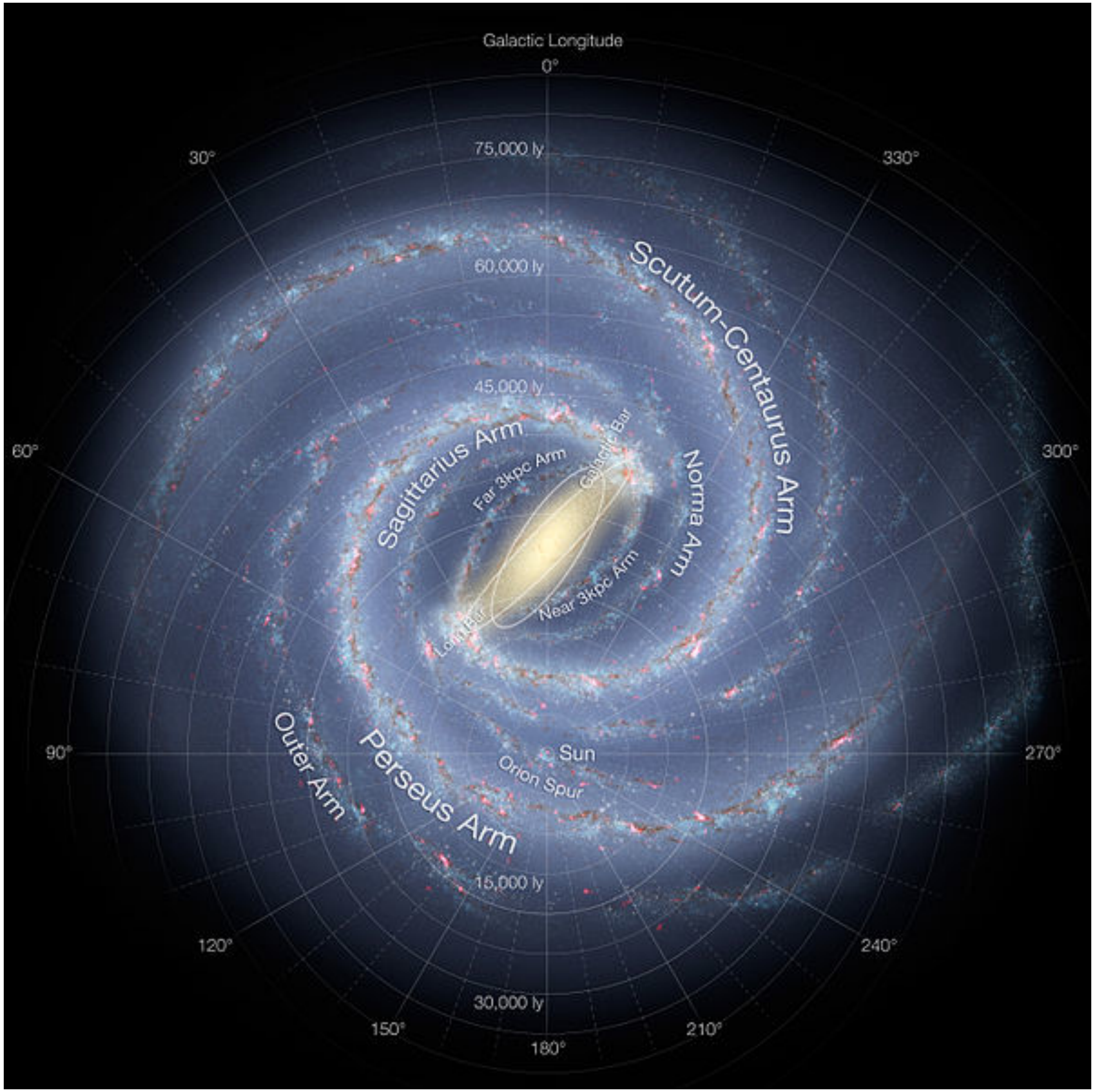
Top-down signals in cortex & sampling

Ralf M. Haefner
Brain & Cognitive Sciences, University of Rochester (NY)

Oct. 20, 2017: Ralf Haefner, Ph.D.

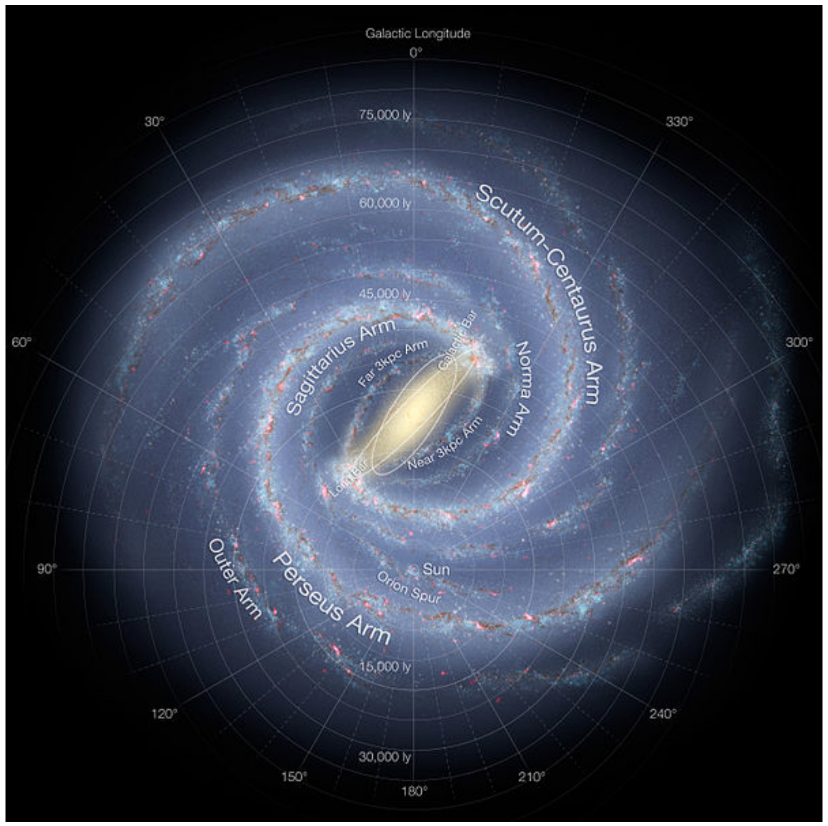
Conference room Alkek N315 at 11 a.m. Host lab Xaq Pitkow. Dr. Haefner is an Assistant Professor of Physics and Astronomy at the University of Rochester in Rochester, New York.

Our Galaxy

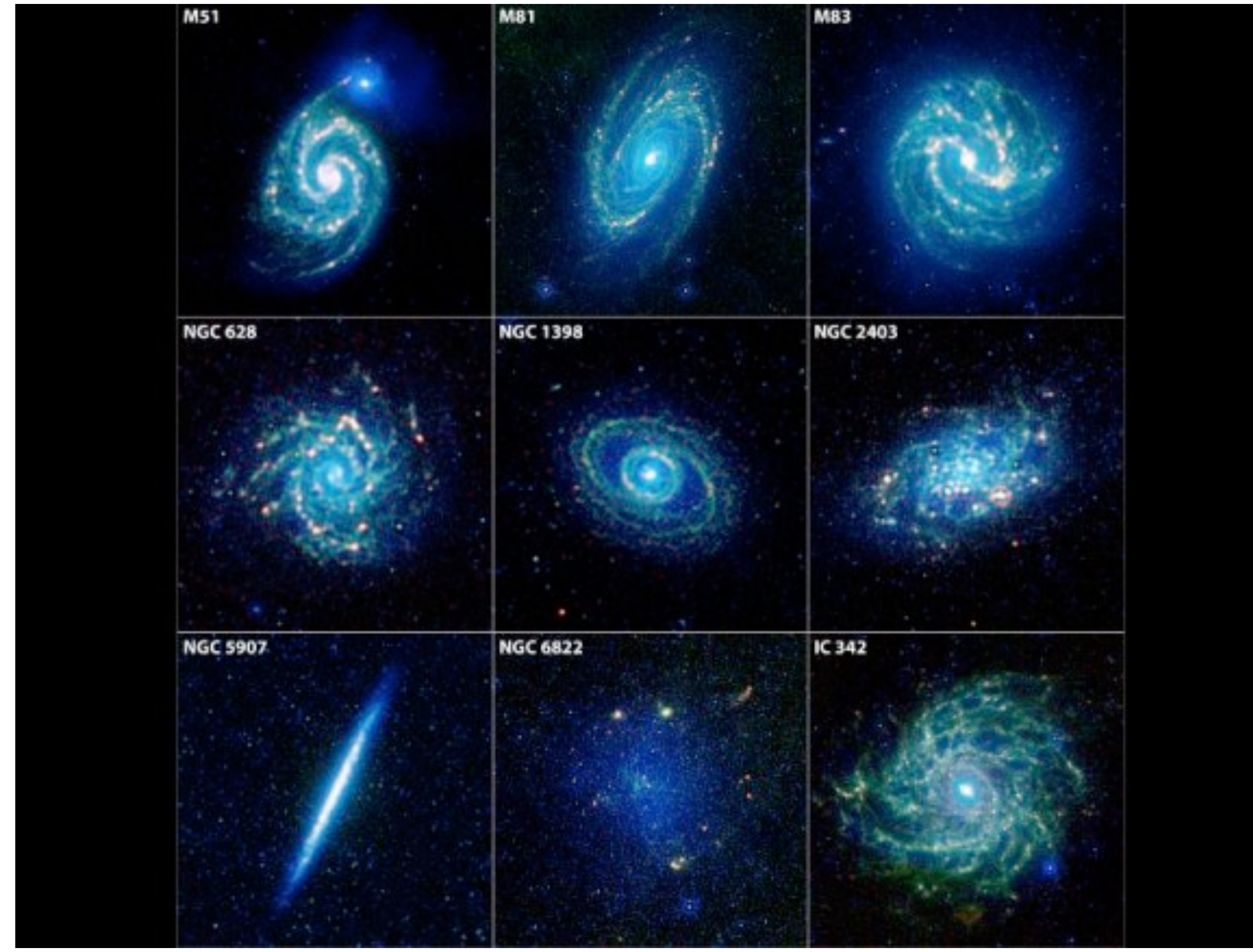


Galactic Longitude
0°
30°
330°
60°
300°
90°
270°
120°
240°
150°
210°
180°
75,000 ly
60,000 ly
45,000 ly
15,000 ly
30,000 ly
Scutum-Centaurus Arm
Sagittarius Arm
Far 3kpc Arm
Galactic Bar
Norma Arm
Near 3kpc Arm
Long Bar
Outer Arm
Perseus Arm
Sun
Orion Spur

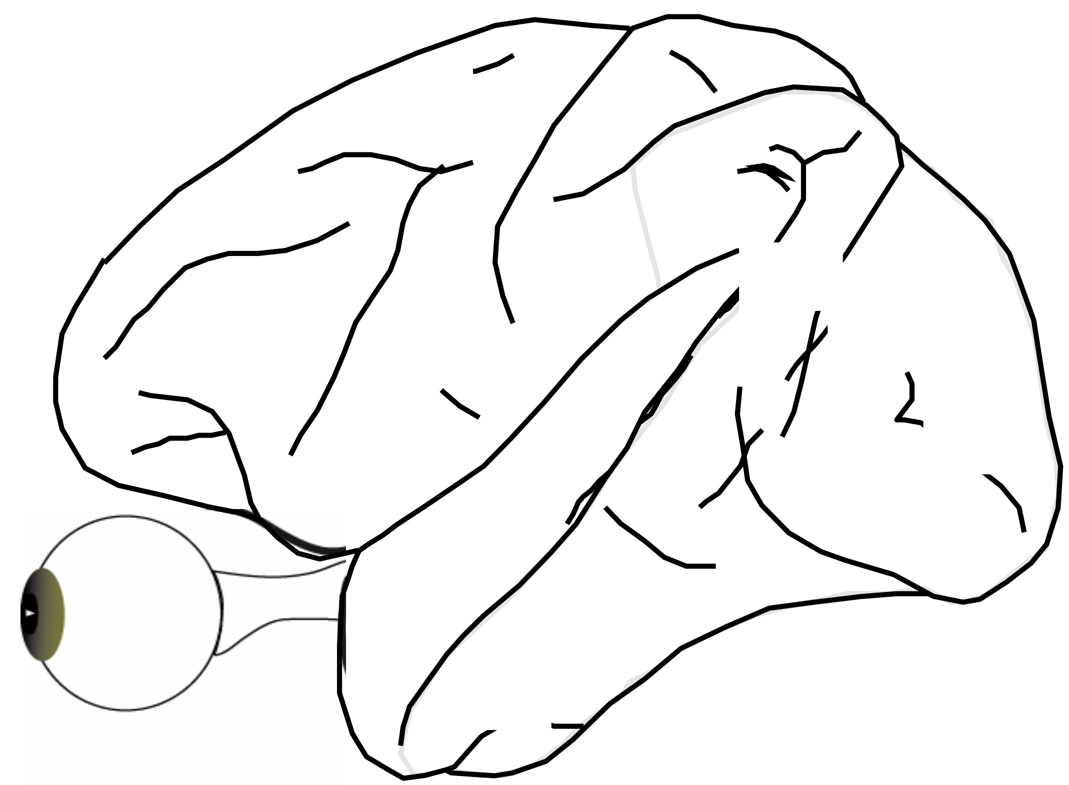
$$F = \gamma \frac{m_1 m_2}{r^2}$$



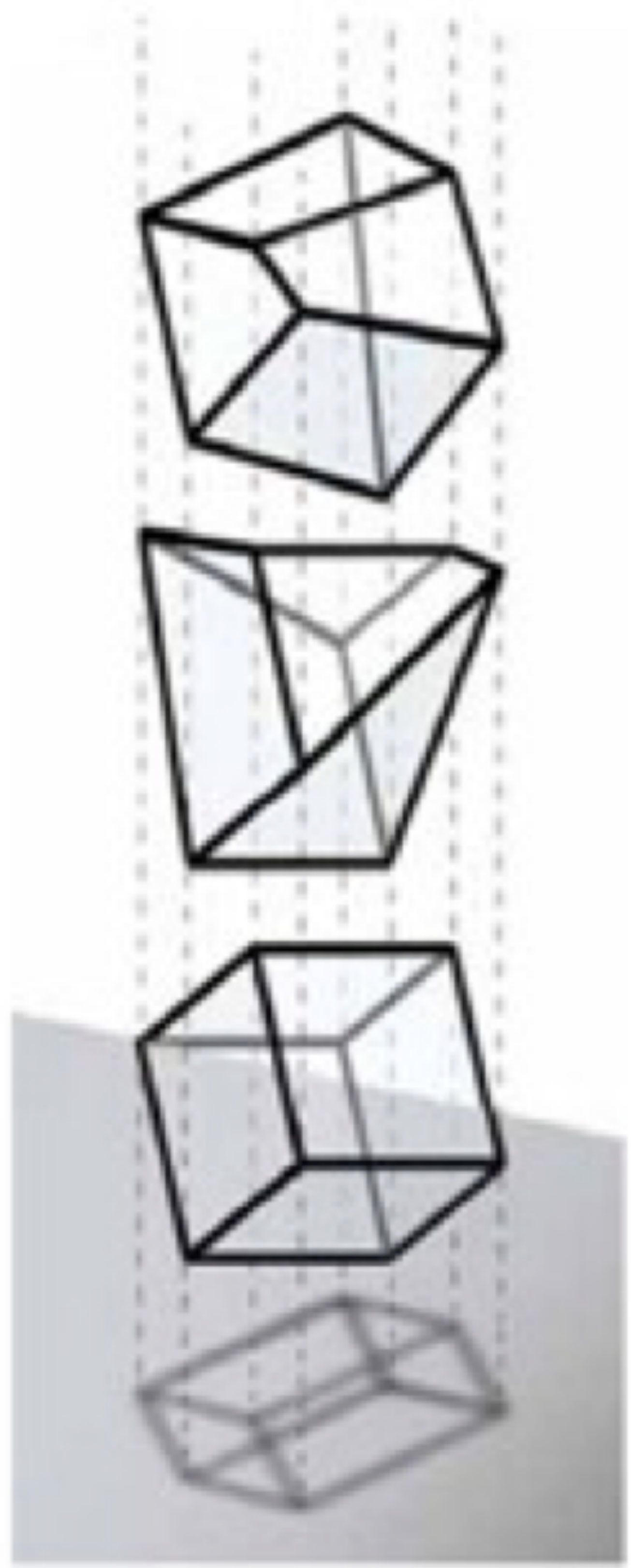
Prior knowledge



Sensory experience



Helmholtz/Mach: Vision as inference



Unlikely

Unlikely

We perceive the most probable cause for the image on the eye

Projection onto retina *I*

Marr’s level of analysis

Computational

Probabilistic inference

Talk Part 2:
Implications for a behavioral bias

Representation/Algorithm

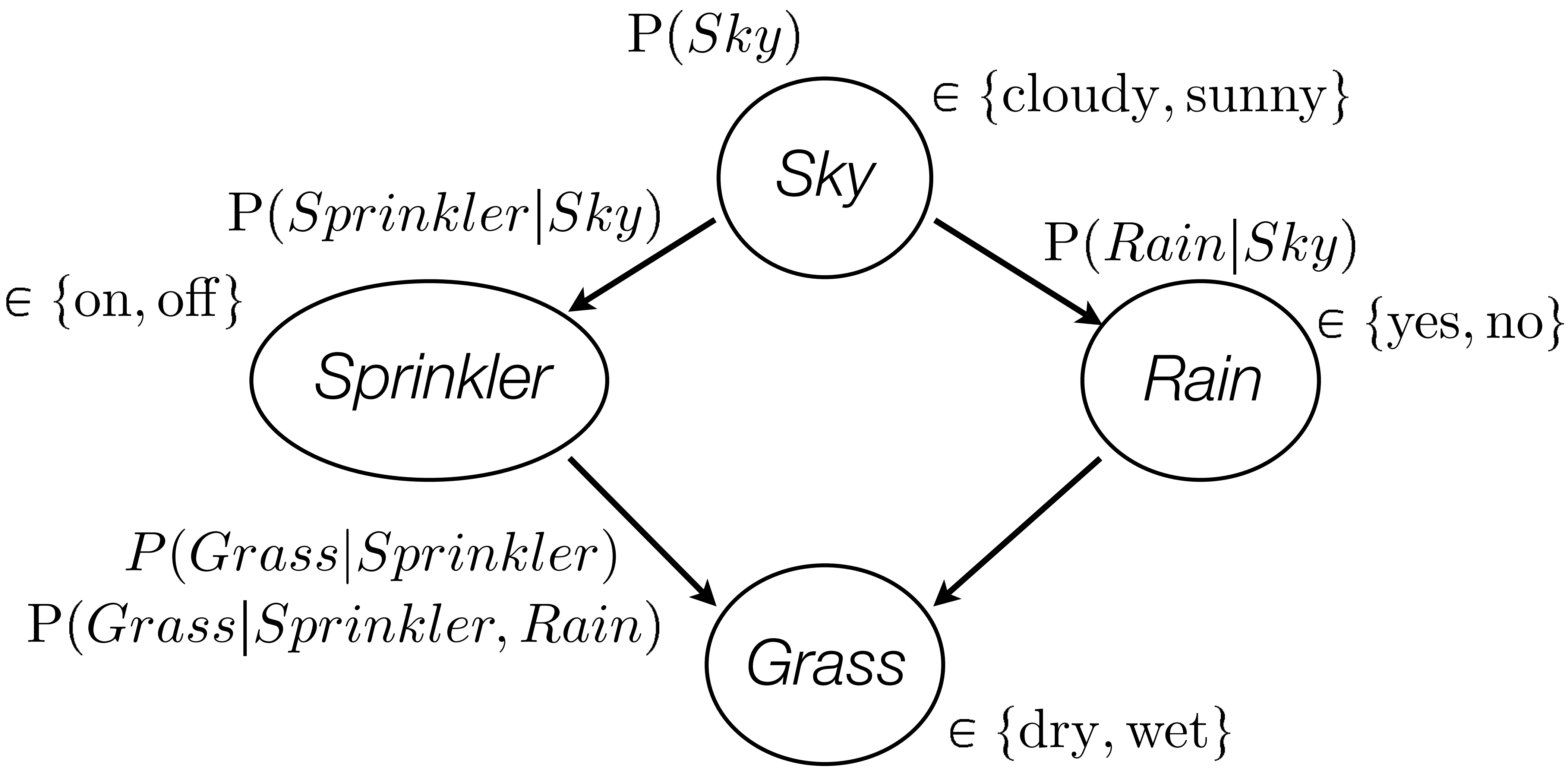
MCMC sampling

Talk Part 1:
Implications for neural responses

Implementation

Spike counts
Spikes
Membrane potentials

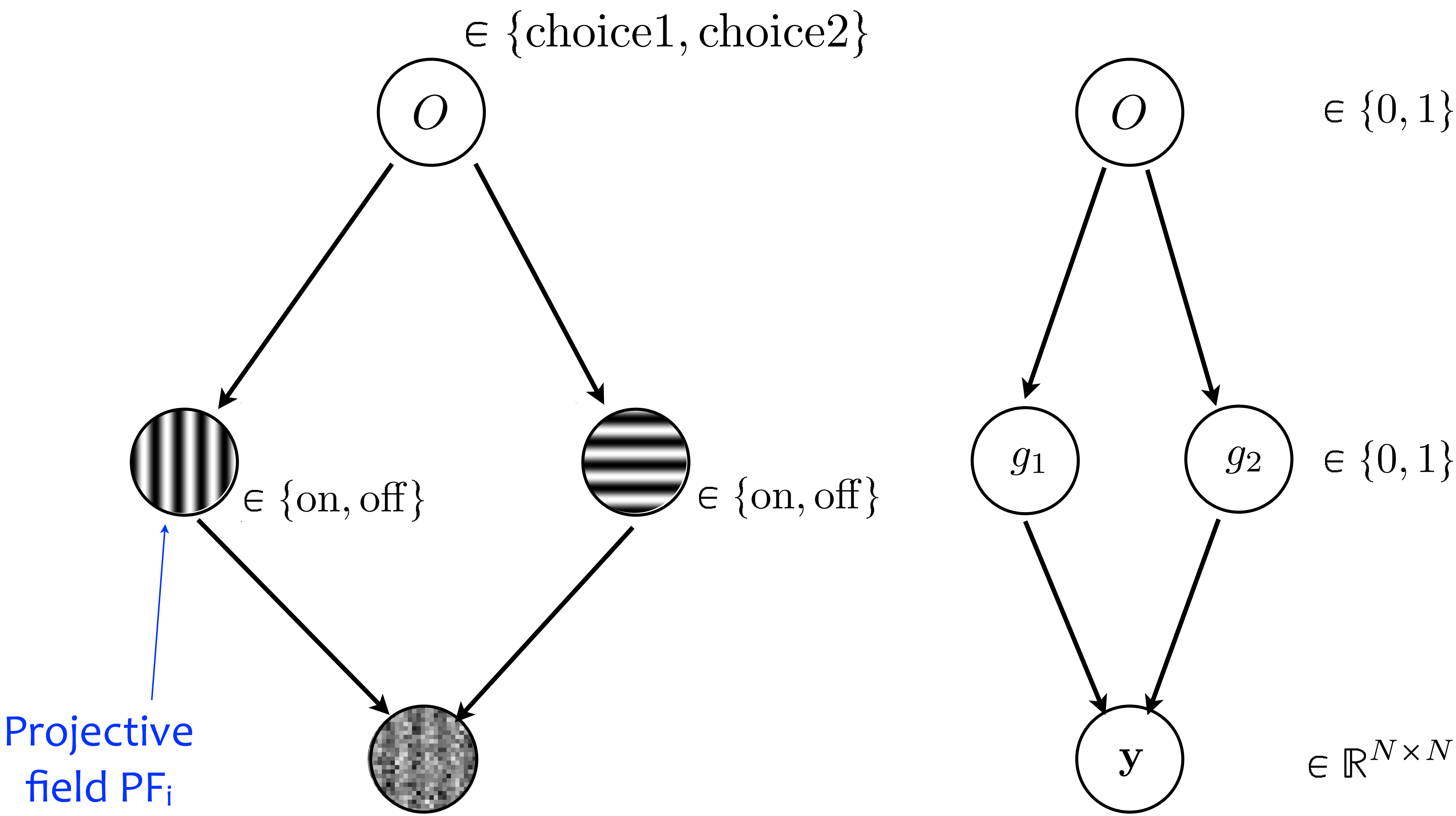
Bayesian network/generative model



$$\mathrm{P}(Sky, Sprinkler, Rain, Grass) = \mathrm{P}(Sky)\mathrm{P}(Sprinkler|Sky)\mathrm{P}(Rain|Sky)\mathrm{P}(Grass|Sprinkler, Rain)$$

- Inference: Act of computing unknown/latent/hidden variables, given observed variables

Orientation discrimination task



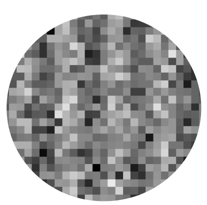
$$p(\mathbf{y}|\mathbf{g}) = \mathcal{N}\left(\mathbf{y} \,:\, \sum_i \mathrm{PF}_i g_i \,,\, \sigma_y^2\right)$$

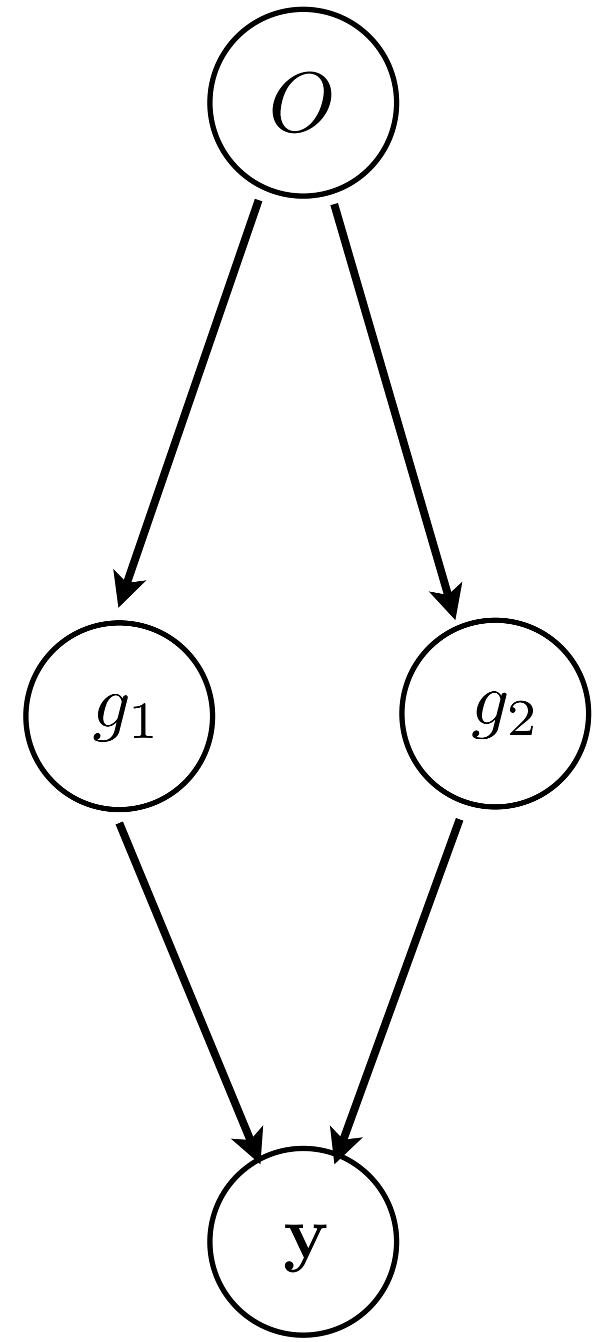
Inference by sampling

$O \sim p(O|g_1, g_2)$ 2 1 1 2 2 2 1 1

$g_1 \sim p(g_1|O, g_2, \mathbf{y})$ 1 1 1 1 0 1 1 1 ...

$g_2 \sim p(g_2|O, g_1, \mathbf{y})$ 0 0 1 1 1 1 1 0

$\mathbf{y} =$ 



$\in \{0, 1\}$

$\in \{0, 1\}$

$\in \mathbb{R}^{N \times N}$

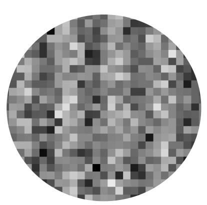
$$p(\mathbf{y}|\mathbf{g}) = \mathcal{N}\left(\mathbf{y} \,:\, \sum_i \mathrm{PF}_i g_i \,,\, \sigma_y^2\right)$$

Inference by sampling

$O \sim p(O|g_1, g_2)$ 2 1 1 2 2 2 1 1 1 1 1 1 $p(O|\mathbf{y})$

$g_1 \sim p(g_1|O, g_2, \mathbf{y})$ 1 1 1 1 0 1 1 1 1 1 1 1 $p(g_1|\mathbf{y})$

$g_2 \sim p(g_2|O, g_1, \mathbf{y})$ 0 0 1 1 1 1 1 0 0 0 0 0 $p(g_2|\mathbf{y})$

$\mathbf{y} =$ 

$$p(O, g_1, g_2|\mathbf{y})$$

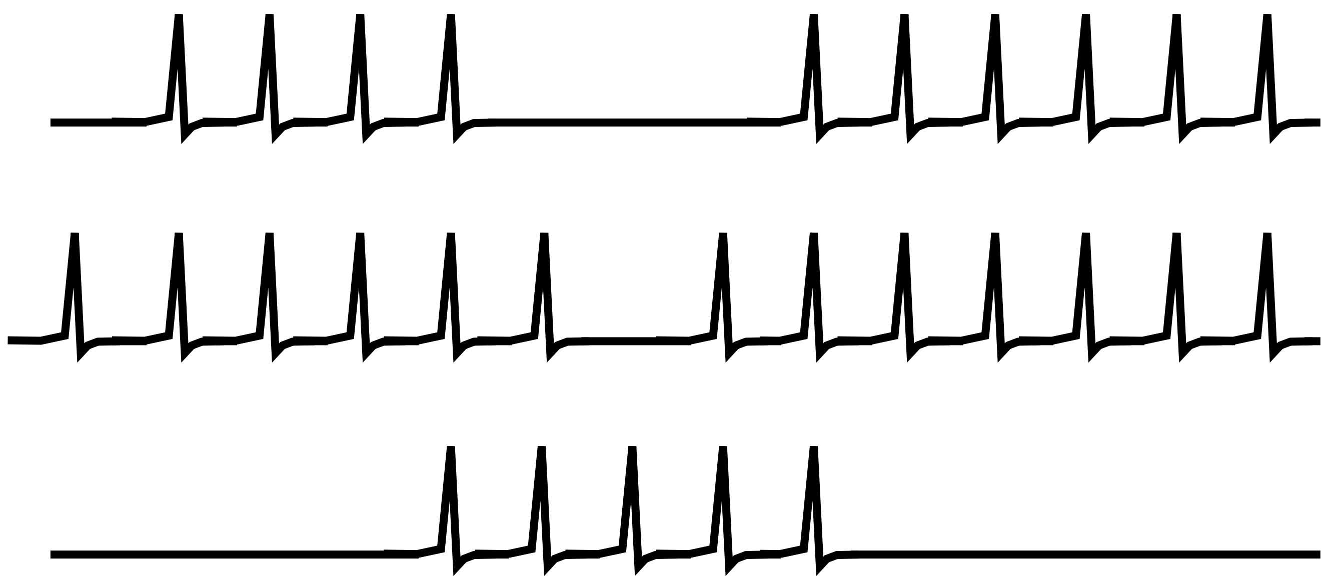
Neural sampling hypothesis

Fiser et al, 2010

$O \sim p(O|g_1, g_2)$

$g_1 \sim p(g_1|O, g_2, \mathbf{y})$

$g_2 \sim p(g_2|O, g_1, \mathbf{y})$



$p(O|\mathbf{y})$

$p(g_1|\mathbf{y})$

$p(g_2|\mathbf{y})$

Realizable in real neurons

(Buesing et al., 2011, Pecevski et al., 2011)

- continuous time
- asynchronous spiking
- refractory period

Implications:

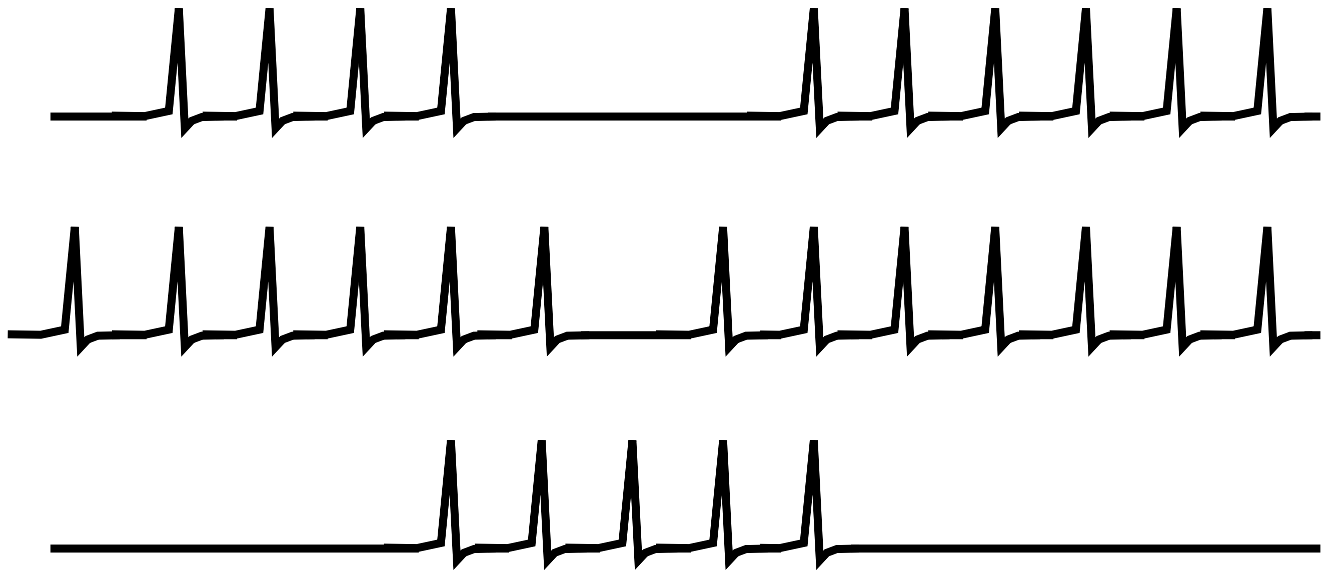
- (Near-)Poisson variability
- Receptive fields
- Tuning curves

Predictions

$O \sim p(O|g_1, g_2)$

$g_1 \sim p(g_1|O, g_2, \mathbf{y})$

$g_2 \sim p(g_2|O, g_1, \mathbf{y})$

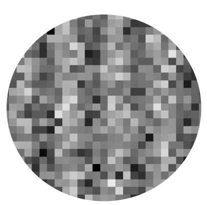


$p(O|\mathbf{y})$

$p(g_1|\mathbf{y})$

$p(g_1, g_2|\mathbf{y})$

$p(g_2|\mathbf{y})$

$\mathbf{y} =$ 

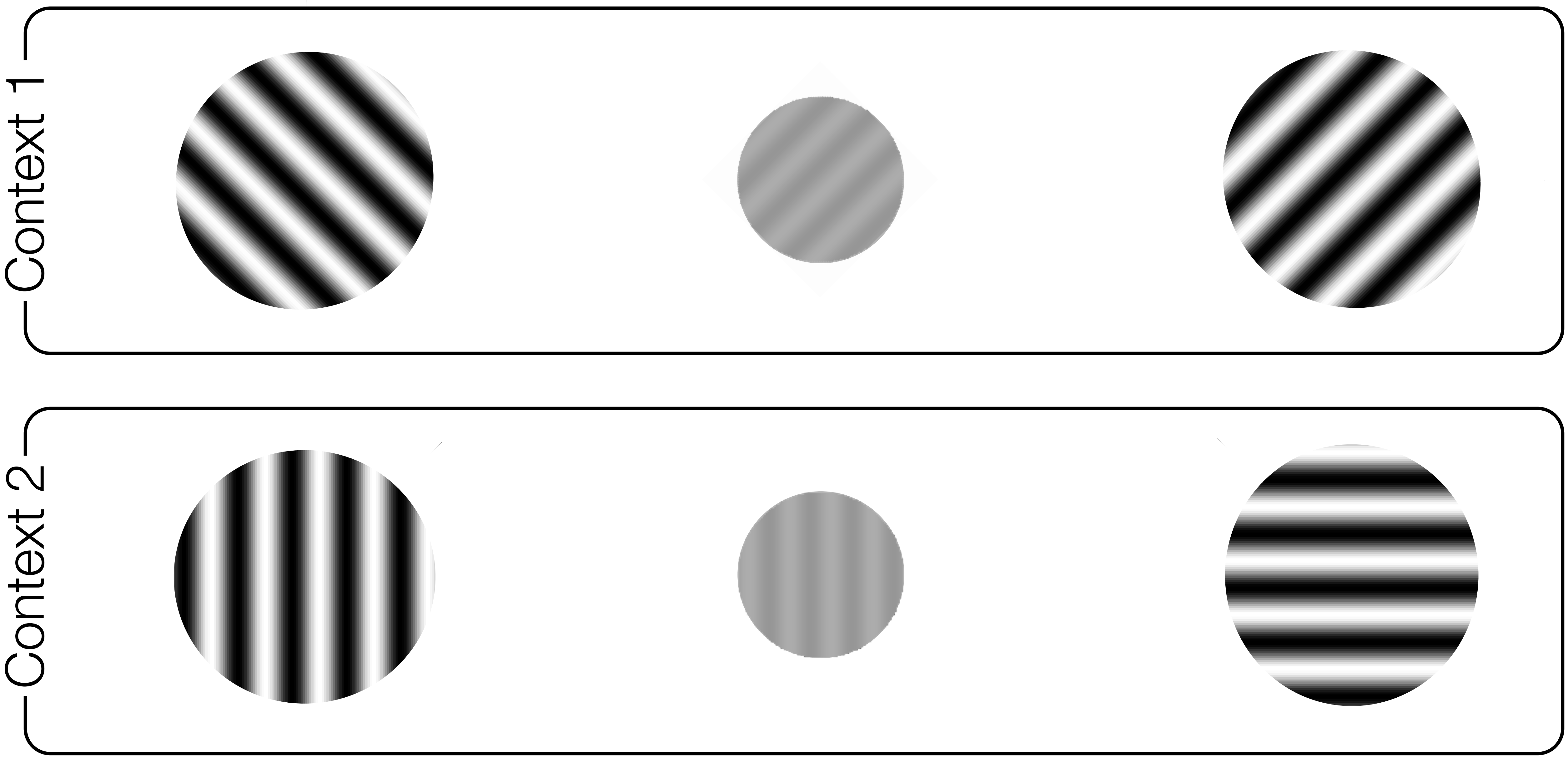
Noise Correlations...

...and their task-dependence

Choice Correlations and Psychophysical Kernels

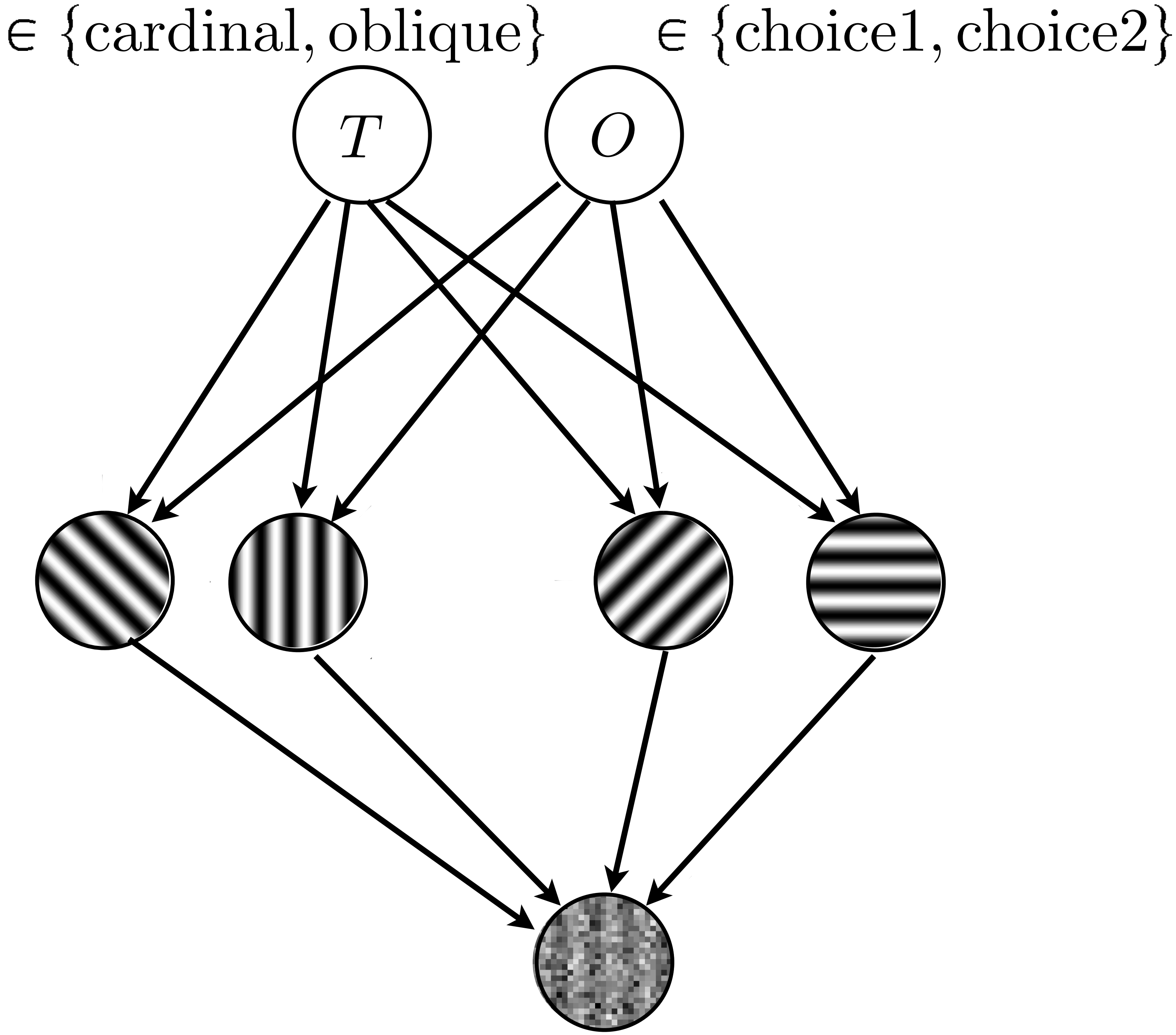
Orientation discrimination task

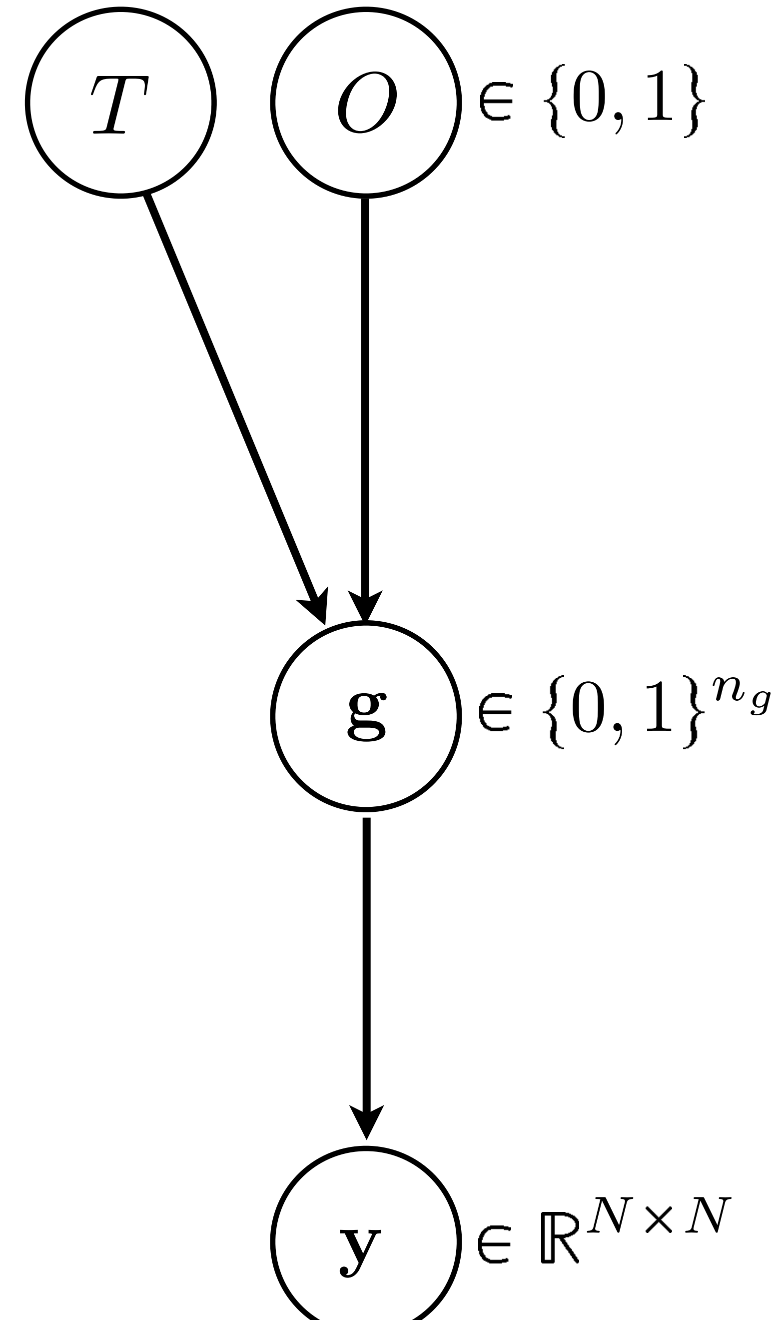
Which of two perpendicular gratings
caused the noisy image on the screen?



Inspired by Cohen & Newsome, Neuron 2008: Motion direction discrimination task/MT

2AFC task: generative model

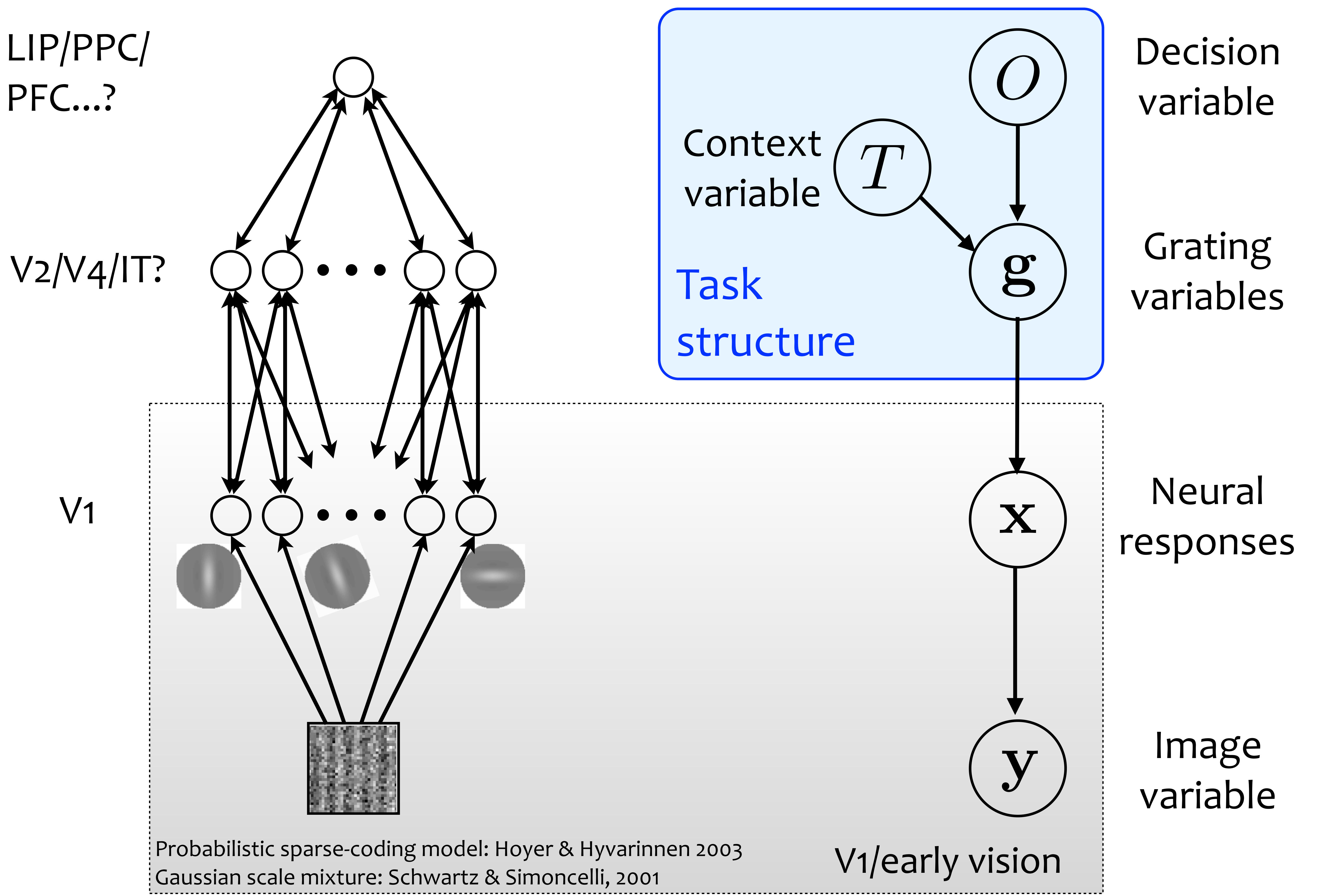




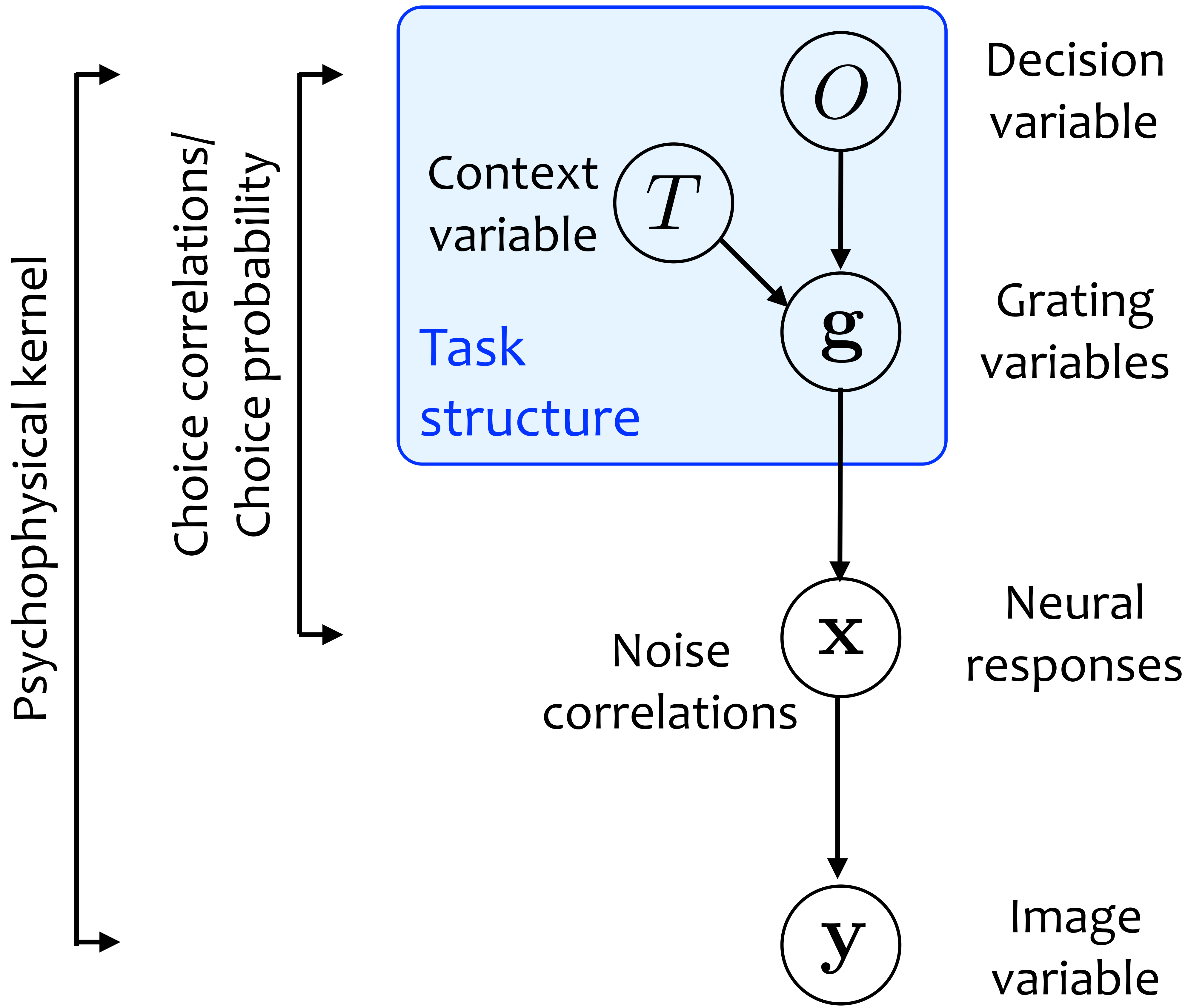
$$p(\mathbf{y}|\mathbf{g}) = \mathcal{N}\left(\mathbf{y} \,:\, \sum_i \mathrm{PF}_i g_i \,,\, \sigma_y^2\right)$$

Haefner, Berkes & Fiser, Neuron 2016

Correspondence to the visual system

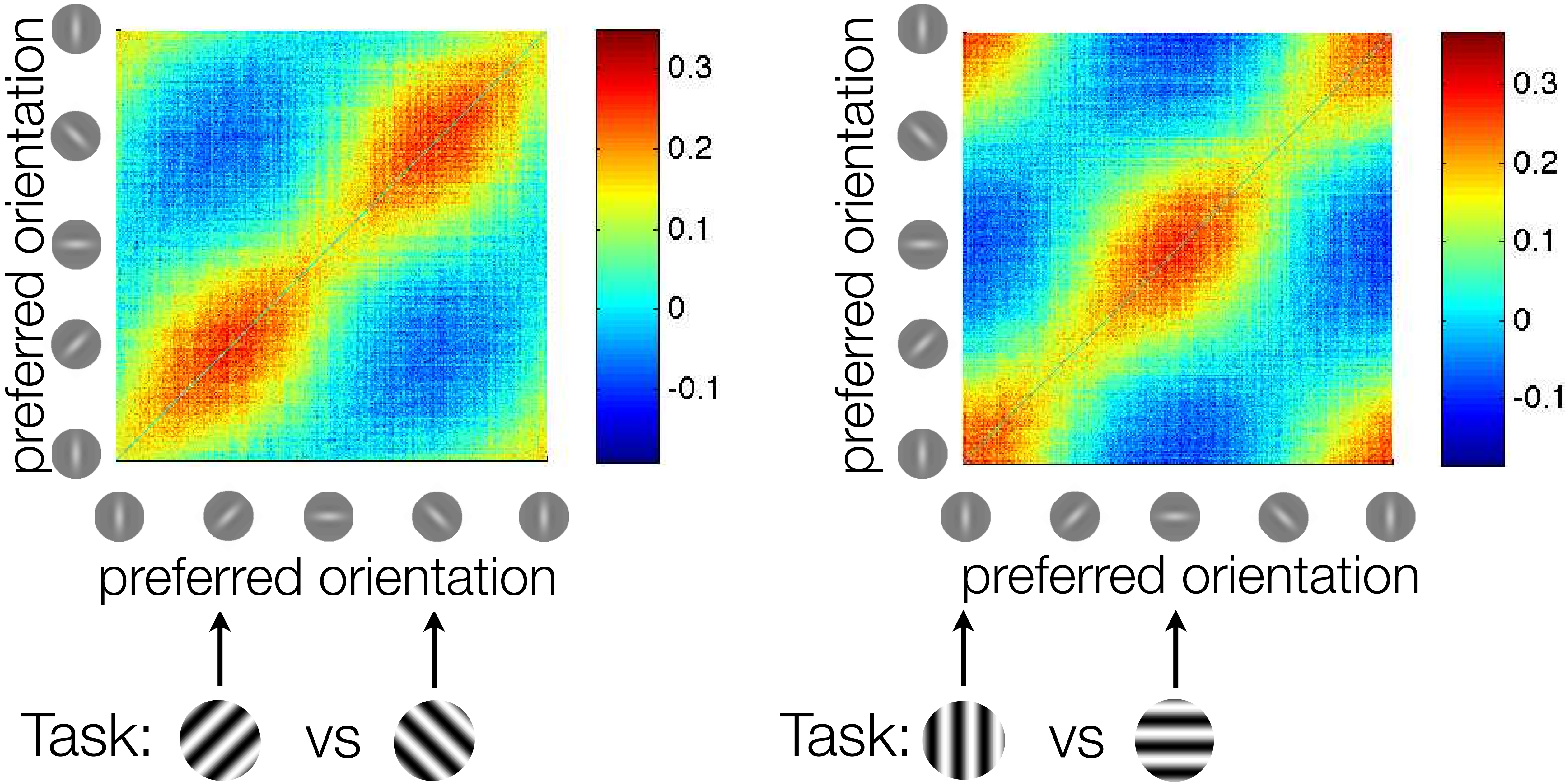


Dependencies and observables



Haefner, Berkes & Fiser, Neuron 2016

Model: noise correlations in $\mathbf{x}$

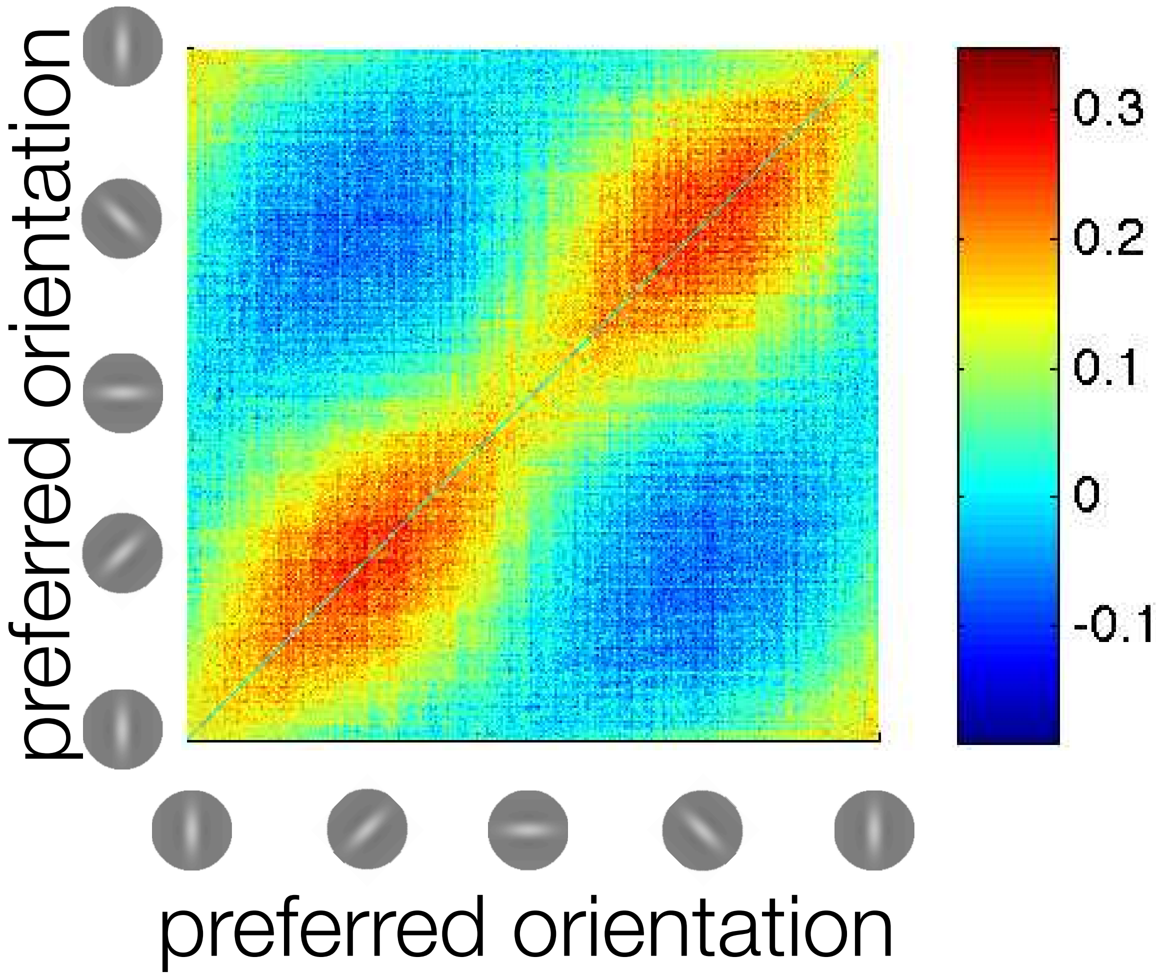


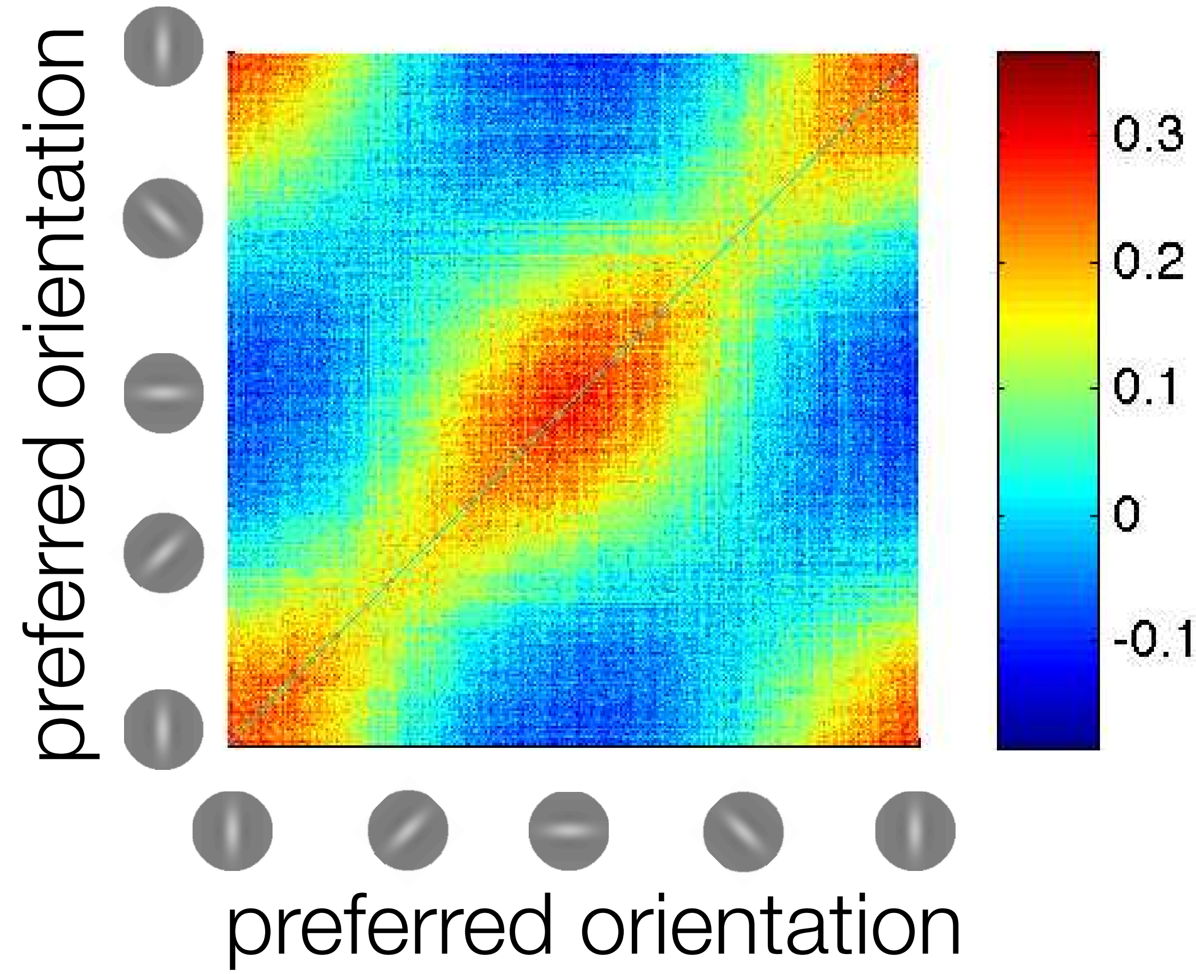
Context 1

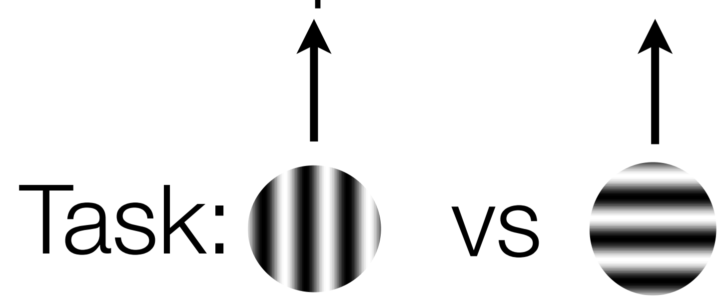
Context 2

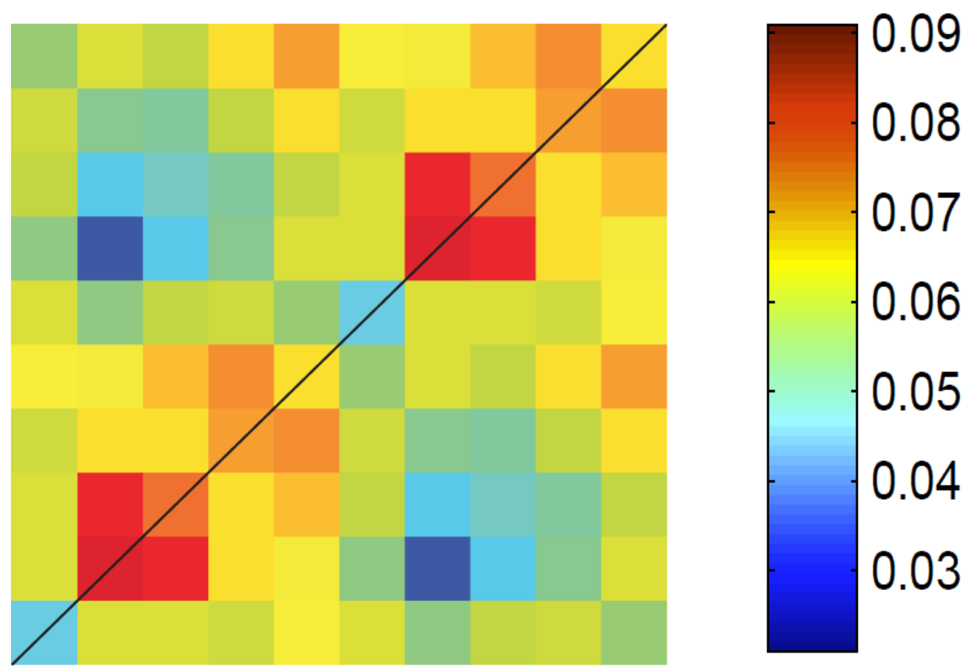
Haefner, Berkes & Fiser, Neuron 2016

Model: noise correlations in $\mathbf{x}$





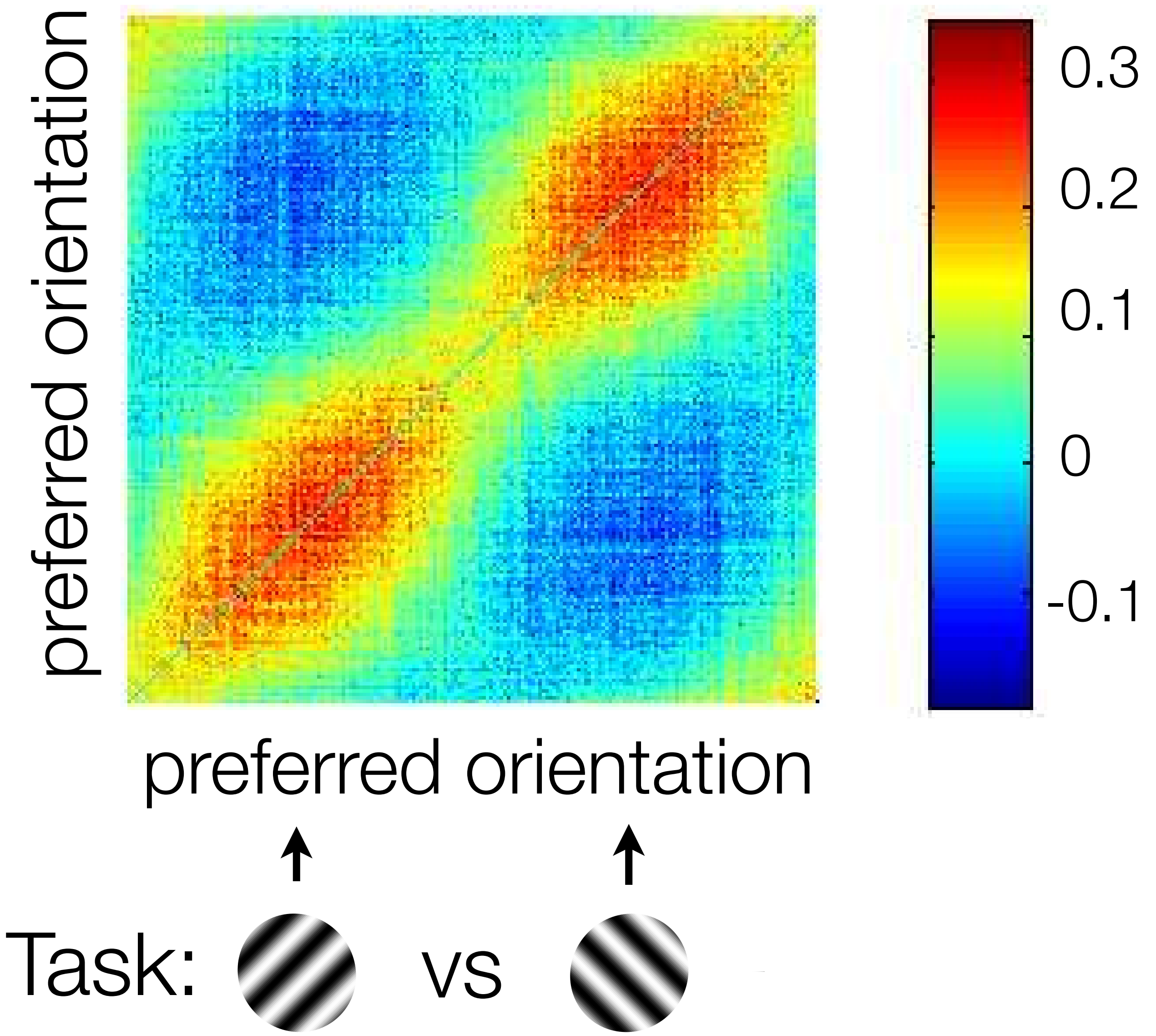


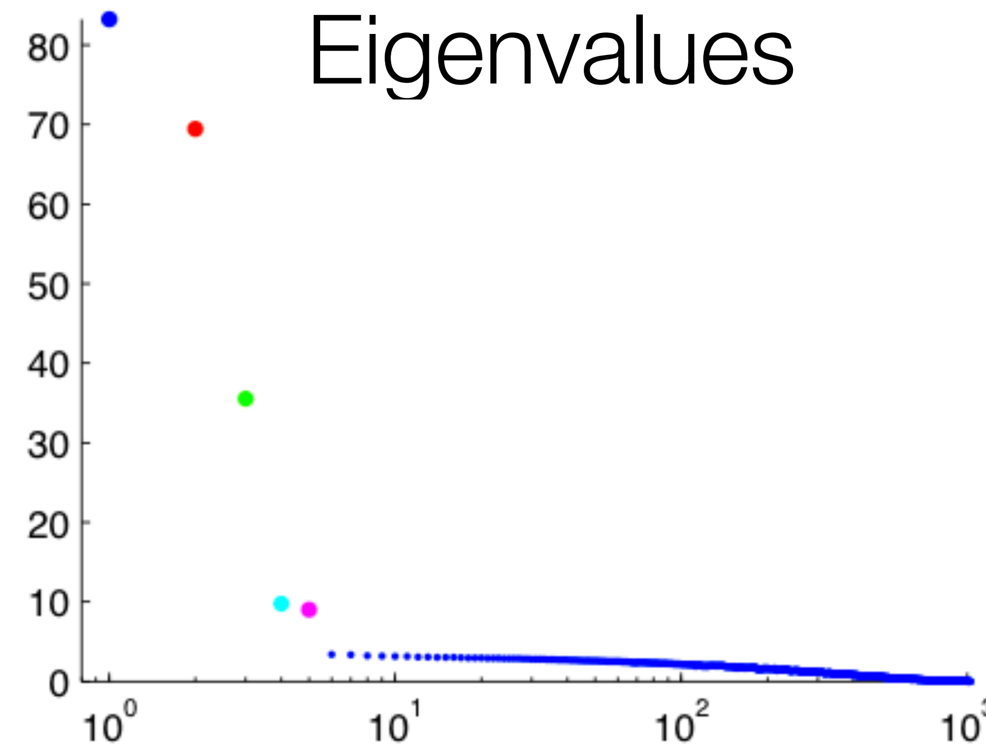


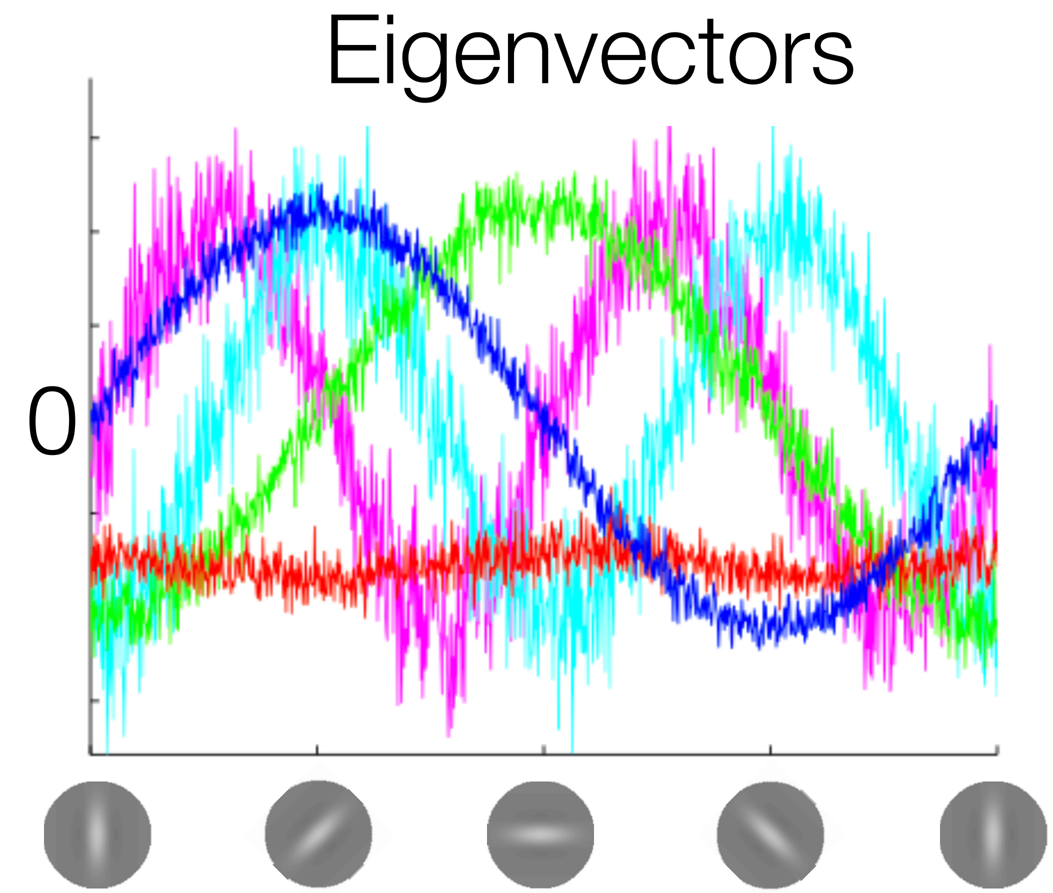
Bondy & Cumming, bioRxiv 2016

Haefner, Berkes & Fiser, Neuron 2016

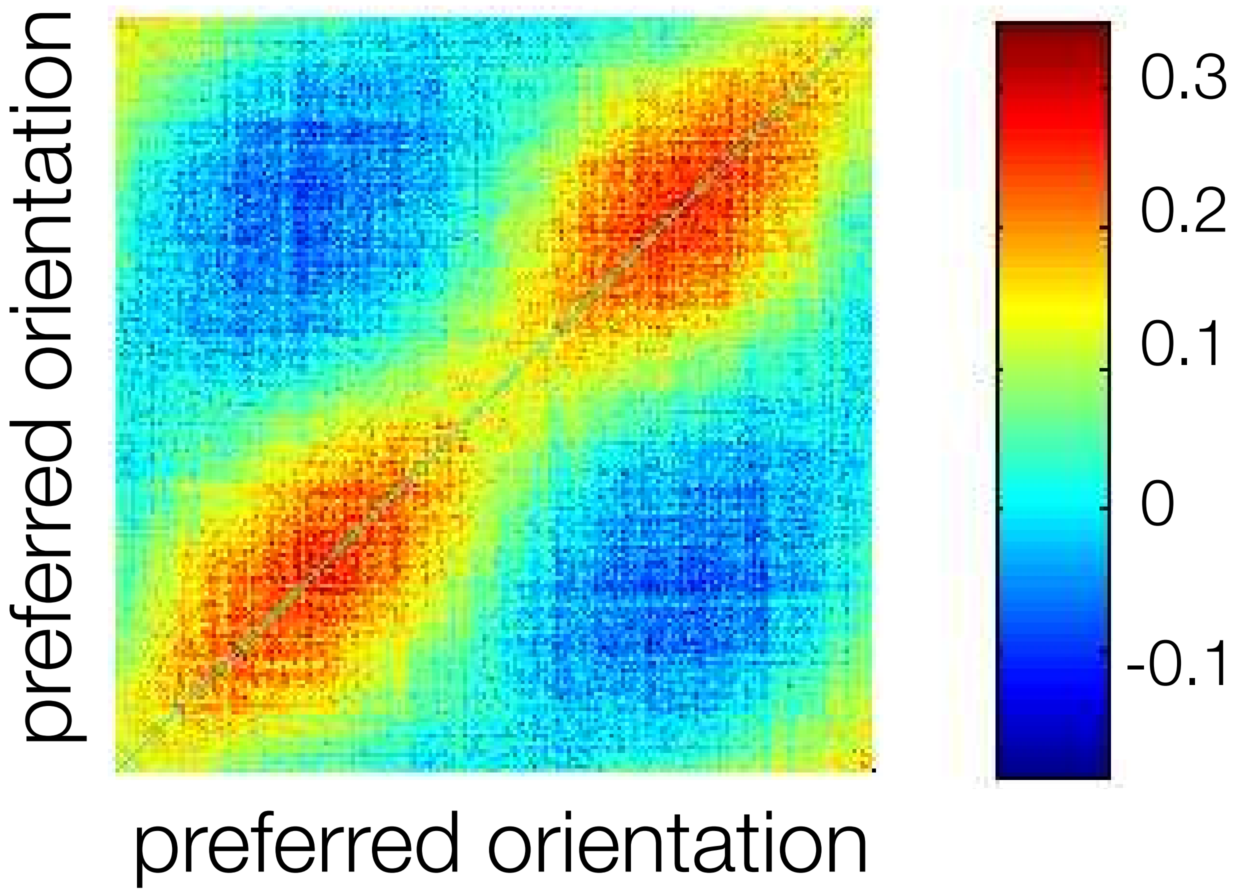
Model: noise correlations in V1

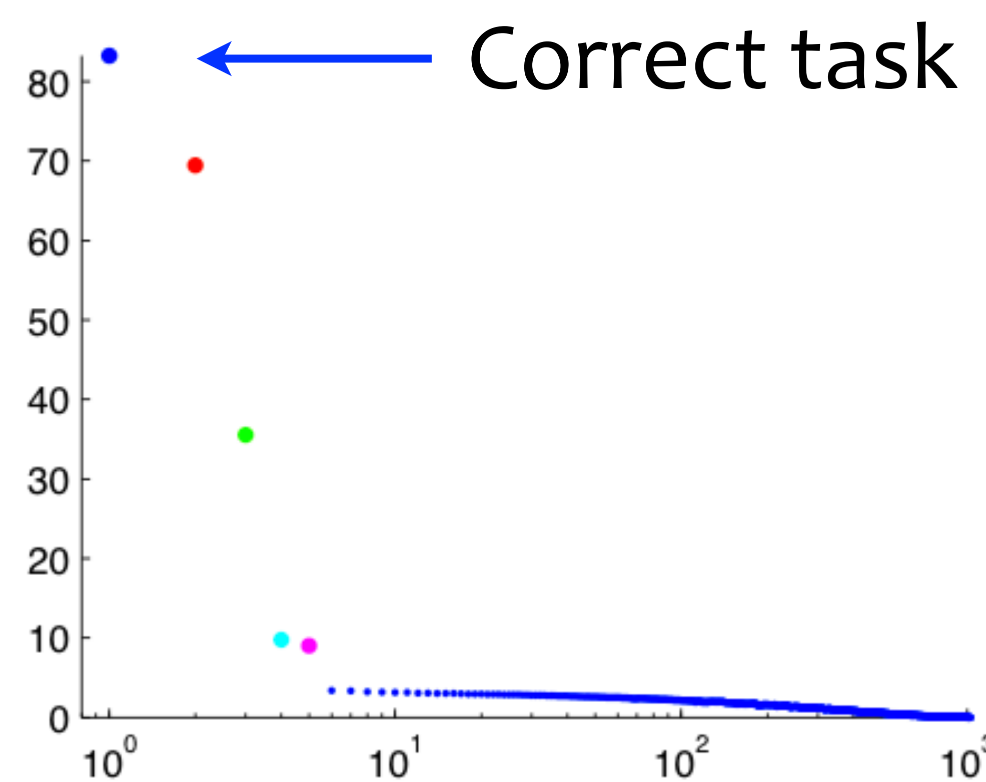


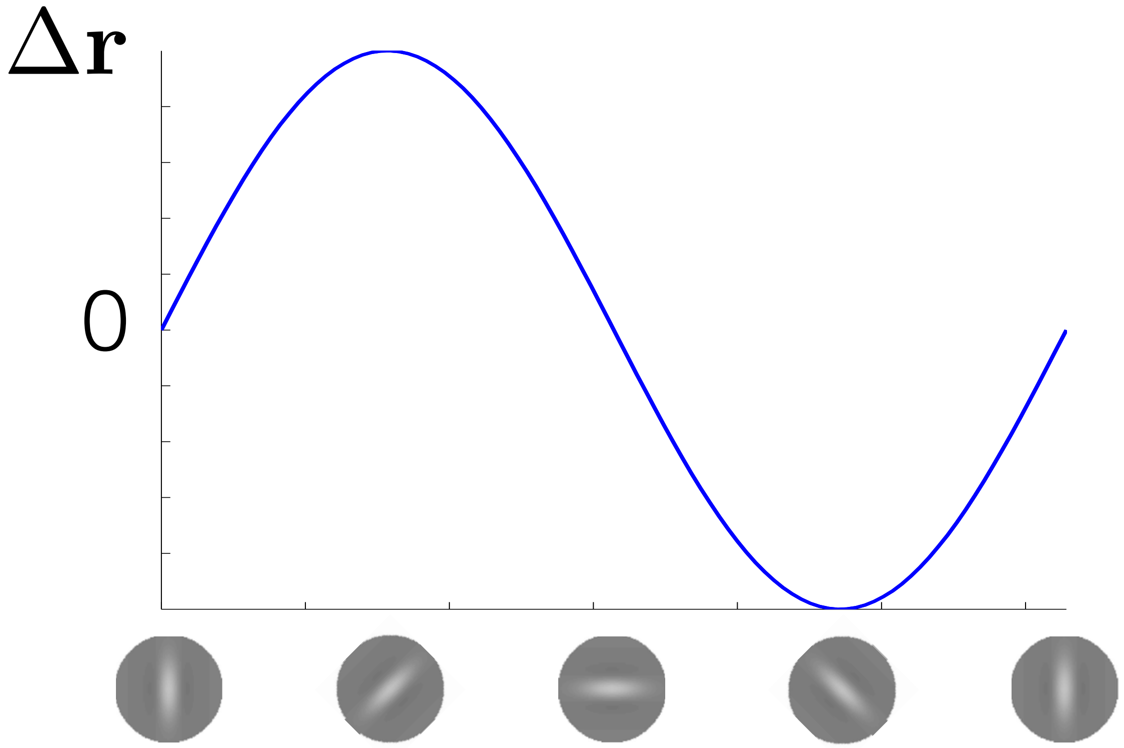


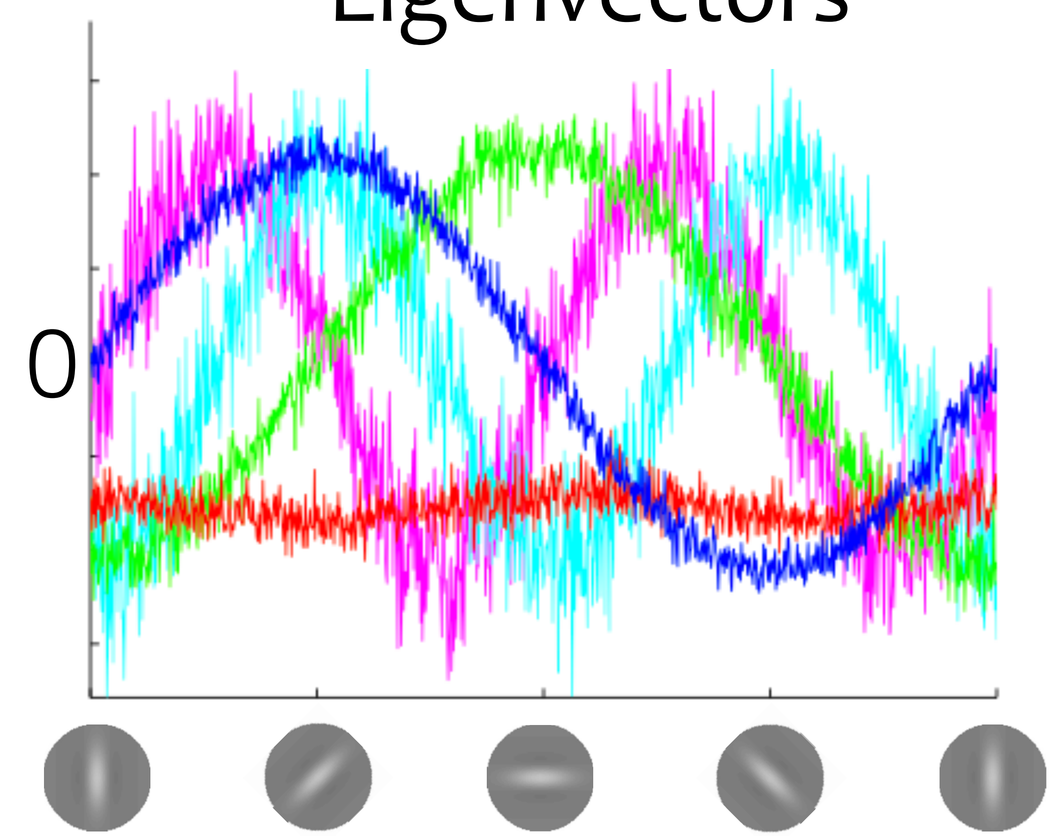


Model: noise correlations in V1

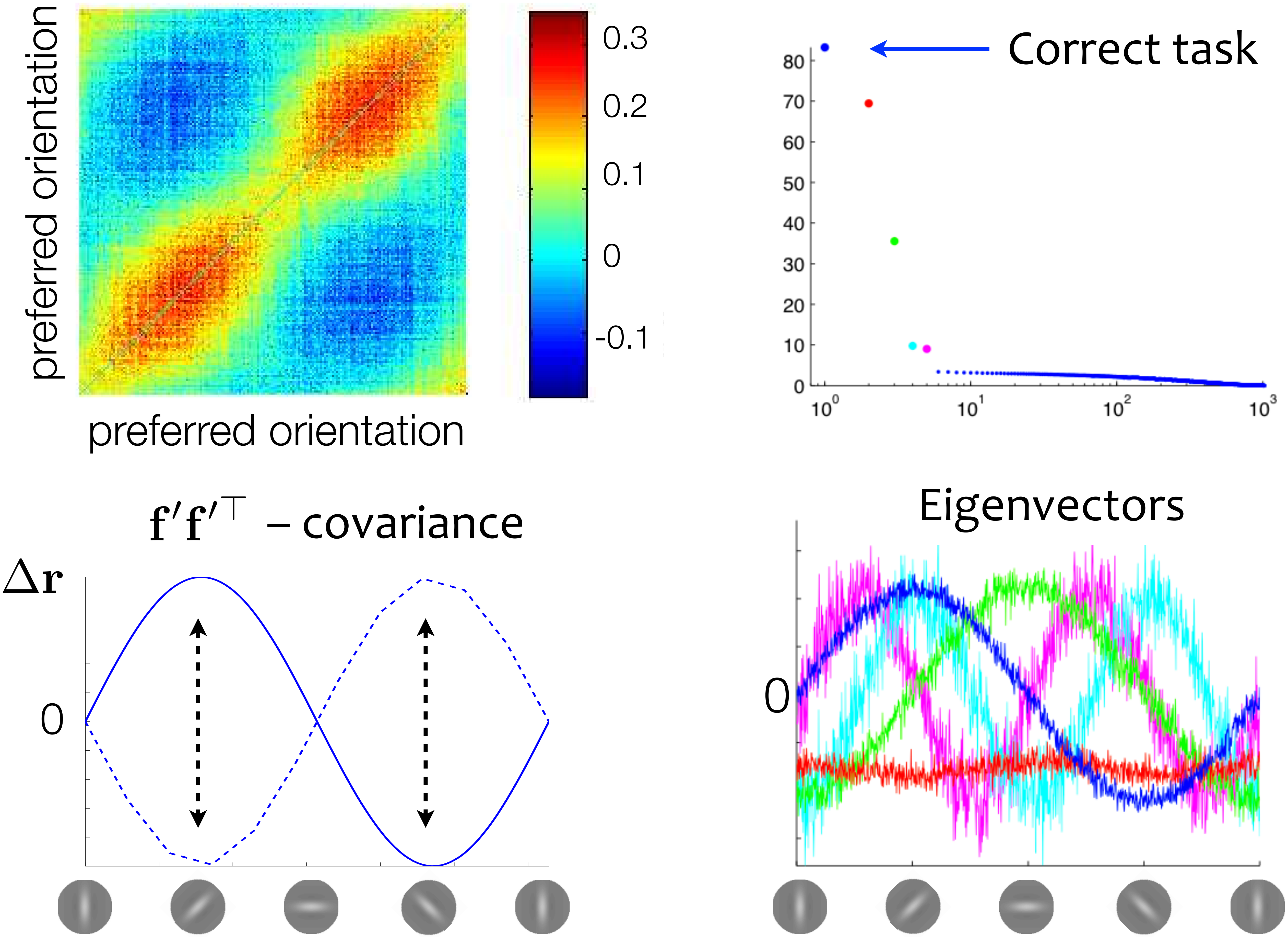




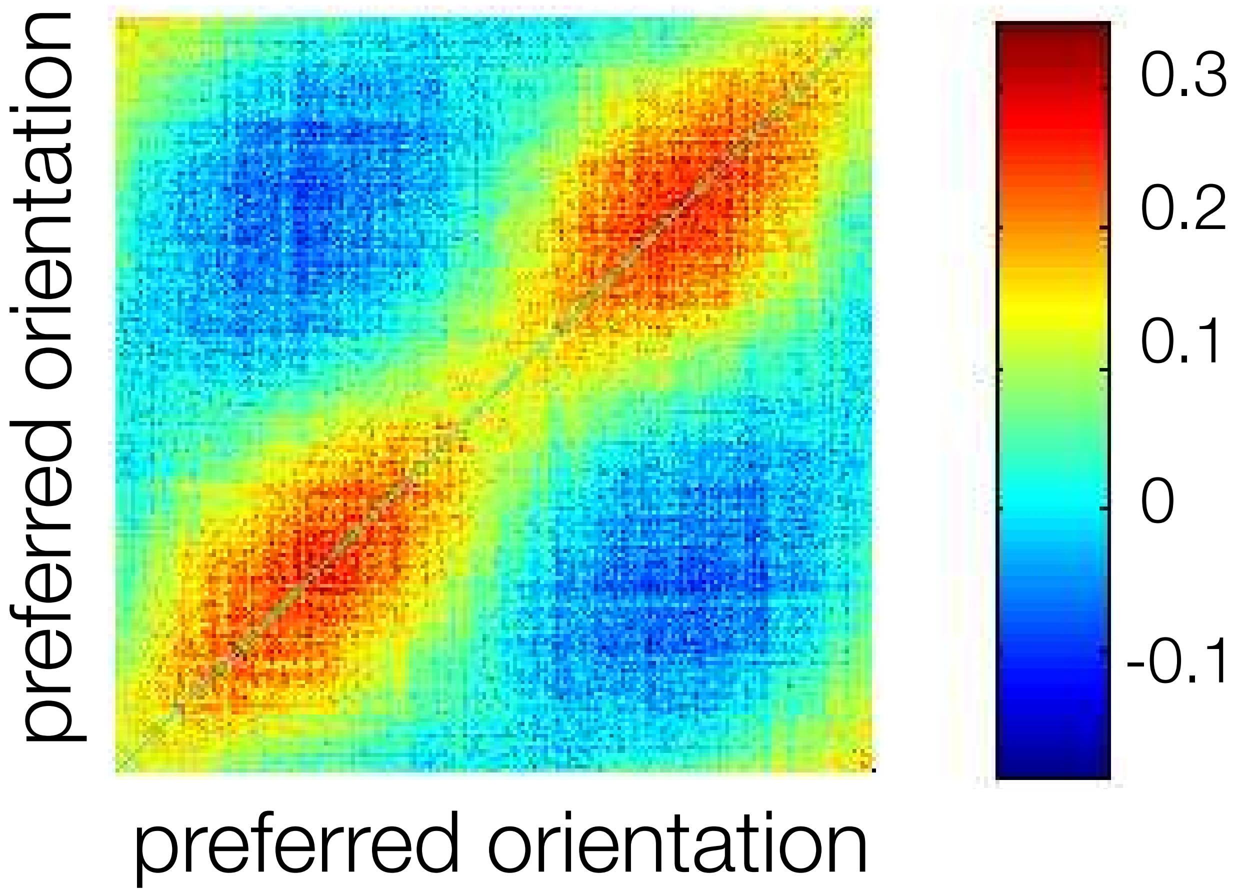


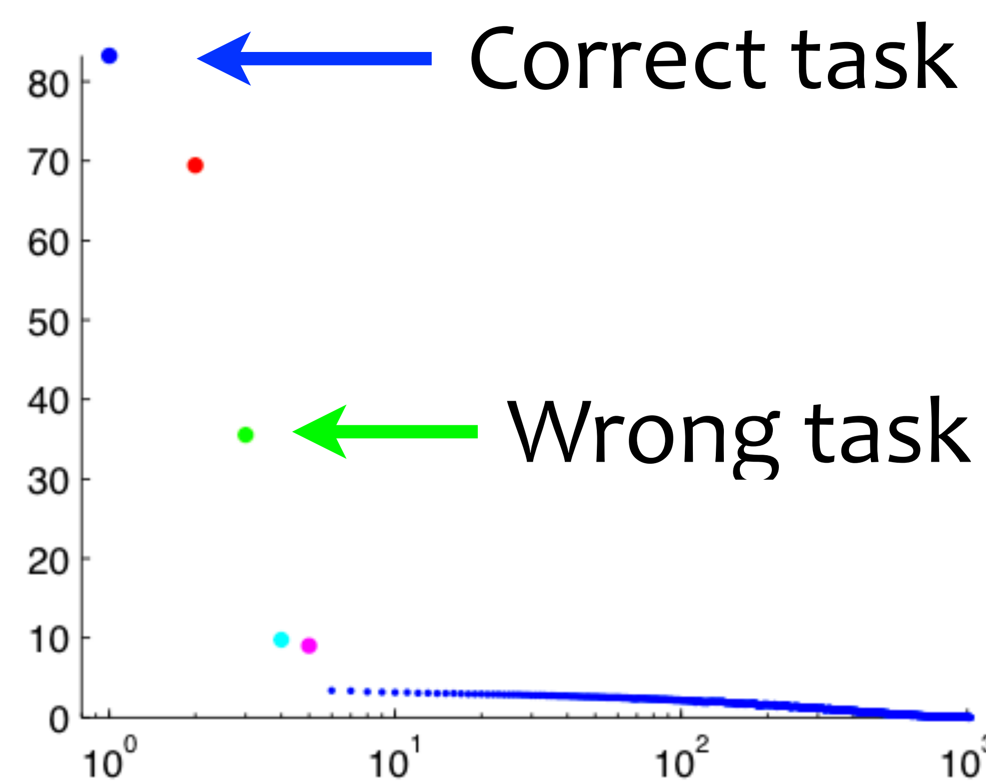


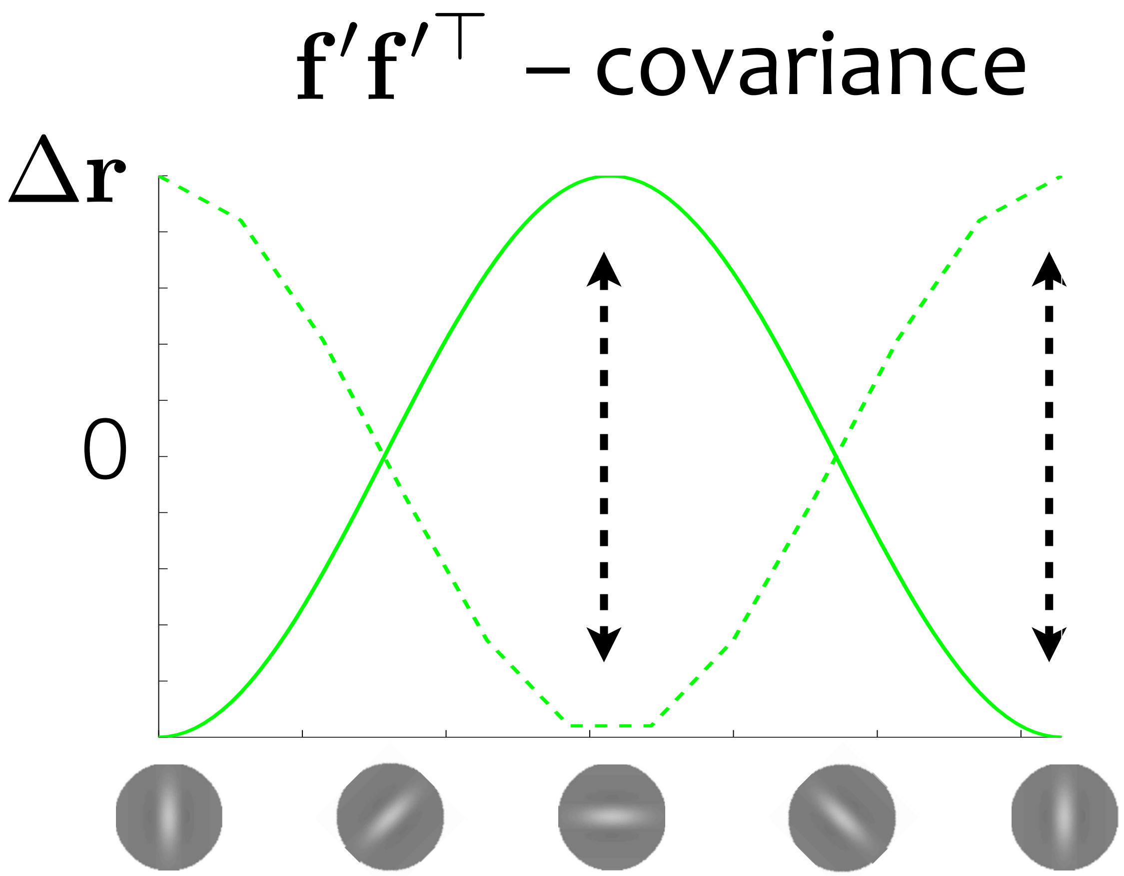
Model: noise correlations in V1

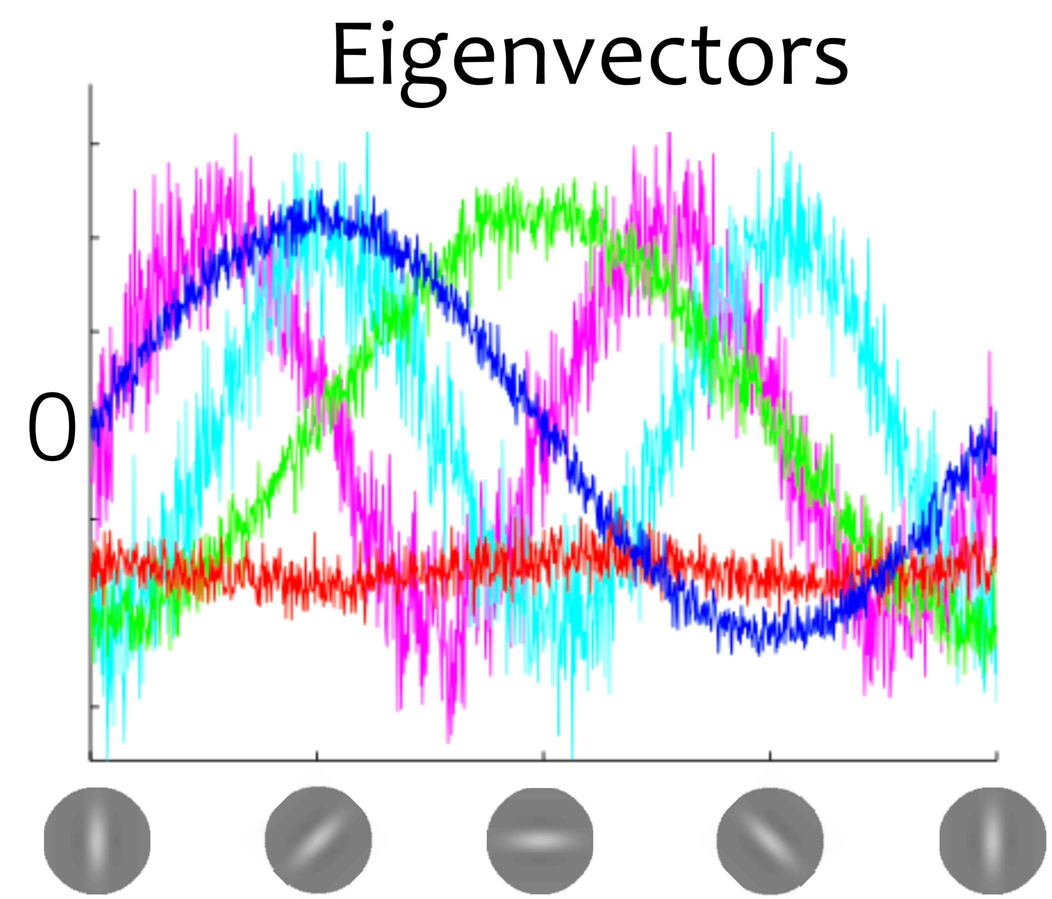


Model: noise correlations in V1

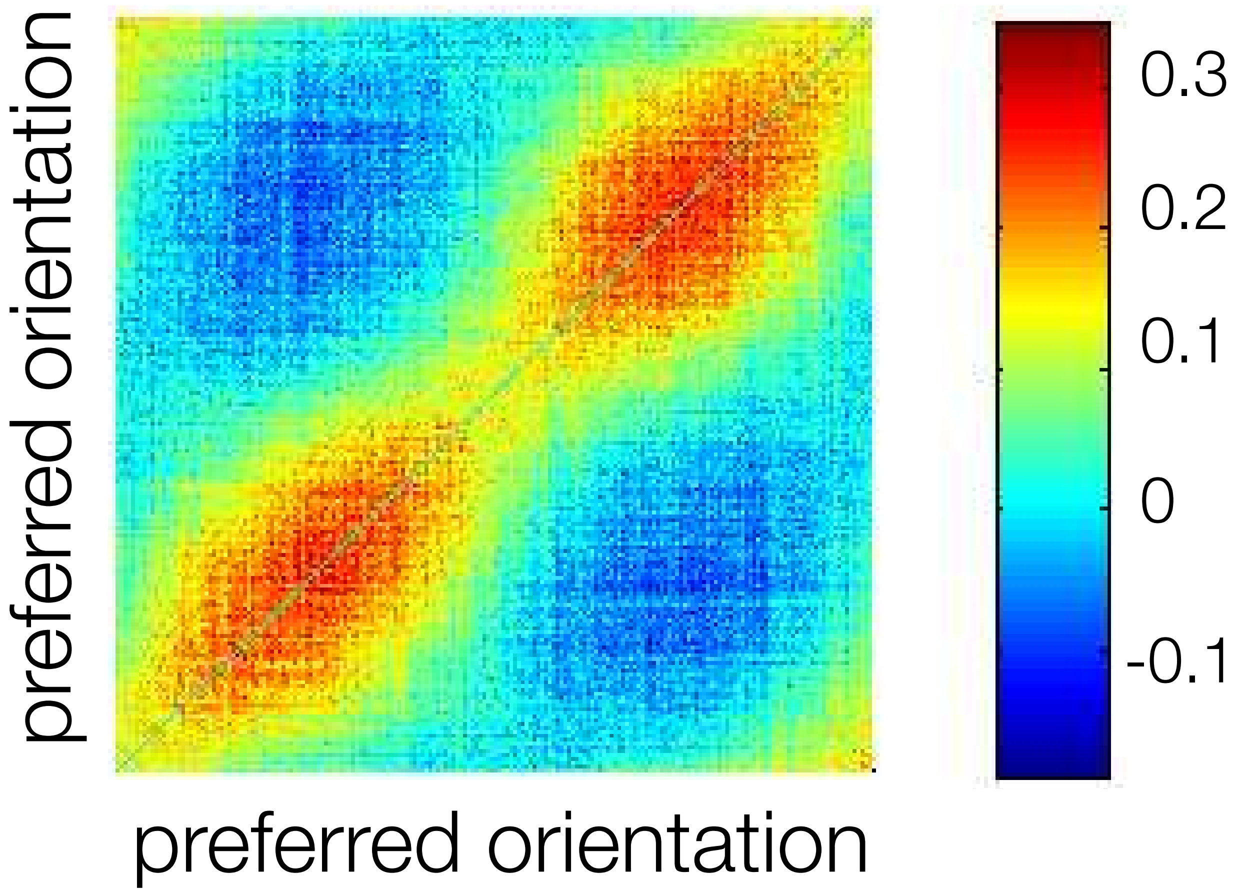


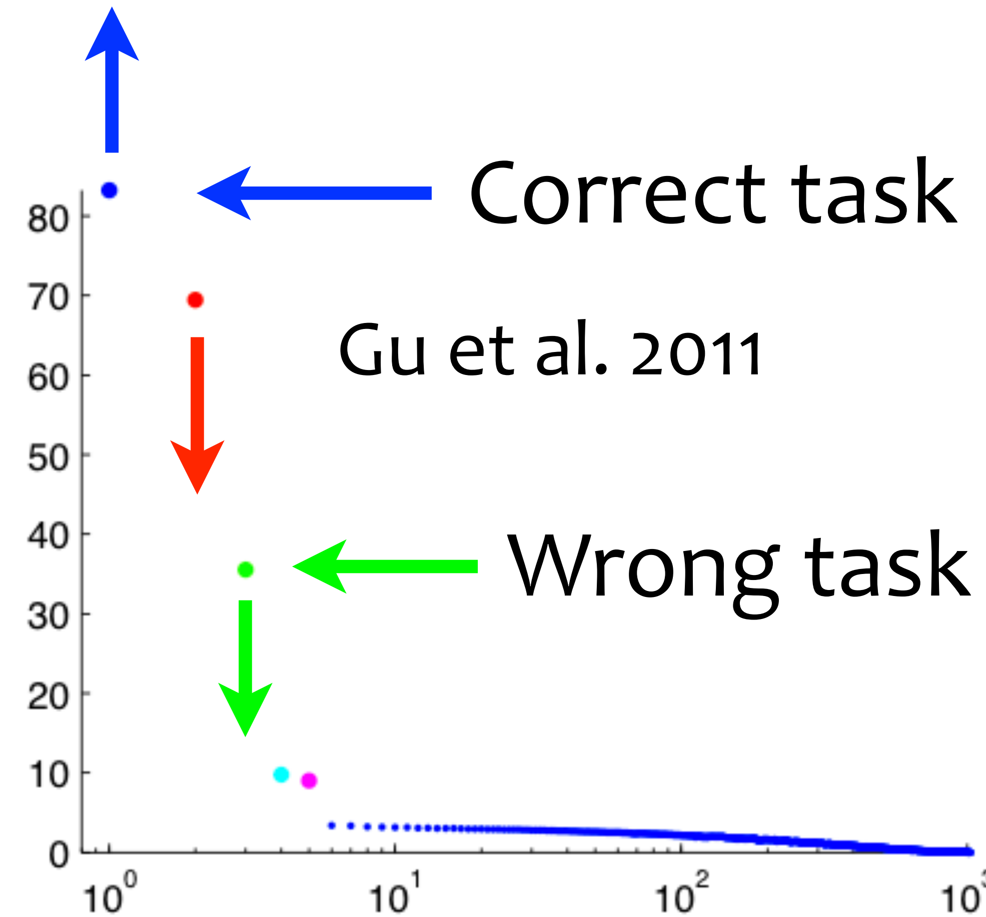


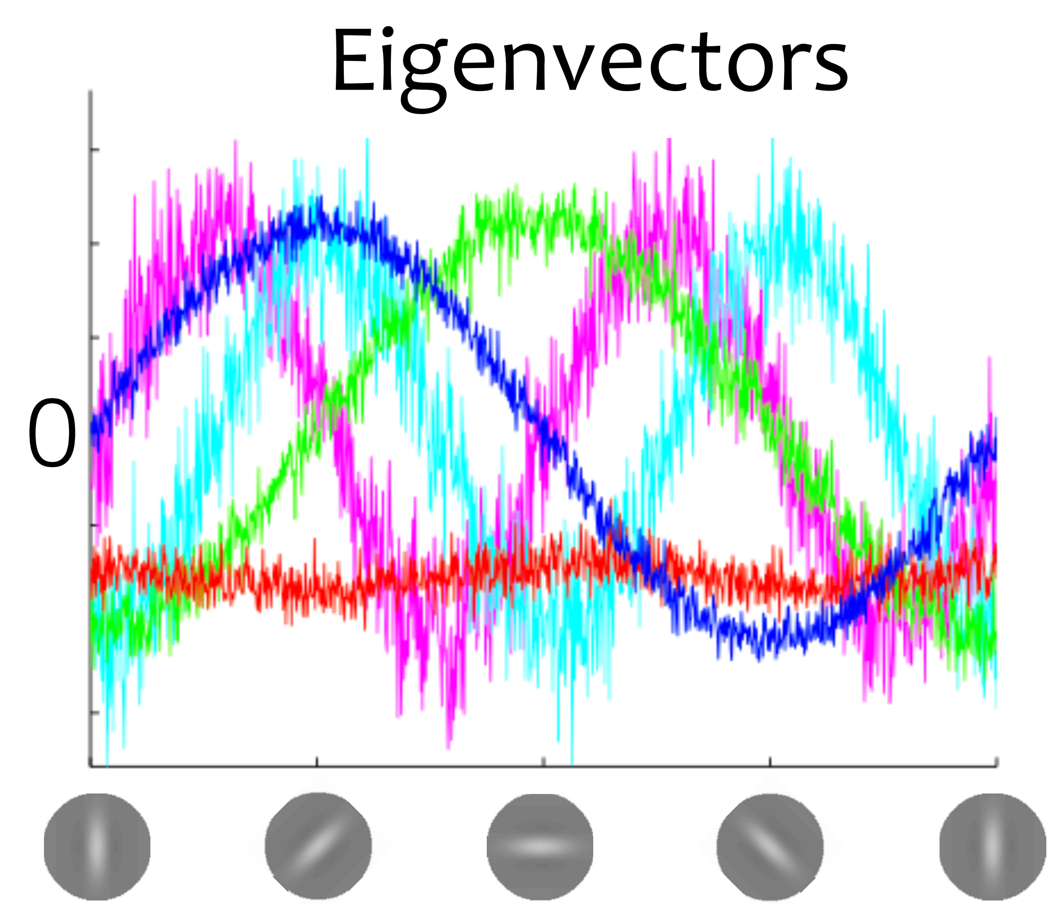




Learning







Learning:

- Increasing EV for correct task
- All other EVs decreasing

Lange & Haefner, bioRxiv 2016

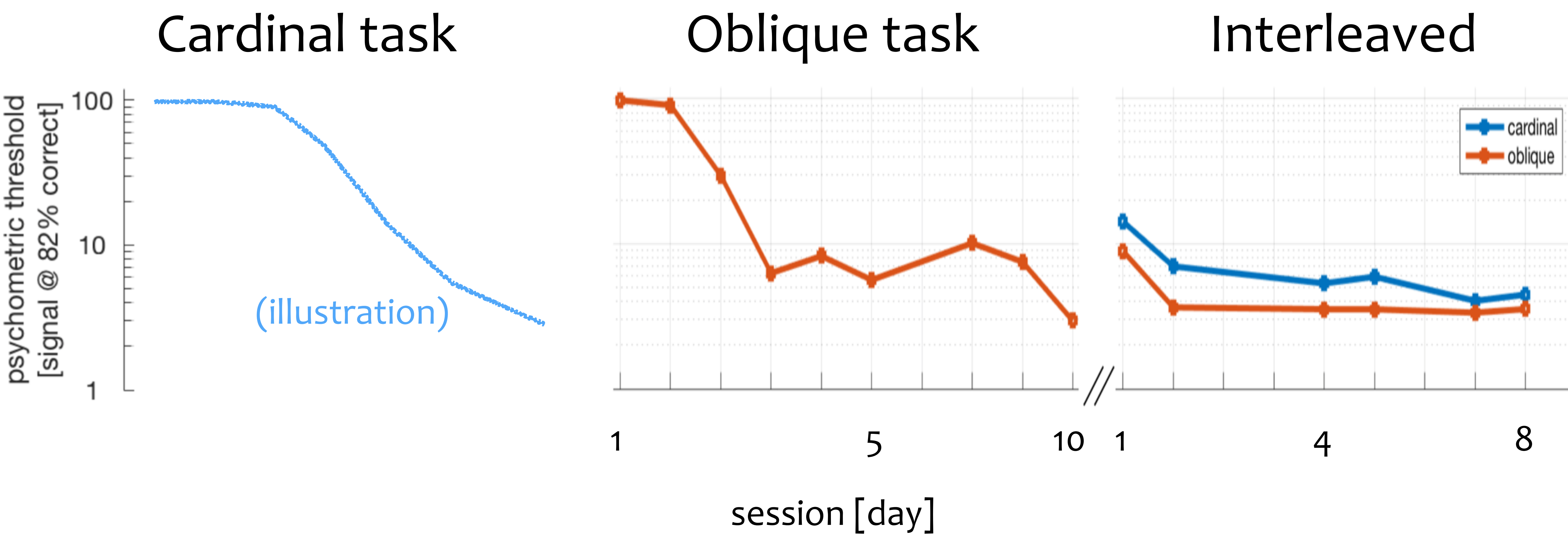
Camille Gomez-Faberge (Harvard)

Rick Born (Harvard)

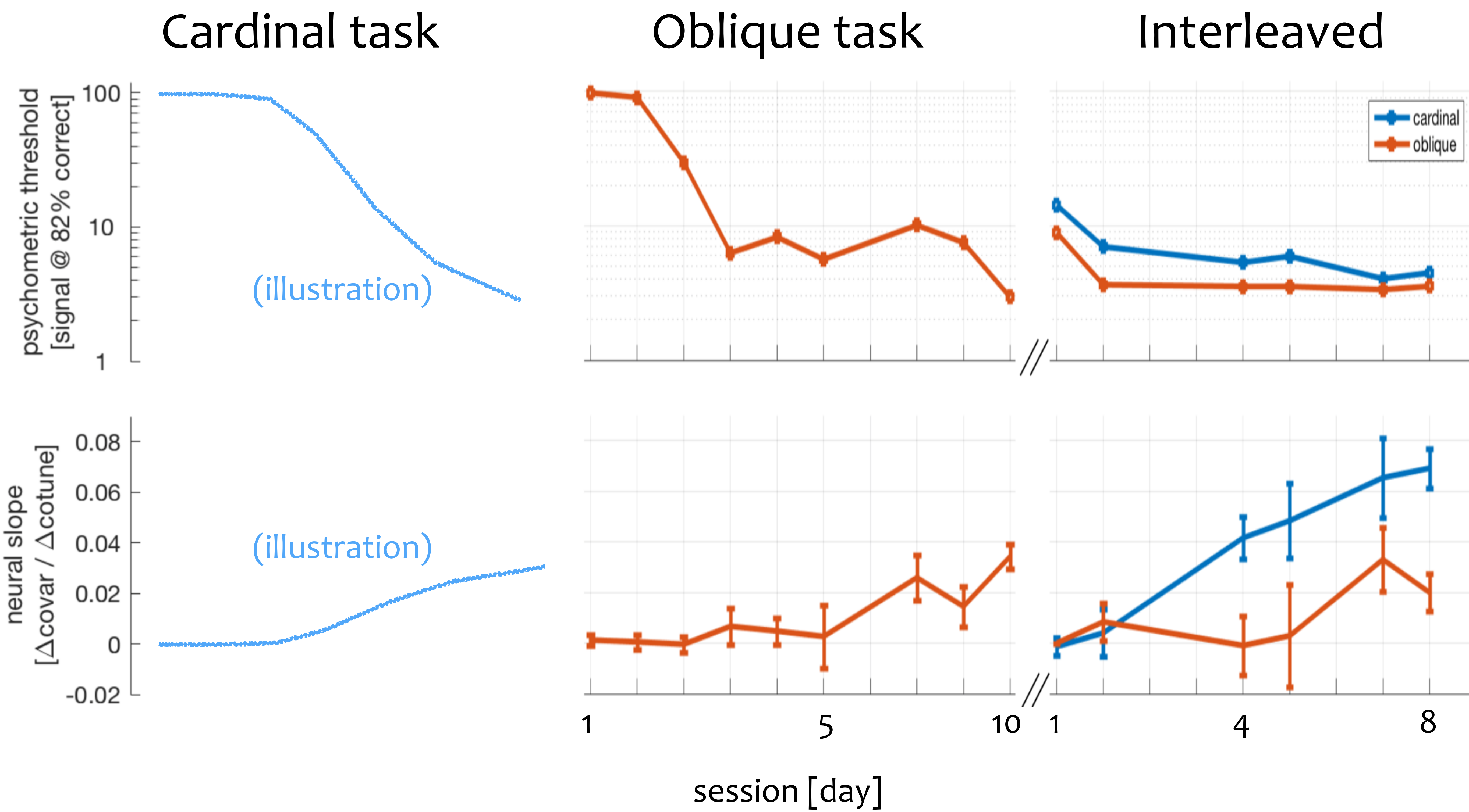
Richard Lange

- Learning
- Inference over the task/multi-tasking
- Reversibly inactivate top-down connections to V1 (by cooling V2)

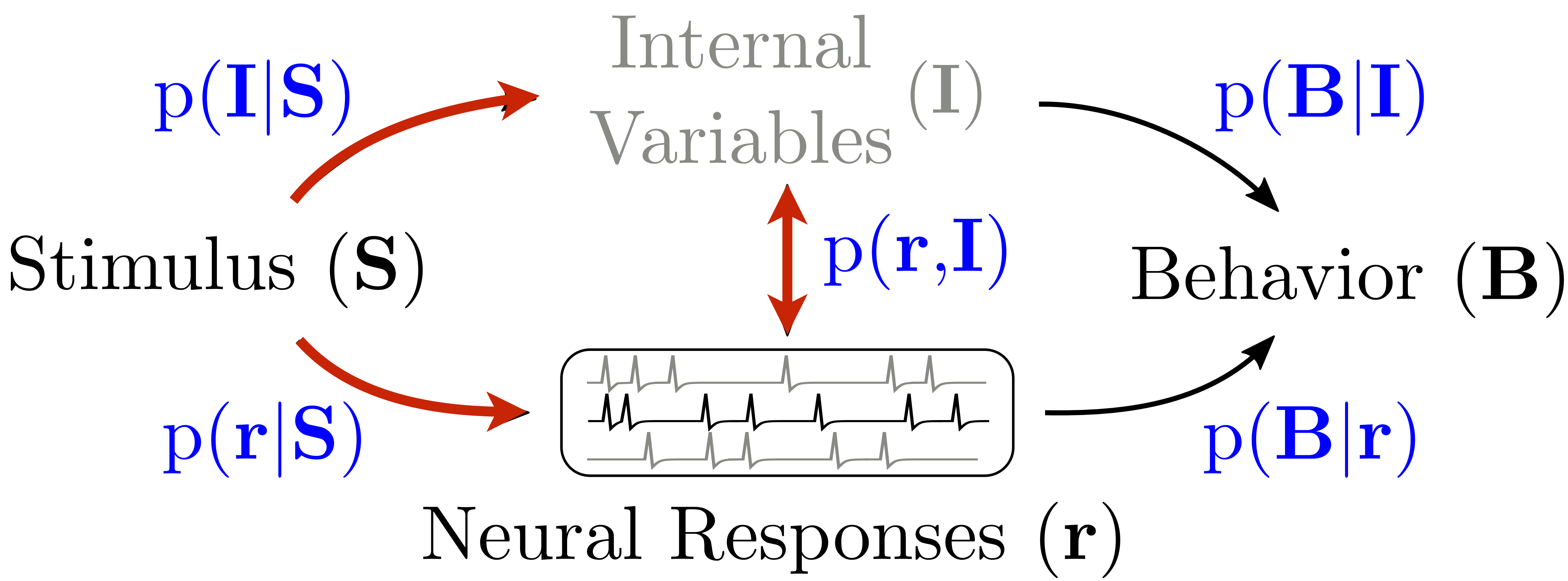
Preliminary results! (1 monkey)



Preliminary results! (1 monkey)



Next steps: correlations for ‘wrong’ task & cooling data



- knowledge of relationships
  stimulus–responses & responses–internal variables

▶ infer relationships: stimulus–internal variables

Lange & Haefner, Current Opinion in Neurobiology 2017

Conclusions so far

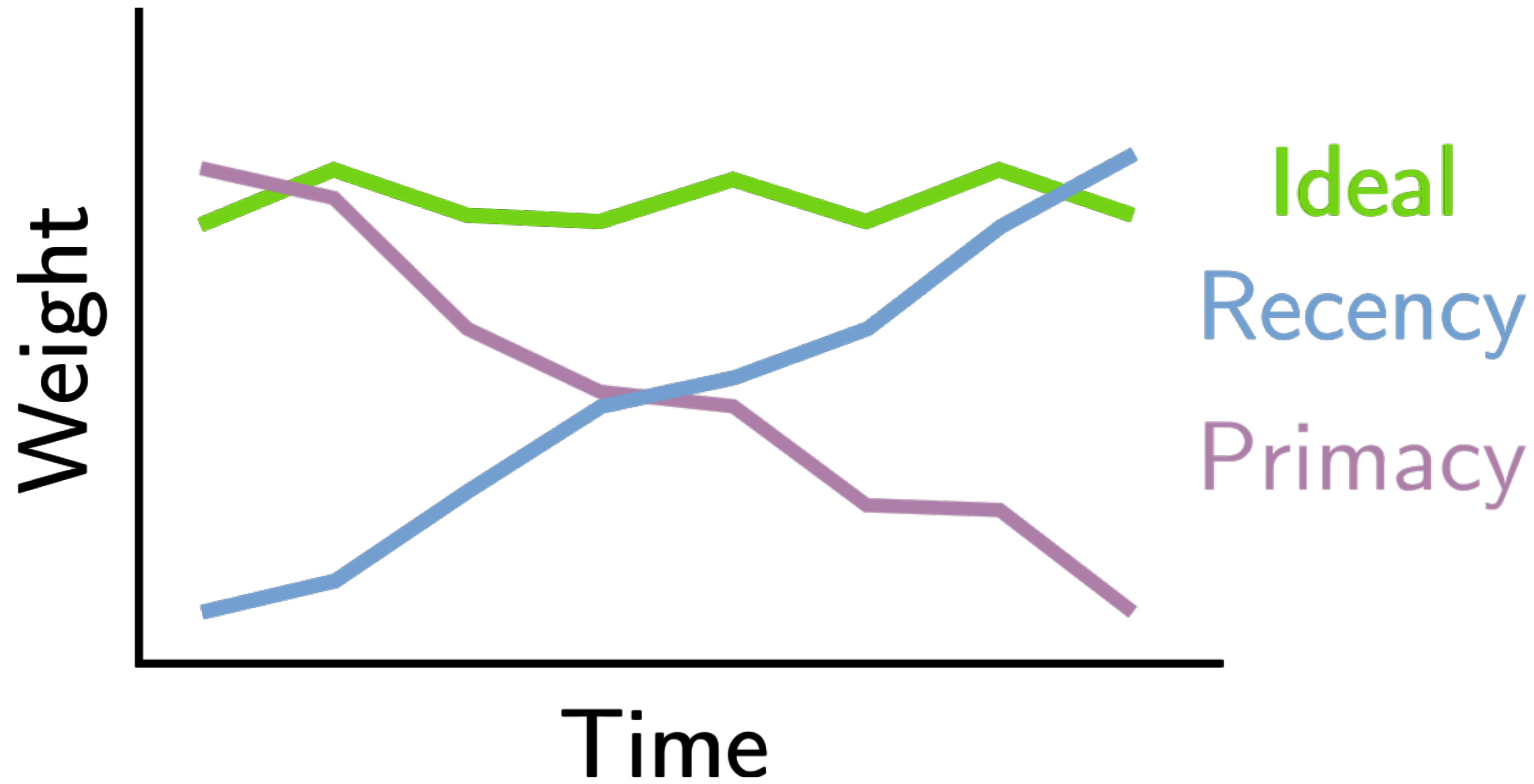
- Probabilistic inference by neural sampling makes strong predictions about the task-dependence of neural correlations

- Data so far confirm those predictions

- Population recordings let us reverse-engineer internal beliefs – e.g. track them over learning

- We can interpret them in stimulus space

Part 2: Confirmation Bias



“weight” = “psychophysical kernel”

Decreasing psycho-physical kernel (PK)

Confirmation bias:
- One sample at a time
- Decision based on inferred, not directly observed variables

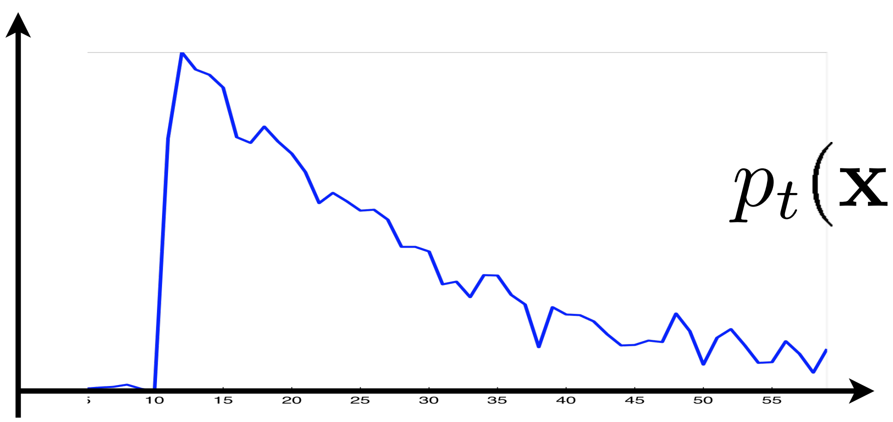
$$p_t(O) \propto p_{t-1}(O)p(O|\mathbf{x}_t)$$

top-down

$$p_t(\mathbf{x}) \propto \cancel{p(\mathbf{x}|\mathbf{y}_t)}p(\mathbf{x}|O_t)$$

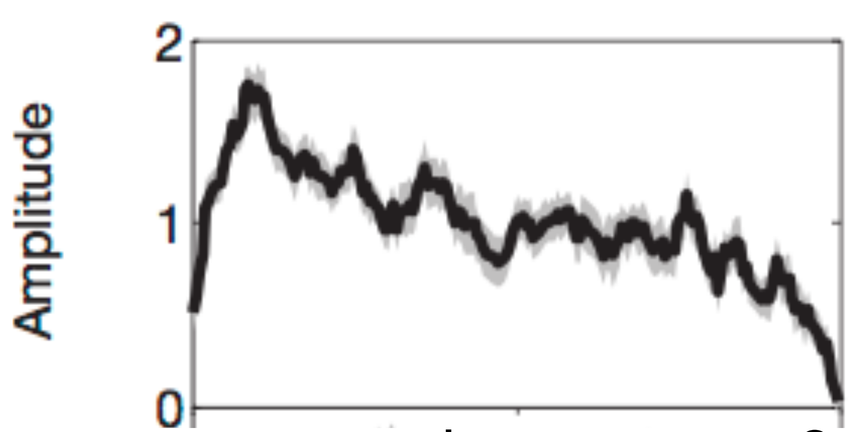
bottom-up

PK



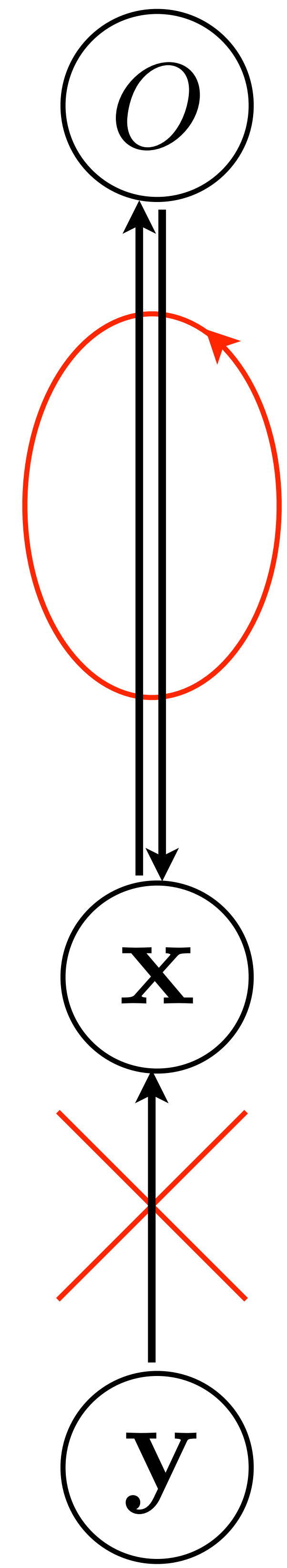
time

PK



time 2s

Nienborg & Cumming, Nature 2009



Decision variable

(Grating variables omitted for simplicity)

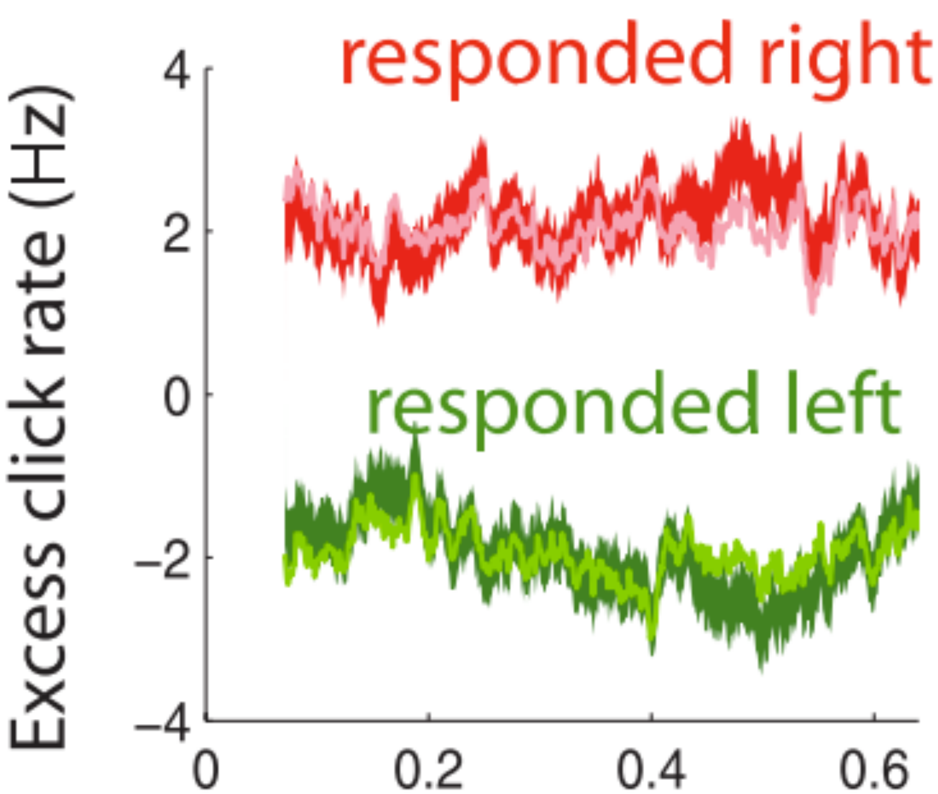
Gabor variables

Image variable

Why constant psycho-physical kernel?

$$p_t(O) \propto p_{t-1}(O)p(O|\mathbf{x}_t)$$

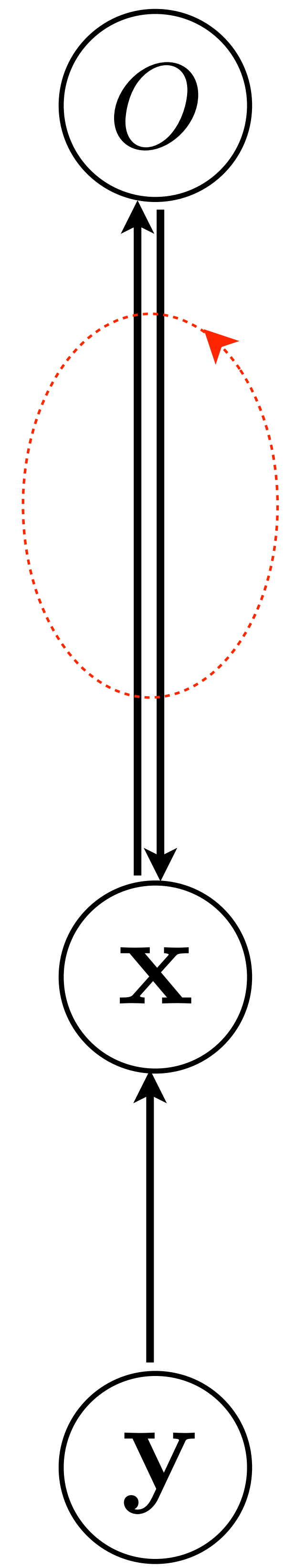
Poisson click task



Brunton et al., Science 2013

$$p_t(\mathbf{x}) \propto p(\mathbf{x}|\mathbf{y}_t)p(\mathbf{x}|O_t)$$

Each click far above threshold, i.e. likelihood dominates prior.



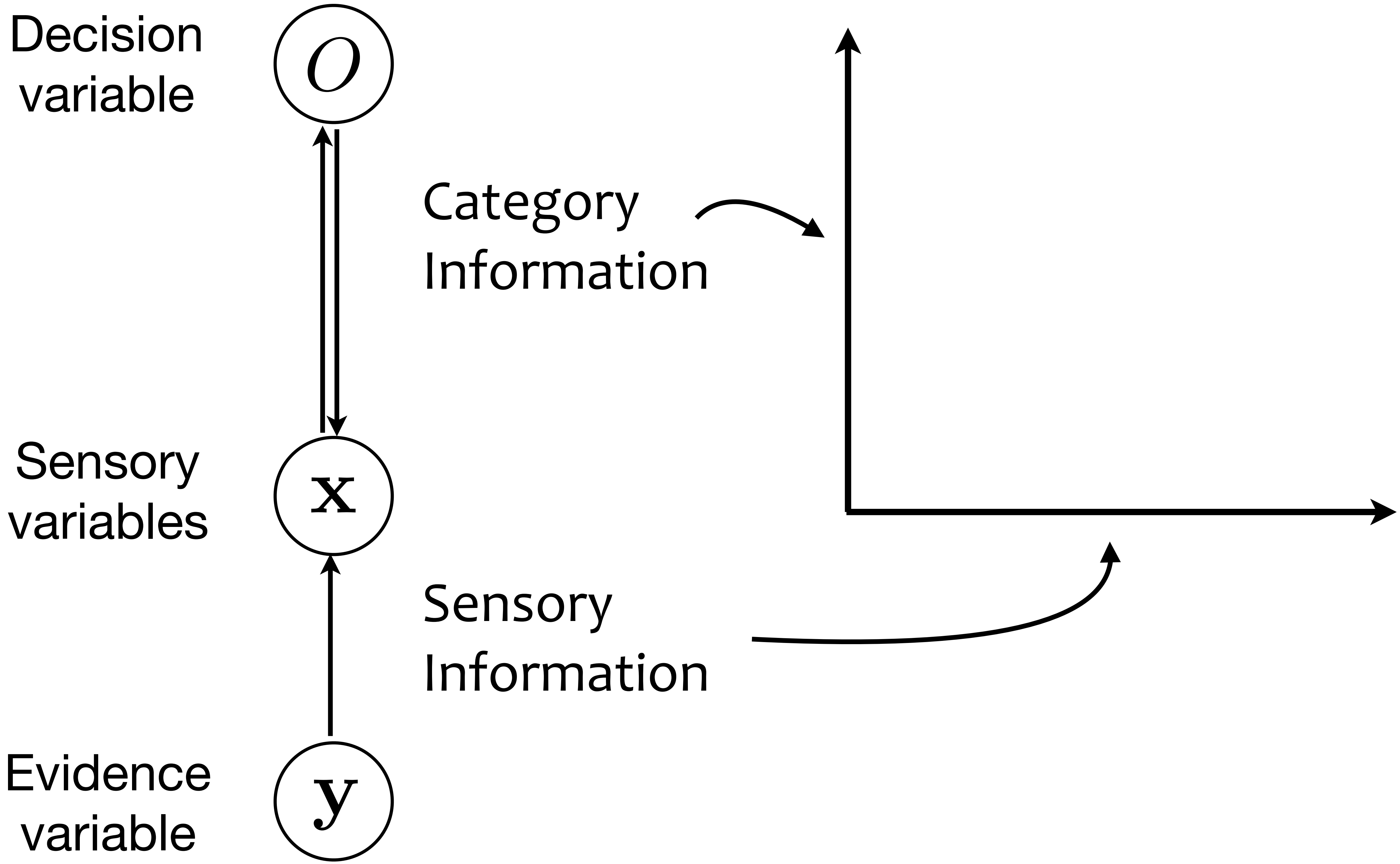
Prediction: Soft clicks -> decreasing PK

Confirmation bias project

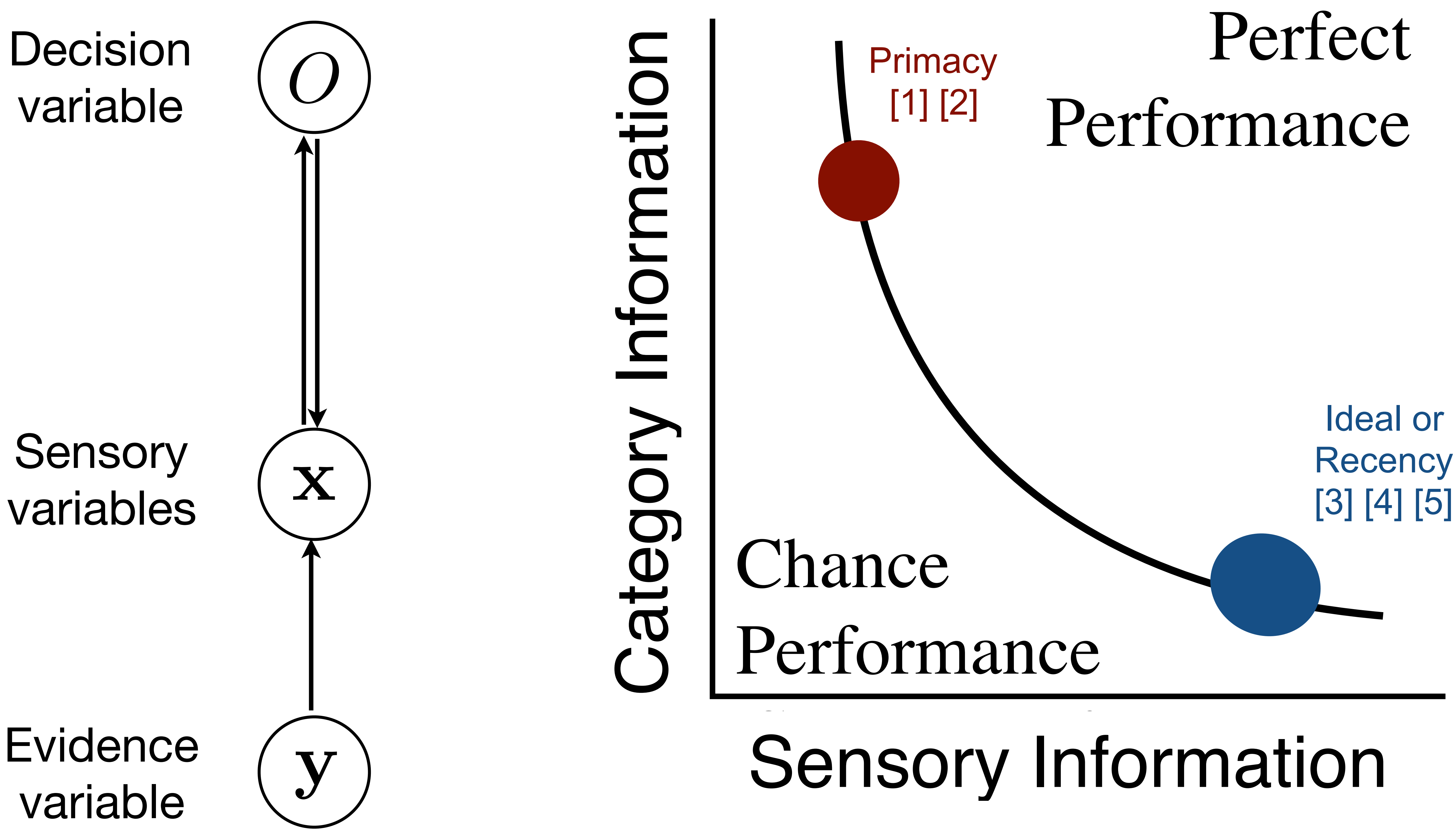
Richard Lange

Ankani Chattoraj

Two kinds of information

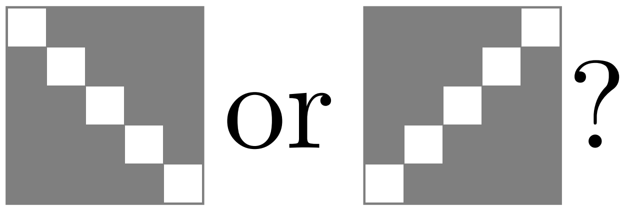


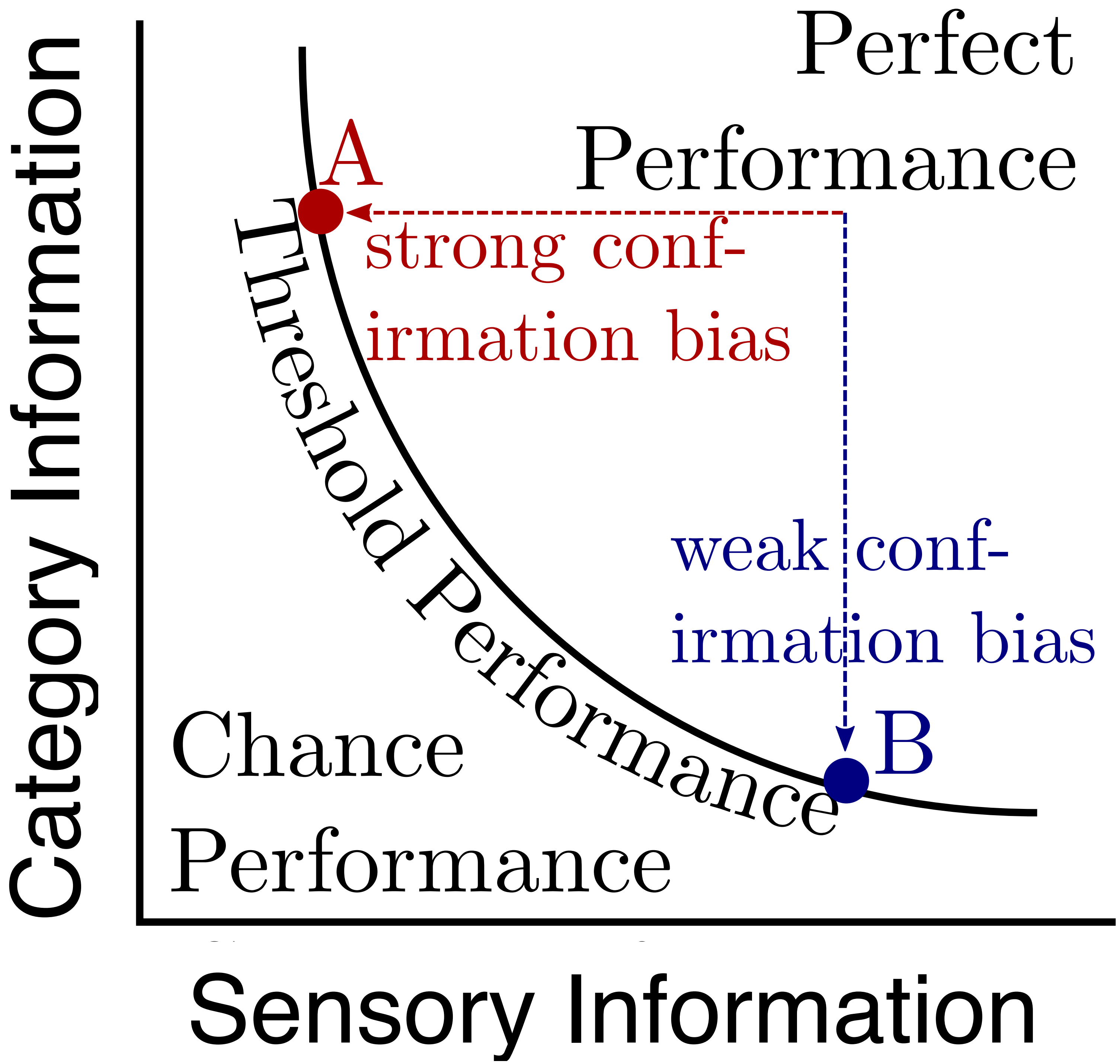
Two kinds of task



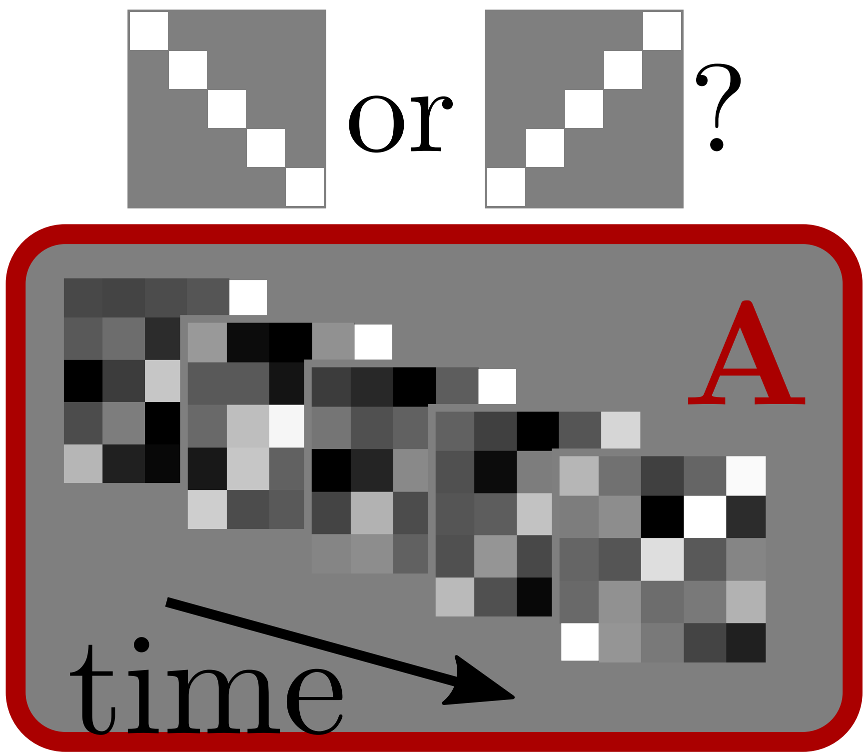
[1] Nienborg & Cumming, Nature (2009). [2] Kiani et al. JNeurosci (2008).
[3] Wyart et al. Neuron (2012) [4] Brunton et al. Science (2013). [5] Drugowitsch et al. Neuron (2016).

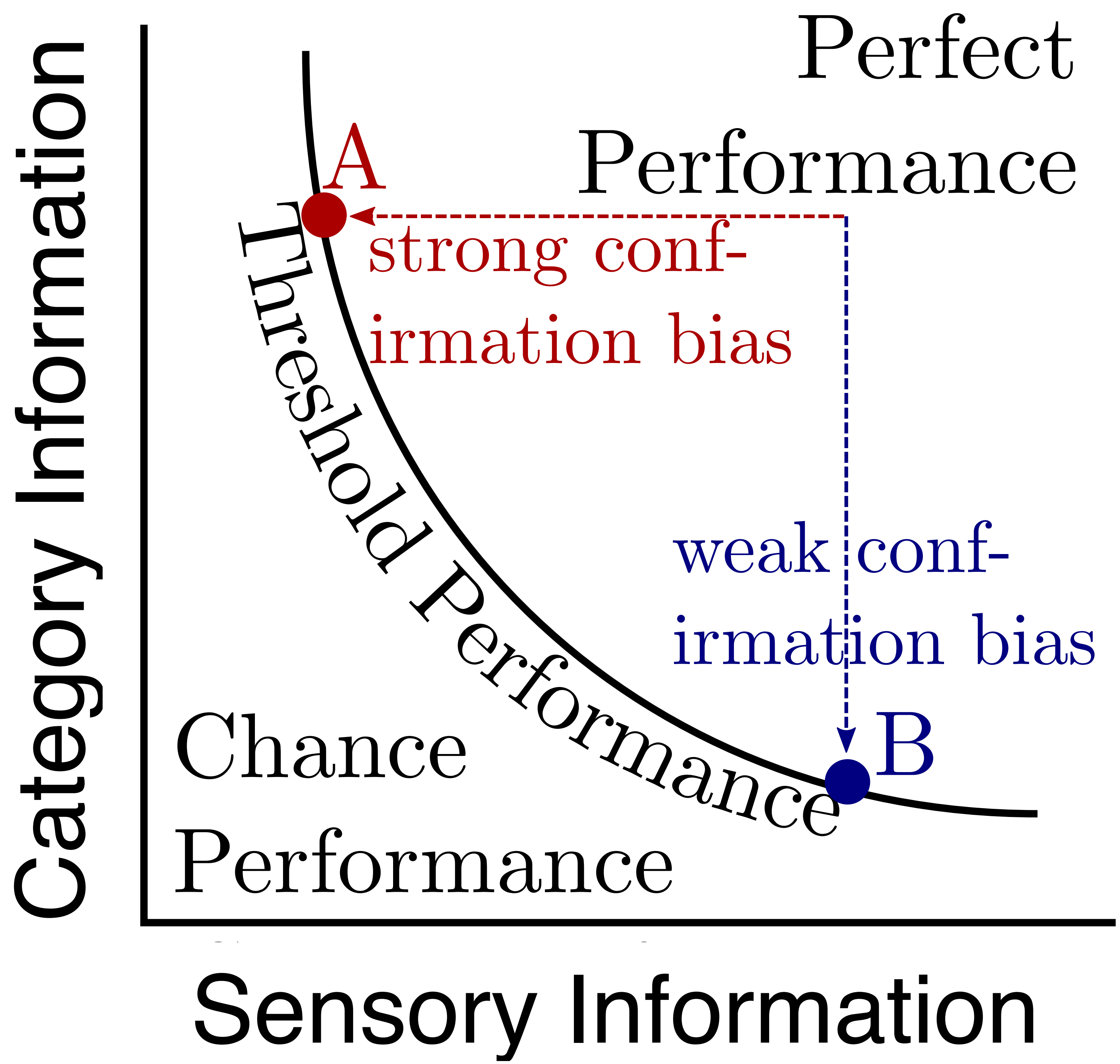
Two different tasks



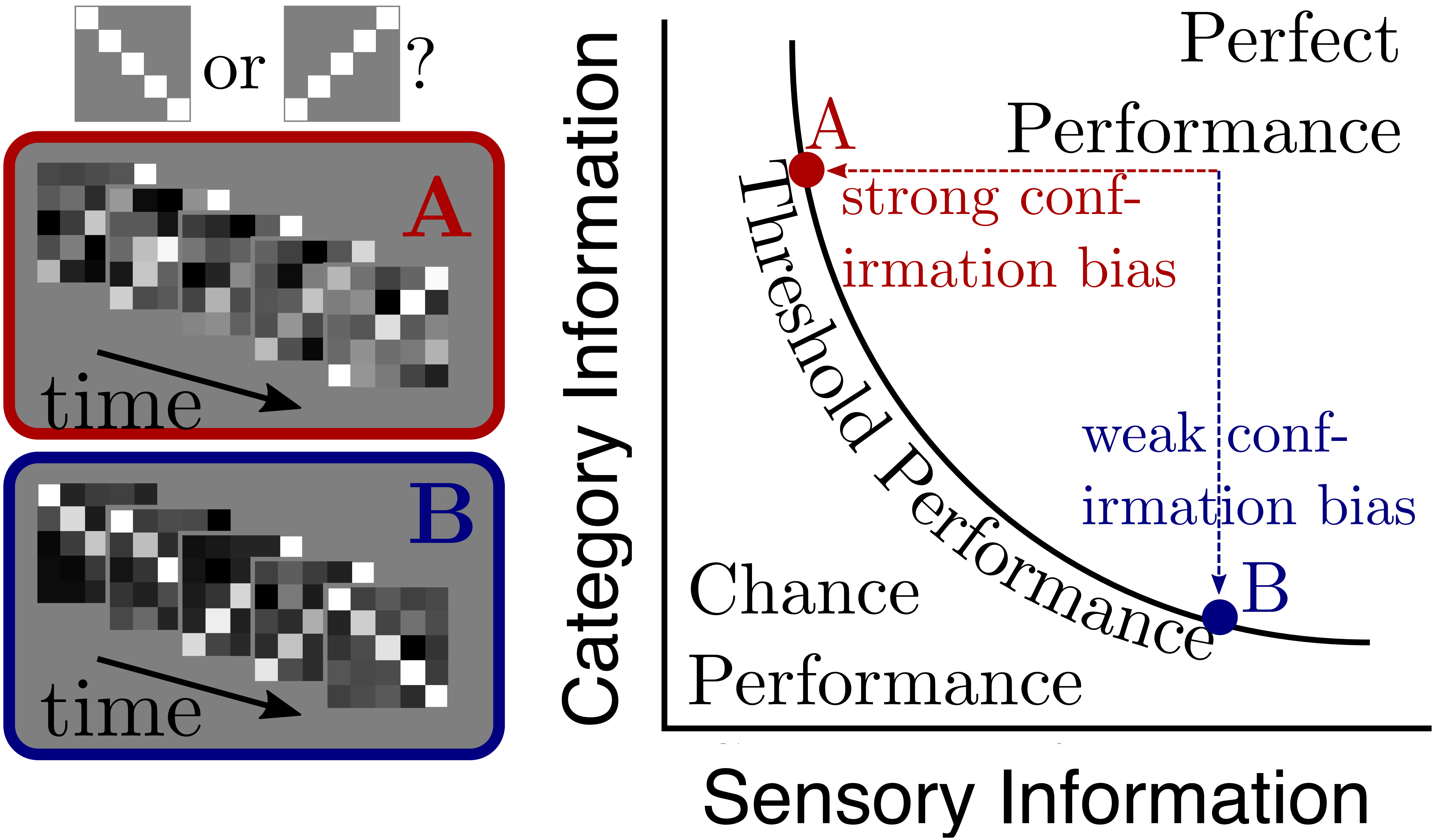


Two different tasks

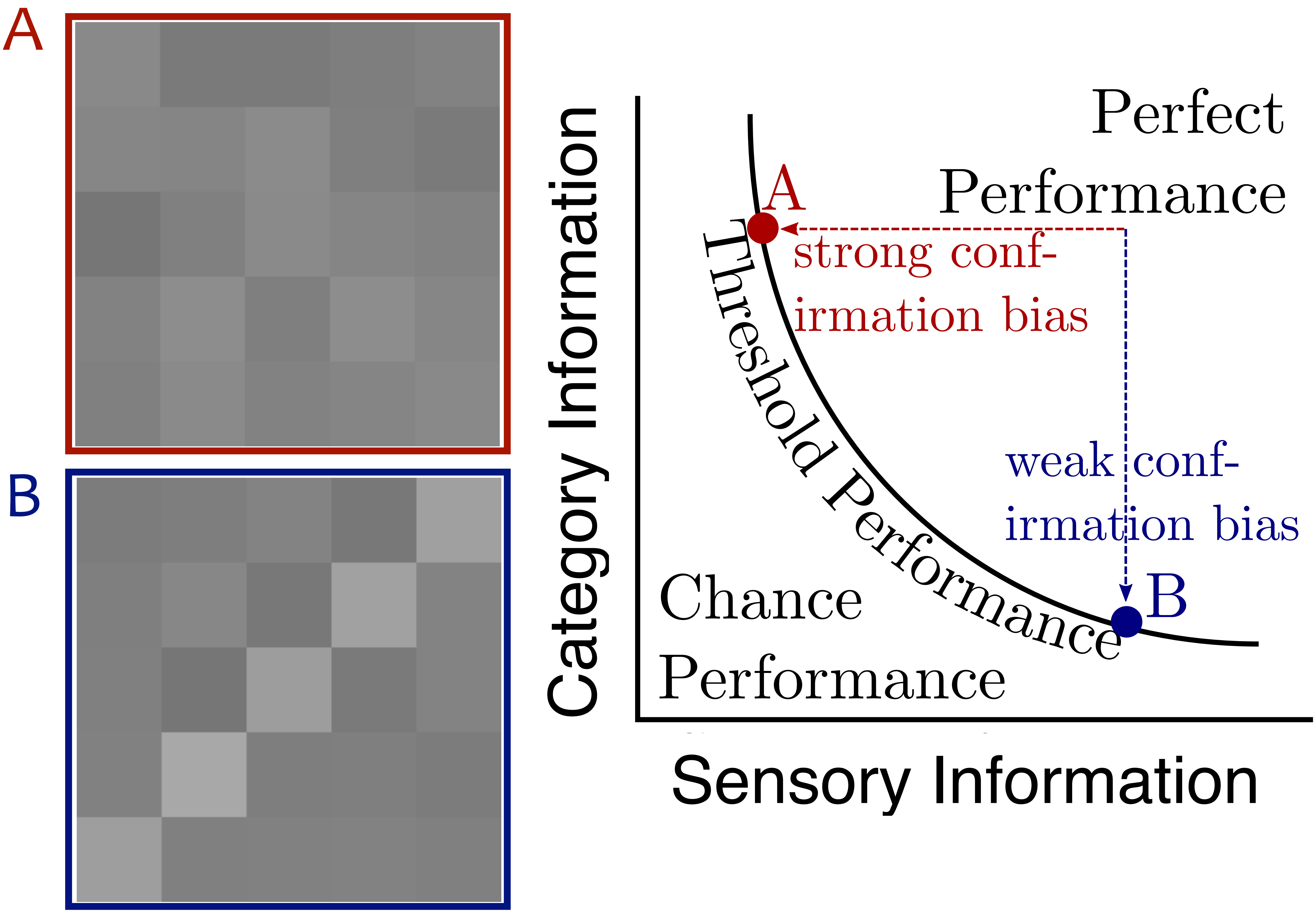


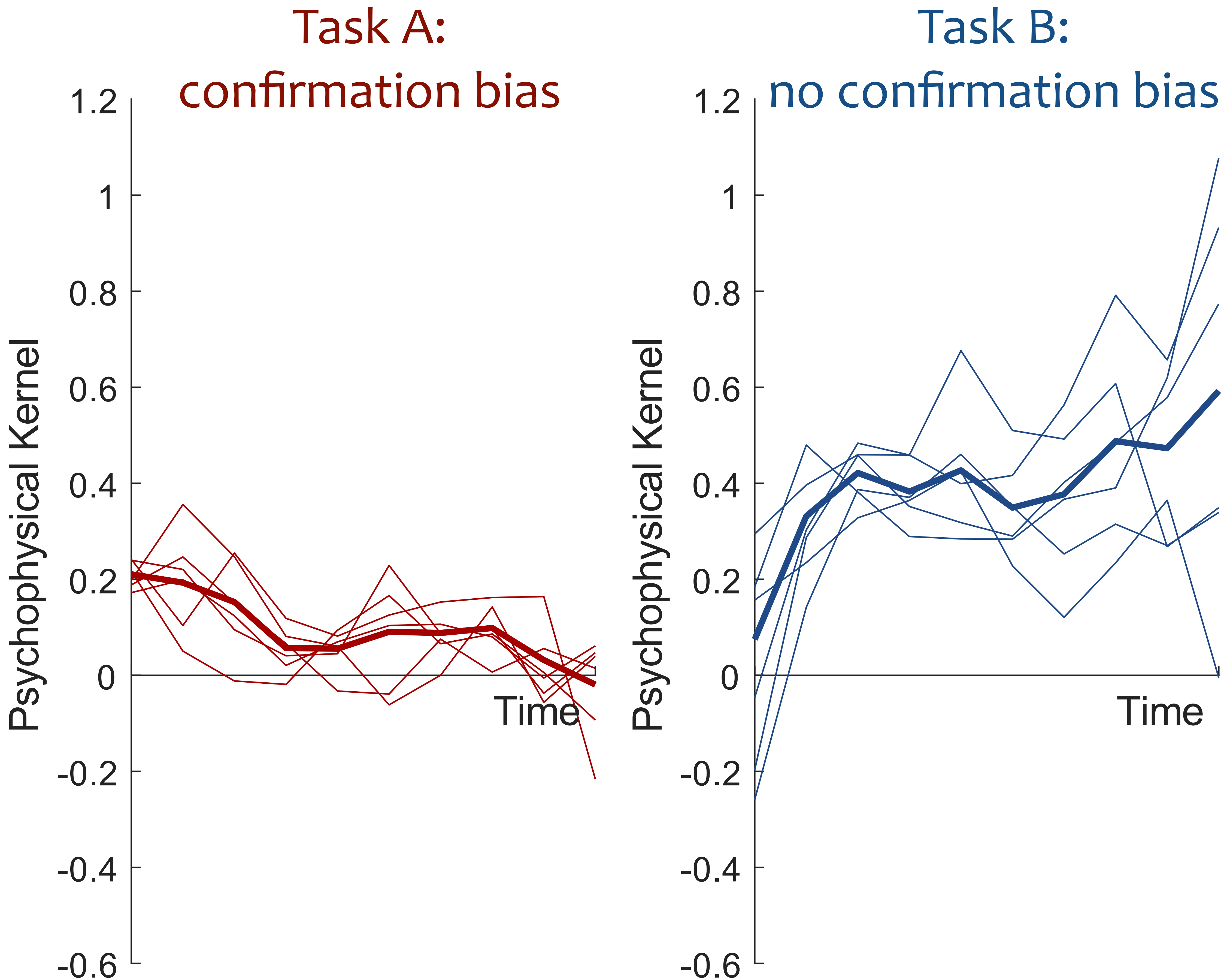


Two different tasks

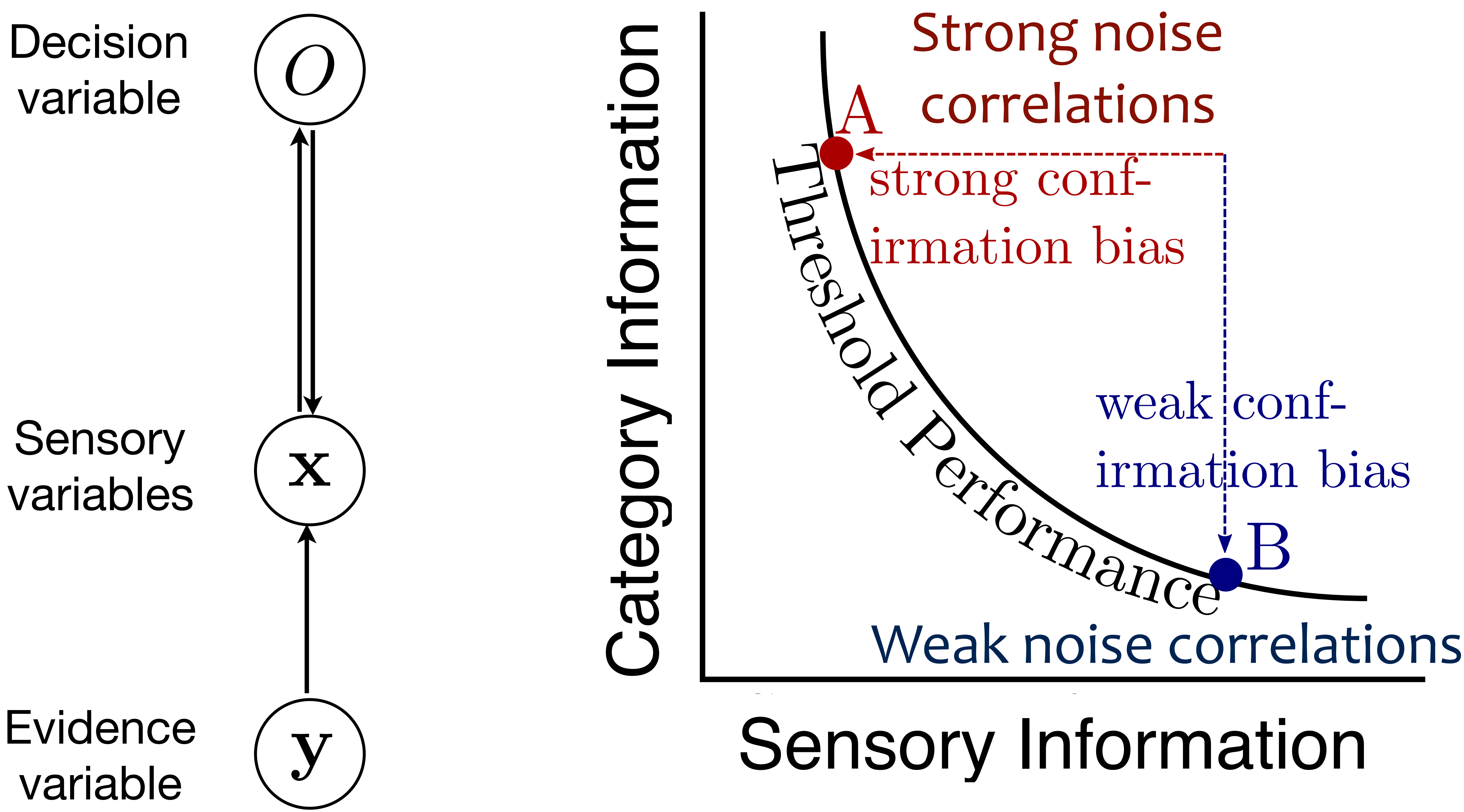


Two different tasks




Task A:
confirmation bias
Task B:
no confirmation bias
Psychophysical Kernel
Time
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6

Neurophysiological predictions



Conclusions: confirmation bias

- Explain different weighting of evidence over time by neural sampling in a hierarchical model

- Explains existing discrepancies in the literature

- Our data confirms model predictions in the same kind of task, same subjects, varying only the 1 factor that matters in our model

  –> Insights for dealing with general confirmation bias?

Overall thoughts

- Responses of sensory neurons represent posterior beliefs about the outside world
- Integrate prior knowledge/expectations via top-down/ feedback signals
- Task-training allows us to manipulate those signals
- Neural population recordings allow us to infer internal beliefs
- Behavioral biases can be used to infer aspects of the inference algorithm used

Acknowledgements

Lab

Richard Lange
Ankani Chattoraj
Sabya Shivkumar

Theoretical collaborators

Pietro Berkes
Jozsef Fiser (CEU)

Experimental collaborators

Adrian Bondy (Princeton)
Bruce Cumming (NIH)

Camille Gomez-Faberge (Harvard)
Till Hartmann (Harvard)
Rick Born (Harvard)

Jacob Yates (Rochester)

University of Rochester (NY)

Looking for postdocs!