Predictive Coding

Computational Psychiatry Course
09.09.2020

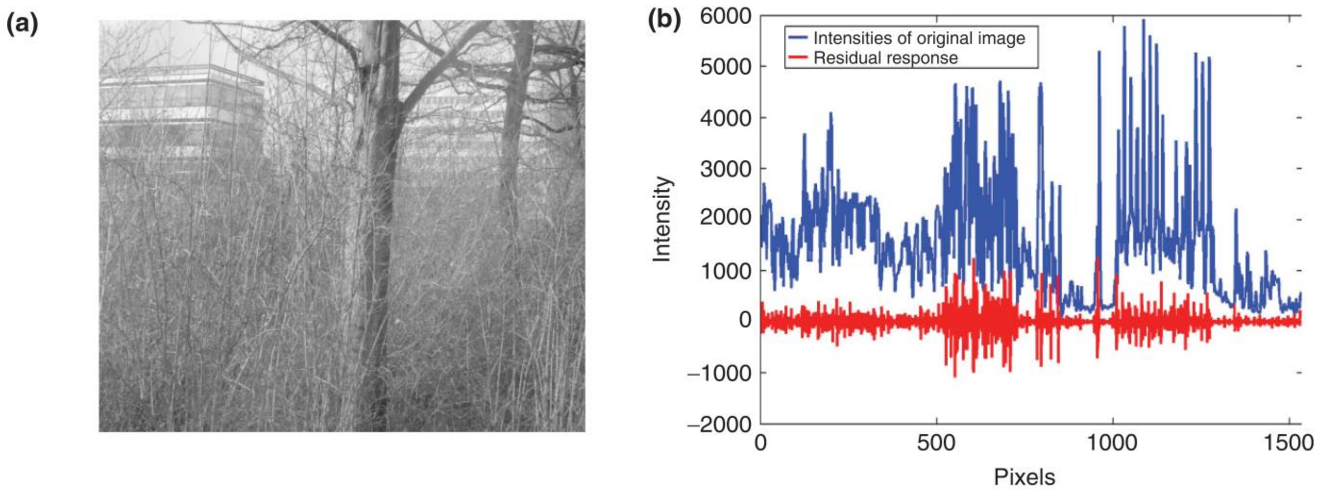
Lilian A. Weber

COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

Predictive coding

prediction error = prediction – input



Huang & Rao 2011 *WIREs Cogn Sci*

Laplace approximation

free energy

Bayesian brain

model inversion

predictive processing

predictive coding

active inference

redundancy reduction

neural networks

sampling

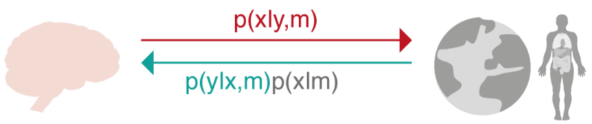
variational Bayes

Bayesian inference

generative model

Levels of analysis

Petzschner et al. 2017 *Biol Psychiatry*



computational

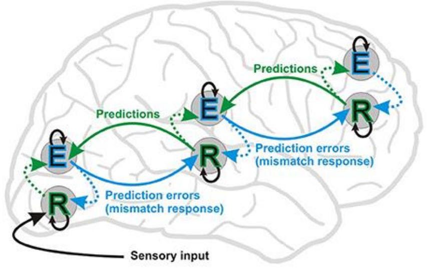
(approximate) Bayesian inference

redundancy reduction

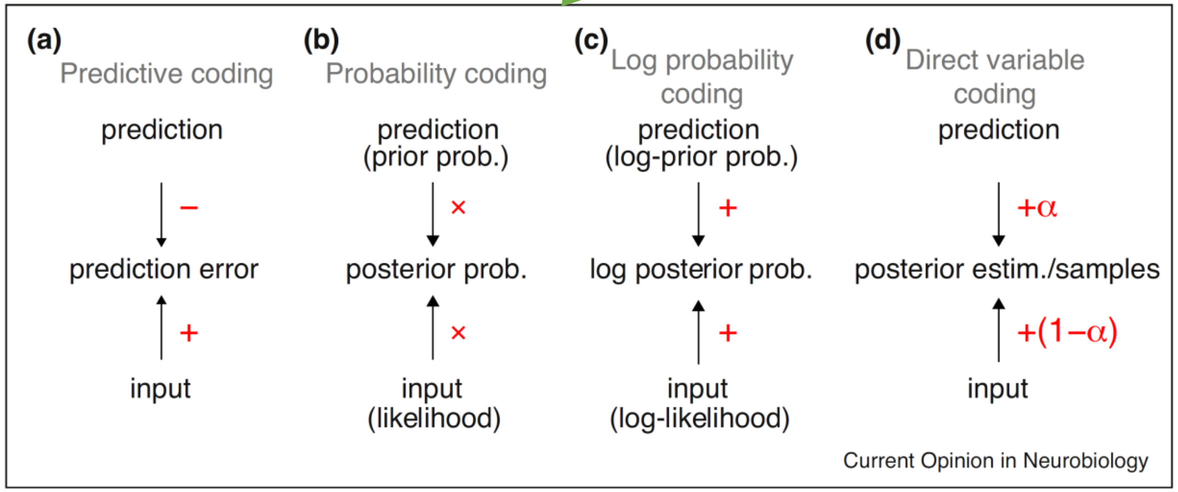
...

algorithmic

predictive coding



Stefanics et al. 2014 *Front Hum Neurosci*



Aitchison & Lengyel 2017 *Curr Op Neurobiol*

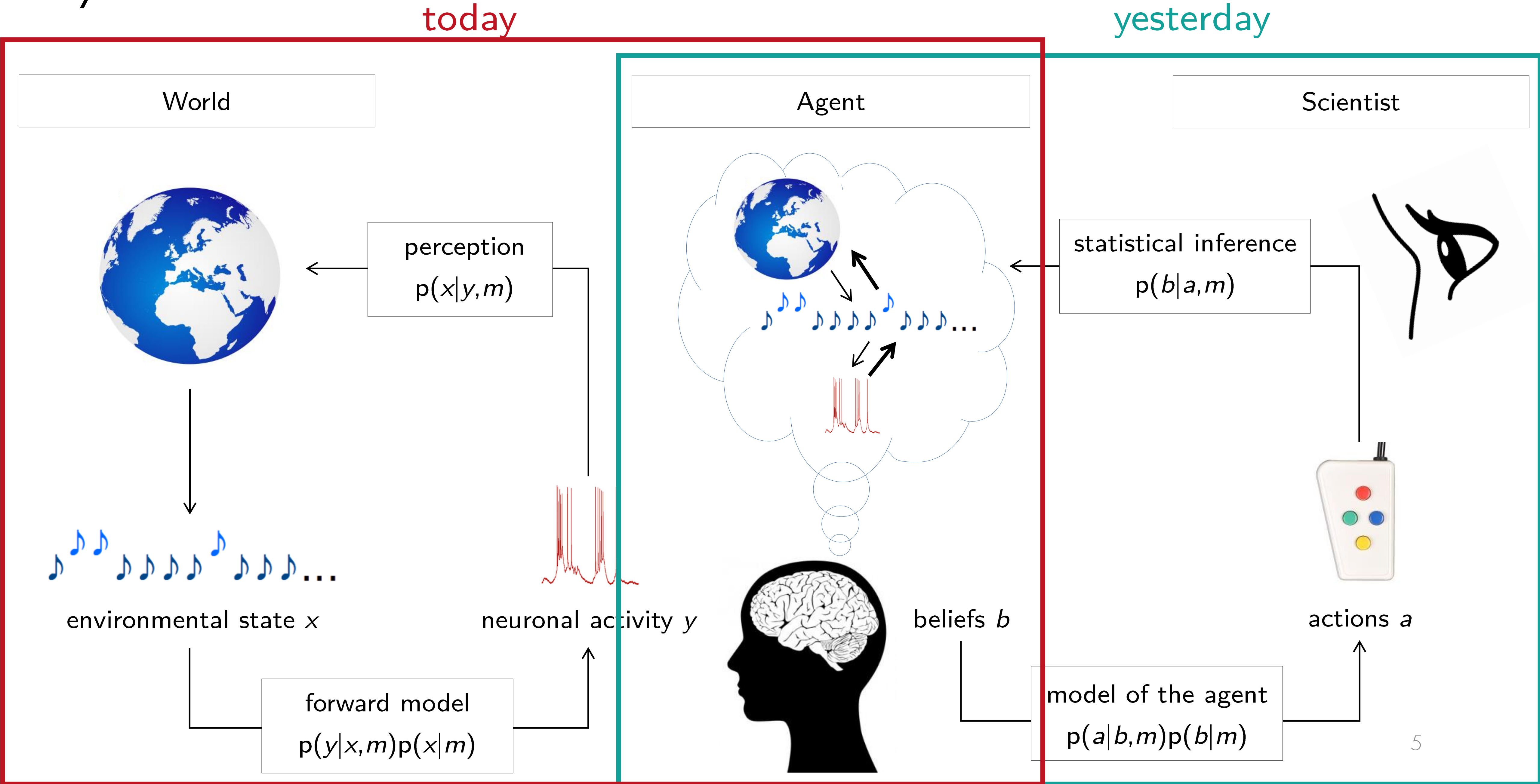
implementational

"Bayesian predictive coding"

COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

Bayesian observers



I. Bayesian predictive coding

From redundancy reduction in signal processing to approximate Bayesian inference in the visual cortex: *Rao & Ballard, 1999*

COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

1. Predicting the input

Firing rates/ Pixel values

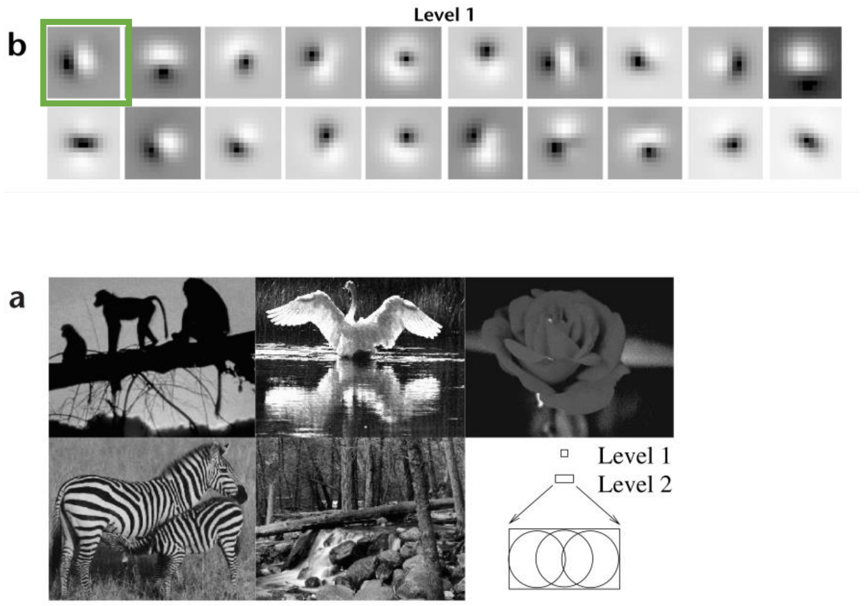
Model (possible patterns)

noise

Inferred causes (which patterns are present)

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

$$\begin{bmatrix} i_1 \\ i_2 \\ i_3 \\ \vdots \\ i_K \end{bmatrix} = f\left(\begin{bmatrix} u_{1,1} & u_{1,2} & \cdots & u_{1,N} \\ u_{2,1} & u_{2,2} & \cdots & u_{1,N} \\ u_{3,1} & u_{3,2} & \cdots & u_{3,N} \\ \vdots & \vdots & \ddots & \vdots \\ u_{K,1} & u_{K,2} & \cdots & u_{K,N} \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{bmatrix} \right) + \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ \vdots \\ n_K \end{bmatrix}$$



Rao & Ballard 1999 *Nat Neurosci*

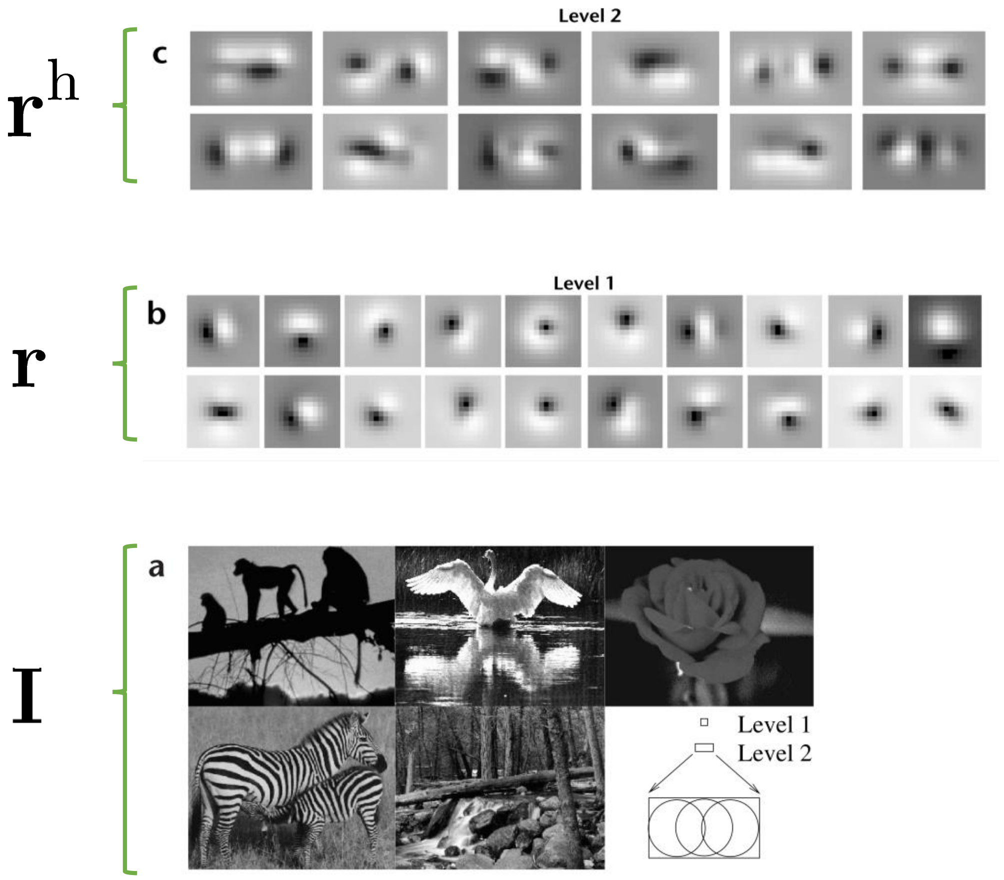
1. Predicting the input

Introducing hierarchy

$$\mathbf{r} = \mathbf{r}^{\mathrm{td}} + \mathbf{n}^{\mathrm{td}}$$

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

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



1. Predicting the input

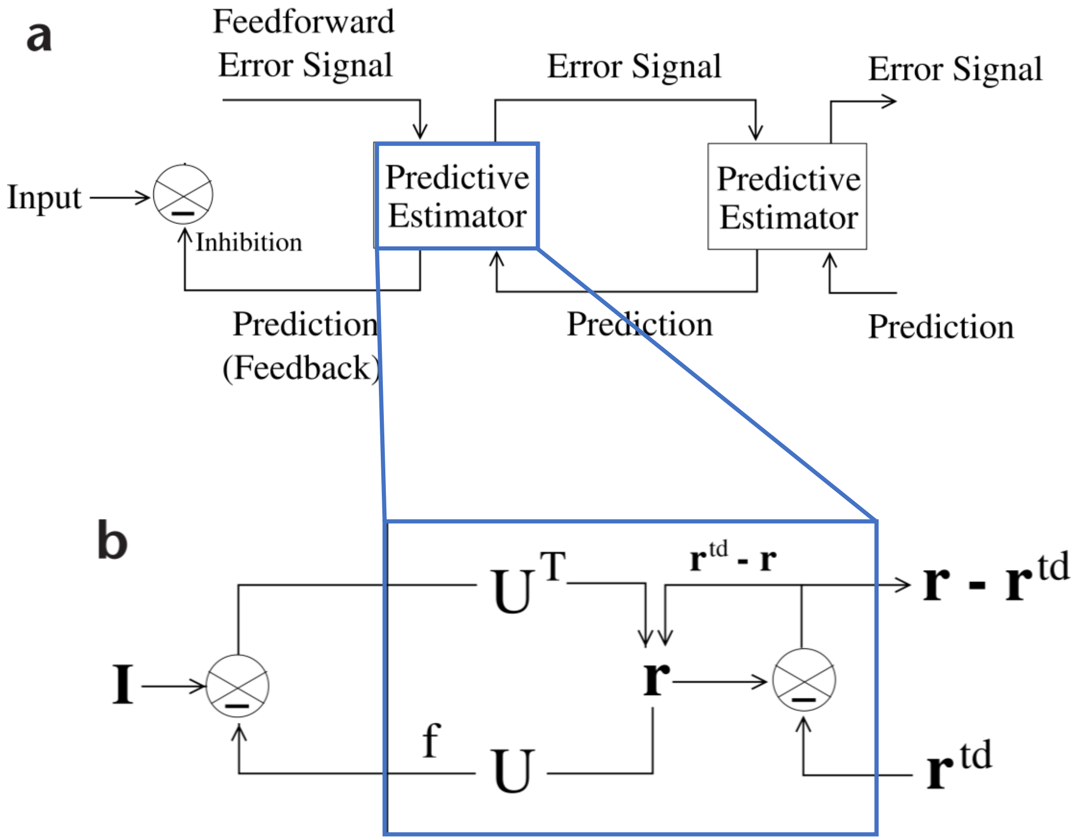
Implementation

$$\mathbf{r} = \mathbf{r}^{\mathrm{td}} + \mathbf{n}^{\mathrm{td}}$$

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

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

Question: How to find *r* and *U*?



Translational Neuromodeling Unit

Rao & Ballard 1999 *Nat Neurosci*

2. Inverting the model

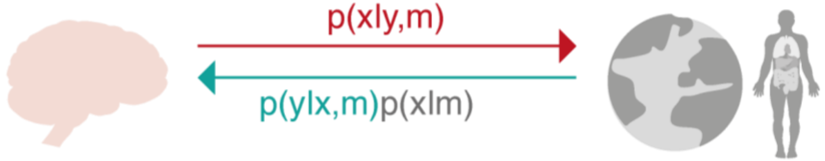
Optimization: Finding causes and associations

$$E = E_1 - \log p(\mathbf{r}) - \log p(U)$$

$$E_1 = \underbrace{\frac{1}{\sigma^2}(\mathbf{I} - f(U\mathbf{r}))^\intercal(\mathbf{I} - f(U\mathbf{r})) + \frac{1}{\sigma_{\mathrm{td}}^2}(\mathbf{r} - \mathbf{r}^{\mathrm{td}})^\intercal(\mathbf{r} - \mathbf{r}^{\mathrm{td}})}_{-\log p(\mathbf{I}|\mathbf{r},U) + C}$$

$$\begin{aligned} E &= -\log p(\mathbf{I}|\mathbf{r},U) - \log p(\mathbf{r}) - \log p(U) \\ &= -\log \boxed{\left(p(\mathbf{I}|\mathbf{r},U)\, p(\mathbf{r})\, p(U)\right)} \end{aligned}$$

Optimization looks like finding the most likely cause (MAP estimate) in Bayesian inference.



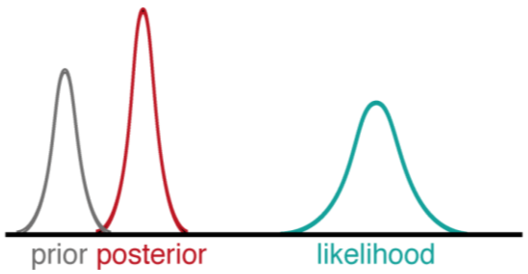
Bayes' Rule

posterior~ likelihood · prior

p(xly,m)~ p(ylx,m)p(xlm)

likelihood: p(ylx,m)
prior: p(xlm)
posterior: p(xly,m)

x = state of world & body
y = sensory data
m = model



COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

Petzschner et al. 2017 *Biol Psychiatry*

3. (Precision-weighted) Prediction errors

Inference: Estimating the causes based on inputs and expectations

Learning: Improving the model to reflect the statistics of the environment

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

Inputs

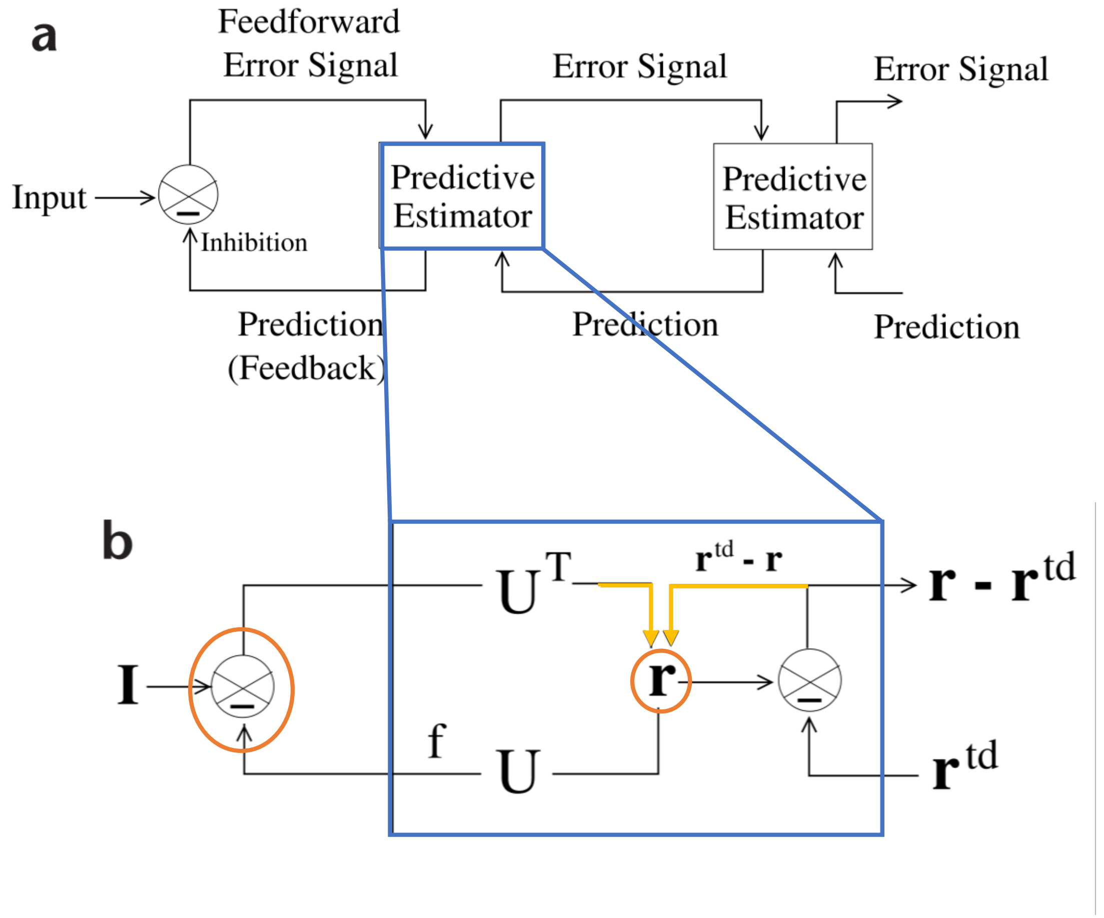
Expectations

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

precision-weighting

$$\frac{\mathrm{d}U}{\mathrm{d}t} = -\frac{k_2}{2}\frac{\partial E}{\partial U} = \frac{k_2}{\sigma^2} \frac{\partial f}{\partial U\mathbf{r}}^{\mathsf{T}} (\mathbf{I} - f(U\mathbf{r}))\mathbf{r} - \frac{k_2}{2}\lambda U$$

Hebbian learning



Rao & Ballard 1999 *Nat Neurosci*

COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

3. (Precision-weighted) Prediction errors

Inference: Estimating the causes based on inputs and expectations

Learning: Improving the model to reflect the statistics of the environment

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

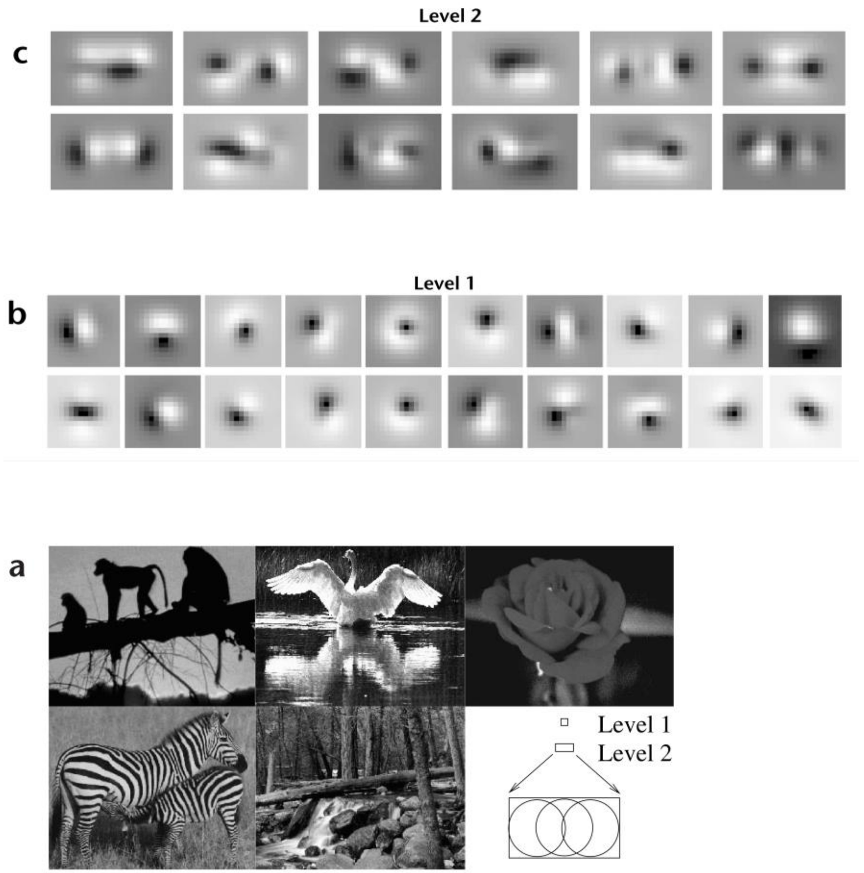
Inputs Expectations

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

precision-weighting

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

Hebbian learning



Rao & Ballard 1999 *Nat Neurosci*

Predictive coding as approximate inference

Inference: Estimating the causes based on inputs and expectations

Learning: Improving the model to reflect the statistics of the environment

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

Inputs

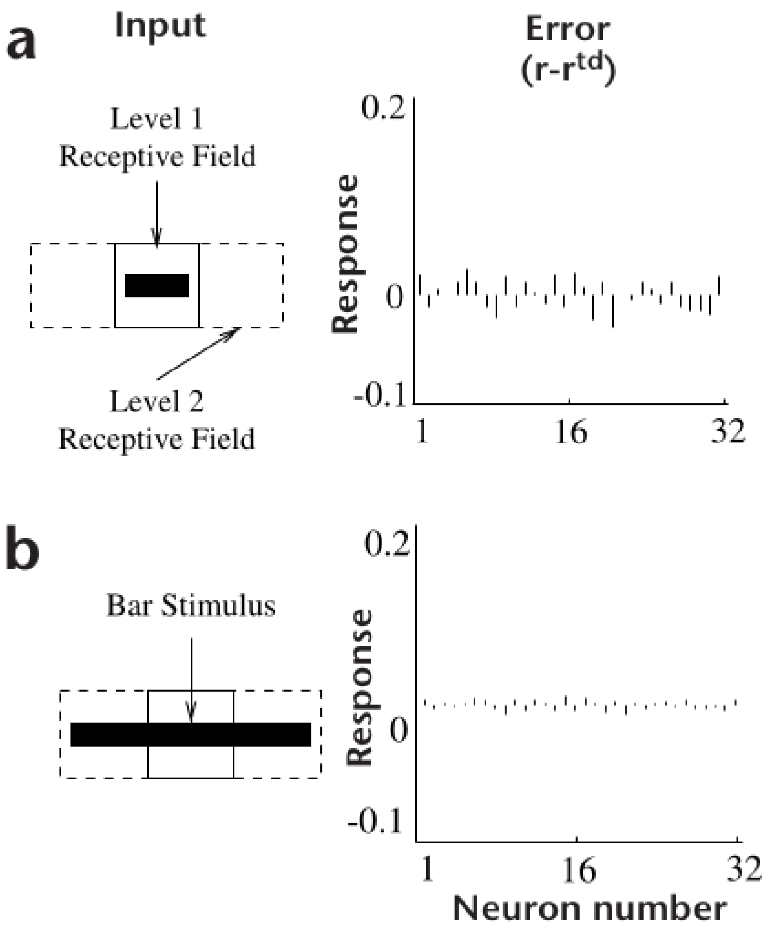
Expectations

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

precision-weighting

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

Hebbian learning



Rao & Ballard 1999 *Nat Neurosci*

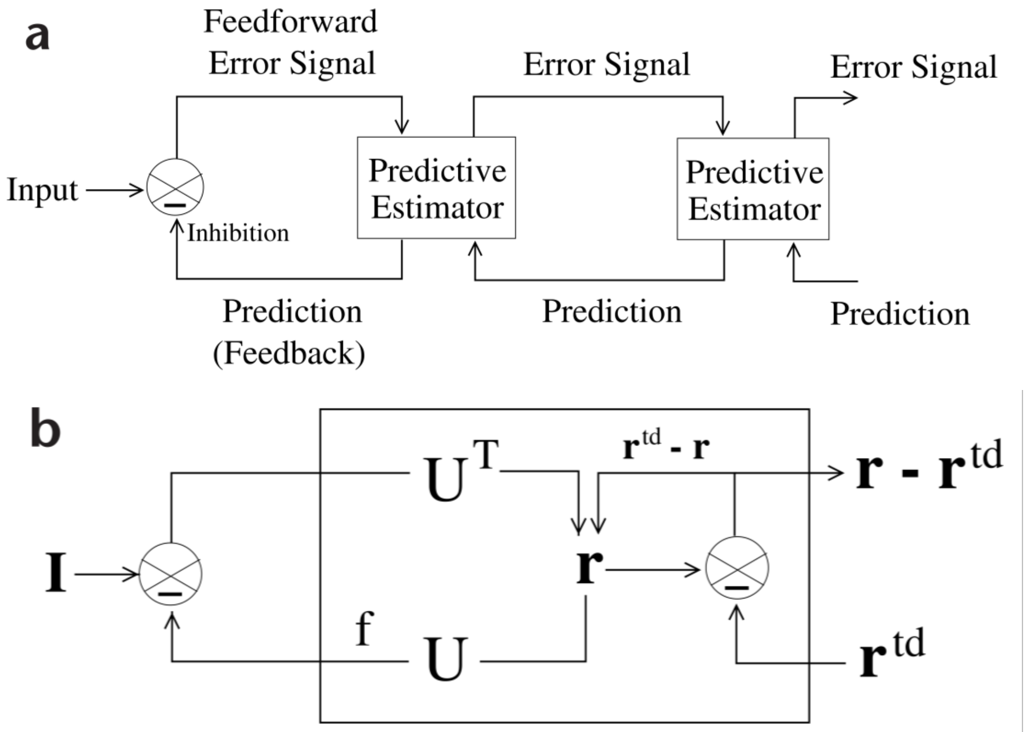
COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

Predictive coding as approximate 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



Rao & Ballard 1999 *Nat Neurosci*

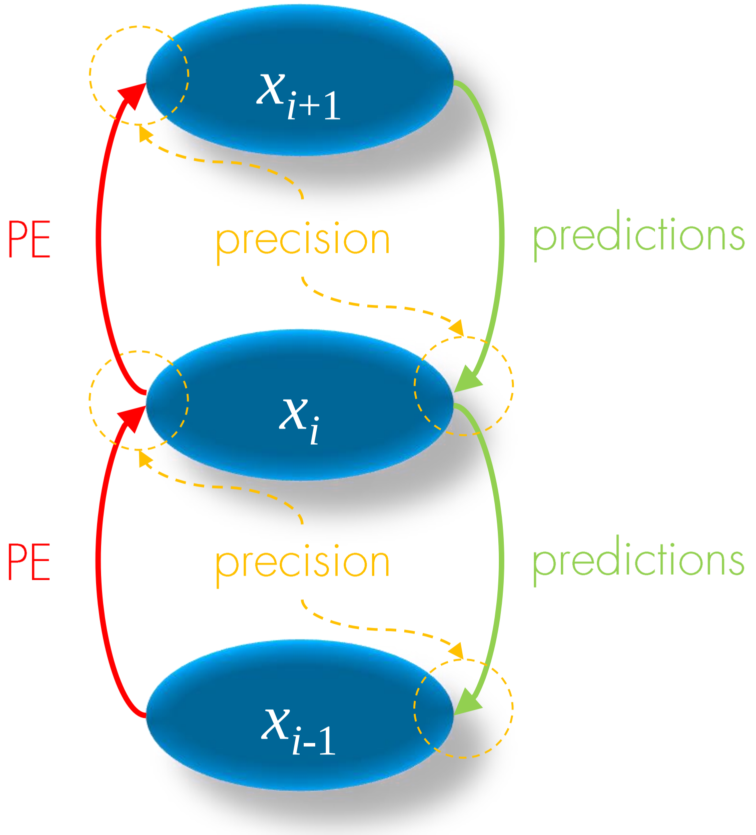
COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

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)

$$\Delta\text{belief} \sim \text{precision} \times \text{PE}$$



COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

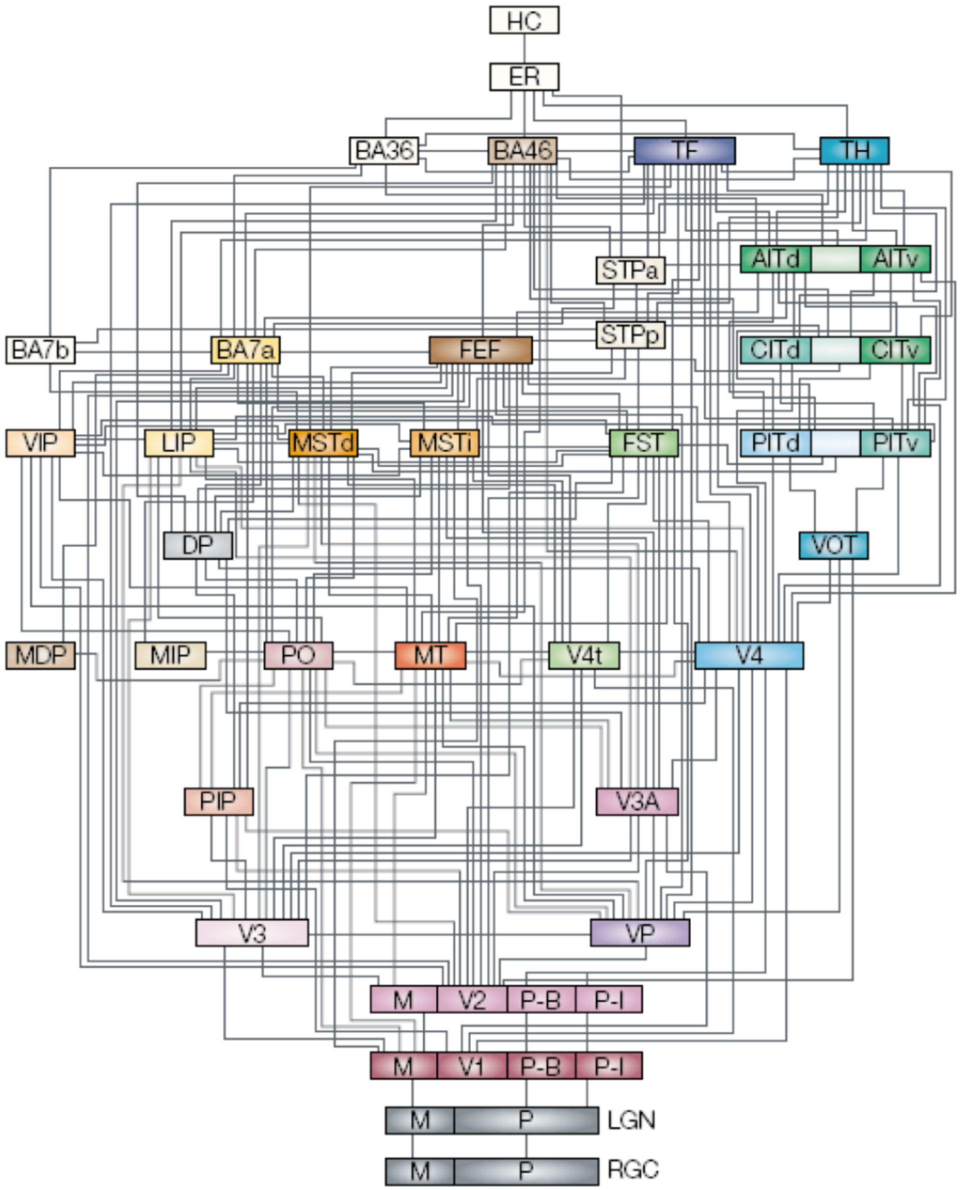
Adapted from Stephan et al. 2016 *Brain*

II. Implementing (and testing) PC

Mapping the ingredients of predictive coding onto neuroanatomy and neurophysiology

Mapping predictive coding onto the brain

- 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*

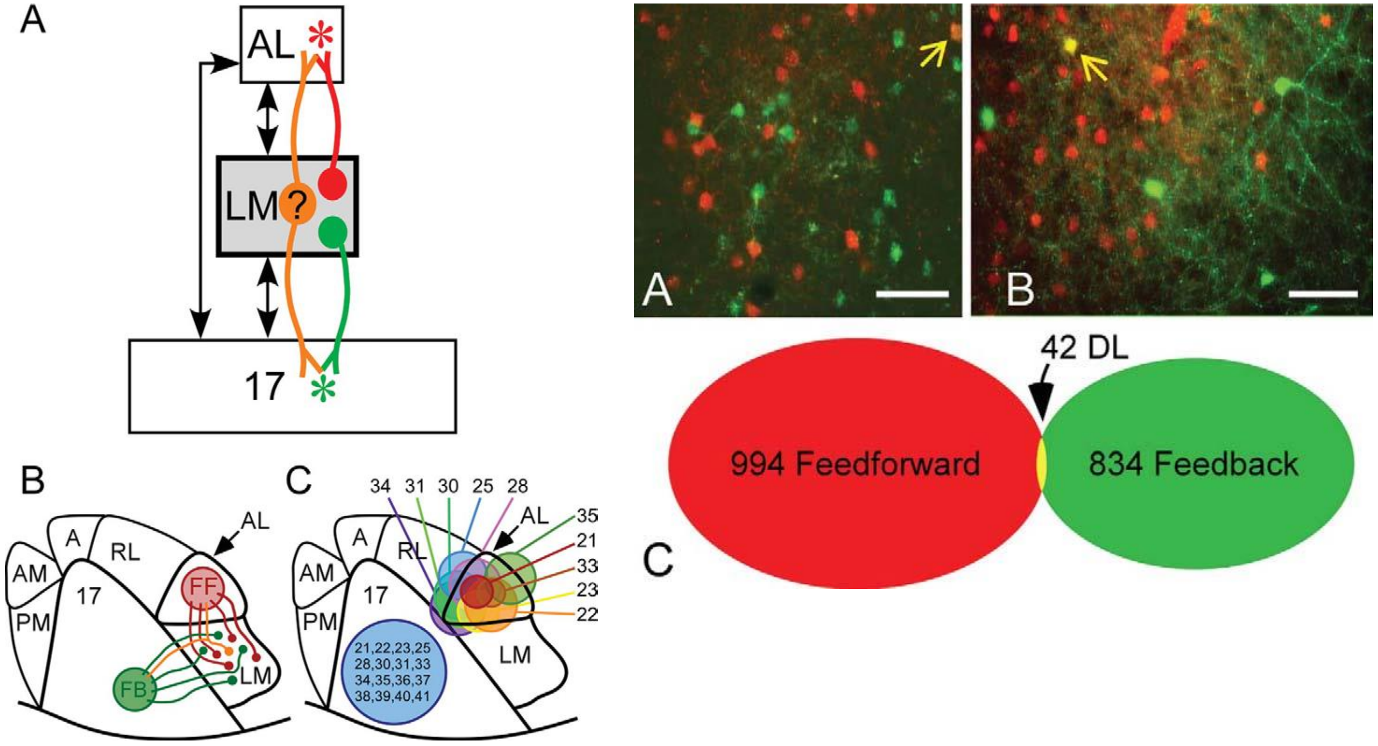
COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

Mapping predictive coding onto the brain

- 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)

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



Berezovskii et al. 2011 *J Comp Neurol*

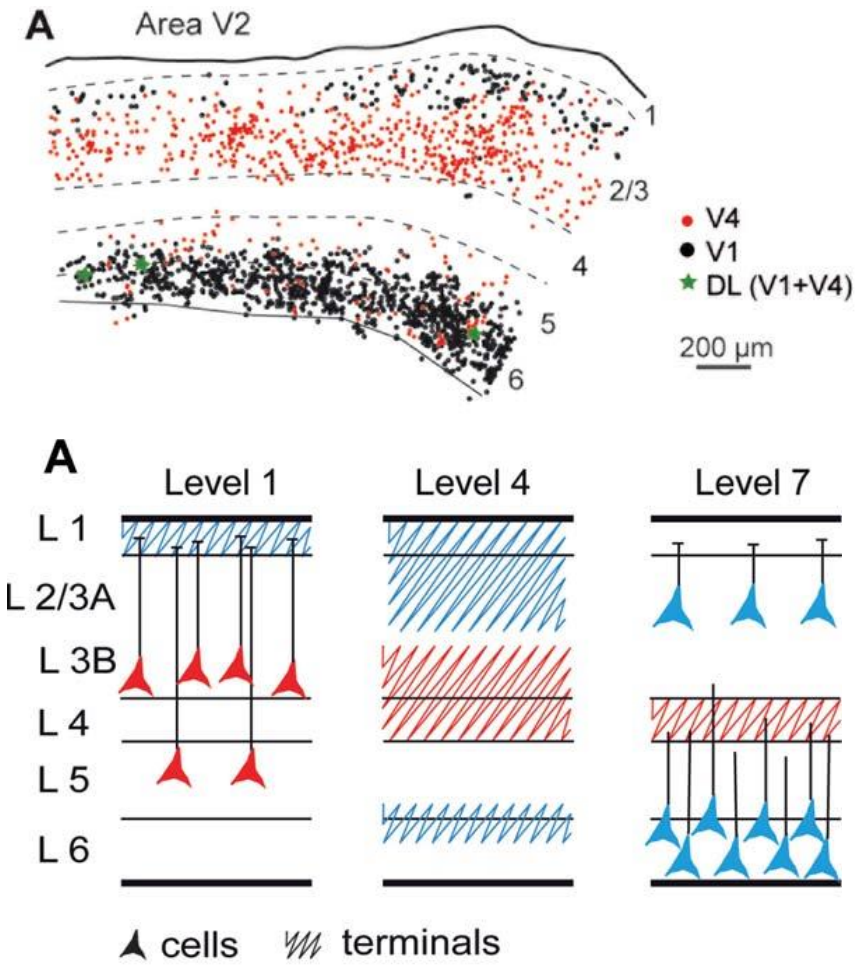
COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

Mapping predictive coding onto the brain

- 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)

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



Markov et al. 2014 *J Comp Neurol*

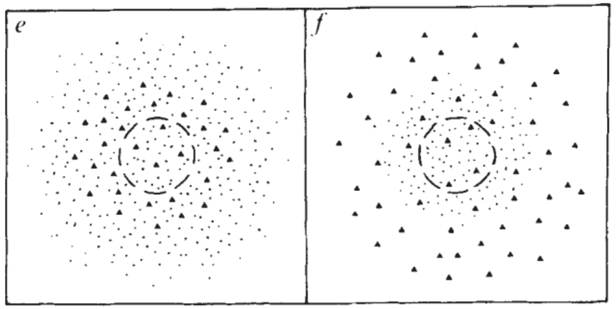
Translational Neuromodeling Unit

COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Mapping predictive coding onto the brain

- 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*

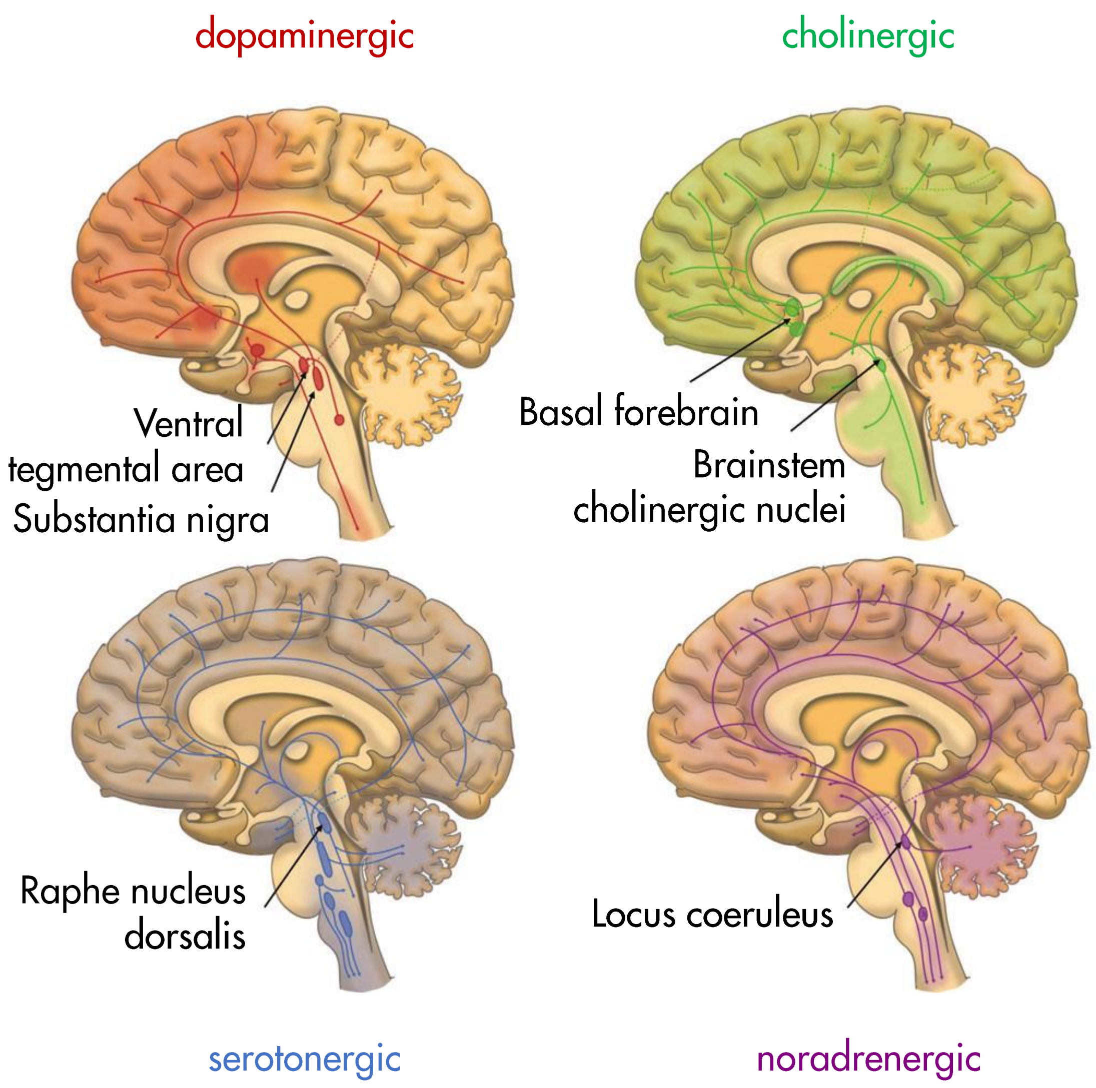
Buffalo et al. 2011 *PNAS*
Bosman et al. 2012 *Neuron*
Bastos et al. 2015 *NeuroImage*

COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

Mapping predictive coding onto the brain

- 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)



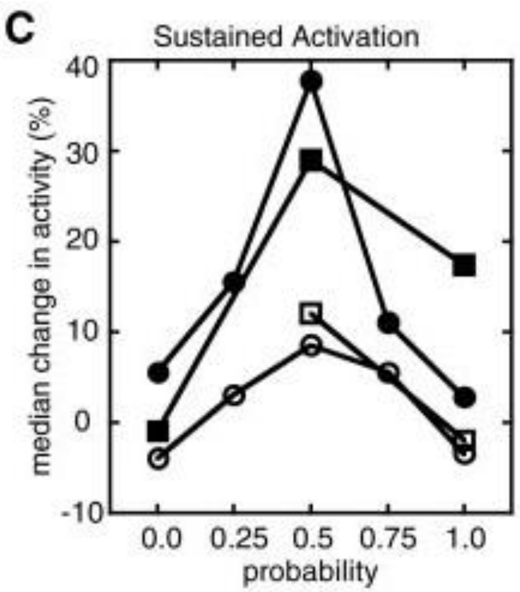
Adapted from Iglesias et al. 2017 *WIRE Cogn Sci*

COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

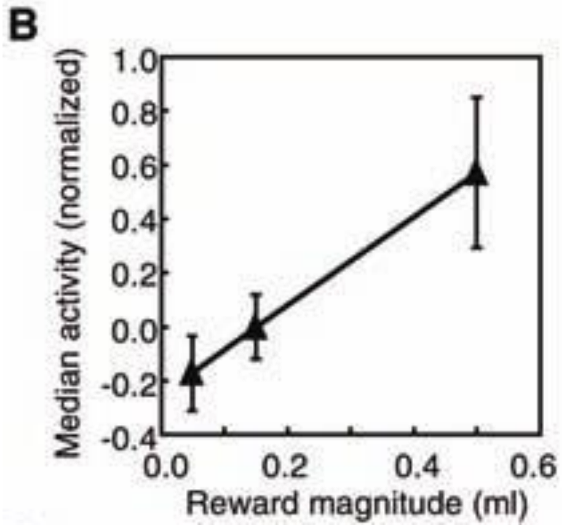
Translational Neuromodeling Unit

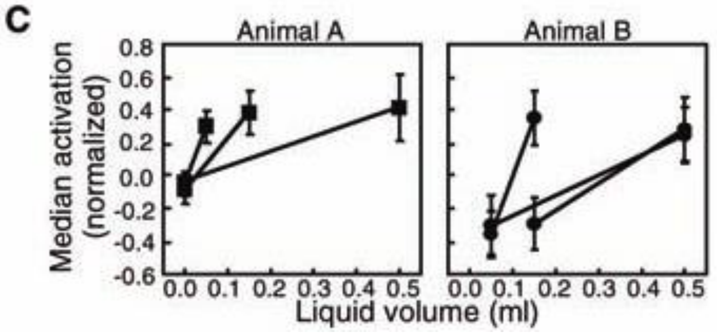
Mapping predictive coding onto the brain

- 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)



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.*

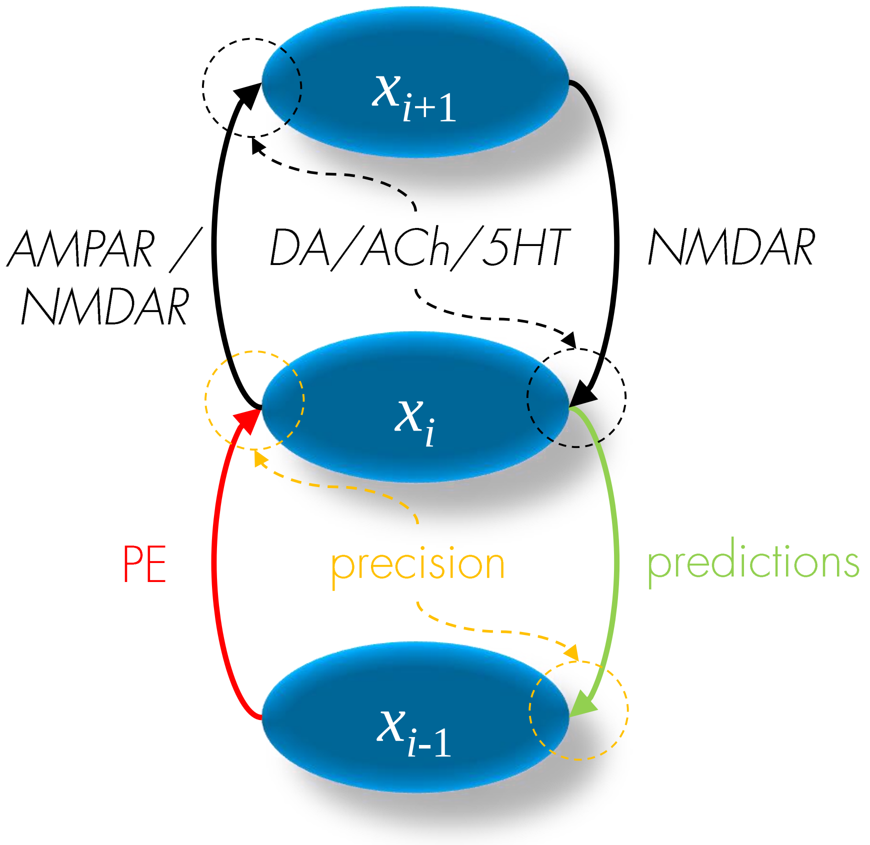
Achieved via cholinergic modulation?
Naudé et al. 2018 *bioRχiv*

COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

Mapping predictive coding onto the brain

- 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)



COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

Stephan et al. 2016 *Brain*

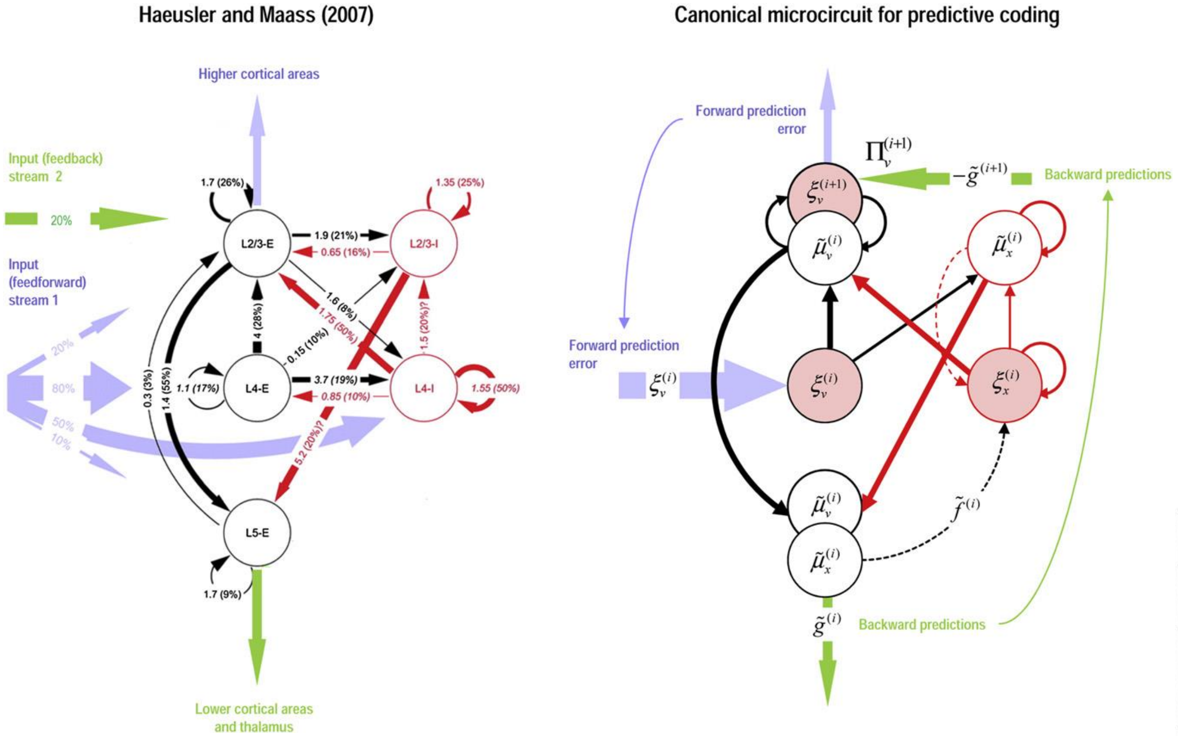
frontiers
in Computational Neuroscience

OPINION
published: 08 September 2015
doi: 10.3389/fncom.2015.00111

CrossMark

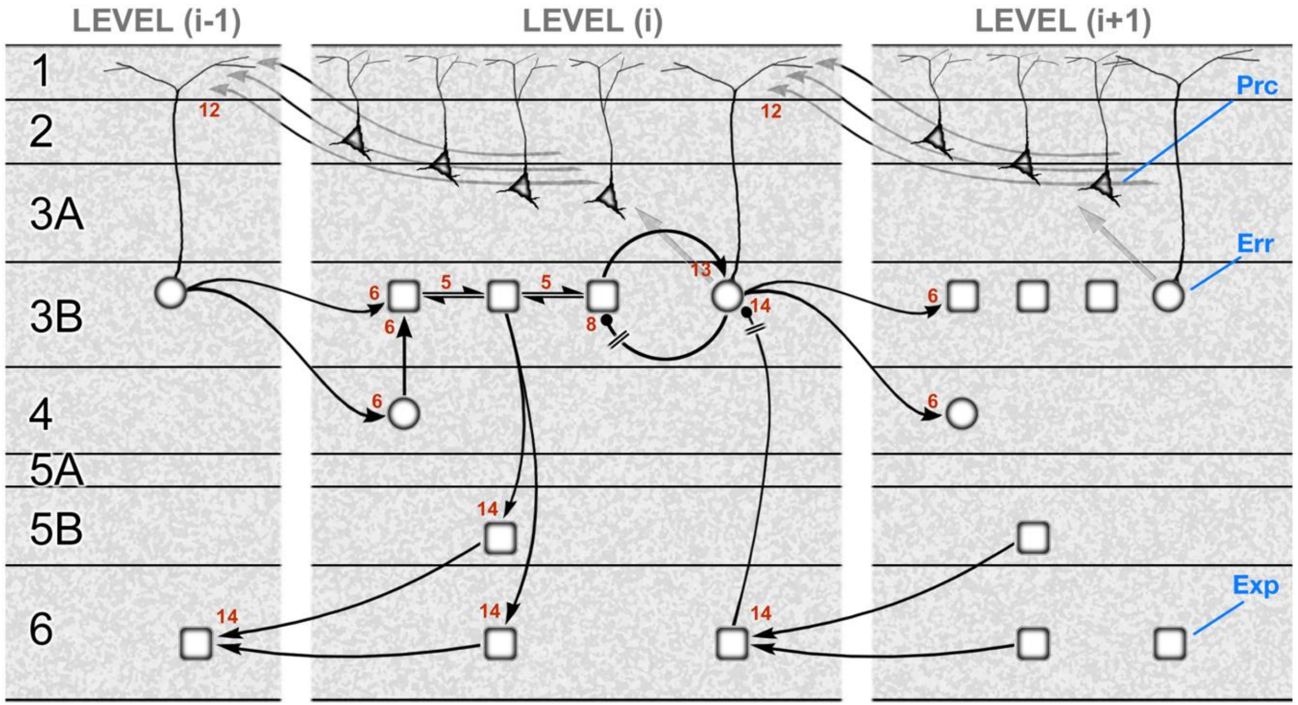
Is predictive coding theory articulated enough to be testable?

*Naoki Kogo * and Chris Trengove*



Bastos et al. 2012 *Neuron*

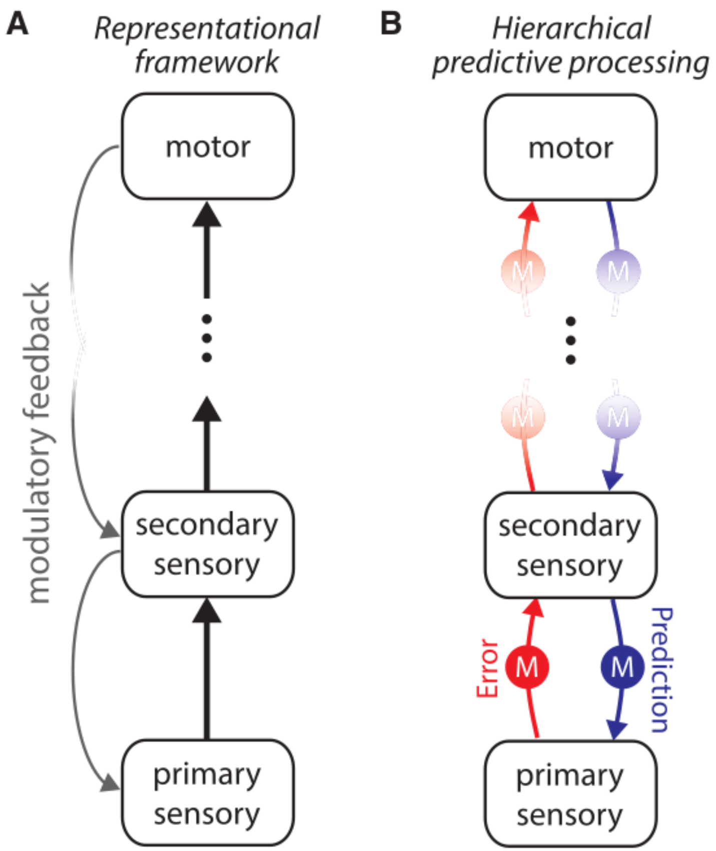
Mapping predictive coding onto the brain … *advanced*



Shipp 2016 *Frontiers in Psychology*

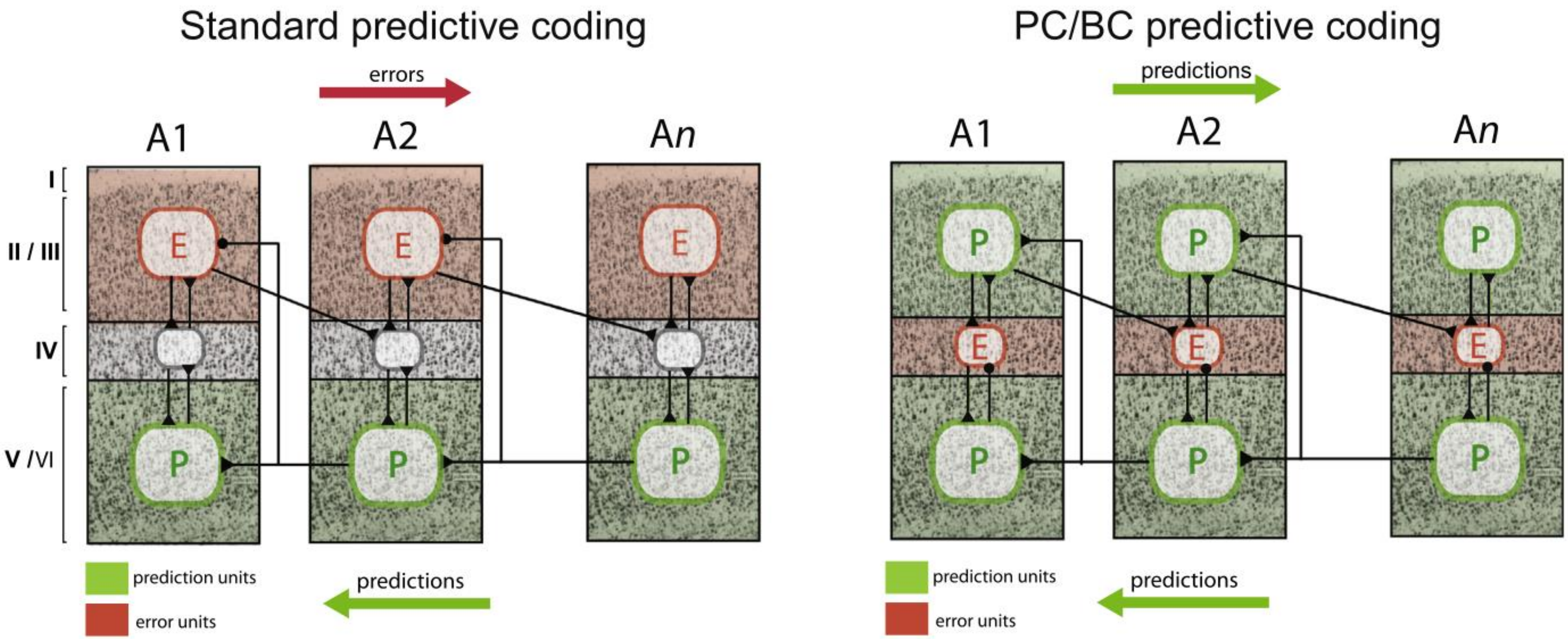
Testing the *relevant* predictions

- 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)



Keller & Mrsic-Flogel 2018 *Neuron*

Testing the *relevant* predictions



COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

Heilbron & Chait 2018 *Neurosci*

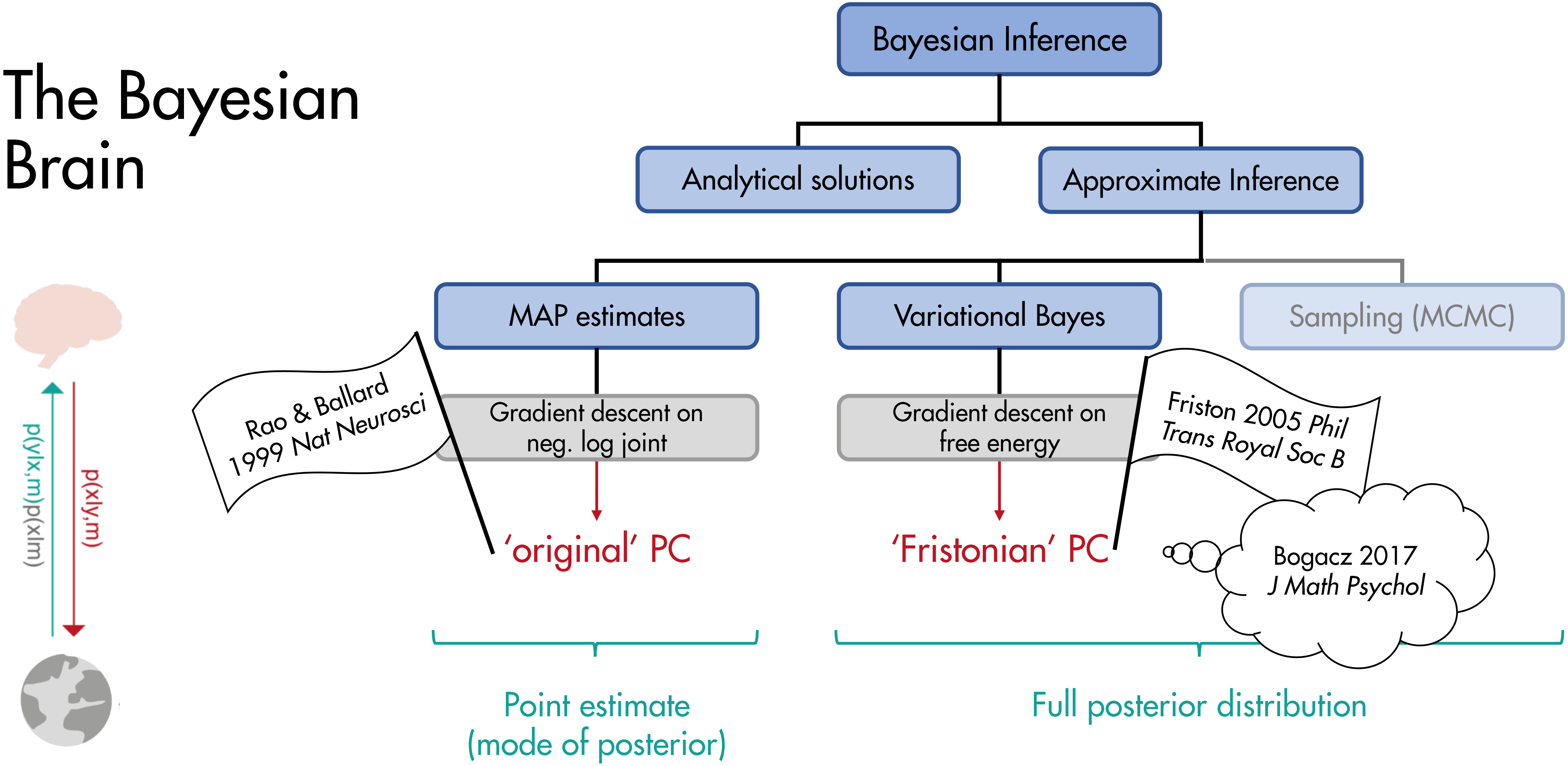
III. Predictive coding and free-energy

From representing probability distributions (and uncertainty) and the Laplace assumption to predictions on cortical infrastructure and plasticity: *Friston, 2005*

COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

The Bayesian Brain



Translational Neuromodeling Unit

Representing uncertainty

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

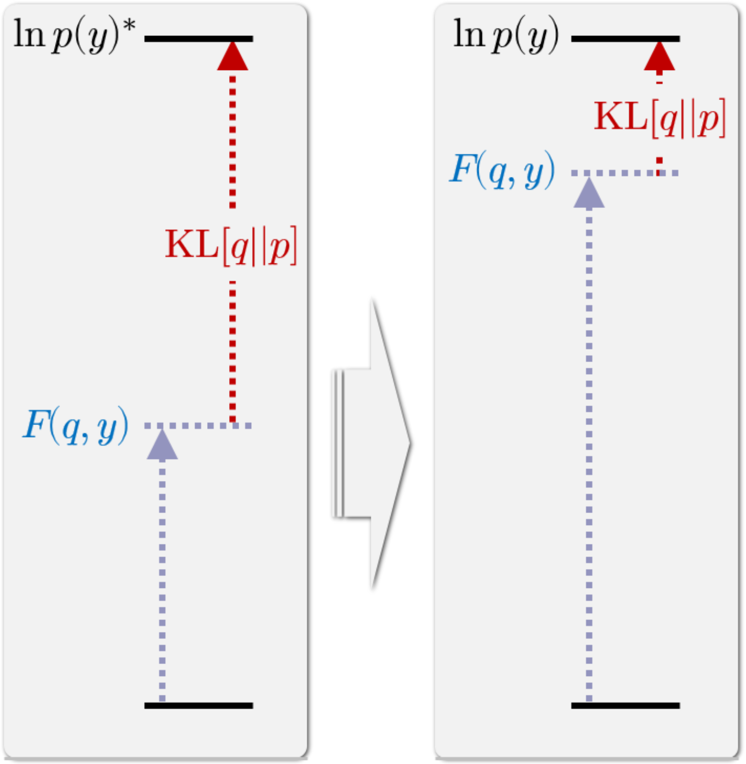
So far, we've only computed the MAP.
To be fully Bayesian: care about your uncertainty!

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

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

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

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



Translational Neuromodeling Unit

COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

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

‘Fristonian’ PC

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

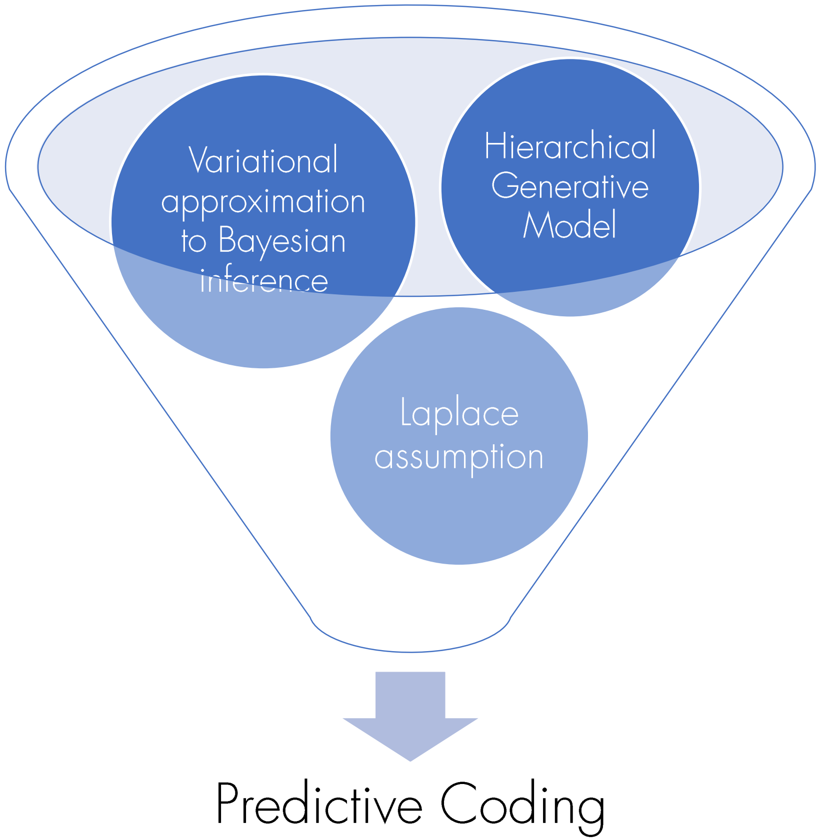
$$F = \int q(x) \ln \frac{p(y,x)}{q(x)} \, \mathrm{d}x$$

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*



Translational Neuromodeling Unit

COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY news & views

PREDICTIVE CODING

Does predictive coding have a future?

In the 20th century we thought the brain extracted knowledge from sensations. The 21st century witnessed a ‘strange inversion’, in which the brain became an organ of inference, actively constructing explanations for what’s going on ‘out there’, beyond its sensory epithelia. One paper played a key role in this paradigm shift.

Karl Friston

IV. PC in CP

Stratifying psychiatric diseases using hierarchical Bayesian models

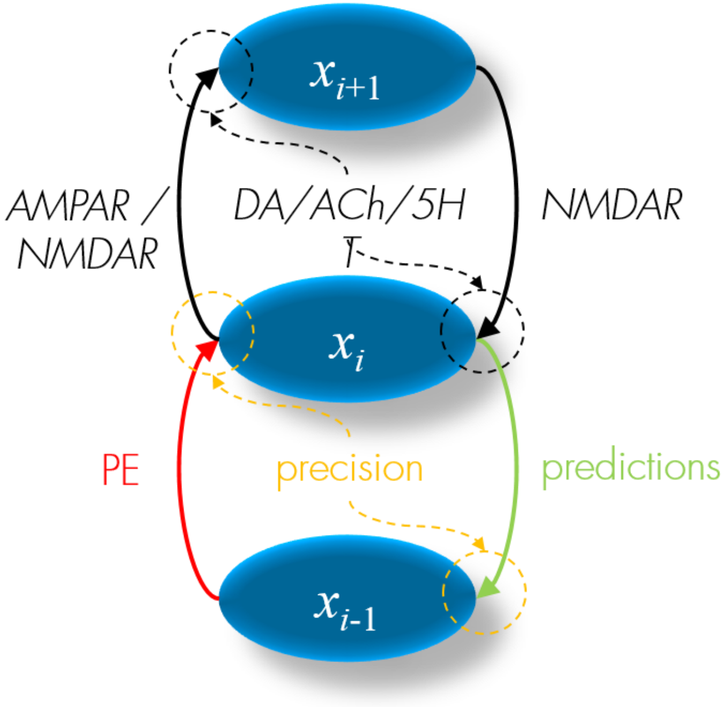
Hierarchical Bayesian Inference in Computational Psychiatry

Possible primary disruption at:

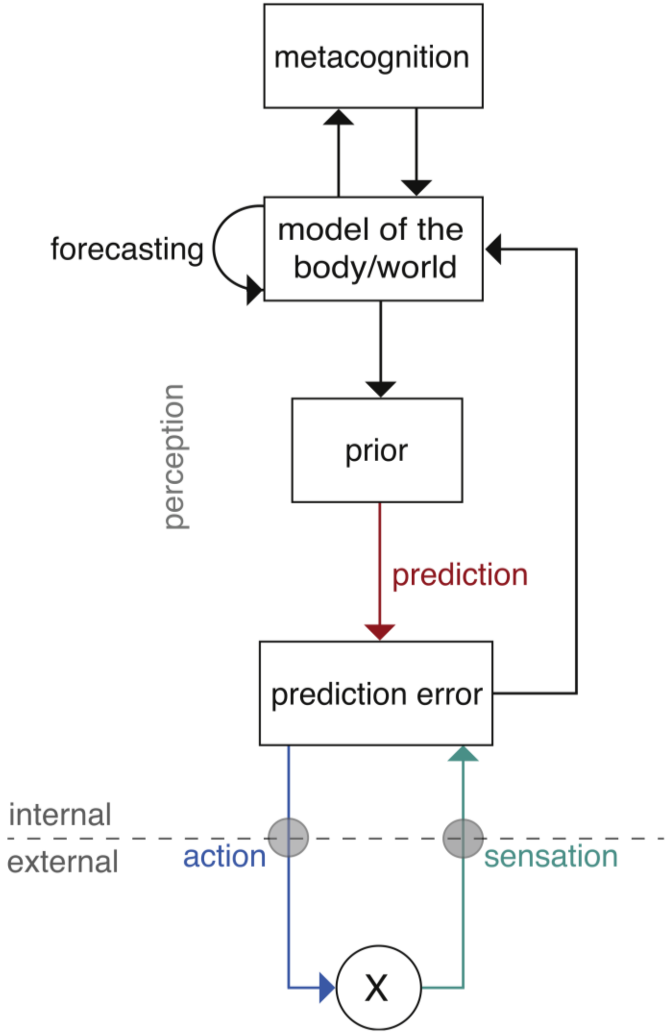
1. Sensory inputs (sensations)
2. Inference (perception)
3. Forecasting
4. Control (action)
5. Metacognition

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*

COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

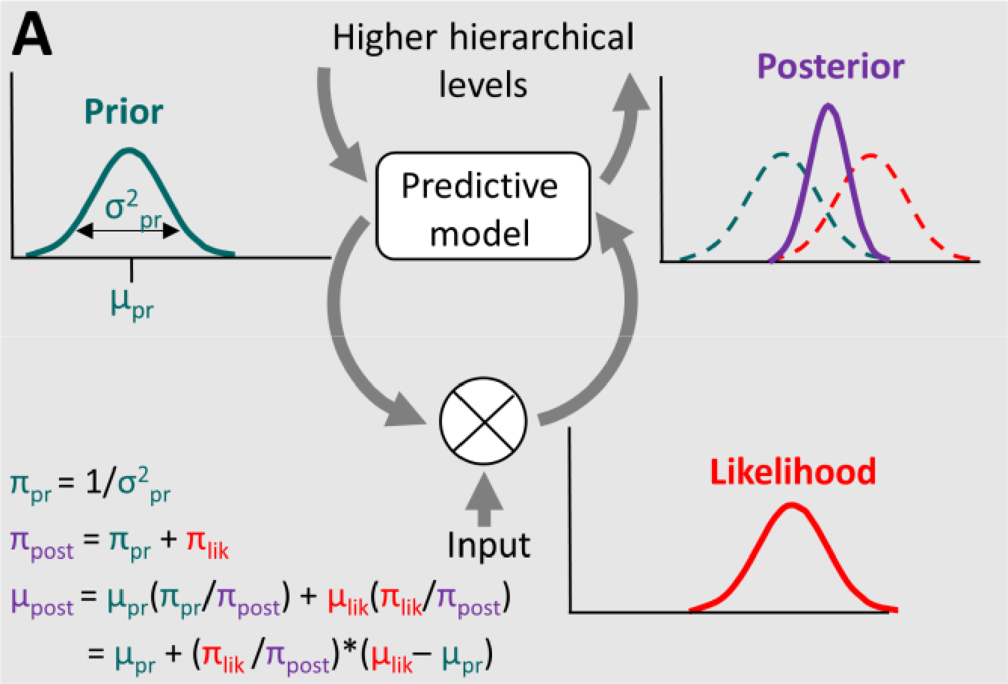
Precision: The right balance

Schizophrenia/Psychosis

Sterzer et al. 2018 *Biol Psychiatry*; Adams et al. 2013 *Front Psychiatry*; Corlett et al. 2011 *NPP*; Stephan et al. 2006 *Biol Psychiatry*; 2009; Powers et al. 2017 *Science*

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*



Interoceptive predictive coding

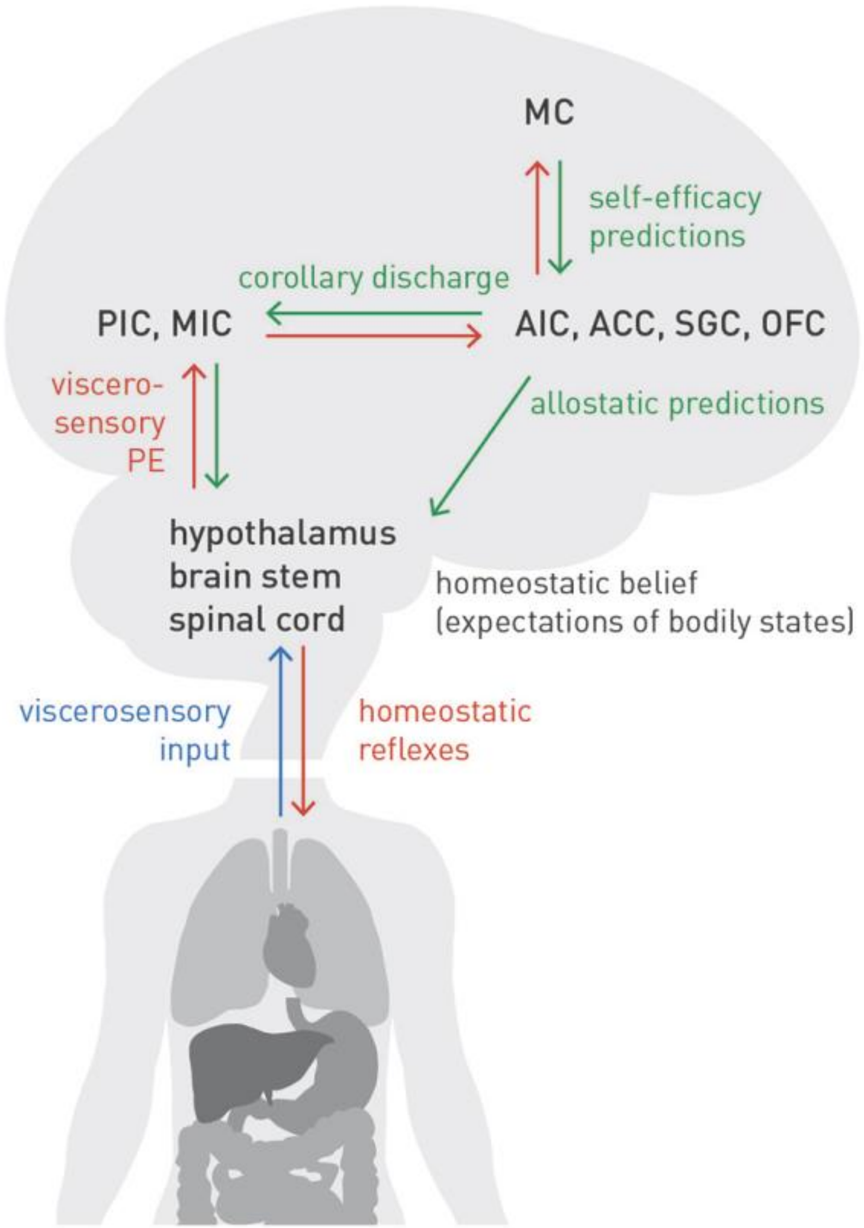
Seth 2013 *TiCS*

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*



COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

Testing further

- Sources of modulatory (precision) signals
  - Neuromodulators and how they interact

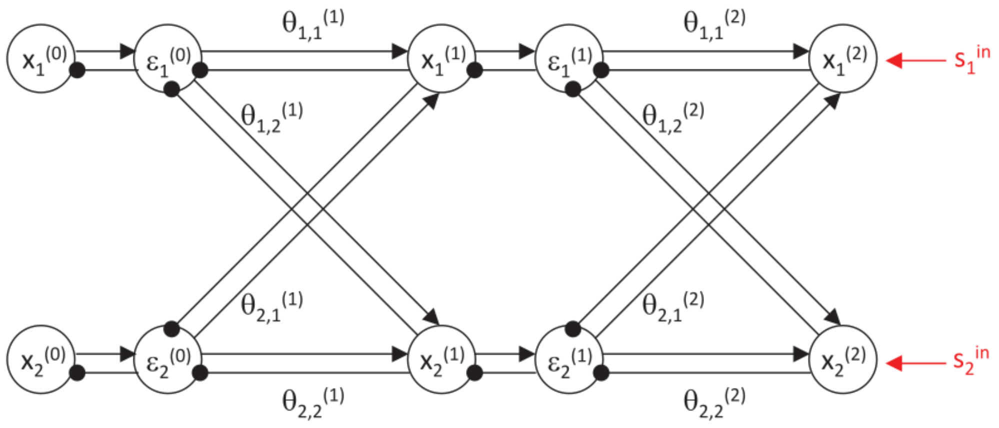
- Separate error units and prediction (representation) units
  - Laminar resolution, especially in the auditory domain (vision: Kok et al. 2016 *Curr Biol*)
  - High temporal resolution (predictions can precede a stimulus, PEs can only follow)

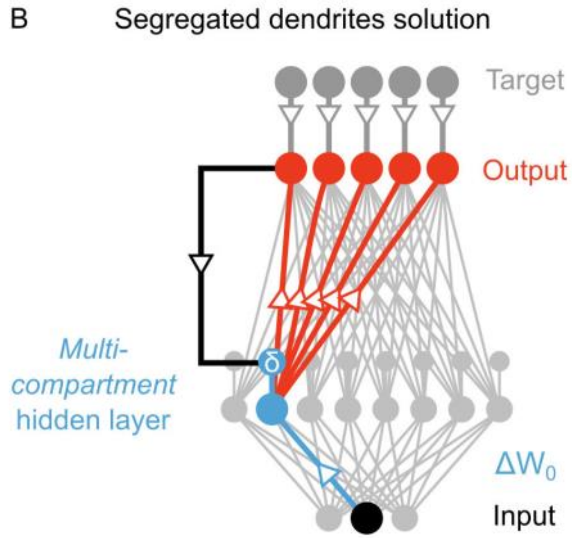
- Computationally explicit analyses ('model-based')
  - to test crucial theoretical distinctions, e.g. PEs vs. precision-weighted PEs
  - e.g., Iglesias et al. 2013 *Neuron*; Sedley et al. 2016 *eLife*; Bell et al. 2016 *Curr Biol*

PC and ANNs

- A predictive coding network can perform supervised learning autonomously using simple Hebbian learning and approximates the backpropagation algorithm **(Whittington & Bogacz 2017 *Neural Computation*)**

- A single pyramidal neuron can implement predictive coding via its apical dendrites **(Urbanczik & Senn 2014 *Neuron*; Sacramento et al. 2018 *NIPS*)** which can help to solve the credit assignment problem… **(Guerguiev et al. 2017 *eLife*)**





COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

Translational Neuromodeling Unit

V. How can I apply this model?

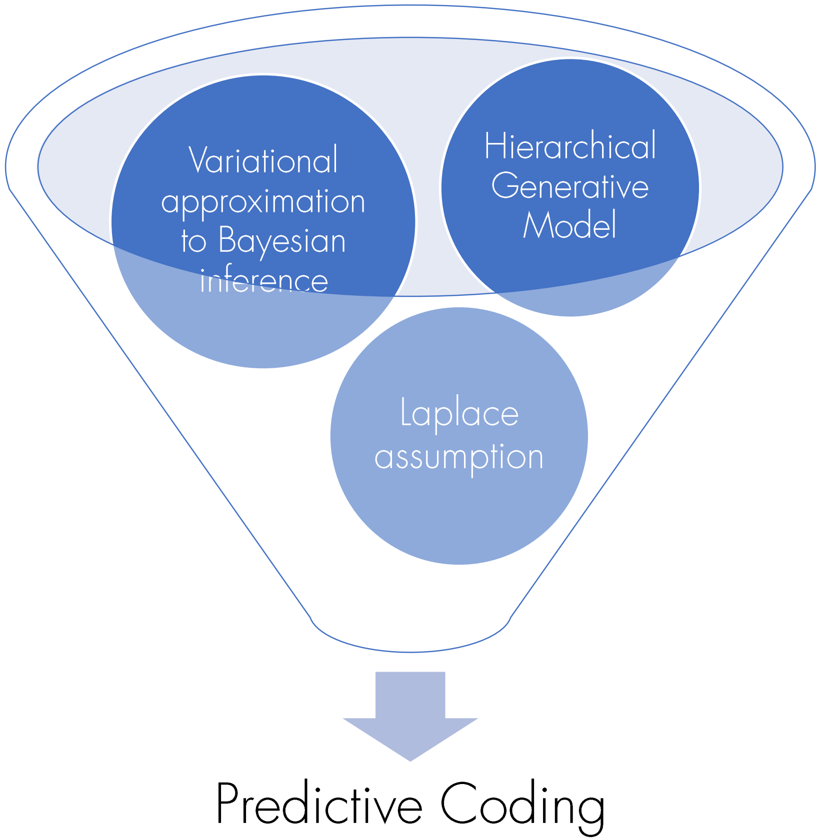
Why there is no predictive coding toolbox (yet)

COGNITIVE COMPUTATIONAL NEUROSCIENCE LABORATORY

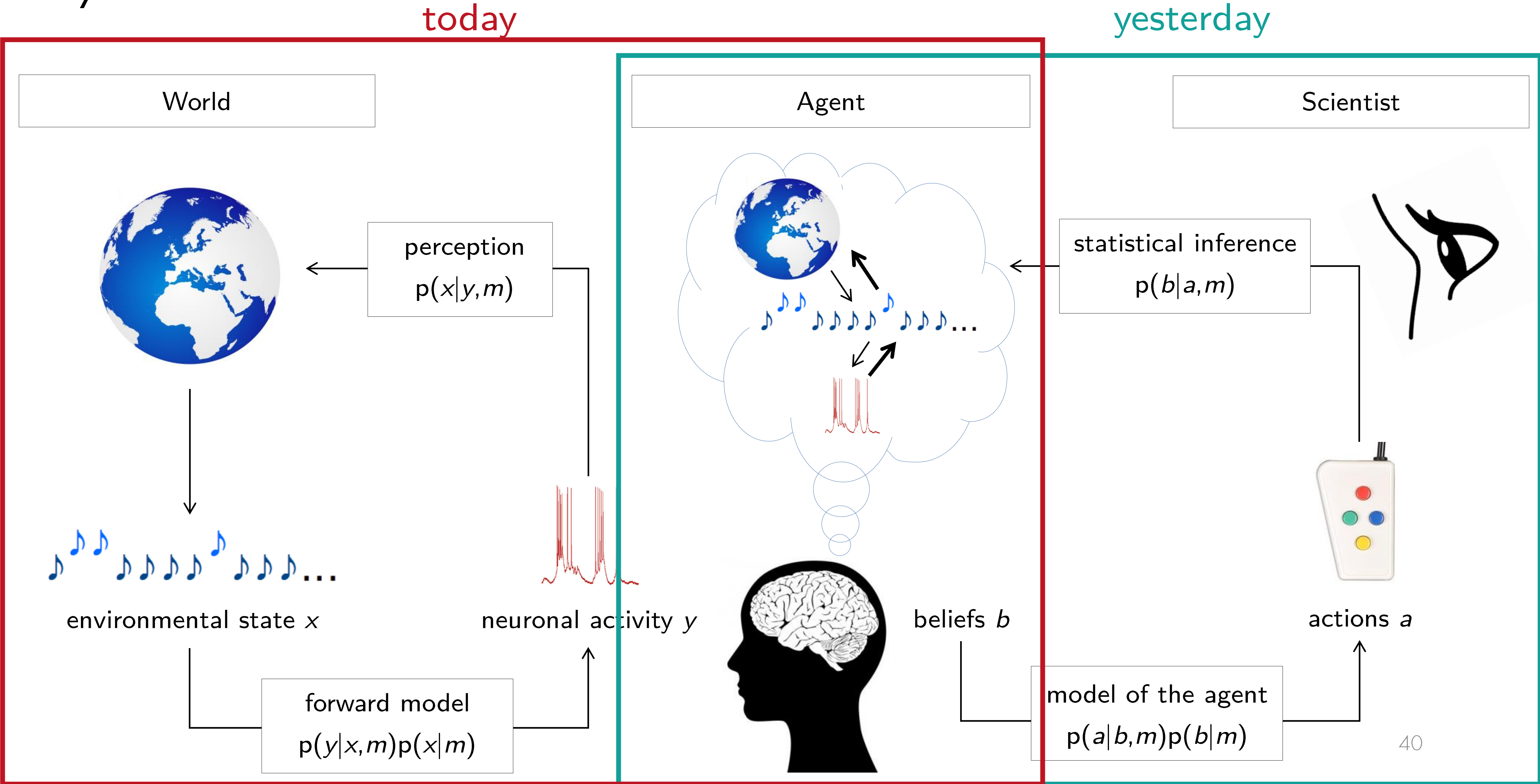
Translational Neuromodeling Unit

Building a model with PC

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



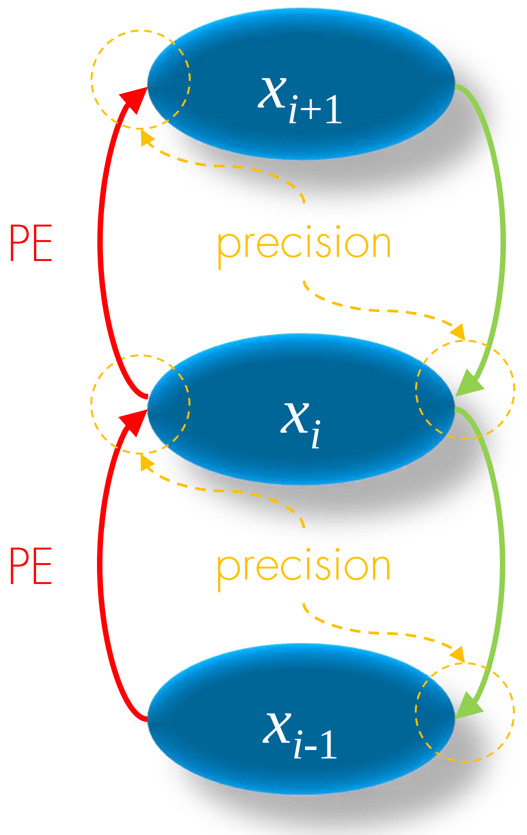
Bayesian observers



Thank you

Translational Neuromodeling Unit

Rick Adams
Rafal Bogacz



Δbelief ~ precision × PE