1 kHz
- Isoluminant conditions in a booth that minimized fluctuations in ambient light sources

Behavior



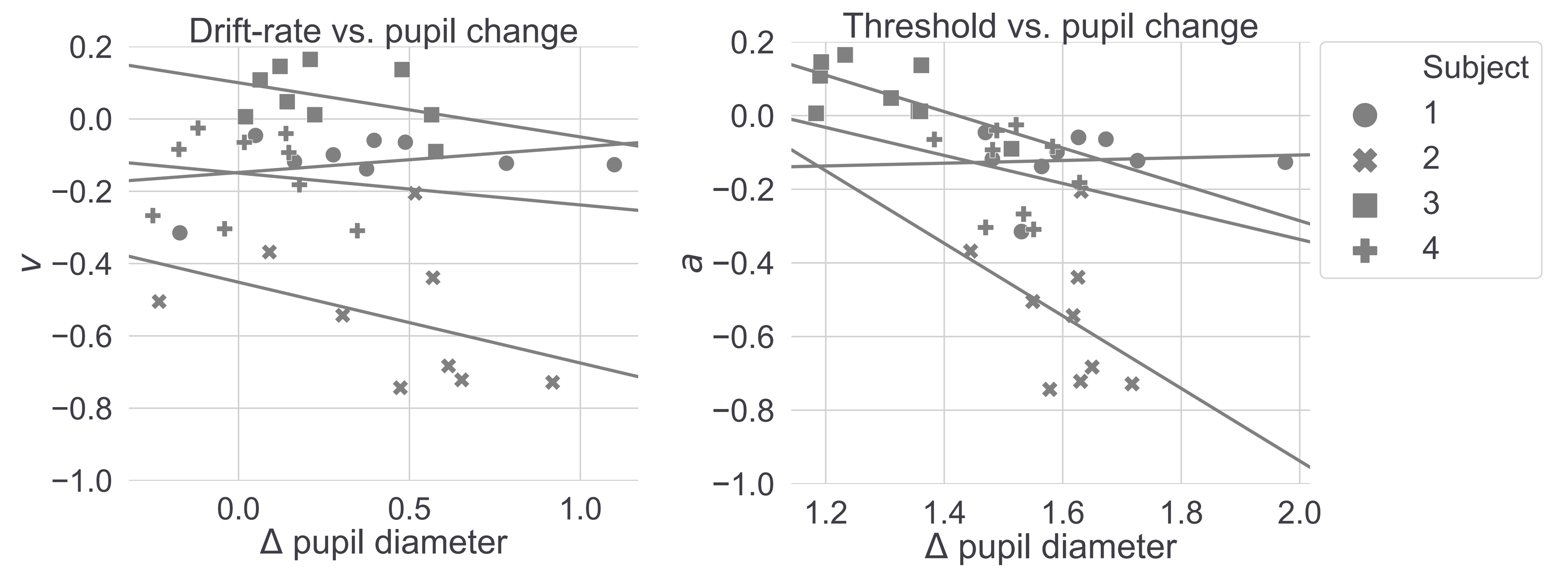
- As change point probability (CPP) increased, accuracy decreased ($p<0.01$ in 3/4 replicates, $\beta=-0.55\pm0.24$)
- Accuracy increased with belief in optimal target value ($p<0.03$, $\beta=0.15\pm0.1$)
- RTs decreased as CPP increased ($p<0.03$ in 3/4 replicates, $\beta=-0.02\pm0.01$)
- The belief in the value of the optimal target minimally impacted RTs ($\beta=0.00\pm0$ in 4/4 subjects)

Decision parameters



- Drift-rate increased with the belief in the value of the optimal target (observed $p(\beta_v<0) < .01$ in all cases; mean $p(\beta_v<0) = 0\pm0$)
- Change point probability weakly decreased the decision threshold (mean observed $p(\beta_a>0) = 0.27\pm0.11$)

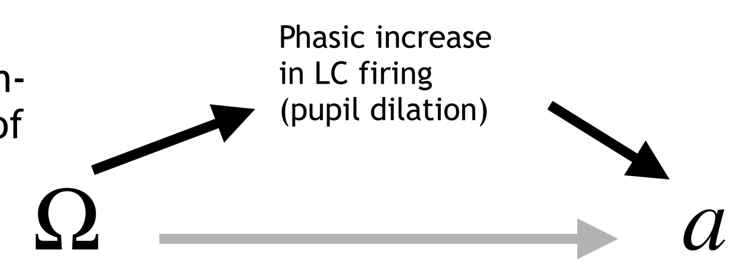
Preliminary pupillometry analysis



Preliminary analysis shows a decrease in threshold and drift-rate as the amplitude of the phasic pupillary response increases (3/4 replicates).

Conclusions and Future Directions

- Belief in the value difference between options had the greatest influence on decision processes, impacting drift-rate, while estimates of environmental change had a weak but detectable influence on the decision threshold.
- Taken together, these findings validate our previous model of adaptive decision-making showing how separate environmental signals impact different aspects of the decision algorithm. These previous findings were supported using:
  - new data
  - a more expansive sampling of the conflict and volatility space
  - a task with different superficial features
- Future work will test how phasic LC firing mediates the relationship between environmental instability and the amount of information needed to make a decision.



Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. Nature neuroscience, 15(7), 1040.
Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: hierarchical bayesian estimation of the drift-diffusion model in python. Frontiers in neuroinformatics, 7, 14
Kim, T. D., Kabir, M., & Gold, J. I. (2017). Coupled decision processes update and maintain saccadic priors in a dynamic environment. Journal of Neuroscience, 3078-16.
Humphries, M. D., Khamassi, M., & Gurney, K. (2012). Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia. Frontiers in neuroscience, 6, 9.

K.B. was funded by the Behavioral Brain Research Training Program at the Center for the Neural Basis of Cognition at Carnegie Mellon University (National Institutes of Health grant T32GM081760). This project was partially funded by National Science Foundation Career Award 1351748.