Predictive Coding

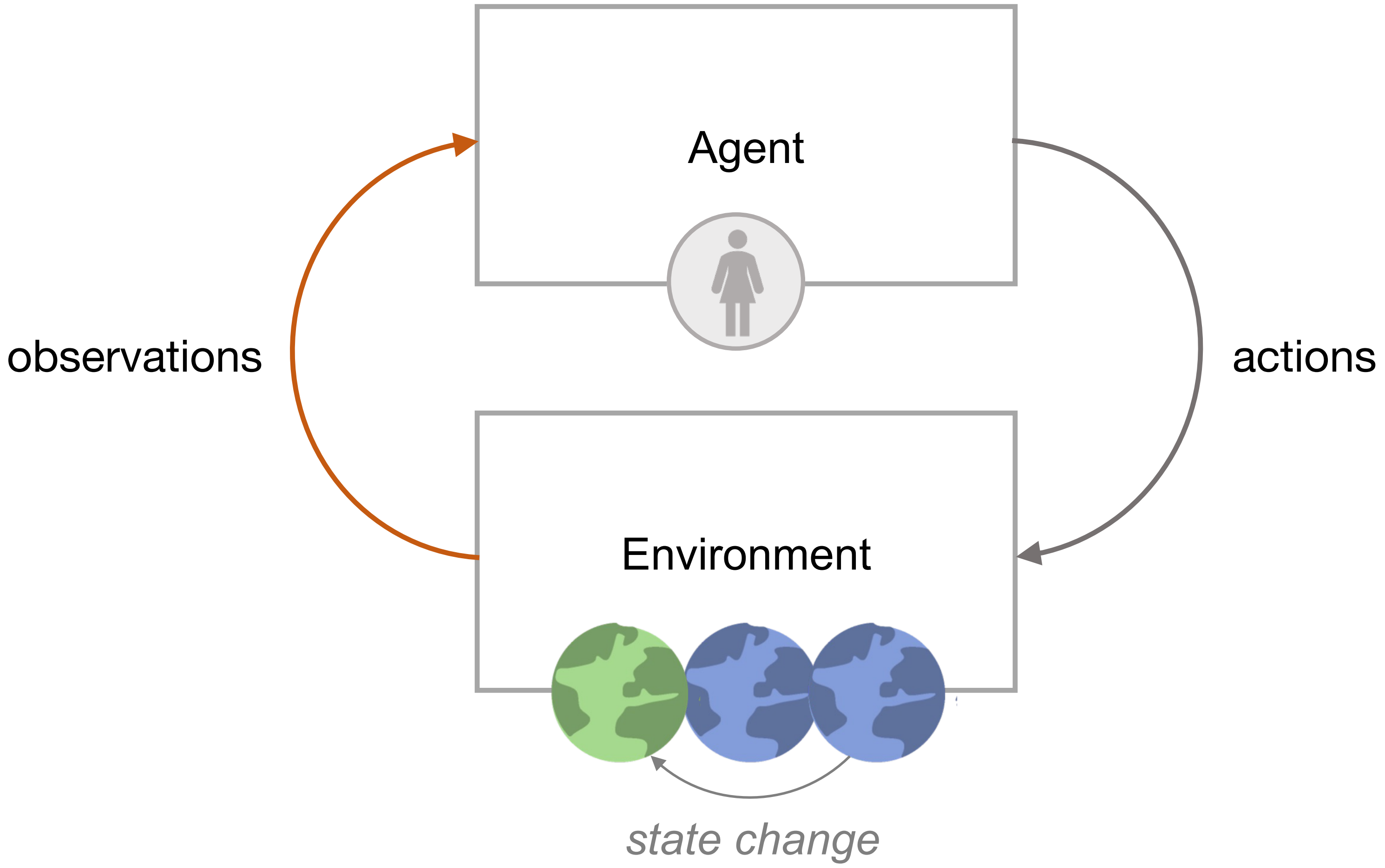
Computational Psychiatry Course

11.09.2024

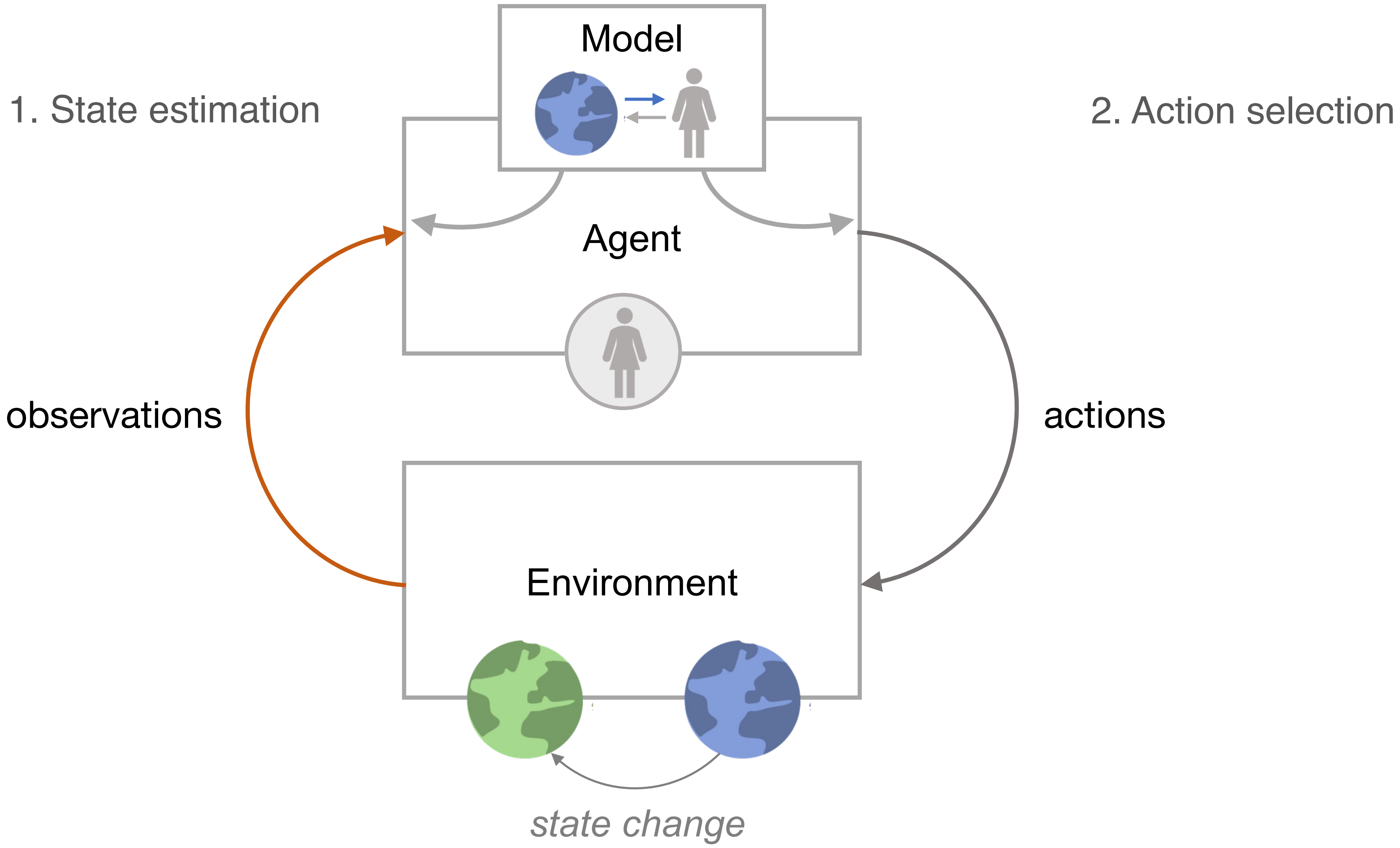
Lilian A. Weber

What is predictive coding?

The action-perception loop



Predictive processing: using internal models

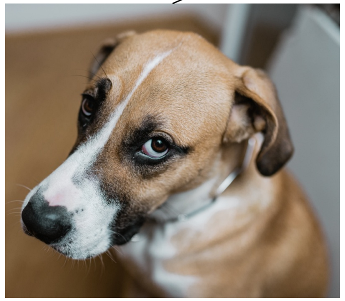


Internal models

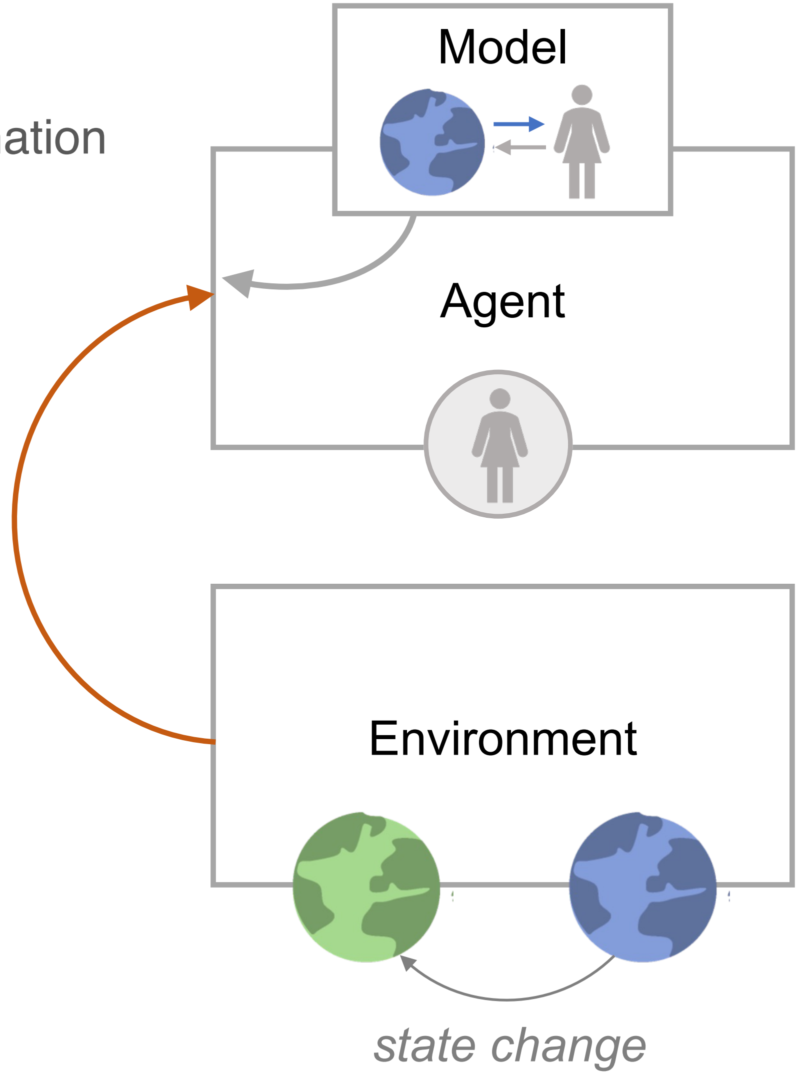
y = a mess

x = dog guilty

$$p(x|y,m) = \frac{p(y|x,m)p(x|m)}{p(y)} = \frac{p(y|x,m)p(x|m)}{\int p(y|x,m)p(x,m)dx}$$

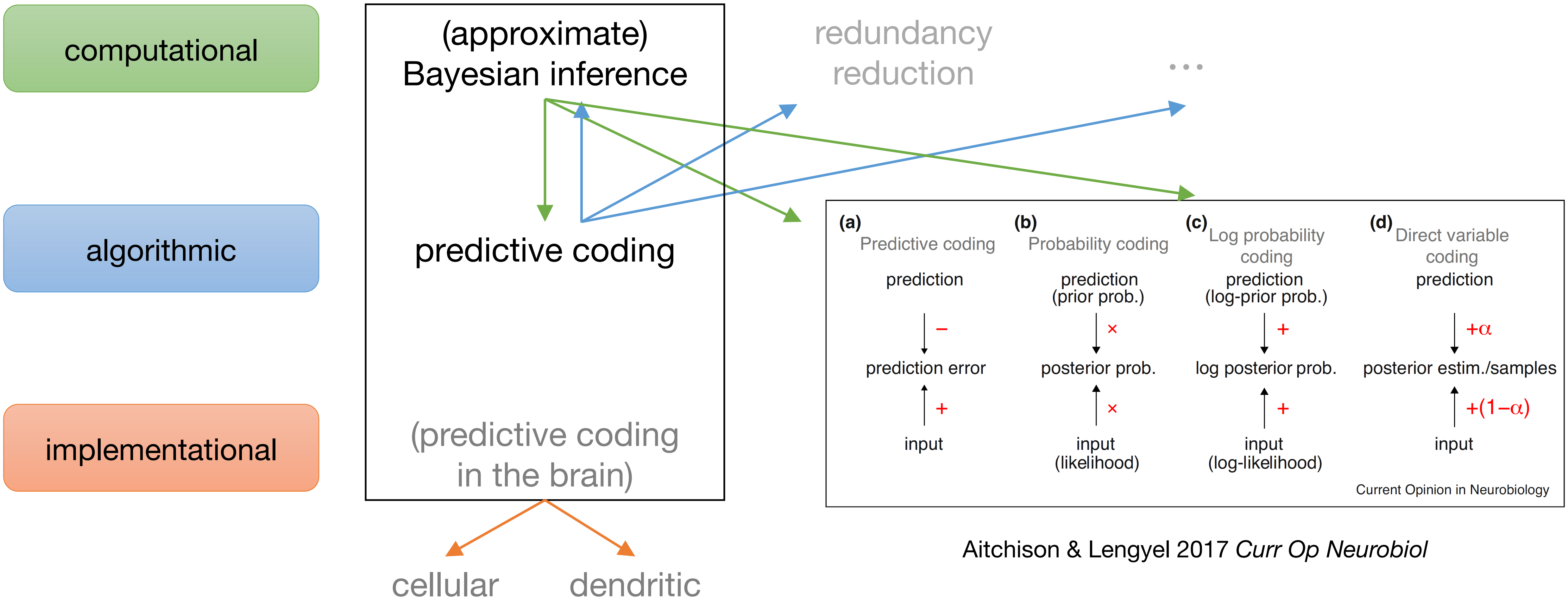


1. State estimation



Levels of analysis

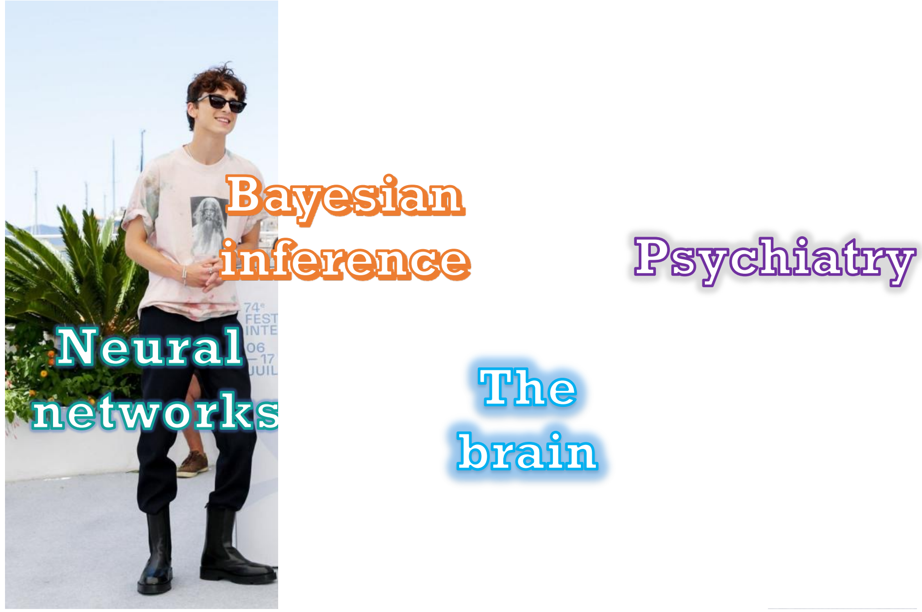
"Bayesian predictive coding"



Aitchison & Lengyel 2017 *Curr Op Neurobiol*

Structure of this talk

- [x] 1. What is predictive coding?
- [ ] 2. How does predictive coding work?
- [ ] 3. How do brains do predictive coding?
- [ ] 4. Why is it useful for computational psychiatry?

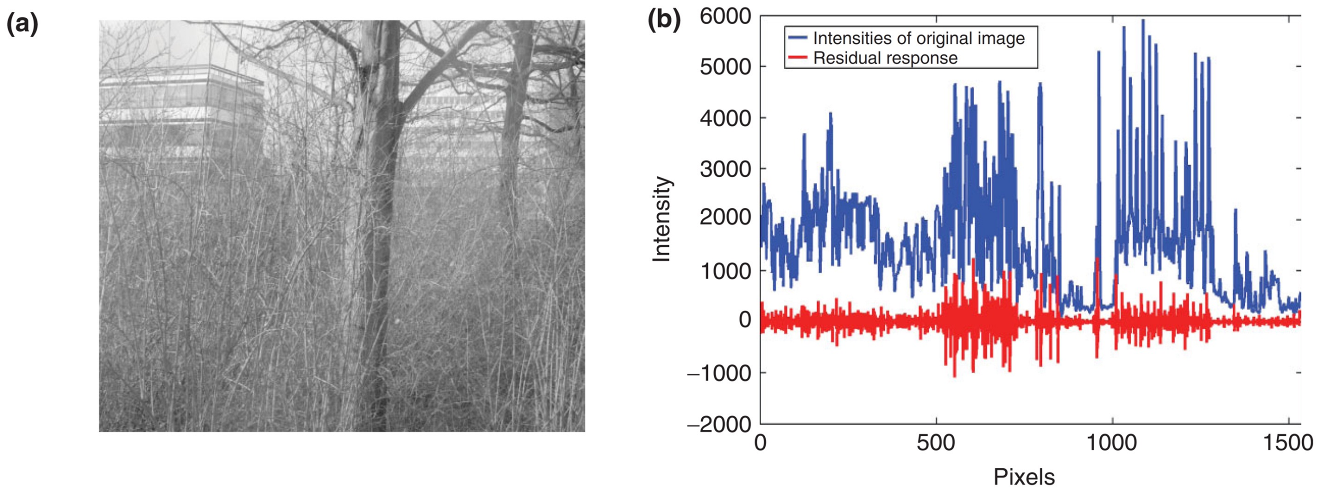


How does predictive coding work?

And how does it solve Bayesian inference?

How does predictive coding work?

prediction error = $\boxed{\text{prediction}}$ – input

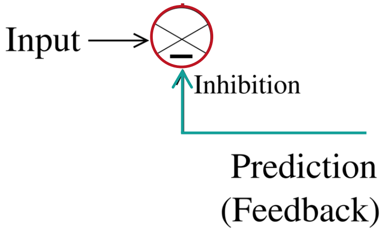


Huang & Rao 2011 *WIREs Cogn Sci*

How does predictive coding work?

prediction error = prediction – input

$$\begin{bmatrix} i_1 \\ i_2 \\ i_3 \\ \vdots \\ i_K \end{bmatrix}$$



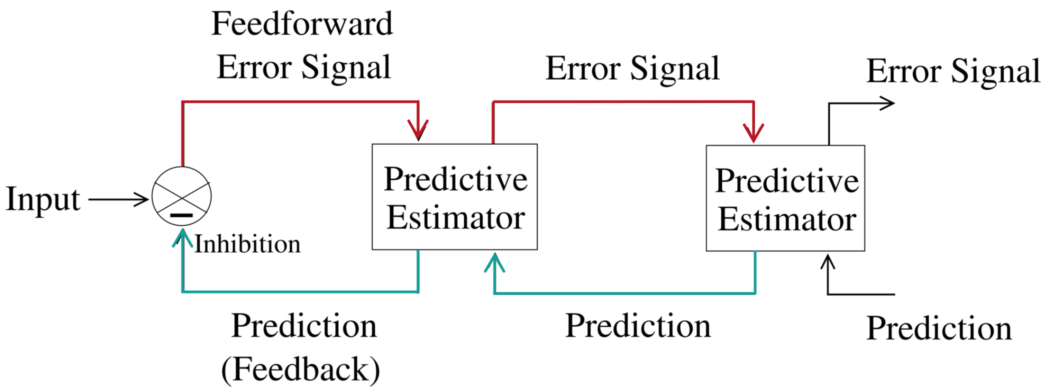
$$\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{bmatrix}$$

$$\mathbf{I} = \boxed{f(U\boxed{\mathbf{r}}) + \mathbf{n}}$$

Rao & Ballard 1999 *Nat Neurosci*

How does predictive coding work?

$\boxed{\text{prediction error}} = \boxed{\text{prediction}} - \text{input}$

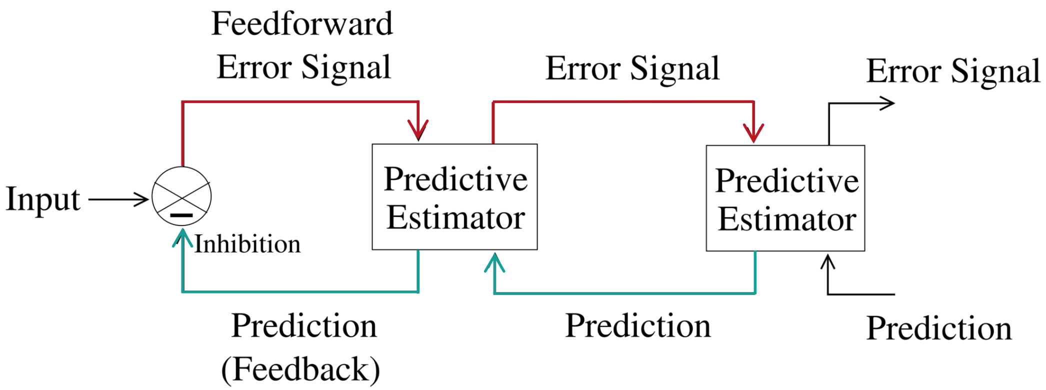


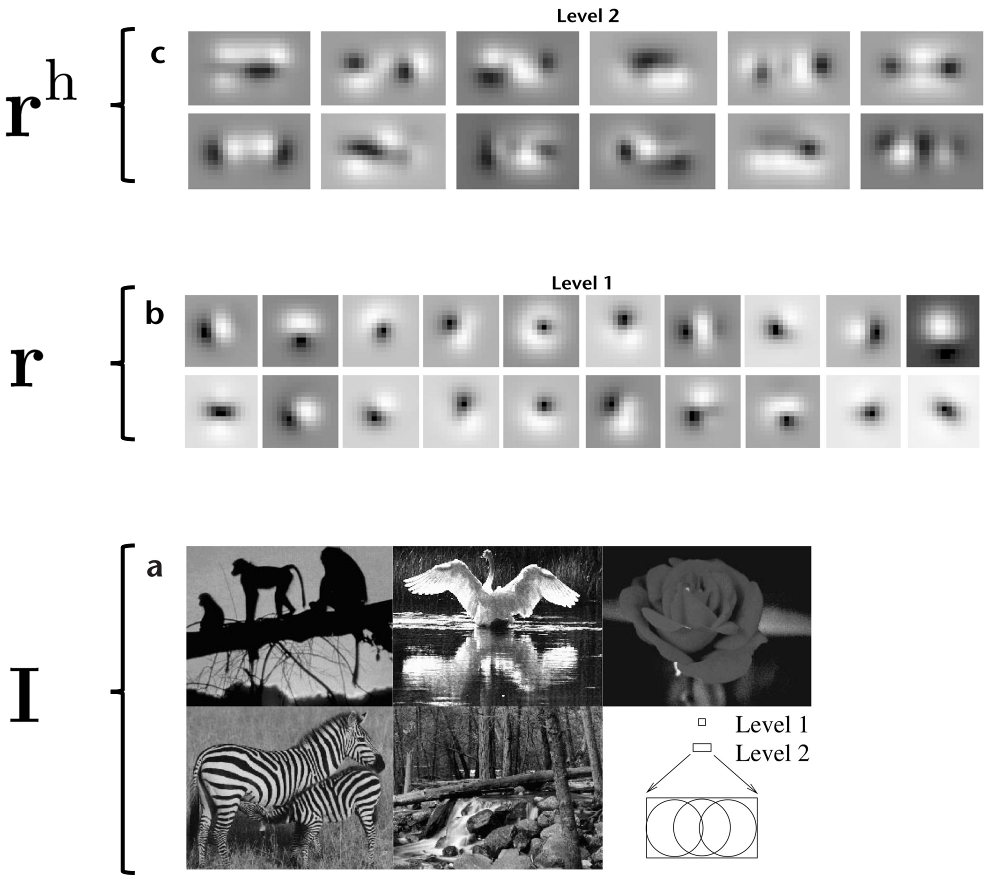
$$\mathbf{I} = f(U\mathbf{r}) + \mathbf{n} \qquad \mathbf{r} = f(U^{\mathrm{h}}\mathbf{r}^{\mathrm{h}}) + \mathbf{n}^{\mathrm{td}}$$

Rao & Ballard 1999 *Nat Neurosci*

How does predictive coding work?

$\boxed{\text{prediction error}} = \boxed{\text{prediction}} - \text{input}$

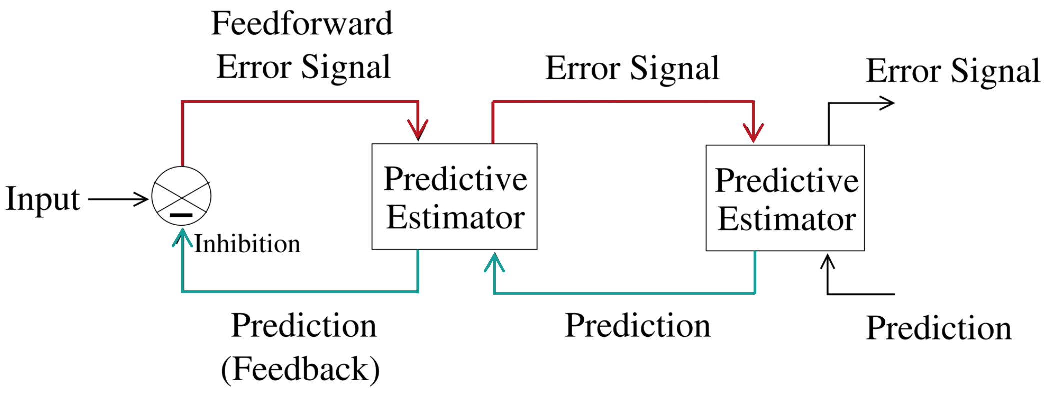


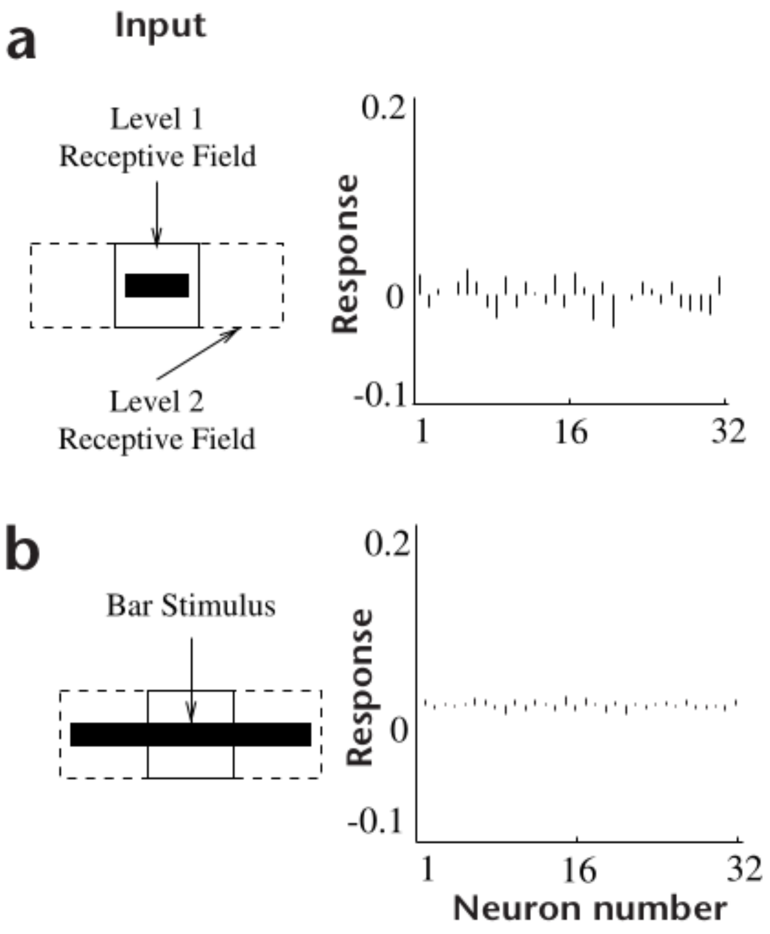


Rao & Ballard 1999 *Nat Neurosci*

How does predictive coding work?

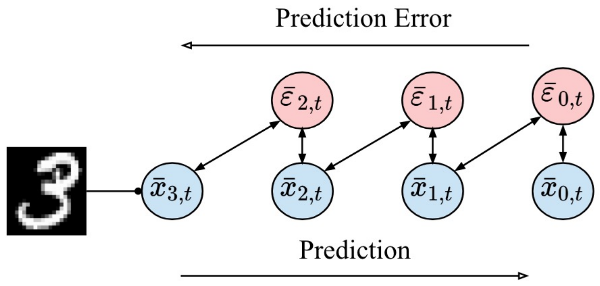
$\boxed{\text{prediction error}} = \boxed{\text{prediction}} - \text{input}$



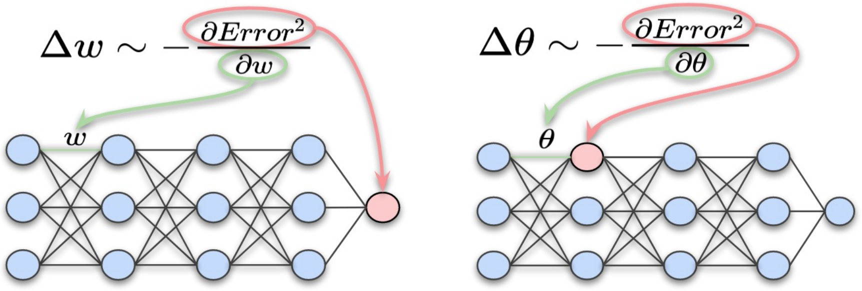


Rao & Ballard 1999 *Nat Neurosci*

How does predictive coding work?



(a) Sketch of a PCN



(b) Difference between BP (left) and PC (right)

Millidge et al. 2022 *arXiv*

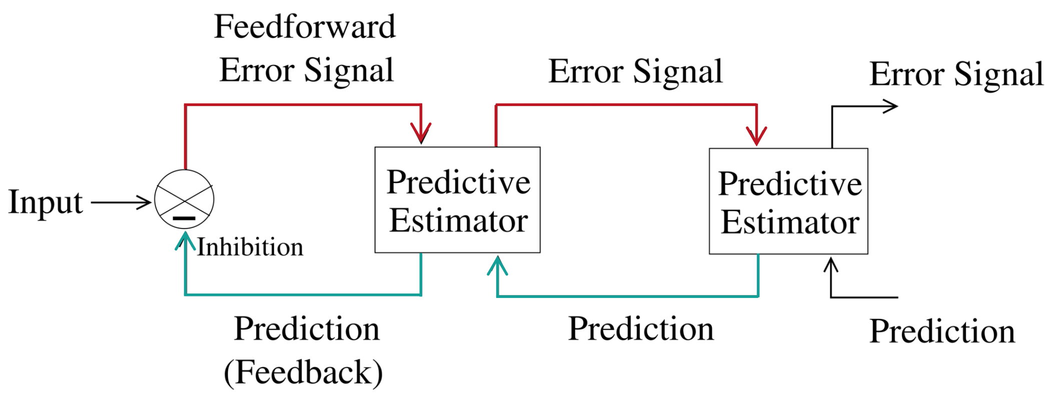
- Recurrent neural networks outperform purely feedforward networks in the presence of noise (e.g., Alamia et al. 2023 *Neural Networks*)
- Predictive coding networks can approximate the backpropagation of error algorithm (Whittington & Bogacz 2017 *Neural Comput*; Song et al. 2020 *Adv Neural Inf Process Syst*)
- Predictive coding confers advantages over BP especially when data is scarce (Millidge et al. 2022 *arXiv*; Song et al. 2024 *Nat Neurosci;* Innocenti et al. 2024 *arXiv*)

How does predictive coding solve Bayesian inference?

1. The architecture represents a hierarchical generative model of how sensory inputs are caused.

$$\mathbf{I} = f(U\mathbf{r}) + \mathbf{n}$$

$$\rightarrow p(\mathbf{I}|\mathbf{r}, U)$$



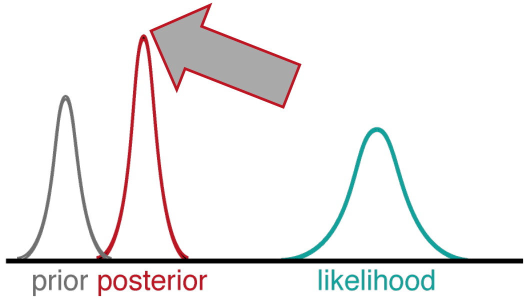
2. The cost function is the negative log joint: maximising the joint means finding the MAP estimates in Bayesian inference.

$$E = -\log p(\mathbf{I}|\mathbf{r}, U) - \log p(\mathbf{r}) - \log p(U)$$
$$= -\log \boxed{(p(\mathbf{I}|\mathbf{r}, U)\, p(\mathbf{r})\, p(U))}$$

Bayes' Rule

posterior~ likelihood · prior

p(x|y,m)~ p(y|x,m)p(x|m)



How does predictive coding solve Bayesian inference?

3. The prediction errors that update the estimated causes are weighted by their inverse variance.

$$\frac{d\mathbf{r}}{dt} = -\frac{k_1}{2}\frac{\partial E}{\partial \mathbf{r}}$$

Inputs

Expectations

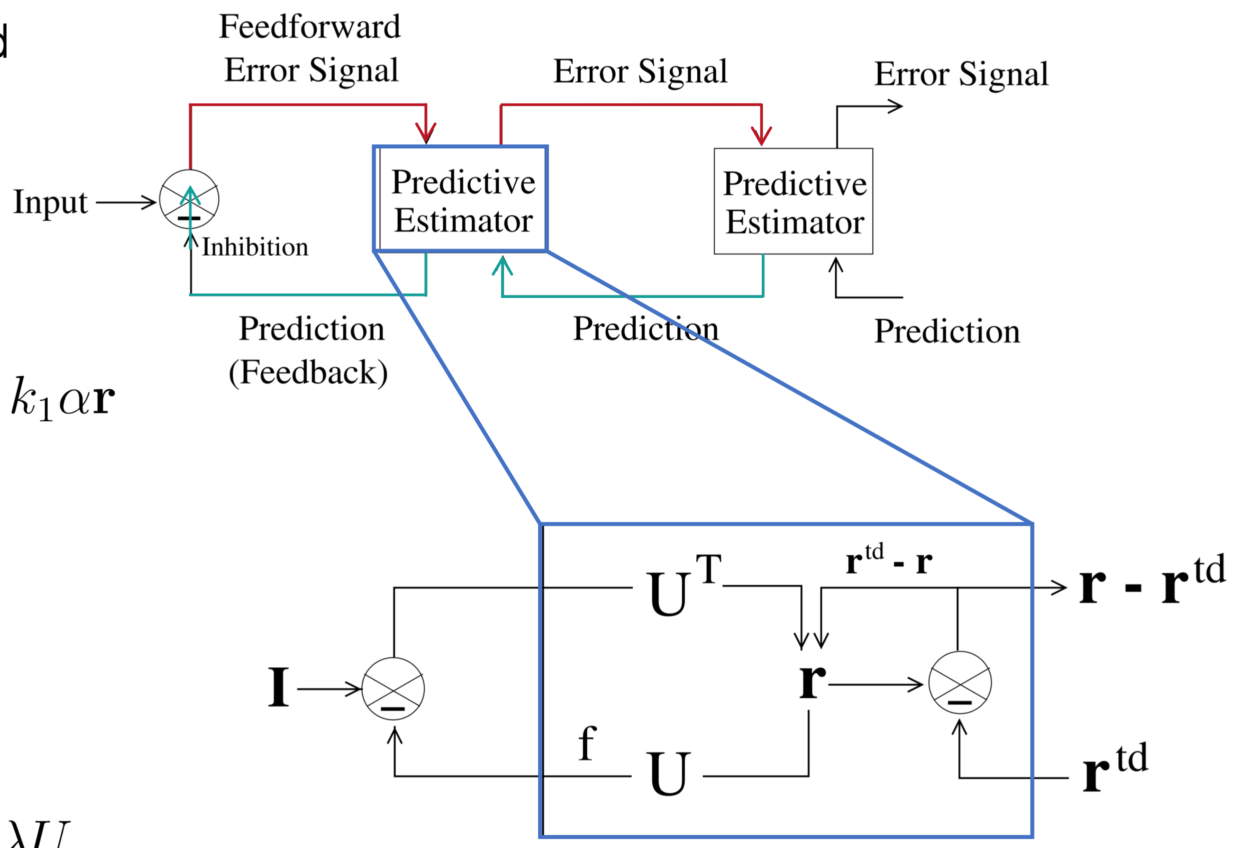
$$= \frac{k_1}{\sigma^2} U^{\intercal} \frac{\partial f}{\partial U\mathbf{r}}^{\intercal} (\mathbf{I} - f(U\mathbf{r})) + \frac{k_1}{\sigma_{\mathrm{td}}^2}(\mathbf{r}^{\mathrm{td}} - \mathbf{r}) - k_1 \alpha \mathbf{r}$$

precision-weighting

4. A single cost function accounts for inference (udating **r**) and learning (updating U)

$$\frac{dU}{dt} = -\frac{k_2}{2}\frac{\partial E}{\partial U} = \frac{k_2}{\sigma^2}\frac{\partial f}{\partial U\mathbf{r}}^{\intercal}(\mathbf{I} - f(U\mathbf{r}))\mathbf{r} - \frac{k_2}{2}\lambda U$$

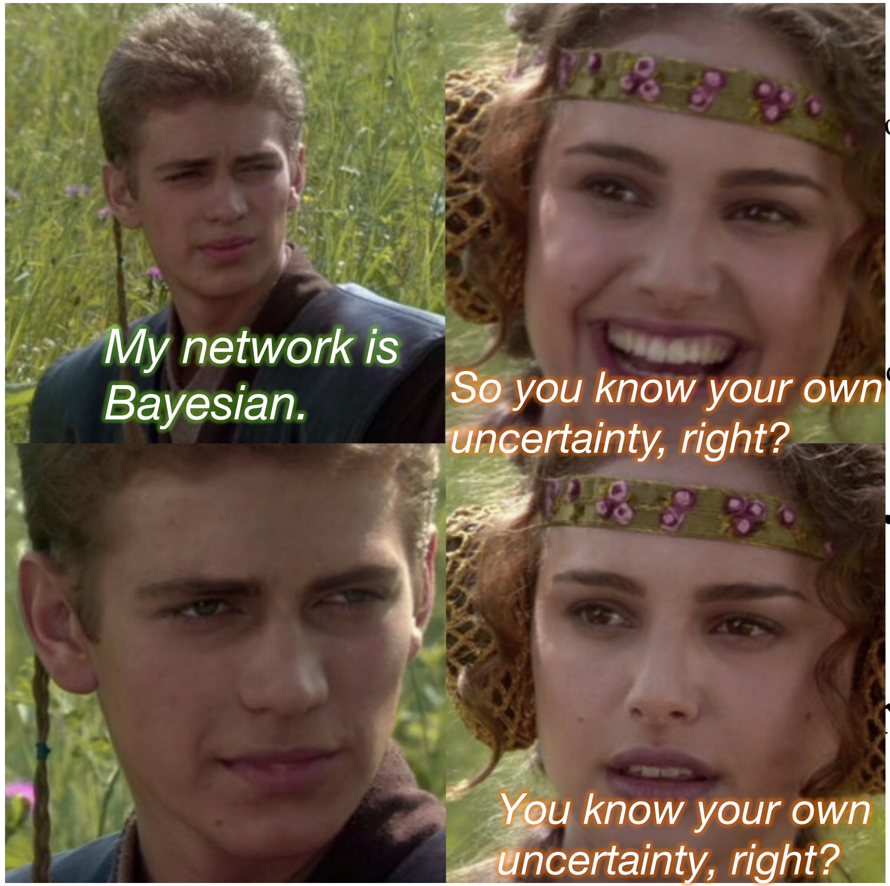
Hebbian learning



How does predictive coding solve Bayesian inference?

Summary

- A hierarchy of causes (predictive estimators and PE units)
- Recurrent message passing: PEs are signaled upwards (bottom-up) and predictions are signaled top-down
- Inference: finding the most appropriate causes to describe the data
- Learning: finding the most appropriate model to describe the data
- Both involve minimizing precision-weighted prediction errors
- This looks like Bayesian inference for finding the most likely causes of sensory inputs (MAP estimates) and like Hebbian learning, respectively



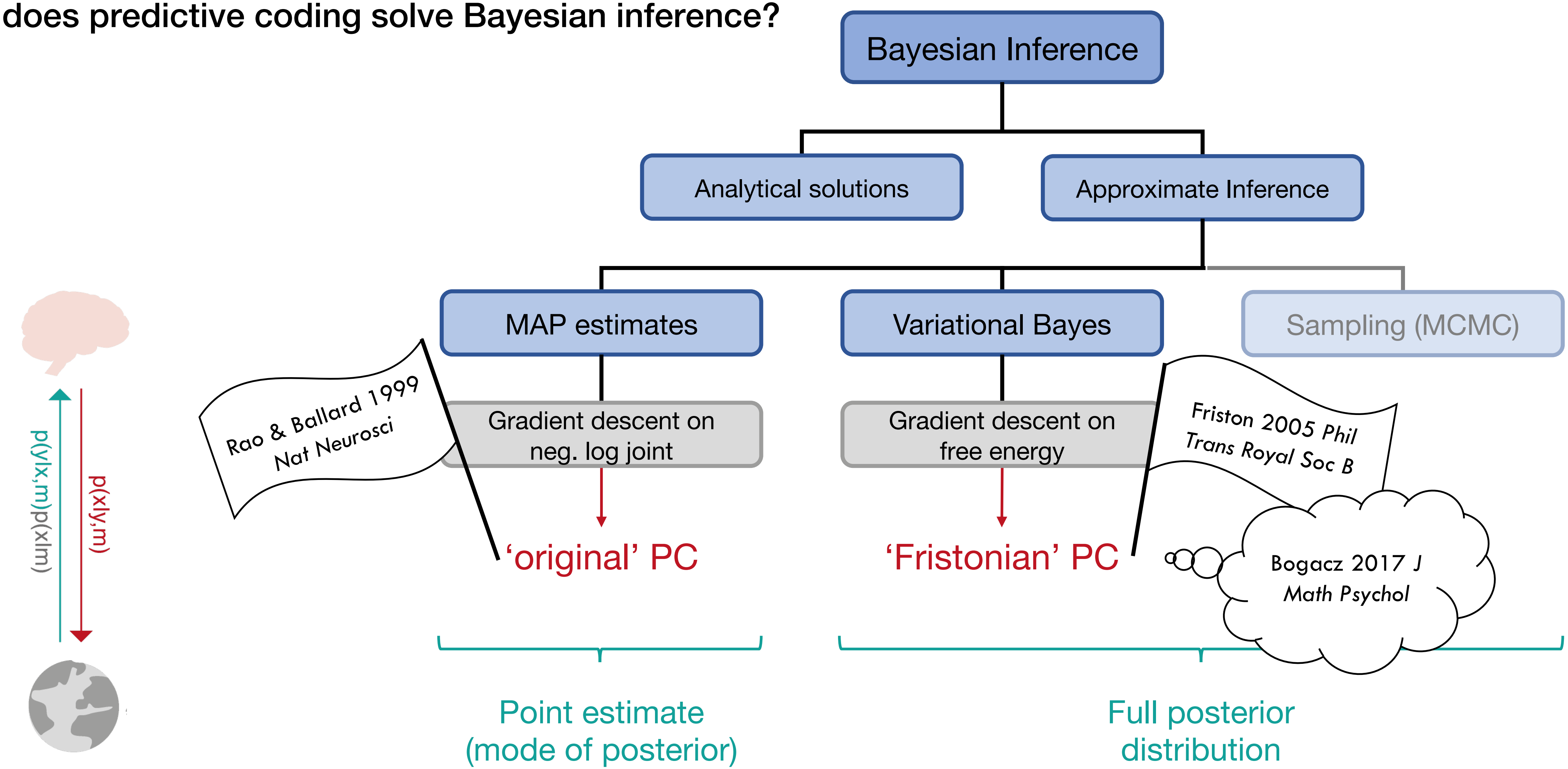
or Signal

diction

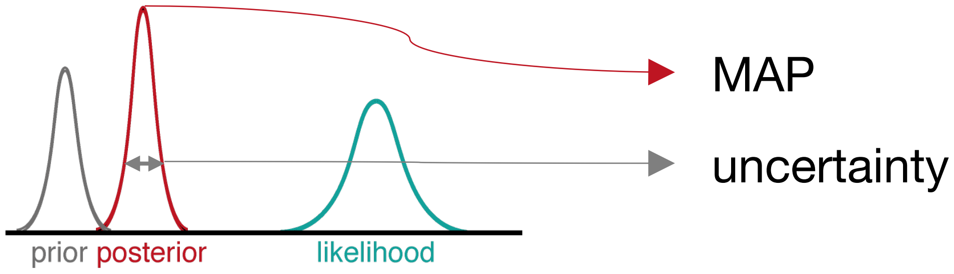
- $\mathbf{r}^{td}$

td

How does predictive coding solve Bayesian inference?



What about uncertainty?



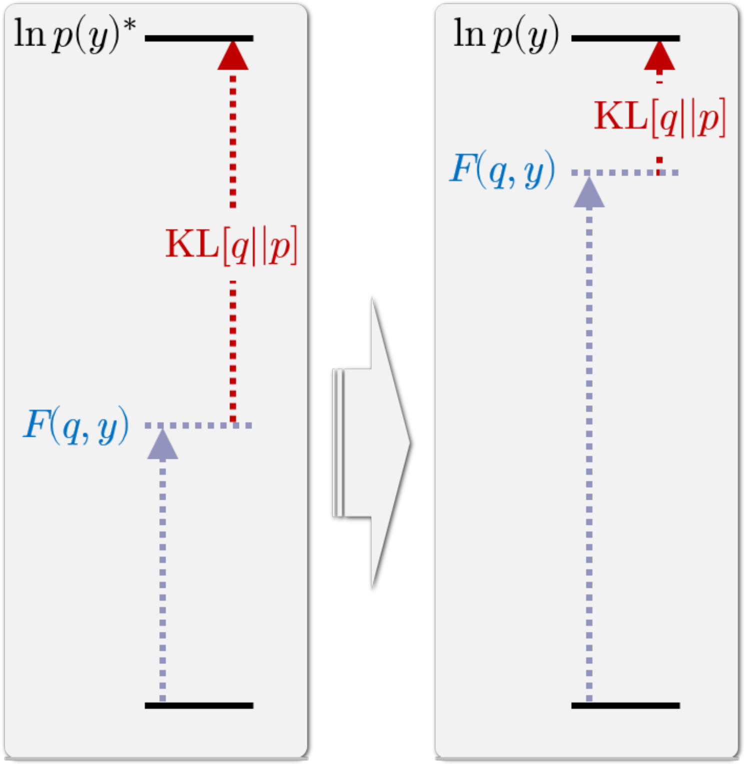
$$p(x|y) = \frac{p(x)p(y|x)}{p(y)} \qquad p(y) = \int p(x)p(y|x)\,\mathrm{d}x$$

$$q(x): \text{ our best guess about } p(x|y)$$

$$\begin{aligned}\mathrm{KL}(q(x), p(x|y)) &= \ln p(y) - \int q(x) \ln \frac{p(y,x)}{q(x)}\,\mathrm{d}x \\ &= \ln p(y) - F\end{aligned}$$

$$\ln p(y|m) = \mathrm{KL}(q(x), p(x|y,m)) + F(q(x), p(x,y|m))$$

$$F = \int q(x) \ln \frac{p(y,x)}{q(x)}\,\mathrm{d}x$$



Bishop, 2008; Friston et al. 2007 *NeuroImage*

How does predictive coding solve Bayesian inference?

> Predictive Coding can be viewed as the implementation of a particular scheme of approximate Bayesian inference in the brain.

The free energy formulation of predictive coding

$$F = \int q(x) \ln \frac{p(y,x)}{q(x)} \, \mathrm{d}x$$

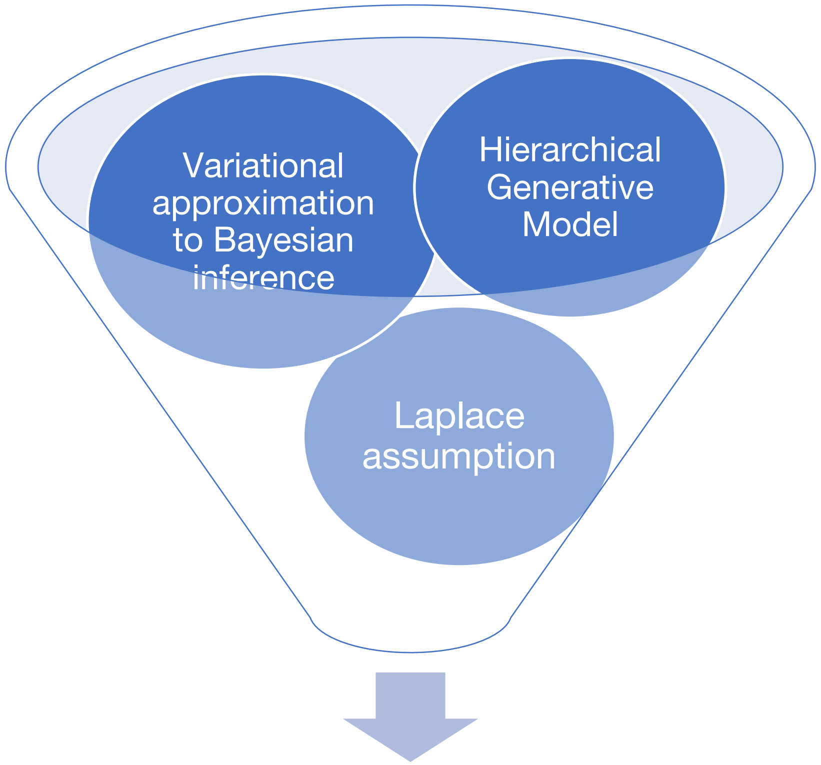
Inference: maximizing F to find the approximate posterior

Learning: maximizing F to improve the model and thereby reduce average surprise over time.

Additional assumption: *q(x)* is a Gaussian. We only need to represent the mean and the variance.

→ *precision-weighted prediction errors*

→ *learning about the **precision** of beliefs*



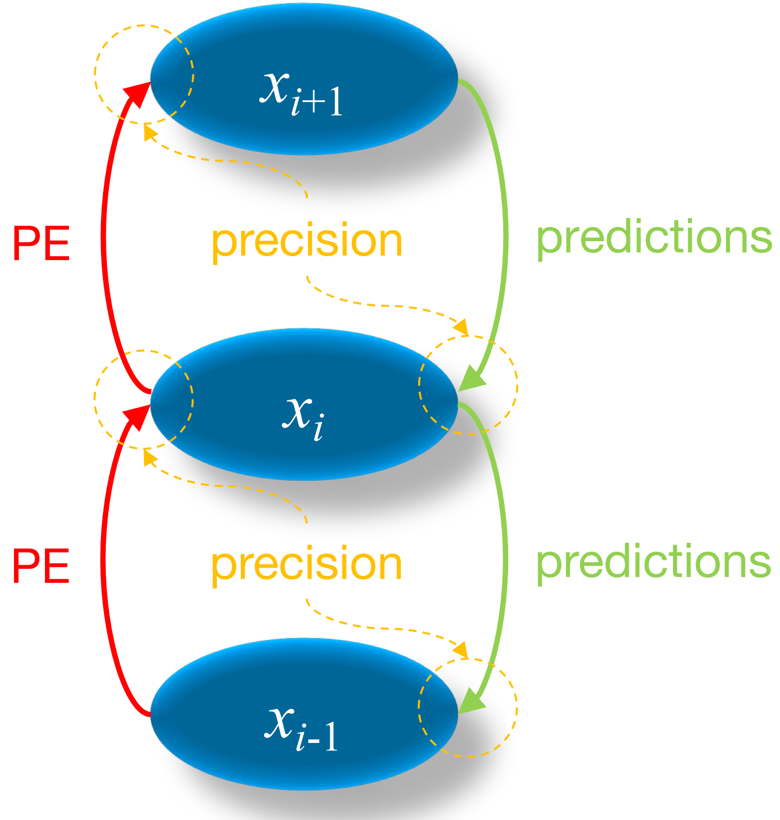
Predictive Coding

The main ingredients of predictive coding (from a computational point of view)

- A hierarchical generative model of sensory inputs
- The estimates of the causes generate predictions of sensory input
- The estimates of causes are updated in response to prediction errors (mismatches)
- The relative influence of PEs and predictions is determined by their relative precision (certainty)
- These computations underlie both perception (inference) and learning (model update)

Δbelief **∼**

precision **×** PE



Adapted from Stephan et al. 2016 *Brain*

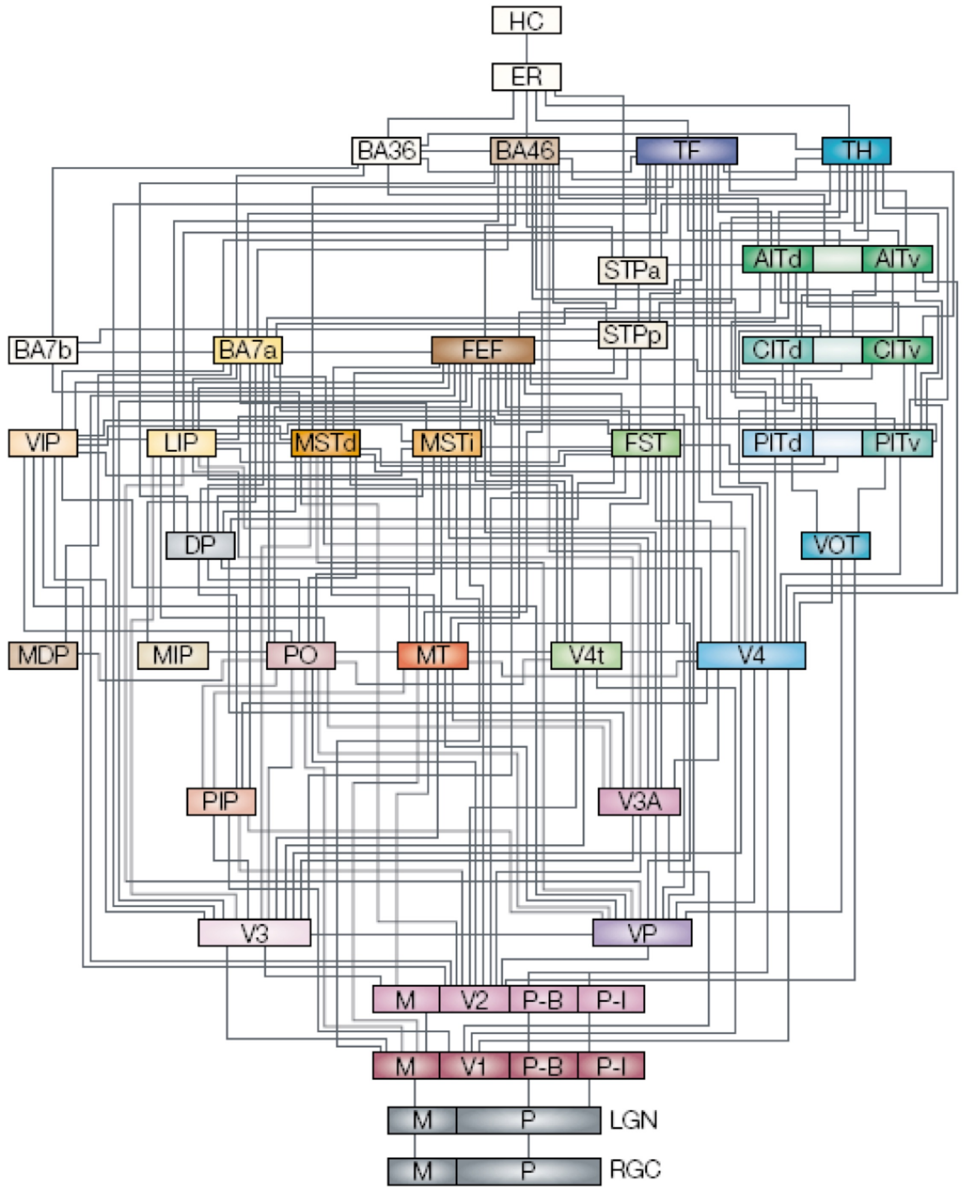
How do brains do predictive coding?

What do brains need for predictive coding?

How can we test predictive coding models of the brain?

What do brains need for predictive coding?

- A hierarchical generative model of sensory inputs
- The estimates of the causes generate predictions of sensory input
- The estimates of causes are updated in response to prediction errors (mismatches)
- The relative influence of PEs and predictions is determined by their relative precision (certainty)
- These computations underlie both perception (inference) and learning (model update)

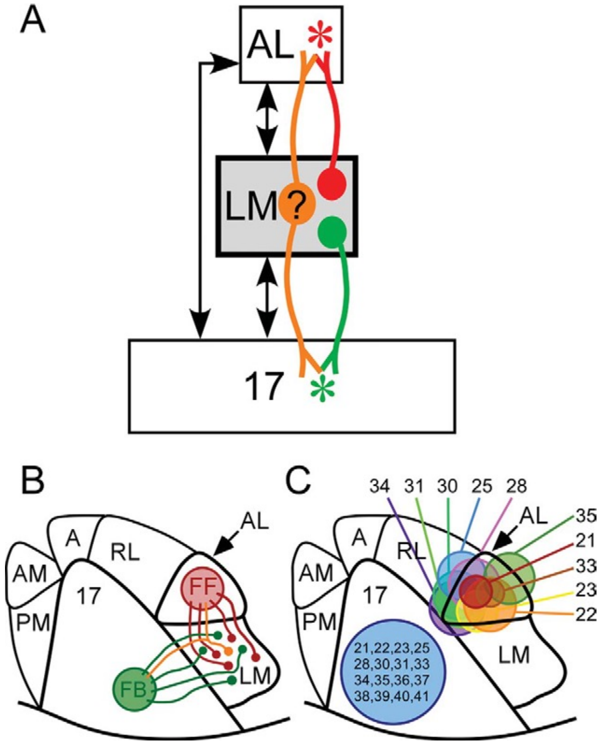


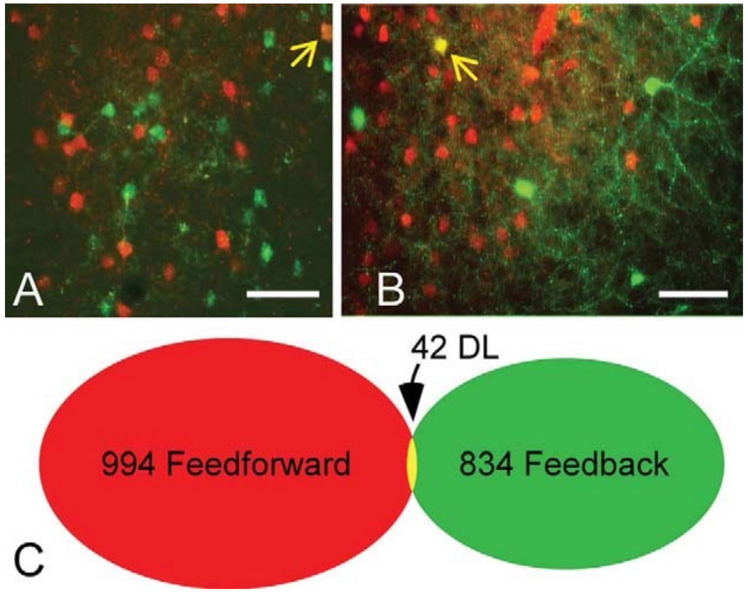
Felleman & Van Essen 1991 *Cereb Cortex*

What do brains need for predictive coding?

1. The source populations of forward and backward pathways should be completely separate, given their functional distinction.

- A hierarchical generative model of sensory inputs
- The estimates of the causes generate predictions of sensory input
- The estimates of causes are updated in response to prediction errors (mismatches)
- The relative influence of PEs and predictions is determined by their relative precision (certainty)
- These computations underlie both perception (inference) and learning (model update)



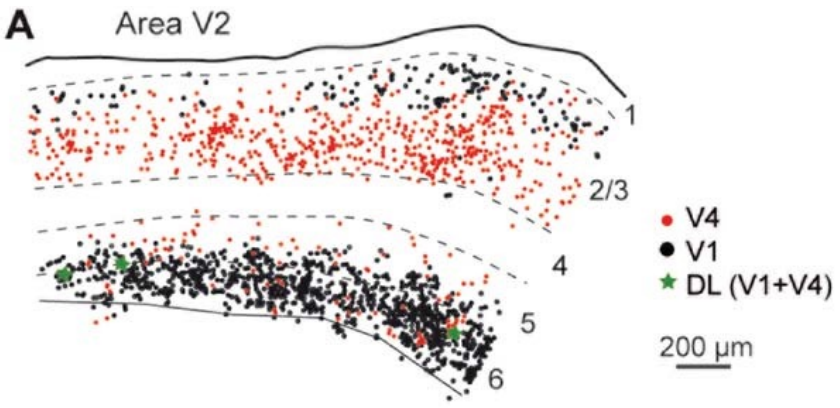


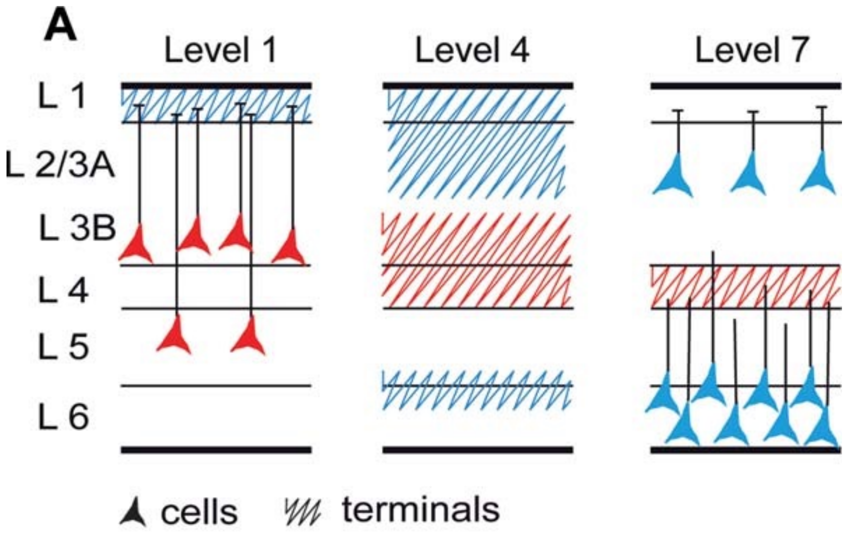
Berezovskii et al. 2011 *J Comp Neurol*

What do brains need for predictive coding?

1. The source populations of forward and backward pathways should be completely separate, given their functional distinction.

- A hierarchical generative model of sensory inputs
- The estimates of the causes generate predictions of sensory input
- The estimates of causes are updated in response to prediction errors (mismatches)
- The relative influence of PEs and predictions is determined by their relative precision (certainty)
- These computations underlie both perception (inference) and learning (model update)



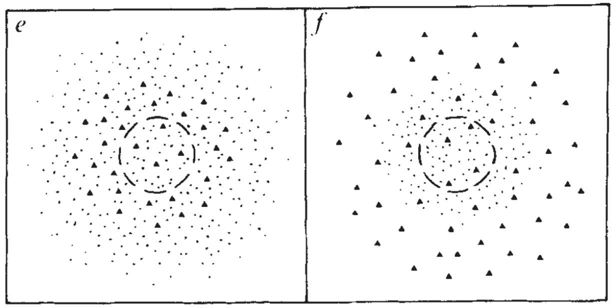


Markov et al. 2014 *J Comp Neurol*

What do brains need for predictive coding?

- A hierarchical generative model of sensory inputs
- The estimates of the causes generate predictions of sensory input
- The estimates of causes are updated in response to prediction errors (mismatches)
- The relative influence of PEs and predictions is determined by their relative precision (certainty)
- These computations underlie both perception (inference) and learning (model update)

2. Backward connections encode predictions and expected precision, so they should be more divergent.



Zeki & Shipp 1988 *Nature*

3. Causes interact non-linearly to generate data, so backward connections should be more modulatory.

Self et al. 2012 *PNAS*
Olsen et al. 2012 *Nature*
Zilles et al. 2004 *J Anat*

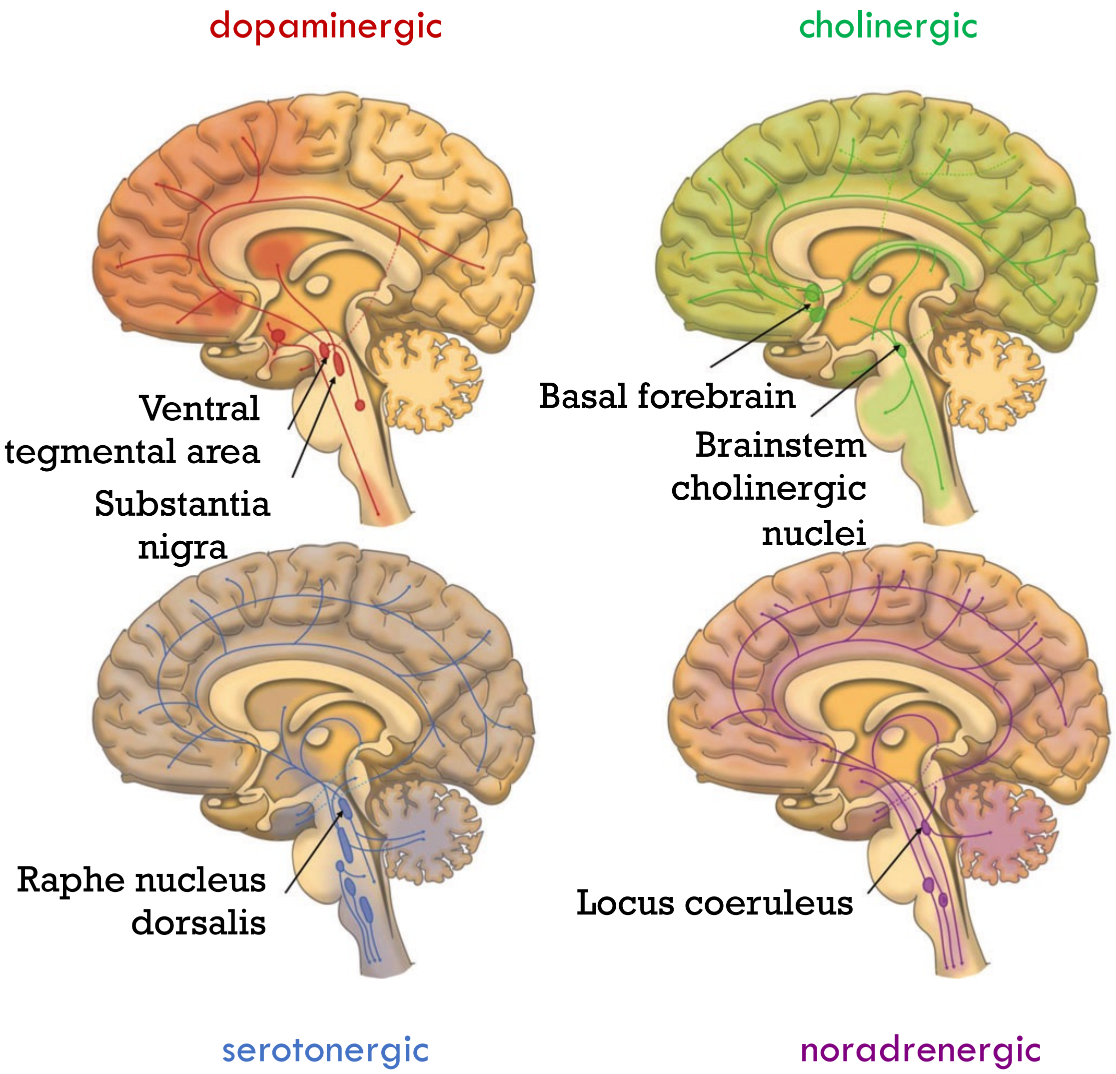
4. Predictions are generated more slowly than prediction errors, suggesting a spectral asymmetry.

Roopun 2006 *PNAS*
Roopun et al. 2008 *Front. Cell. Neurosci*

Buffalo et al. 2011 *PNAS*
Bosman et al. 2012 *Neuron*
Bastos et al. 2015 *NeuroImage*

What do brains need for predictive coding?

- A hierarchical generative model of sensory inputs
- The estimates of the causes generate predictions of sensory input
- The estimates of causes are updated in response to prediction errors (mismatches)
- The relative influence of PEs and predictions is determined by their relative precision (certainty)
- These computations underlie both perception (inference) and learning (model update)



Adapted from Iglesias et al. 2017 *WIRE Cogn Sci*

What do brains need for predictive coding?

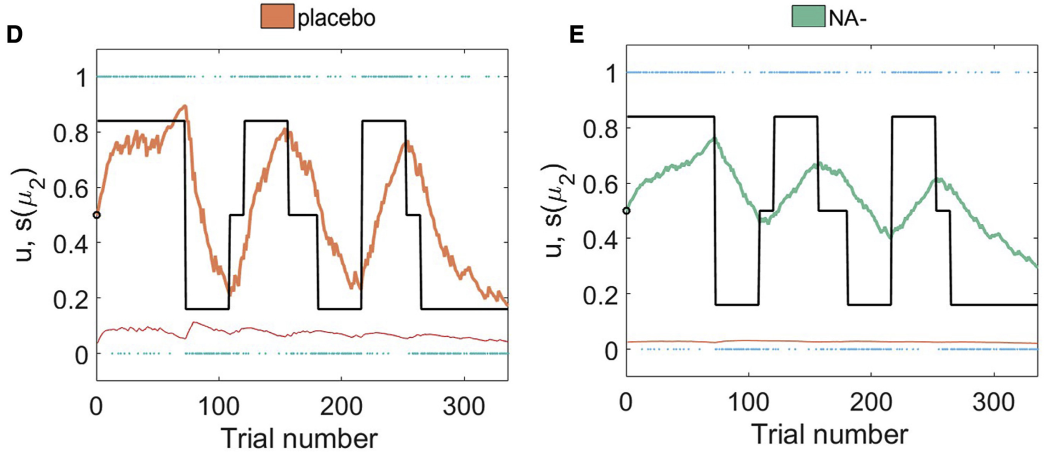
- A hierarchical generative model of sensory inputs
- The estimates of the causes generate predictions of sensory input
- The estimates of causes are updated in response to prediction errors (mismatches)
- The relative influence of PEs and predictions is determined by their relative precision (certainty)
- These computations underlie both perception (inference) and learning (model update)

Dopamine
Fiorillo et al. 2003 *Science*
Tobler et al. 2005 *Science*

see also:
Fiorillo et al. 2005 *Behav Brain Func*
Bunzeck et al. 2010 *Hum. Brain. Map.*

Diederen et al. 2016 *Neuron*
Diederen et al. 2017 *J Neurosci*

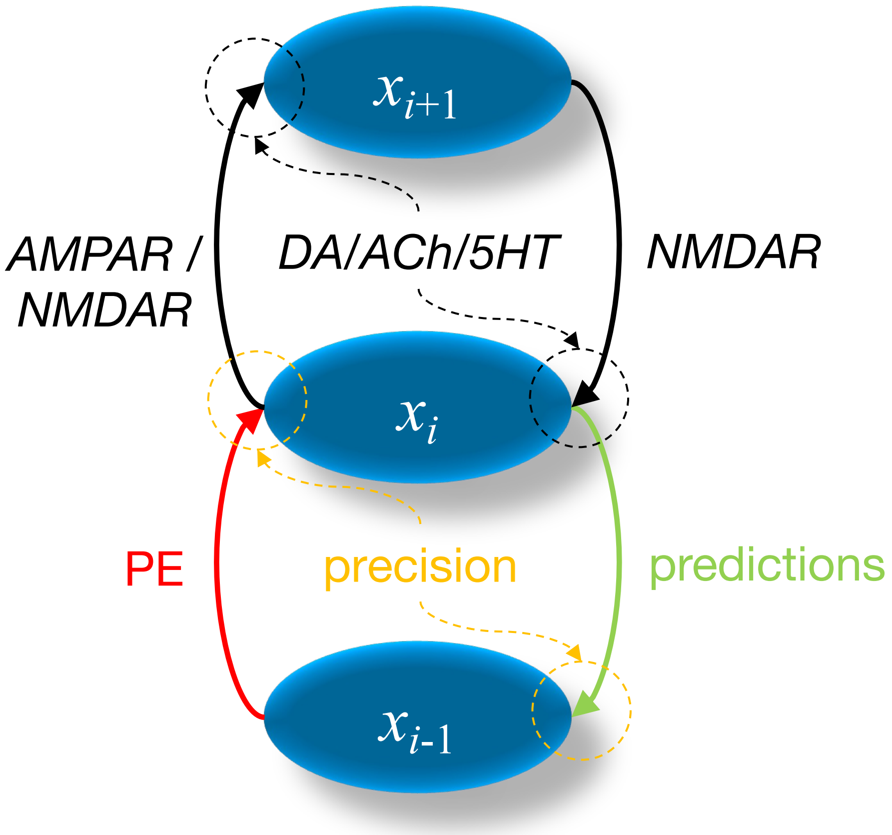
Noradrenaline



Lawson et al. 2021 *Curr Biol*

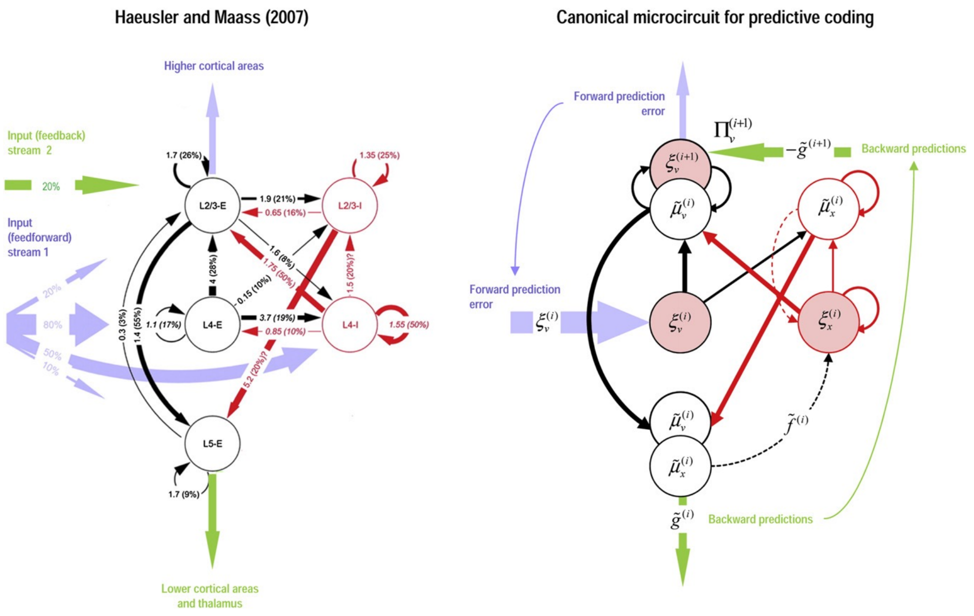
What do brains need for predictive coding?

- A hierarchical generative model of sensory inputs
- The estimates of the causes generate predictions of sensory input
- The estimates of causes are updated in response to prediction errors (mismatches)
- The relative influence of PEs and predictions is determined by their relative precision (certainty)
- These computations underlie both perception (inference) and learning (model update)

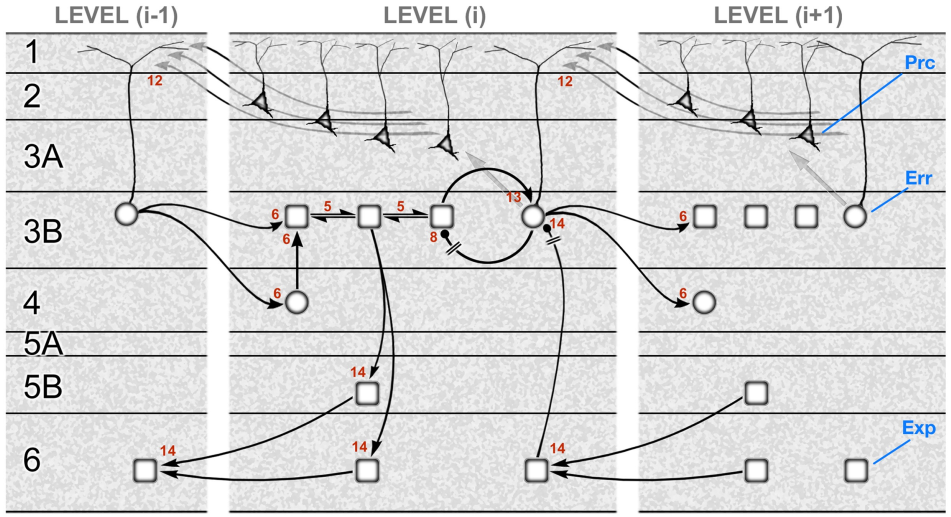


Stephan et al. 2016 *Brain*

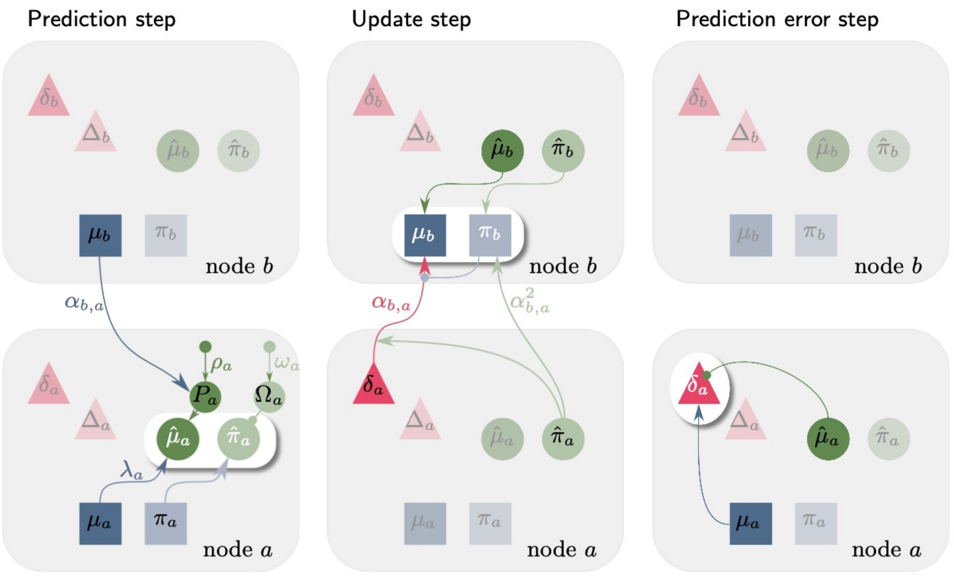
What do brains need for predictive coding?



Bastos et al. 2012 *Neuron*



Shipp 2016 *Frontiers in Psychology*

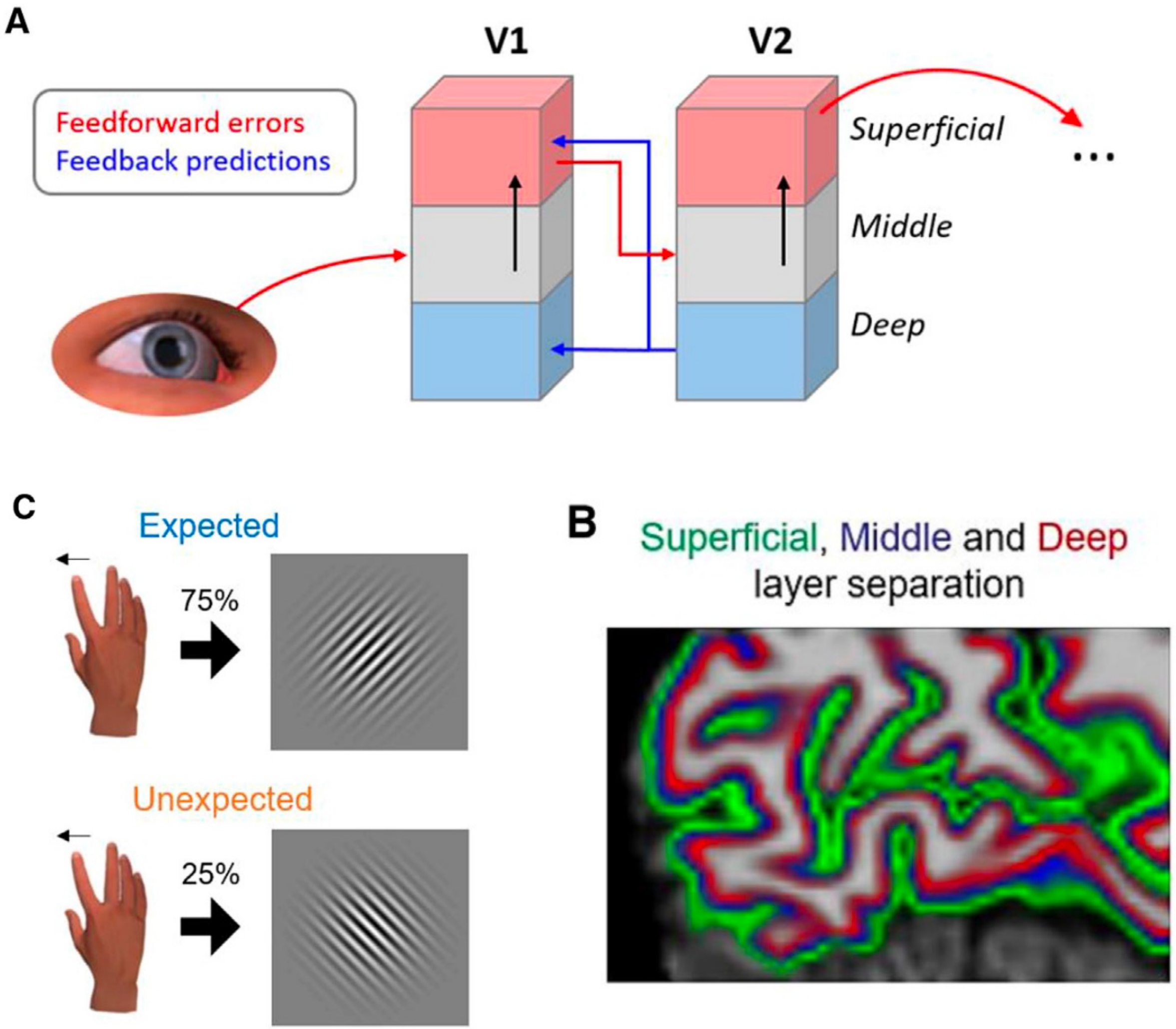


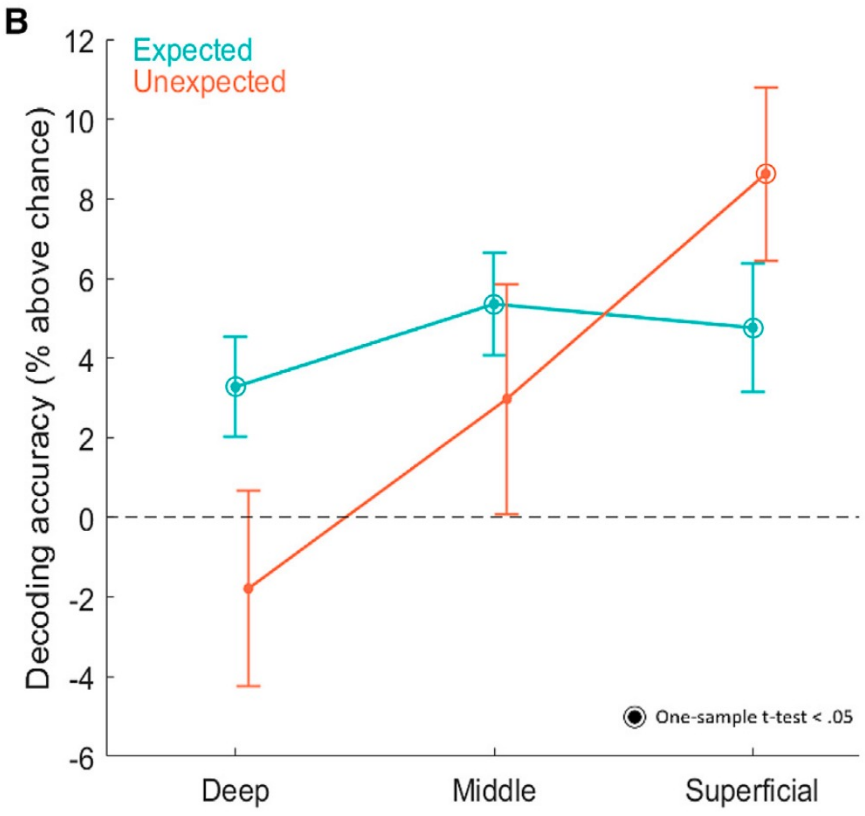
Weber et al. 2023 *arXiv*

Recent overview over evidence:
Walsh al. 2020 *Ann N Y Acad Sci*

How can we test predictive coding theories of the brain? – Layer fMRI

prediction error = $\boxed{\text{prediction}}$ – input



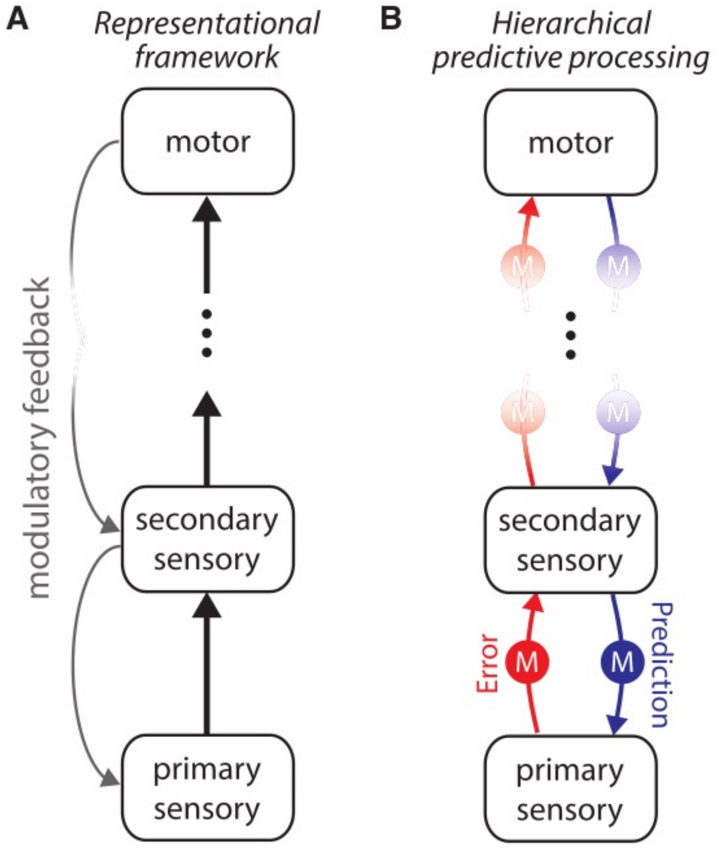


Thomas et al. 2024 *Current Biology*

How can we test predictive coding theories of the brain? - Locomotion

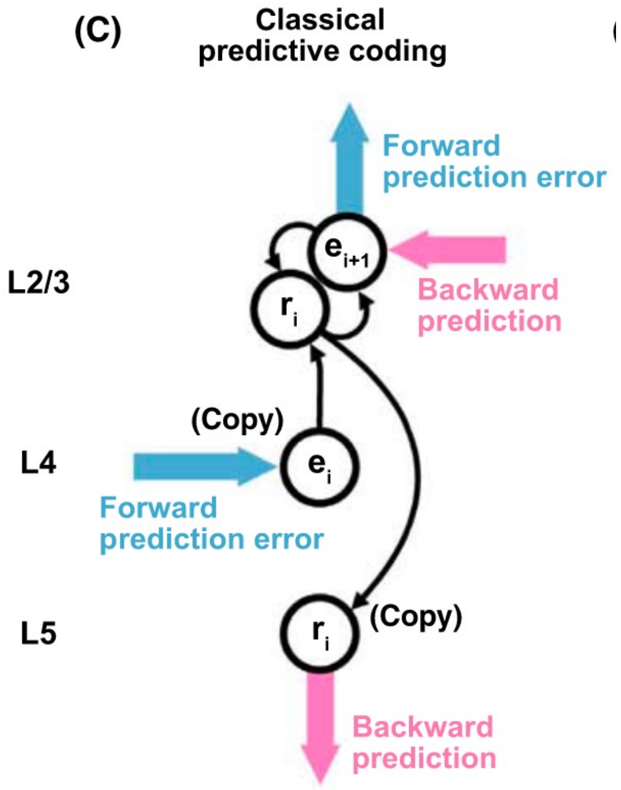
prediction error = $\boxed{\text{prediction}}$ – input

- locomotion is sufficient to drive activity in mouse V1, even in the complete absence of visual input (Keller et al., 2012, Saleem et al., 2013)
- layer 2/3 of sensory cortices signals a mismatch between predicted and actual sensory feedback (Eliades and Wang, 2008, Keller et al., 2012)
- this results from a comparison of an excitatory motor-related input and an inhibitory visual input (Attinger et al., 2017, Zmarz and Keller, 2016)
- A24b/M2 provides a strong and dense projection to V1, which conveys motor-related signals that depend on the mouse's visuomotor experience. **This projection fulfills all the criteria necessary to be interpreted as a prediction of visual flow given a motor output** (Leinweber et al., 2017)
- The role of neuromodulators for predictive coding (Jordan & Keller 2023 *eLife*; Yogesh & Keller 2024 *eLife*)



Keller & Mrsic-Flogel 2018 *Neuron*

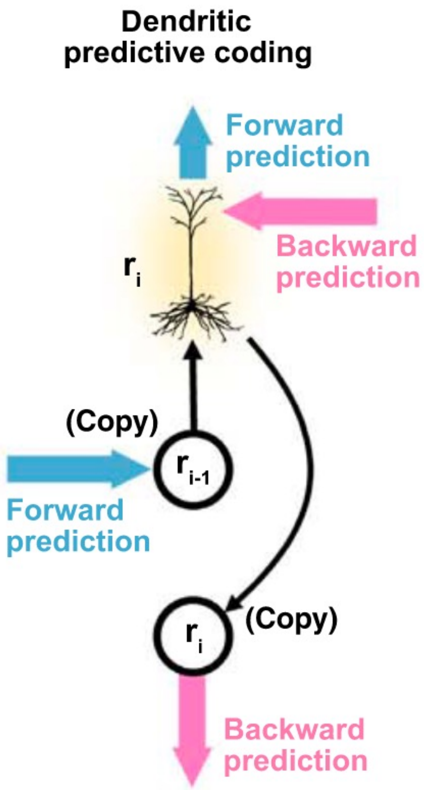
How can we test predictive coding theories of the brain?



Mikulasch et al. 2023 *TiNS*

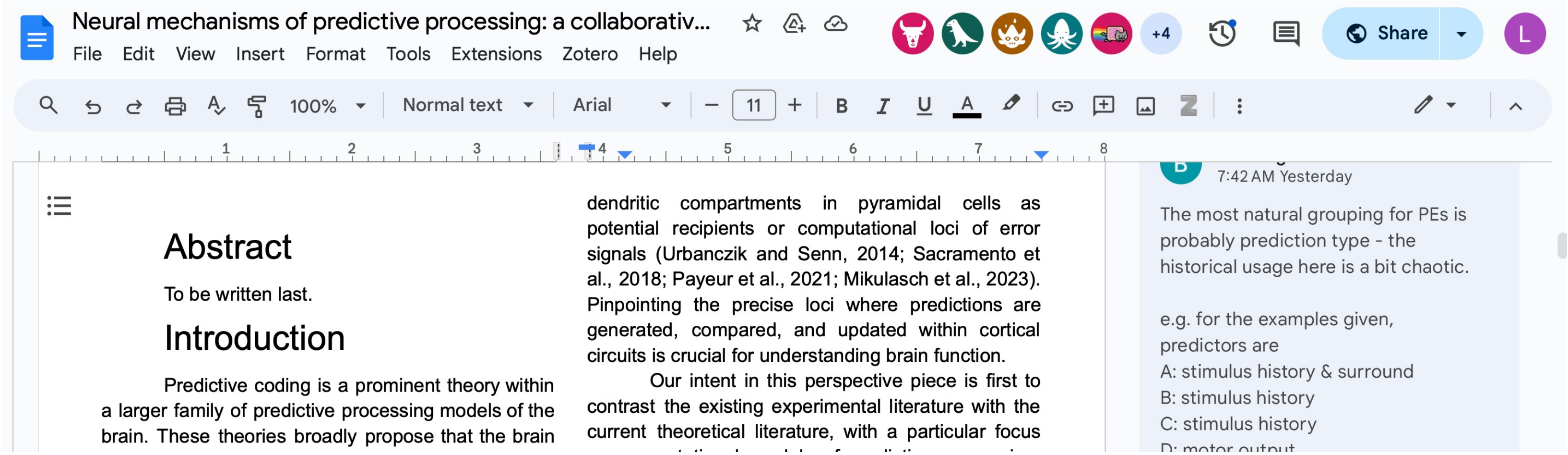
A single pyramidal neuron can implement predictive coding via its apical dendrites

(Urbanczik & Senn 2014 *Neuron*; Sacramento et al. 2018 *NIPS*; Guerguiev et al. 2017 *eLife*)



Mikulasch et al. 2023 *TiNS*

How can we test predictive coding theories of the brain?



Why is PC useful for CP?

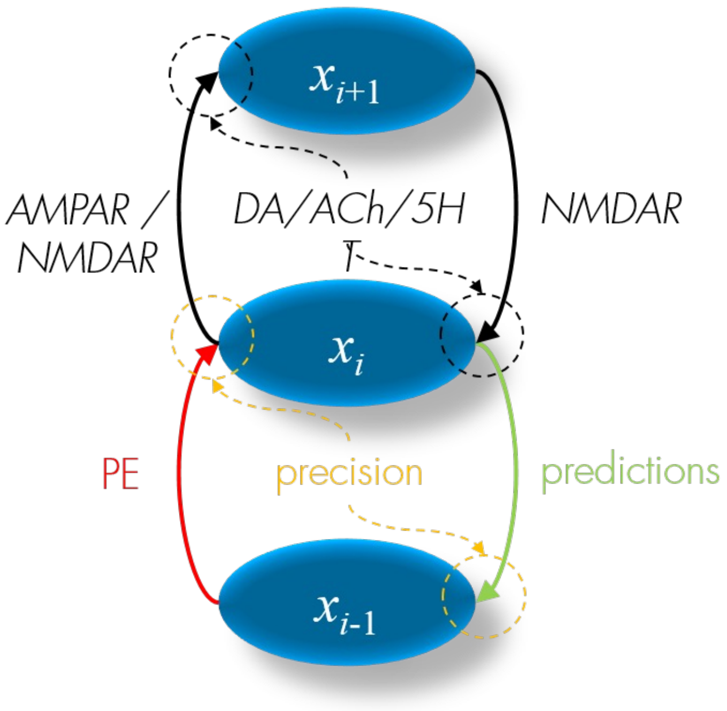
Predictive coding in computational psychiatry

Possible primary disruption at:

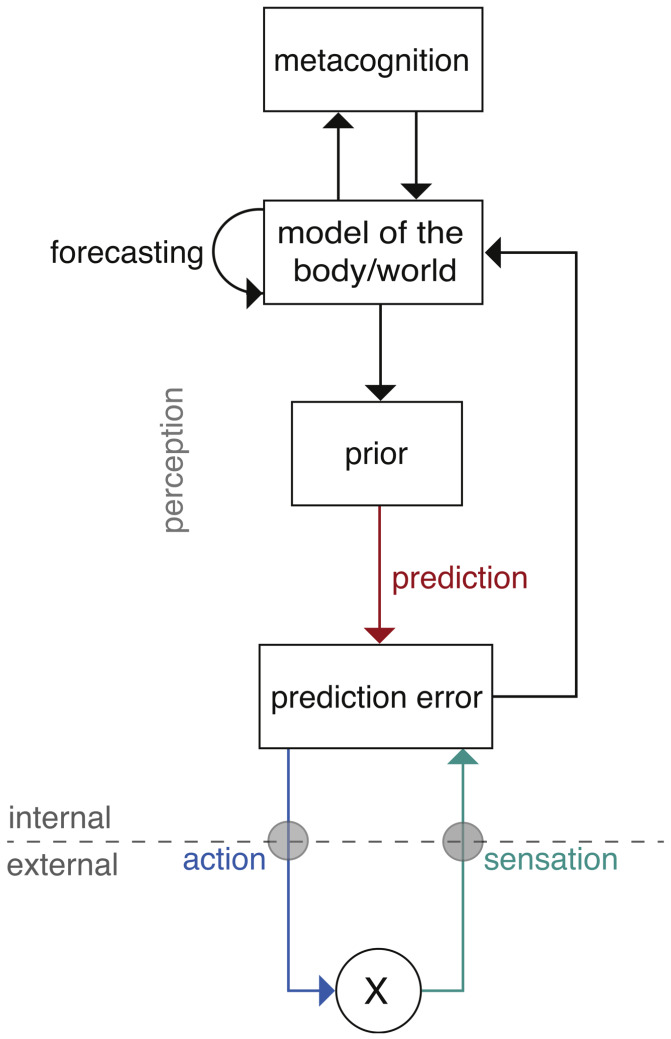
1. Sensory inputs (sensations)
2. Inference (perception)
3. Forecasting
4. Control (action)
5. Metacognition

At any of these, possible disturbance of:

- Prediction error (PE) computation
- Predictions/Expectations
- Estimation of their precision.

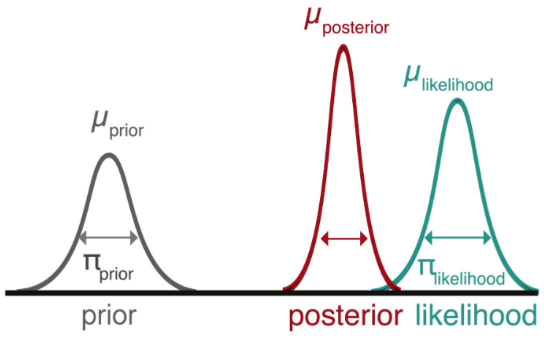


Stephan et al. 2016 *Brain*



Petzschner et al. 2017 *Biol Psychiatry*

PC in CP I: precision-weighting



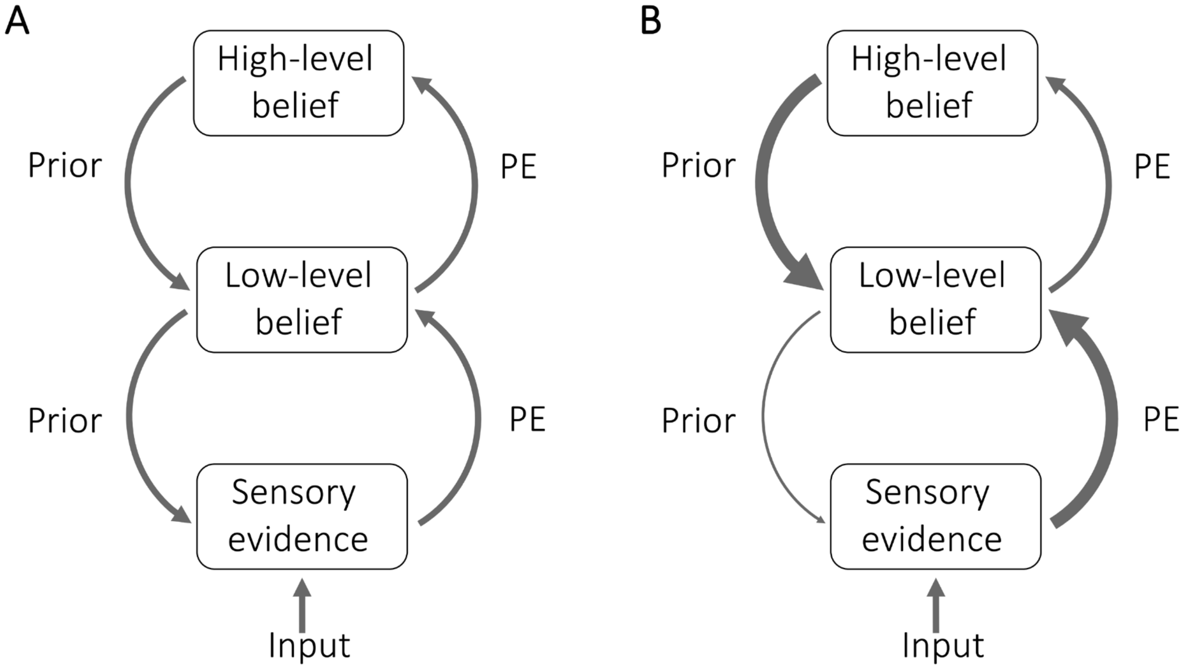
$$\text{change in belief} \propto \frac{\text{precision of data}}{\text{precision of current belief}} \bullet \text{prediction error}$$

Schizophrenia/Psychosis

Adams et al. 2013 *Front Psychiatry*; Corlett et al. 2011 *NPP*; Stephan et al. 2006 *Biol Psychiatry*; 2009; Schmack et al. 2013 *JNeuro*; Powers et al. 2017 *Science*; Sterzer et al. 2018 *Biol Psychiatry*; Corlett et al. 2019 *TiCS*; Petrovich & Sterzer 2023 *Schiz Bull*

Autism spectrum disorder

Pellicano & Burr 2012 *TiCS*; Van de Cruys et al. 2014 *Psychol Rev*; Lawson et al. 2014 *Front Hum Neurosci*; Lawson et al. 2017 *Nat Neurosci*



Petrovich & Sterzer 2023 *Schiz Bull*

PC in CP II: interoceptive predictive coding

Seth 2013 *TiCS*

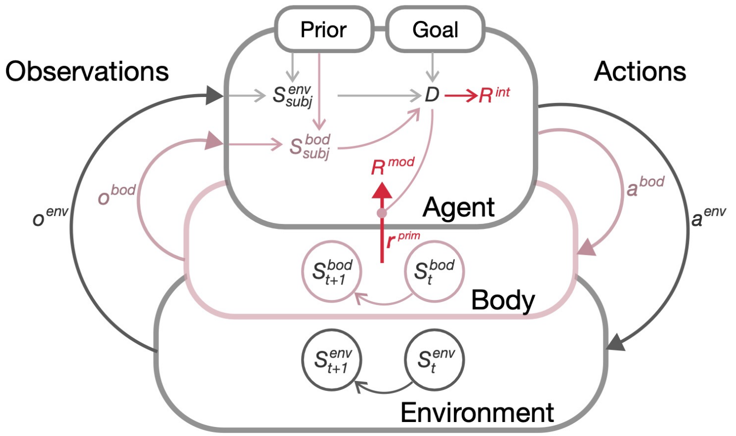
Barrett & Simmons 2015 *Nature Rev Neurosci*

Seth & Friston 2016 *Phil Trans Royal Soc B*

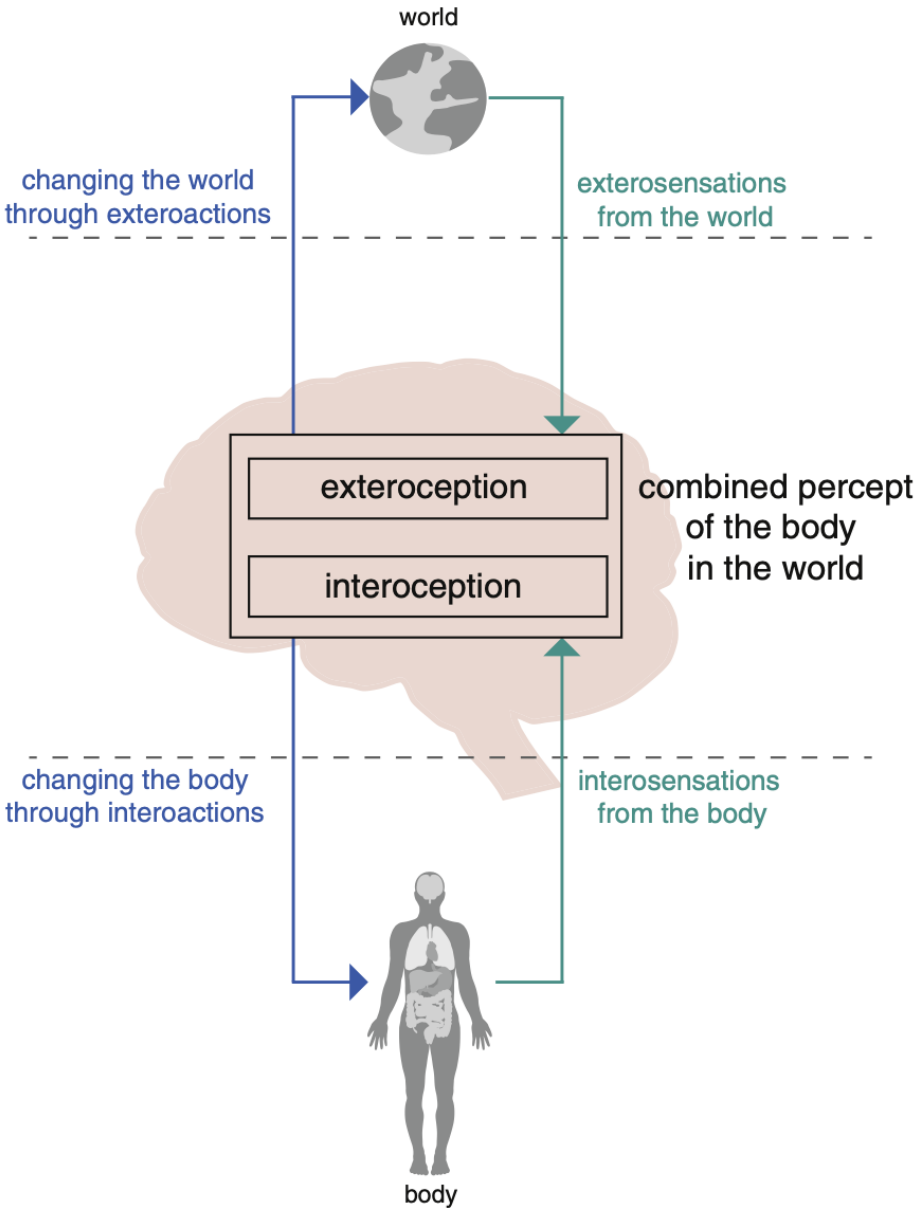
Stephan et al. 2016 *Front Hum Neurosci*

Petzschner et al. 2017 *Biol Psychiatry*

…as a basis for reward?

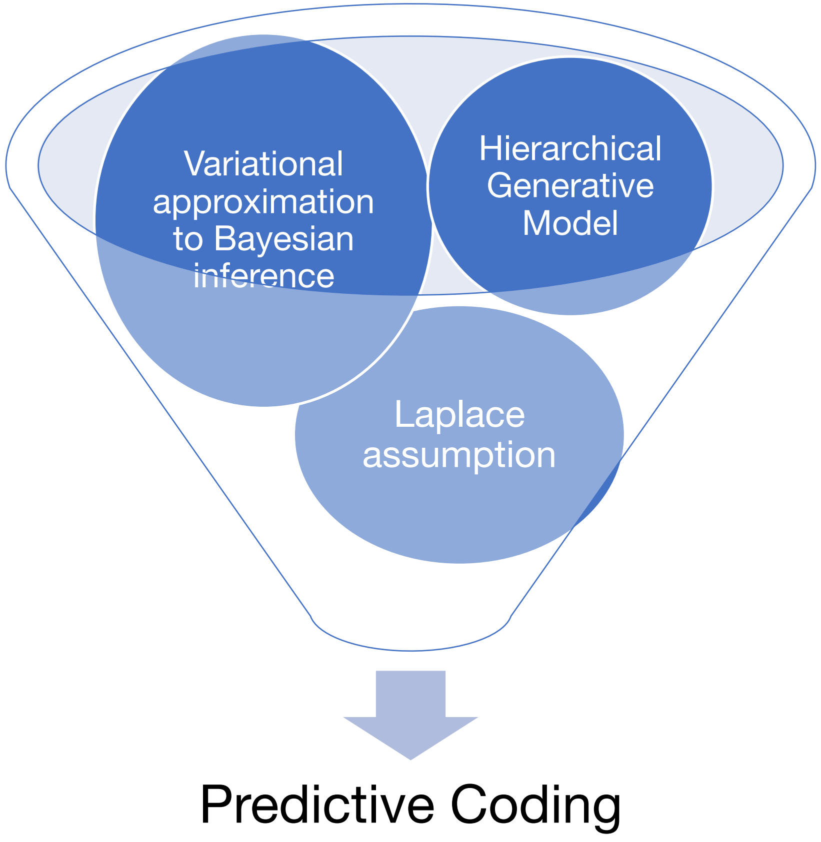


Weber, Yee, Small, Petzschner 2024 *PsyArXiv*



Summary: How to build a PC model

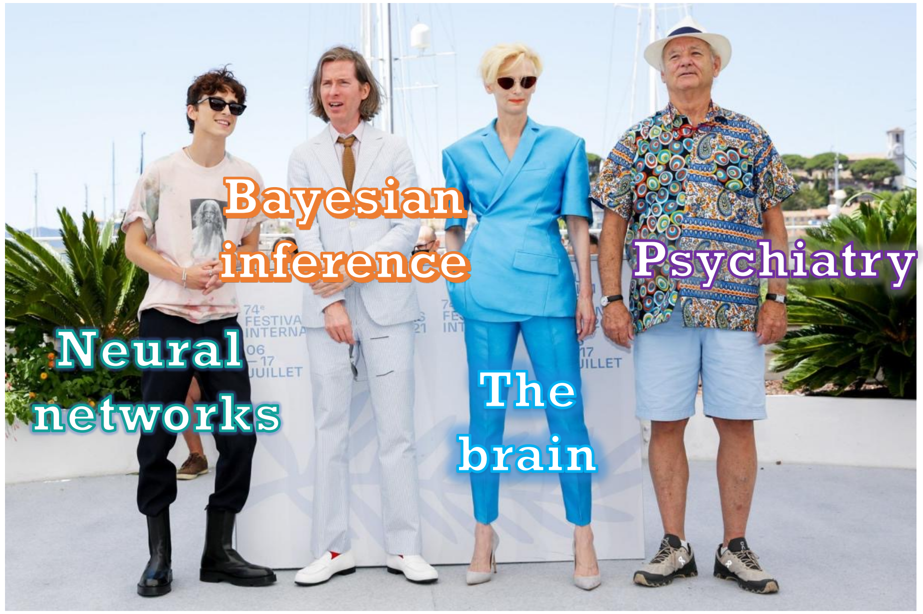
1. Predict the input
   - generative model of the sensory inputs
2. Invert the model
   - gradient descent on the neg. log joint
   - or: be fully Bayesian and use F
3. Simulate!
4. Fitting: Be the scientist
   1. Specify a mapping to observable data
   2. Write down the likelihood, specify your priors
   3. Acquire data & invert your model



Instead of 2.:

Use PyHGF, a neural network library for predictive coding https://ilabcode.github.io/pyhgf/

Thank you



Alternatives

divisive normalization (another canonical algorithmic motif) (Carandini & Heeger 2012 Nat Rev Neurosci)

cells compute a ratio between their direct (bottom-up) inputs and the summed activity of a pool of neurons

different from divisive predictive coding in that all neurons use a single global divisor, rather than each neuron's activity being divided by its own specific prediction

can describe a range of effects, including saturation, cross-orientation suppression, and surround suppression [79], and it is modulated by attention [80], locomotion [81], and even disease [82]

divisive normalization implements inference in a powerful statistical model of natural images [83,84,85], which, critically, assumes a direct variable, rather than a predictive code

inference in such a model not only accounts for the extra-classical receptive field effects commonly characterised by simple laboratory stimuli [84,58], but also the degree of surround suppression observed in response to natural images [85]

probabilistic population codes

neural firing rate of one neuron = posterior probability of one possible value (or a range) of the latent variables

involves multiplication (likelihood * prior), but usually: in log space (--> summation)

other parametric representations (= neural responses representing the parameters of the posterior probability distribution)

(log-)probability codes including probabilistic population codes

Raju & Pitkow 2016 NIPS

direct variable coding
neural activity directly represents latent variables
in sparse coding models of visual images (or image patches), the latent variables typically correspond to the intensity with which a visual feature (such as an oriented Gabor filter) is present in the image [56]
in a direct variable coding representation, neural responses directly encode these intensities: no response implies the feature represented by the neuron is absent, a small or a large response means the feature's intensity is low or high, respectively
predictive coding schemes also use a one-to-one correspondence between latent variables and neurons, but they define neural responses as representing differences between inferred and predicted variable values
Neural responses in direct variable encoding schemes either deterministically converge to the single best setting of the latent variables
or stochastically sample multiple different plausible settings for the latents
[18,57,58 ,59]
merely computing prediction errors does not imply that there must be cells whose responses directly represent these prediction errors: in fact, self-consistent neural circuit dynamics can be constructed using pure direct variable coding
the resulting population activities exhibit an integration of top-down (conveying priors) and bottom-up inputs (conveying stimulus-related information) that often takes the form of a simple weighted average of the a priori expected value and that suggested by sensory evidence
a stochastic extension of these theories, in which the activity of neurons represents latent variable values that are sampled from the posterior distribution [18], accounts for task-dependent [59] and stimulus-dependent variability [58] and for the similarity of evoked and spontaneous activities in V1 [57].
Orbán, Berkes, Fiser, Lengyel 2016 Neuron
Aitchison & Lengyel 2016 PLoS

Resources

- *all things PC (very comprehensive):*
  - Millidge, Seth, Buckley 2021 *arXiv:2107.12979*
- *evidence for PC in the brain:*
  - Walsh et al. 2020 *Ann N Y Acad Sci*
  - Keller & Mrsic-Flogel 2018 *Neuron*
  - Heilbron & Chait 2018 *Neurosci*
- *tutorials on the free energy formulation of PC*
  - Bogacz 2017 *J Math Psychol*
  - Buckley et al. 2017 *J Math Psychol*
- *further thoughts on precision*
  - Yon & Frith 2021 *Curr Biol*
- *Recent PC perspectives on delusions and hallucinations*
  - Corlett & Fletcher 2021 *Cogn Neuropsych* (delusions)
  - Corlett et al. 2019 *TiCS* (hallucinations)