Online Supplementary Material

# Posterior Cingulate Cortex:

# Adapting Behavior to a Changing World

John M. Pearson[1,2], Sarah R. Heilbronner[1,2], David L. Barack[2,3], Benjamin Y. Hayden[1], and Michael L. Platt[1,3,4]

[1]*Department of Neurobiology, Duke University Medical Center, Box 3209, Durham, NC 27710*

[2]*Center for Cognitive Neuroscience, Duke University, B203 Levine Science Research Center, Box 90999, Durham, NC 27708*

[3]*Department of Philosophy, Duke University, Box 90743, Durham, NC 27708*

[4]*Department of Evolutionary Anthropology, Duke University,104 Biological Sciences Building, Box 90383, Durham, NC 27708-0680*

Corresponding author: Platt, M.L. (email: platt@neuro.duke.edu)

**Supplementary Methods:**

We reanalyzed data from a previous publication [S1] consisting of 83 single neurons recorded over 47 behavioral sessions from two monkeys performing a 4-armed bandit task. As in that work, all isolated units collected were analyzed, with no selection criteria for collection. Once again, we fit three behavioral models to monkeys' choices each day: a Bayesian greedy model, a heuristic model, and a reduced Kalman filter [S1]. However, in our reanalysis, we also fit a fourth model based on a full Kalman filter with uncertainty in outcomes. Despite the fact that outcomes in our task had no variance in addition to the random walk, fitting this expanded model allowed for the possibility that monkeys perceived the task in this way.

**Kalman Filter Model**

We implemented a Kalman filter [S2, S3] that performed online Bayesian learning in our task. In this model, all targets values were updated each trial according to a drift model:

$$\mu_i \leftarrow (1-\zeta)\mu_i + \zeta\theta$$
$$\sigma_i^2 \leftarrow (1-\zeta)^2 \sigma_i^2 + D^2,$$

where $\mu_i$ and $\sigma_i$ are the mean and standard deviation of the posterior estimate of each option's value, $\zeta$ is a central tendency of options to drift toward an asymptotic value, $\theta$, and $D$ reflects the growing uncertainty of an unchosen target's value over time due to drift.

In addition, for the chosen target, we calculated learning parameters as follows:

$$\delta_i = r - \mu_i$$
$$\alpha_i = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_0^2},$$

where $r$ is the outcome of the current trial, $\mu_i$ is the mean of option $i$, and $\mu_0$ and $\sigma_0$ are the mean and standard deviation parameters for the target value post-jump. As usual, $\delta$ is the reward prediction error and $\alpha$ the learning rate.

We then updated the parameters of the chosen target:

$$\mu_i \leftarrow \mu_i + \alpha_i \delta_i$$
$$\sigma_i^2 \leftarrow (1-\alpha_i)\sigma_i^2.$$

Thus, each trial yields a single $\delta$ and $\alpha$, along with vectors $\mu$ and $\sigma$.

Finally, we calculated the likelihood of choosing each option according to a softmax model with "inverse temperature" $\beta$ and intrinsic target biases $\gamma_i$:

$$p(i|\{\mu_i\}) = \frac{\gamma_i e^{\beta\mu_i}}{\sum_i \gamma_i e^{\beta\mu_i}} .$$

**Model fitting**

As in [1], we fit models using a maximum likelihood procedure via custom scripts written using the Matlab Optimization Toolbox. For each behavioral session, we fit the Kalman filter (parameters: $\beta$, $\theta$, $\zeta$, $\mu_0$, $\sigma_0$, $D$, $\gamma_i$) using N=5 random starting points in the search space, selecting the set that produced maximum log likelihood. We then compared behavioral models using the Akaike Information Criterion (AIC) [S4]:

$$AIC = -2LL + 2k ,$$

with $LL$ the log likelihood of behavioral choices and $k$ the number of fitted parameters. This we transformed for comparison purposes to an Akaike weight:

$$w_i = \frac{e^{-\frac{1}{2}\Delta_i}}{\sum_i e^{-\frac{1}{2}\Delta_i}} ,$$

where $\Delta_i = AIC_i - \min\{AIC\}$. Since $\sum w_i = 1$, $w_i$ may be thought of as a percentage of evidence in favor of model $i$.

**Neuronal analysis**

Following [S1], we examined the relationship between firing in both the decision and post-reward epochs of the task using partial correlations that controlled for the effect of reward (both chosen and mean reward across option). Partial correlations control for false positives associated with covariance between independent variables, and therefore serve as a more conservative estimate of true correlation. "Variance chosen" refers to the standard deviation of the posterior distribution of the value of the chosen option ($\sigma$ above). "Mean variance" refers to the mean of this quantity across all four options. In Figure 2, we report the results of this analysis in terms of both numbers of significantly correlated neurons and the mean absolute value of correlation for this subset of neurons.

**Supplementary References**


S1 Pearson, J.M., *et al.* (2009) Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task. *Curr Biol* 19, 1532-1537

S2 Anderson, B. and Moore, J. (1979) *Optimal filtering*. Prentice-Hall

S3 Daw, N.D., *et al.* (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441, 876-879

S4 Burnham, K. and Anderson, D. (2002) *Model Selection and Multimodel Inference*. Springer-Verlag