Perceptual confirmation bias from approximate online inference

Richard D. Lange*, Ankani Chattoraj*, Matthew Hochberg, Jacob Yates, Ralf M. Haefner

Brain and Cognitive Sciences, University of Rochester

*authors contributed equally

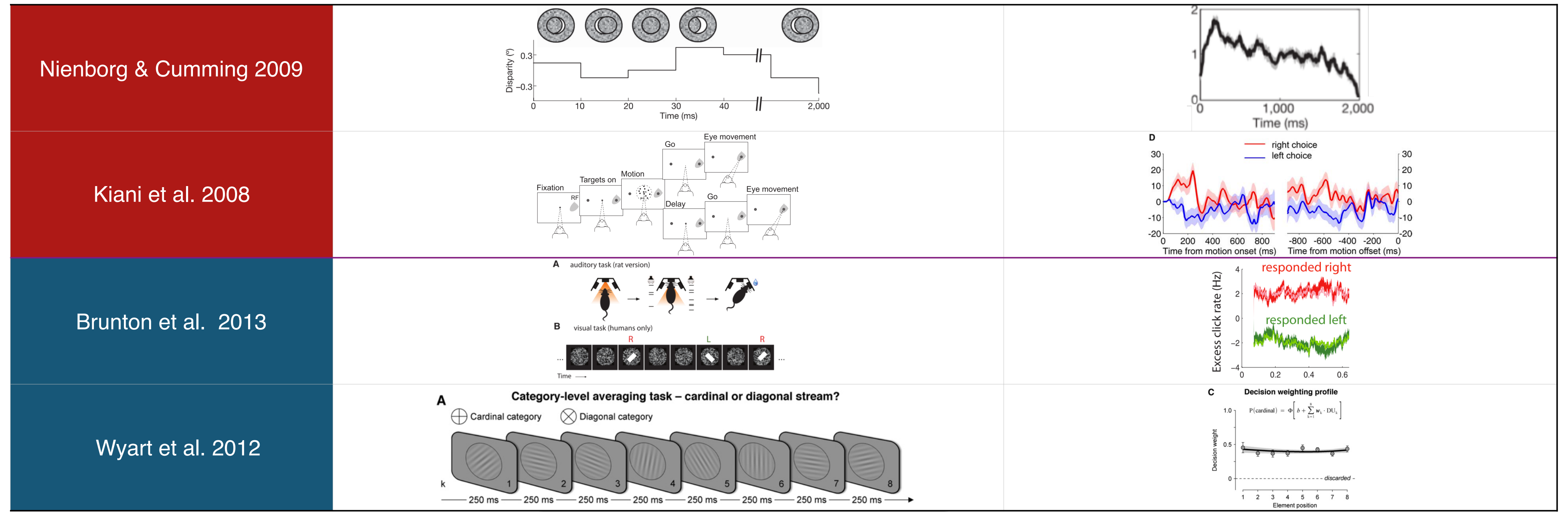
Introduction

In **evidence integration** tasks,subjects make a categorical decision from a sequence of (typically i.i.d.) sensory information.

A **psychophysical kernel (PK)** quantifies the 'weight' subjects give to evidence in space or in time.

A **confirmation bias (CB)** occurs when people upweight information confirming existing beliefs, thus strengthening those beliefs.

A **Perceptual CB** implies a PK that decreases over time.
Different studies have reported different temporal PK shapes, typically flat or decreasing.



Sampling Model[5,6]

Generative model:
C = category / decision-area
x = sensory representation
e = evidence

Goal: compute posterior over **C** given **e**

$$p(C|e_1,\ldots,e_T) \propto p(C)\prod_{t=1}^{T} p(e_t|C)$$

...using **online updates** $p_0(C) \equiv p(C)$

$$\log\frac{p_t(C=+1)}{p_t(C=-1)} \equiv \log\frac{p(C=+1|e_1,\ldots,e_t)}{p(C=-1|e_1,\ldots,e_t)}$$

$$= \log\frac{p_{t-1}(C=+1)}{p_{t-1}(C=-1)} + \boxed{\log\frac{p(e_t|C=+1)}{p(e_t|C=-1)}}$$

update to log posterior odds each frame
(1/S update each sample)

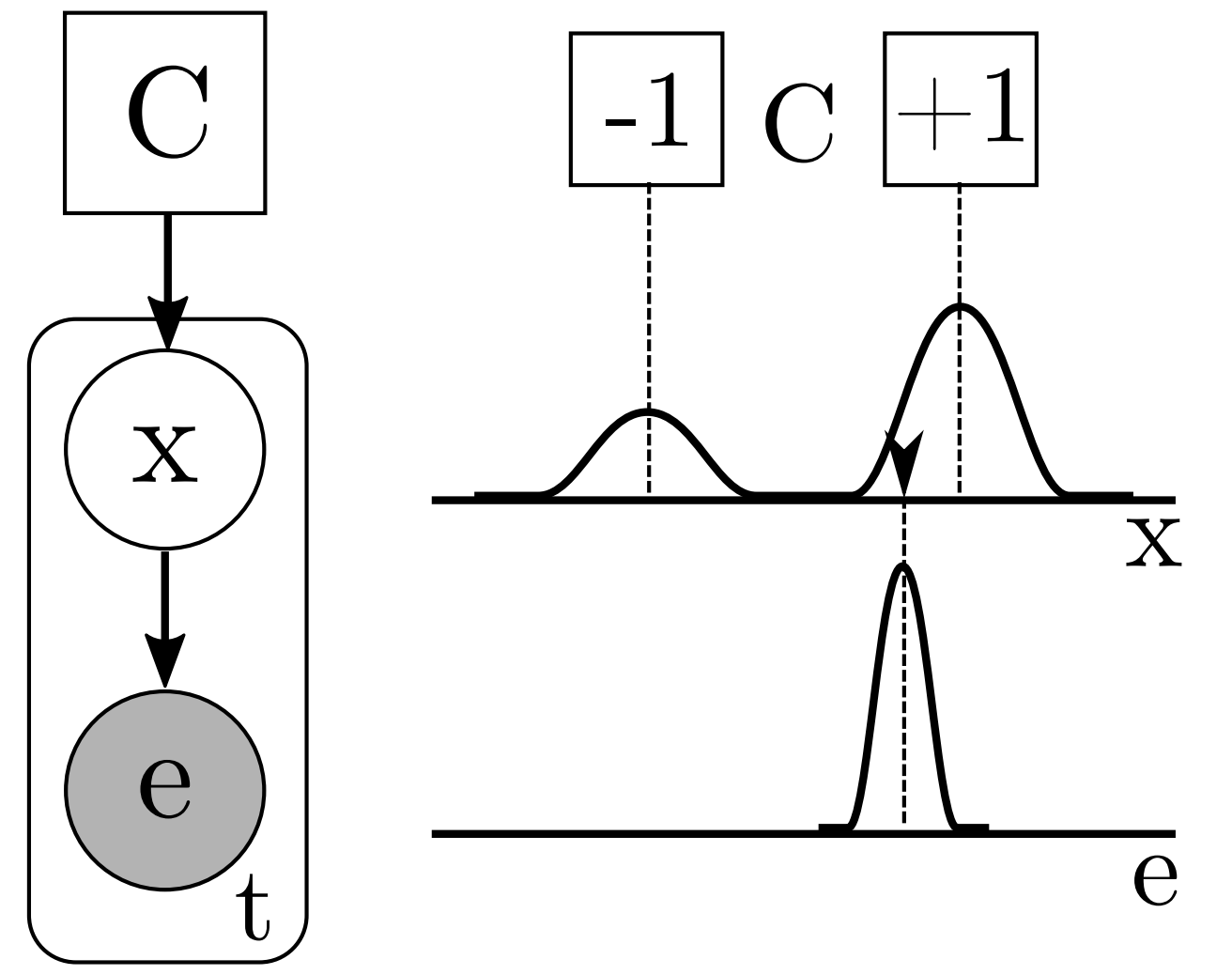
...while **sampling** from the **full posterior** at every moment.

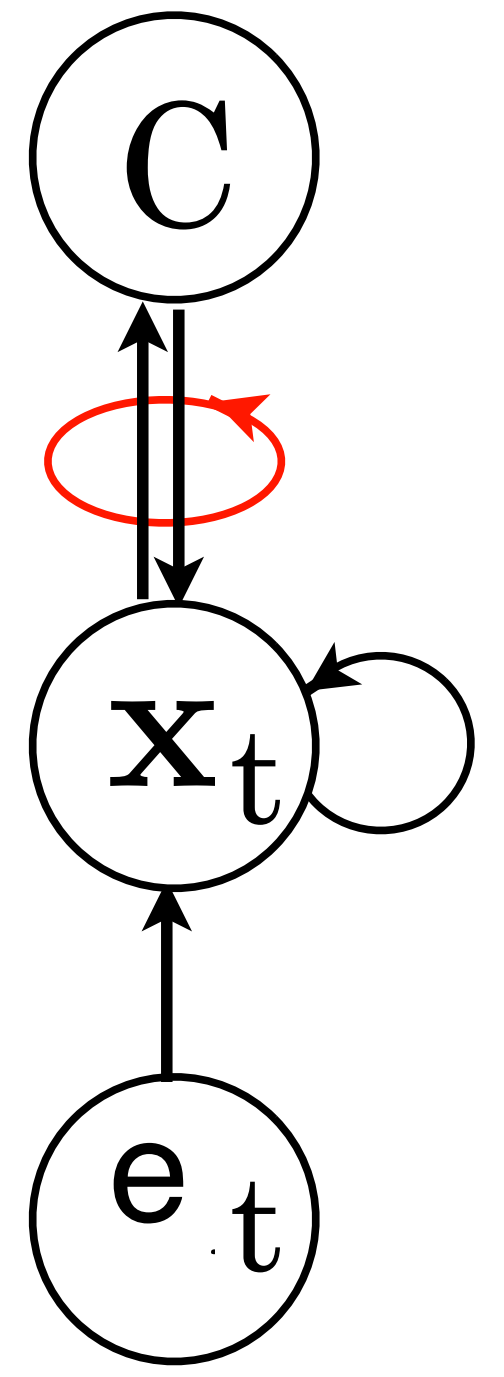
$$\boxed{=\boxed{\log\frac{p_{t-1}(e_t,C=+1)}{p_{t-1}(e_t,C=-1)}} - \boxed{\gamma\log\frac{p_{t-1}(C=+1)}{p_{t-1}(C=-1)}}} \approx \boxed{\log\frac{p(C=+1|x_t^{(i)})}{p(C=-1|x_t^{(i)})}}$$

Subtract prior from the update.

$$0 \leq \gamma \leq 1$$

$$x_t^{(i)} \sim \sum_C p(x_t|e_t,C)p_{t-1}(C)$$





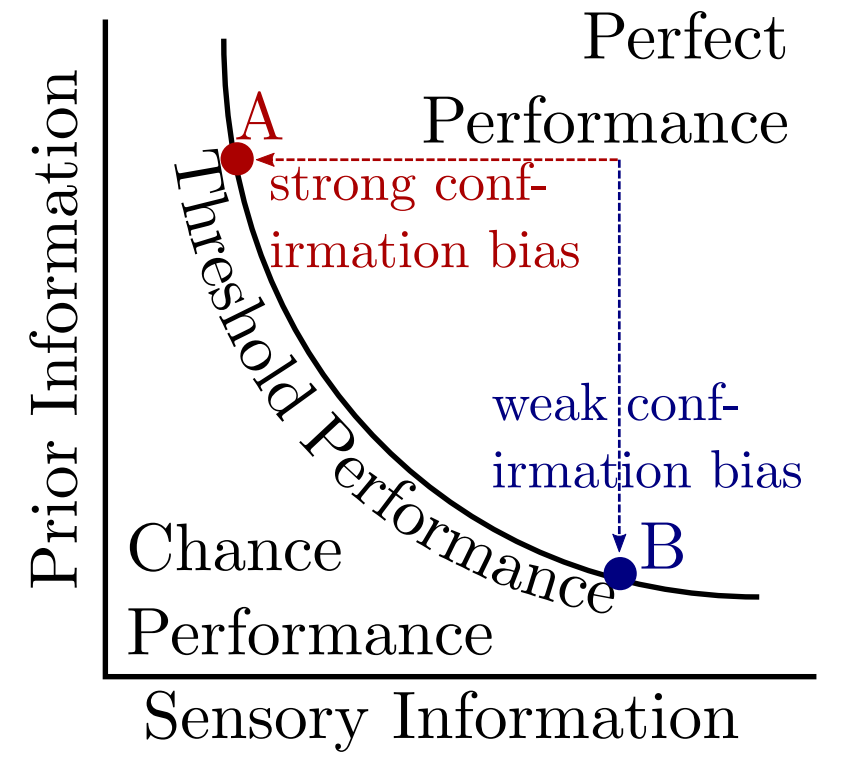
Information flow in our model:
feedforward from the retinal image et to the sensory representation xt and from the sensory representation to the decision area.

Full update rule:

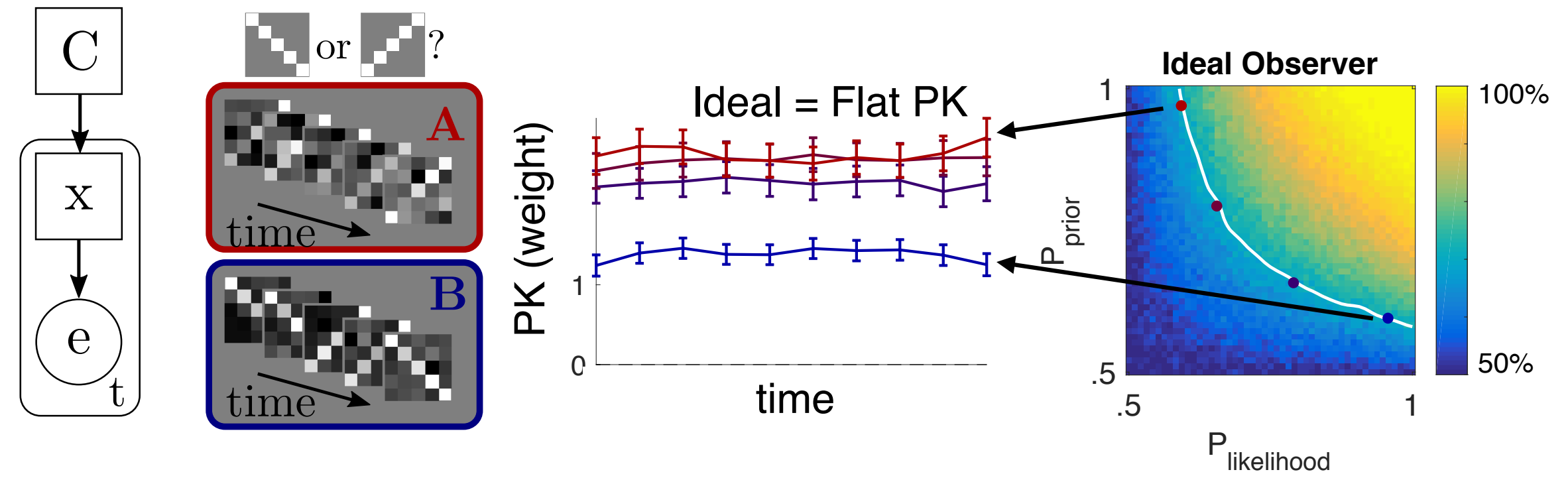
$$\log\frac{p_t^{(i)}(C=+1)}{p_t^{(i)}(C=-1)} = \log\frac{p_t^{(i-1)}(C=+1)}{p_t^{(i-1)}(C=-1)} + \left[\log\frac{p(x_t^{(i)}|C=+1)}{p(x_t^{(i)}|C=+1)} - \gamma\log\frac{p_t^{(i-1)}(C=+1)}{p_t^{(i-1)}(C=-1)}\right]/S$$

Our Framework

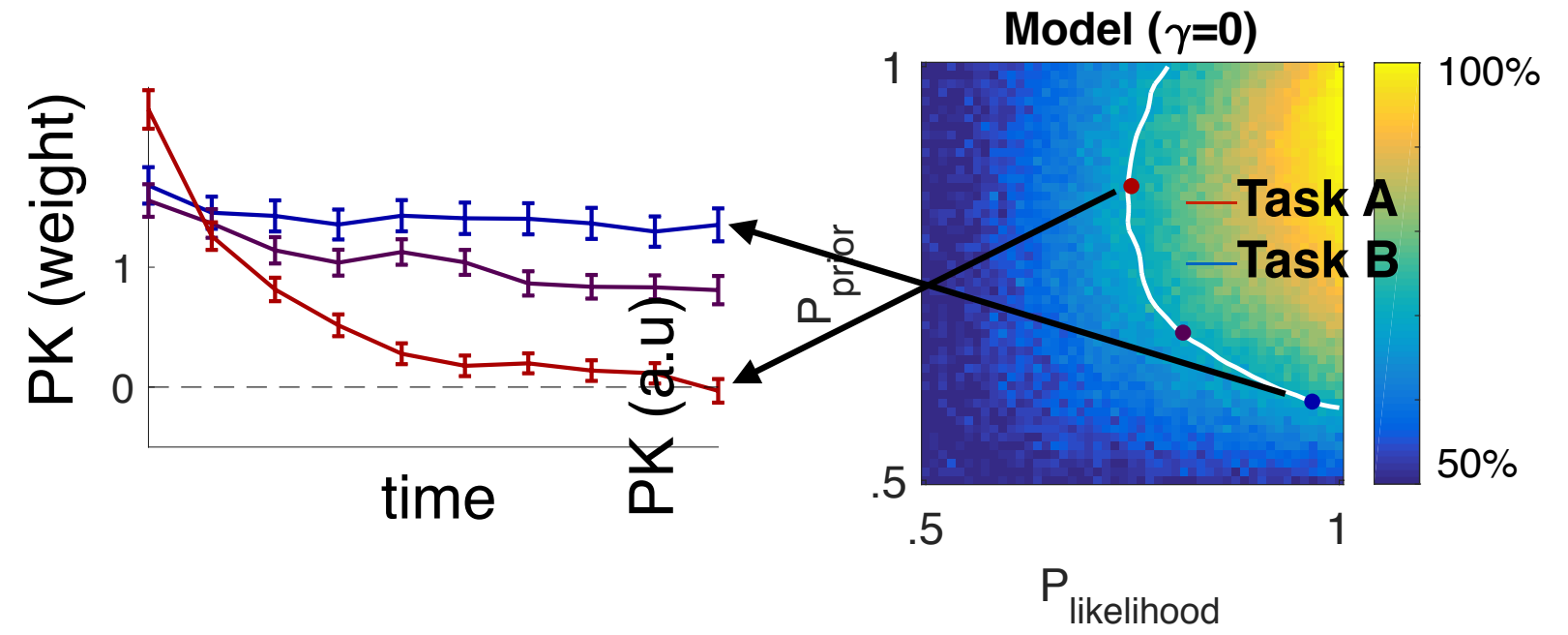
Different **sources of uncertainty:**
With **high-contrast stimuli** that are each **weakly predictive** of the correct choice, **no CB** is observed.
With **low-contrast stimuli** that are each **highly predictive** of the correct choice, a **CB** is observed.
Threshold performance is achieved at a balance between these.



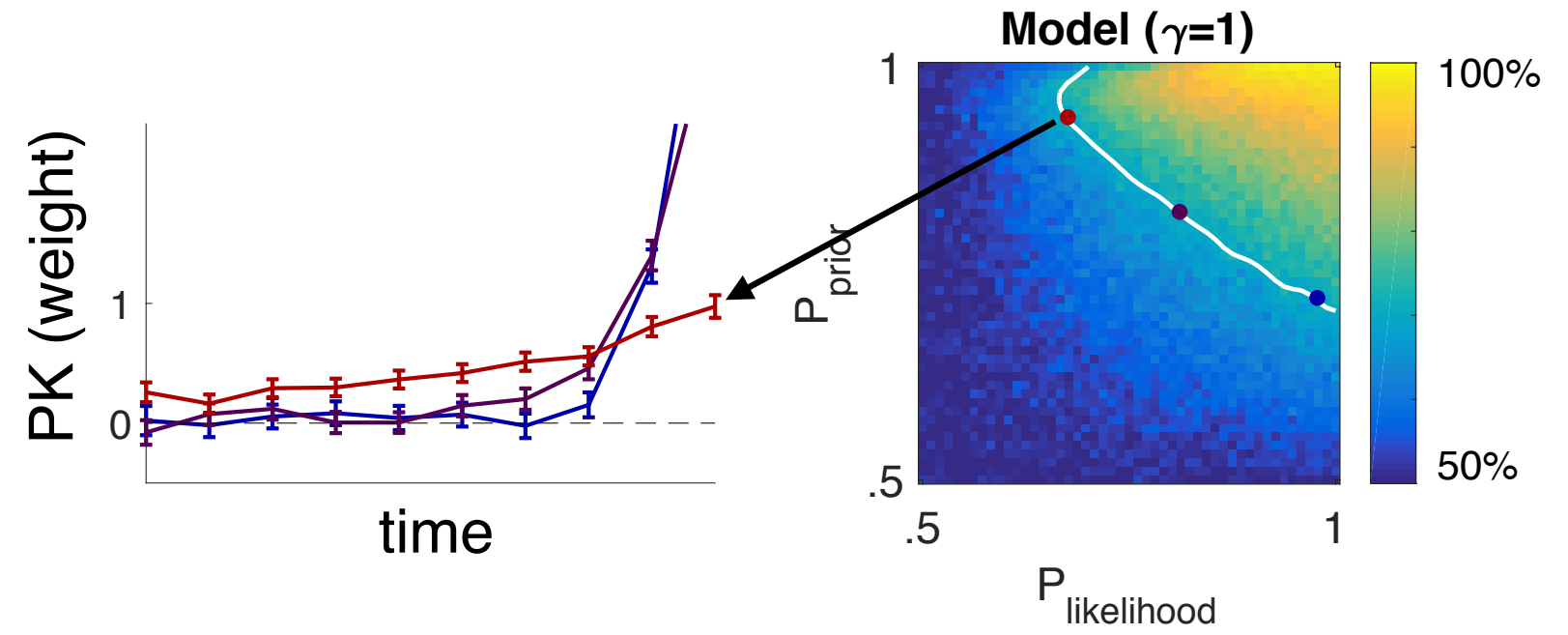
We define **prior information** as the probability of an ideal observer guessing the category of a single 'frame' $\mathbf{x}_t$ given **C**
We define **sensory** or **likelihood information** as the probability of guessing $\mathbf{x}_t$ given $\mathbf{e}_t$



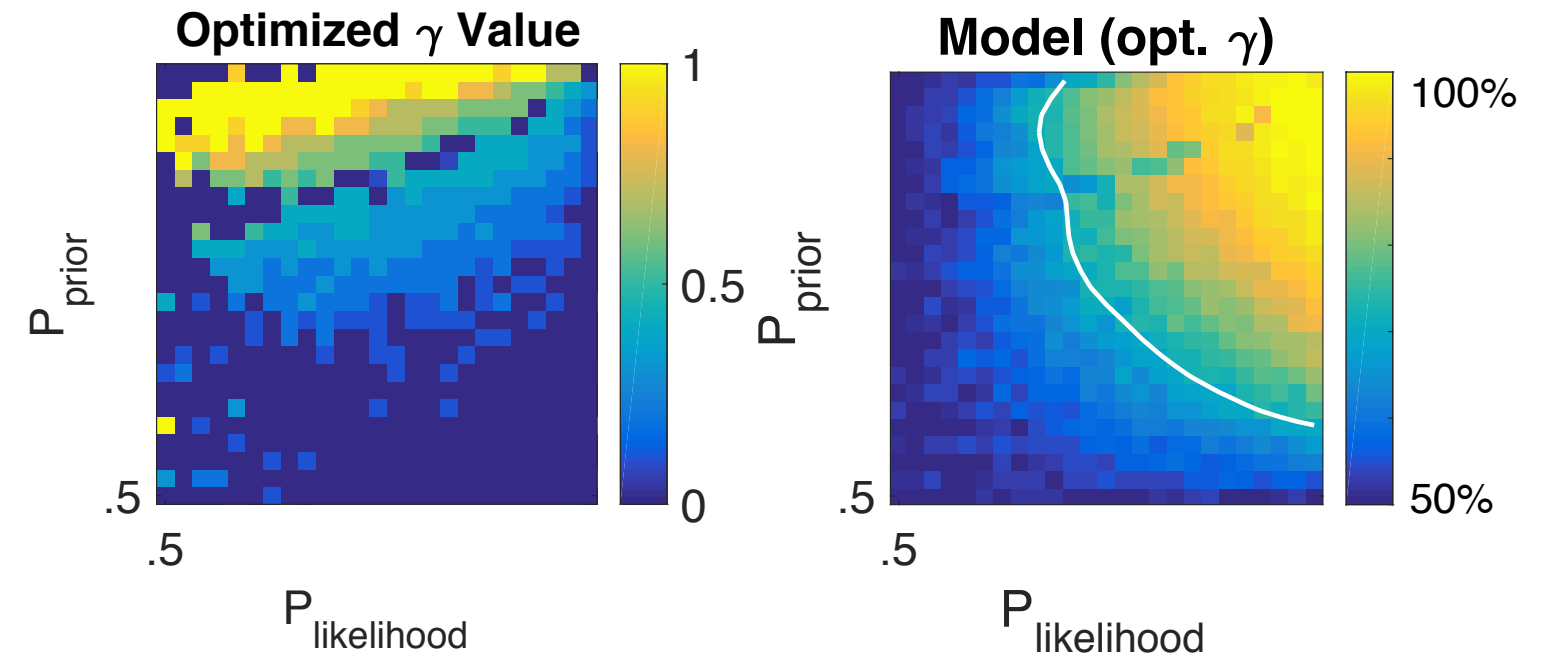
No bias correction (γ=0):
- CB at **task A**
- Ideal at **task B**



Full bias correction (γ=1):
- Recency effect
- Impaired in **task B**



Performance in different tasks can be tuned using different values of γ.
Can subjects adjust γ according to the task?



References

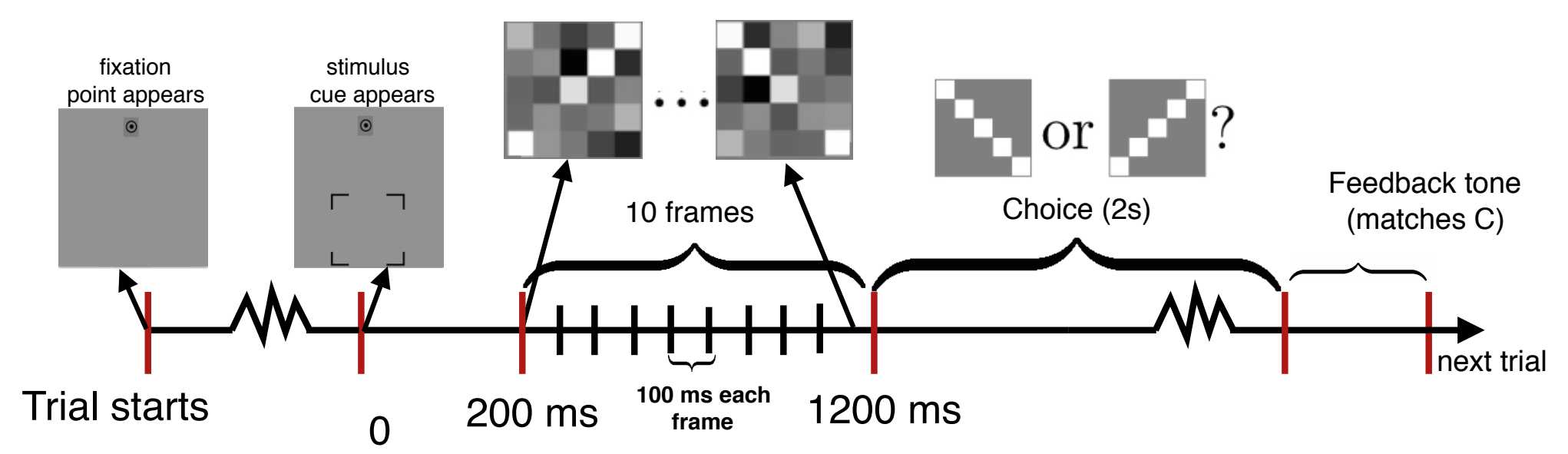
[1] Nienborg, H., & Cumming, B. G. (2009). Decision-related activity in sensory neurons reflects more than a neuron's causal effect. Nature, 459(7243), 89–92.
[2] Kiani, R., Hanks, T. D., & Shadlen, M. N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. The Journal of Neuroscience, 8(12), 3017–3029.
[3] Brunton, B. W., Botvinick, M. M., & Brody, C. D. (2013). Rats and humans can optimally accumulate evidence for decision-making. Science, 340(6128), 95–8.
[4] Wyart, V., Gardelle, V. De, Scholl, J., & Summerfield, C. (2012). Rhythmic Fluctuations in Evidence Accumulation during Decision Making in the Human Brain. Neuron, 76(4), 847–858.
[5] Haefner, R. M., Berkes, P., & Fiser, J. (2016). Perceptual Decision-Making as Probabilistic Inference by Neural Sampling. Neuron, 90, 649–660.
[6] Fiser, J. J., Berkes, P., Orbán, G., & Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. Trends in Cognitive Sciences, 14(3), 119–30.

Experiments

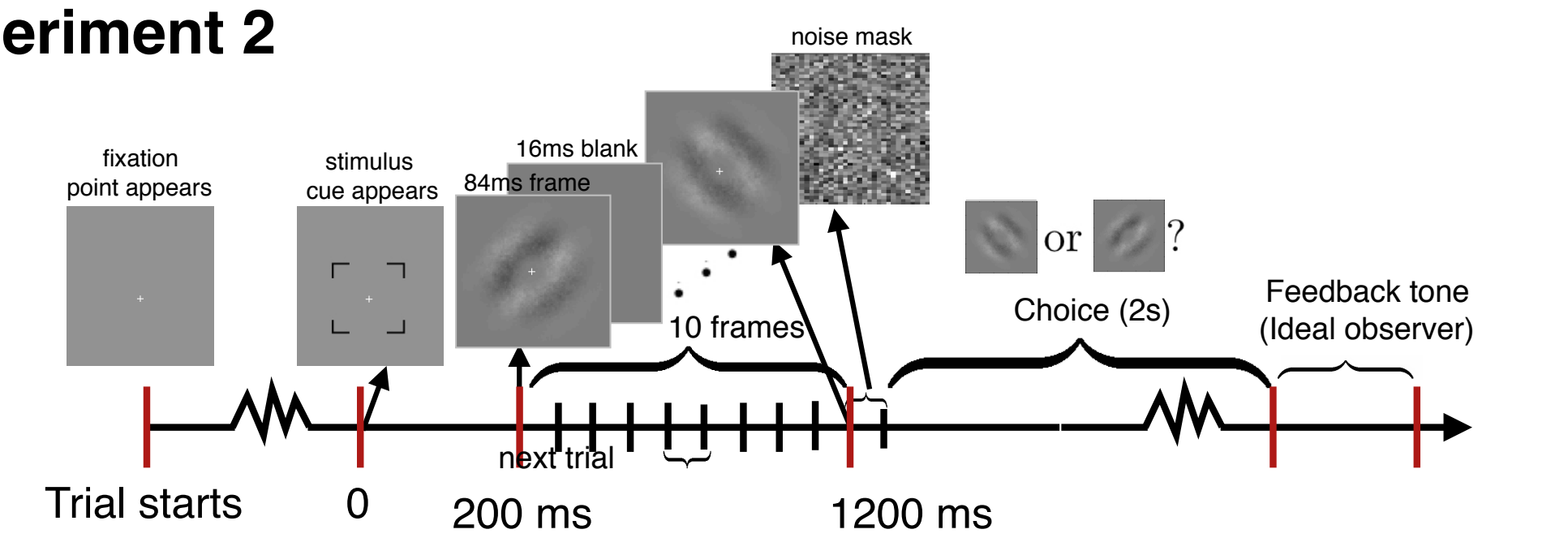
We used an orientation-discrimination paradigm with templates embedded in noise.
- In **task A**, a staircase on the template **contrast** is run to reach 70% performance
- In **task B**, a staircase is run on $\mathbf{p}_{\mathbf{prior}}$
- Both tasks' staircases begin at the same set of parameters.

Change in PK slopes consistent with our framework's predictions, but significant variability between subjects.

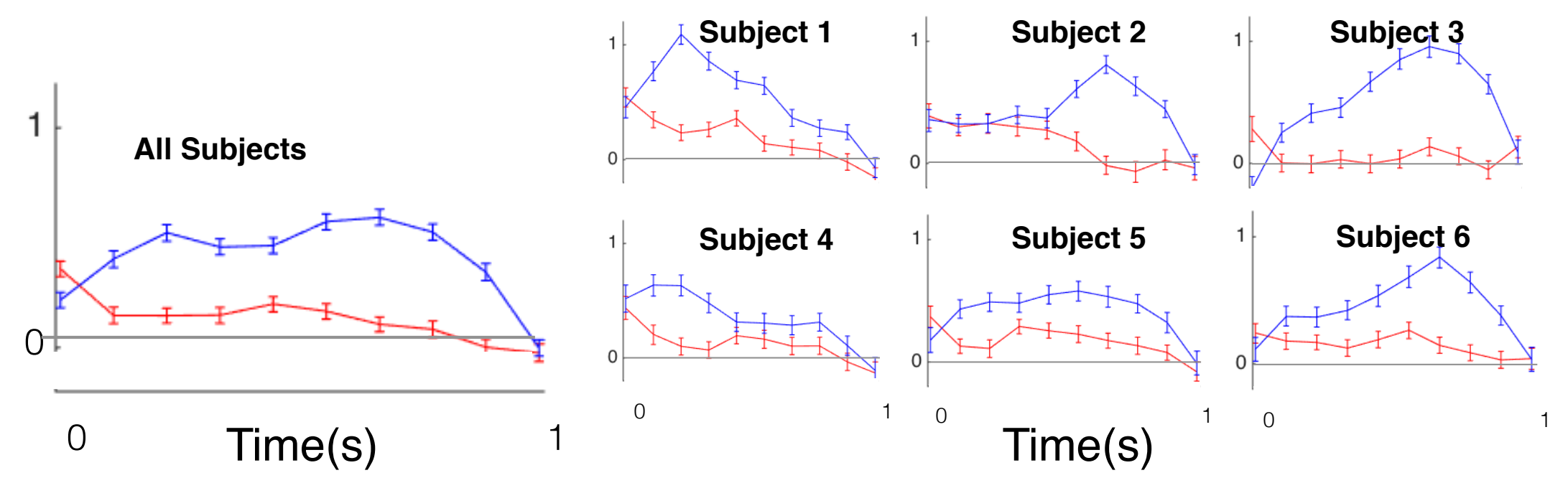
Experiment 1



Experiment 2



Data (Experiment 1)



Summary

- Discrepancy in studies reporting evidence-weighting over time, unexplainable by ideal observer.
- We introduce a 2-dimensional categorization of tasks, separating stimulus-to-sensory and sensory-to-decision correlations.
- Approximate inference by sampling can explain published results.
- Model makes neurophysiological predictions for influence of evolving confidence in decision on sensory area
- Our preliminary experimental results:
  - Test framework within-subject, with a single modality & paradigm.
  - Have qualitative agreement with the model and published results.

Questions

- What paradigm is most likely to obtain flat PK for high contrast?
- Is feedback best probabilistic or based on ideal observer?
- Can gamma explain subject variability?