NACS 645 – Bayesian brains

-

Valentin Guigon

DEPARTMENT OF PSYCHOLOGY

SLD Lab

PROGRAM IN NEUROSCIENCE & COGNITIVE SCIENCE

The problem of induction

Human beings live in **uncertain, ever-changing environments.**

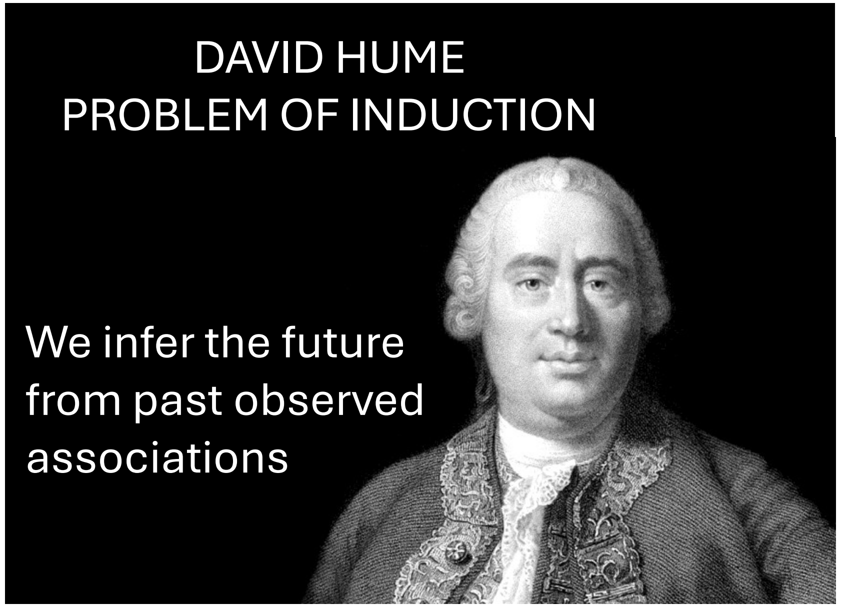
These environnements contain :

- **Systematic regularities**
- **Irregularities**

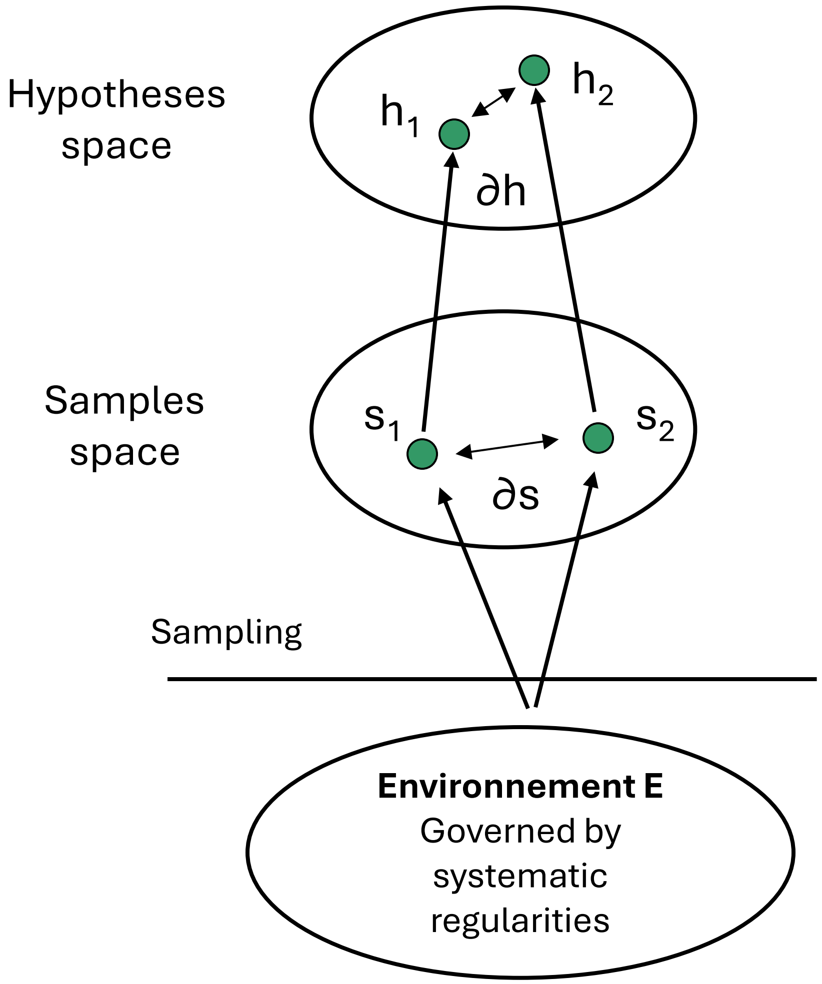
Humans facing a given situation :

- **Start from an initial state**
- **Gather information from the environment** (signal and noise; past and present)
- **Estimate what is probable/possible based on the available information and knowledge**
- **Take action**
- **Learn** from the consequences

Implication: learn from a particular a **generalization** that will be **applied to unobserved events: inductive inference**



From environments to samples to hypotheses



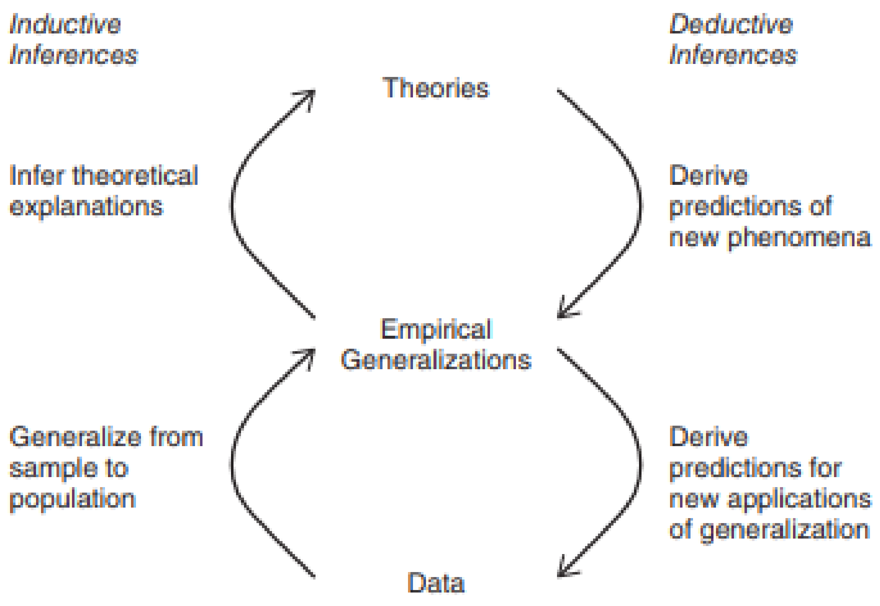
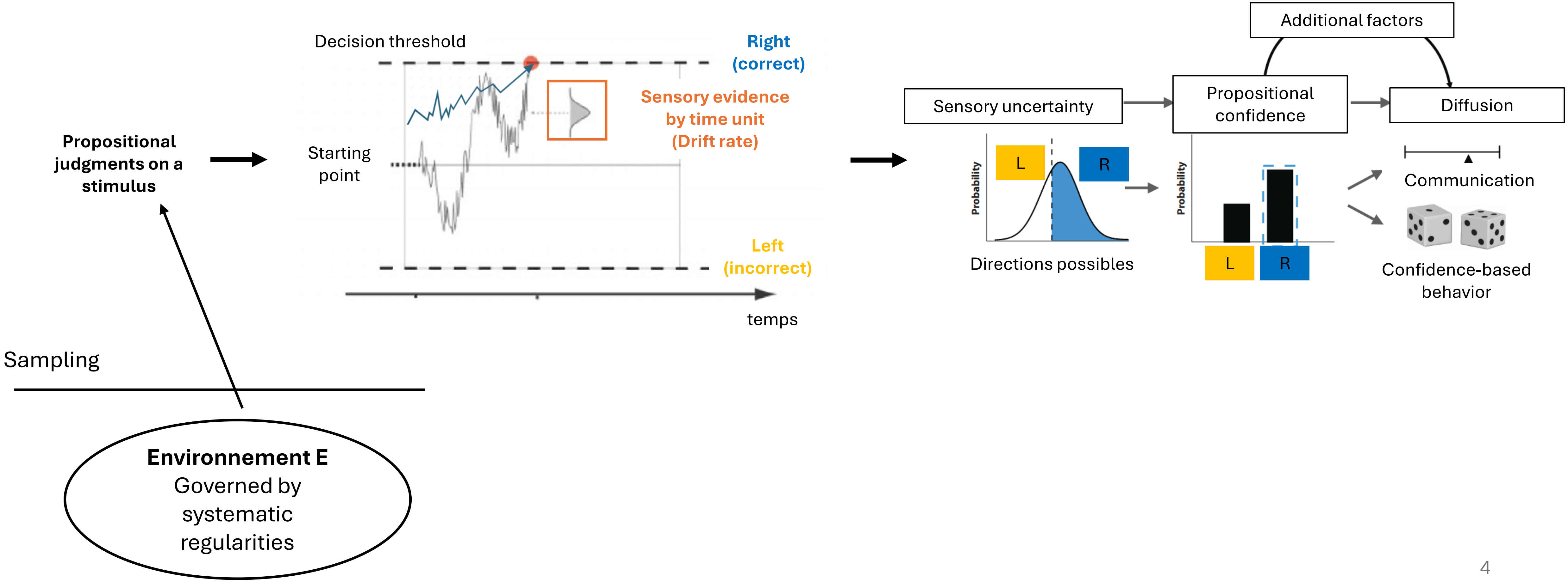


Figure 1.1 Two levels of inferences in science.

Lewandowsky et Oberayern, 2018. John Wiley & Sons, Inc.

Judgments on propositions are linked to the uncertainty on evidence



Judgments: near-optimal or biased performances?

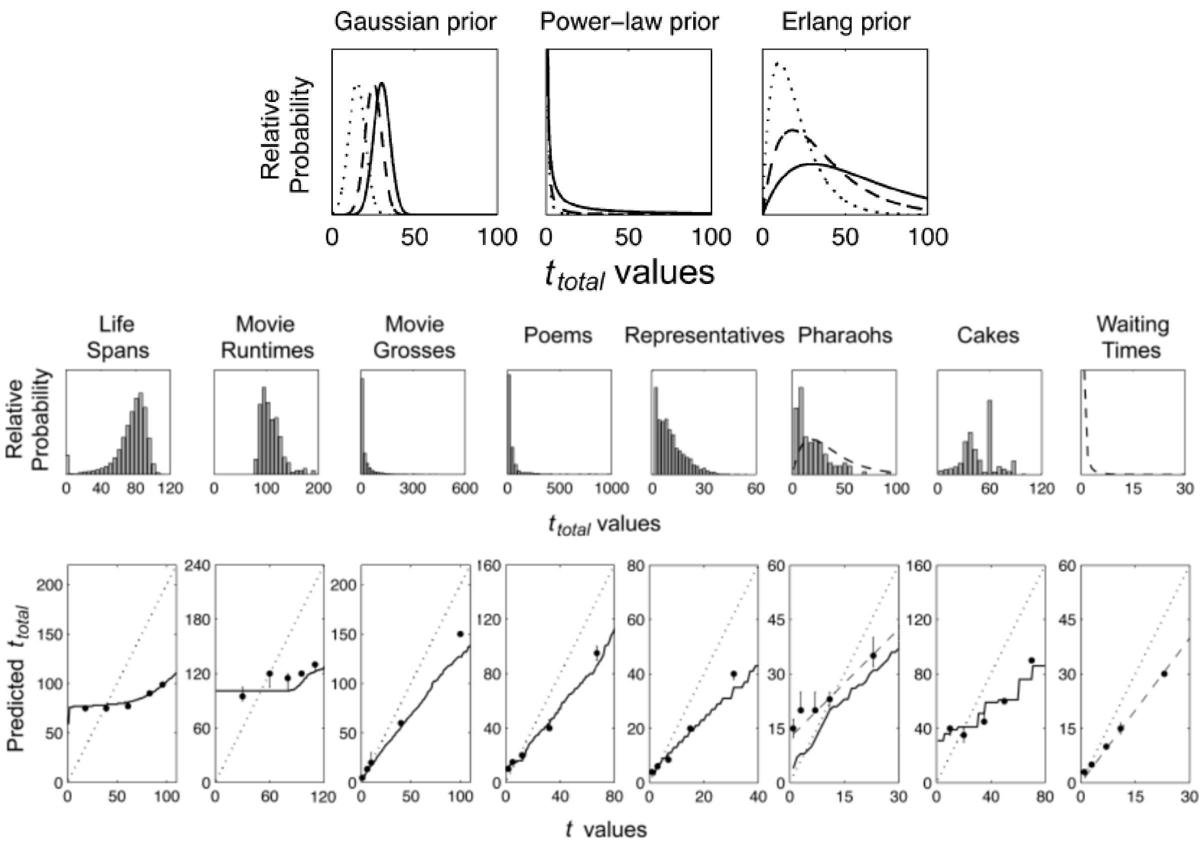
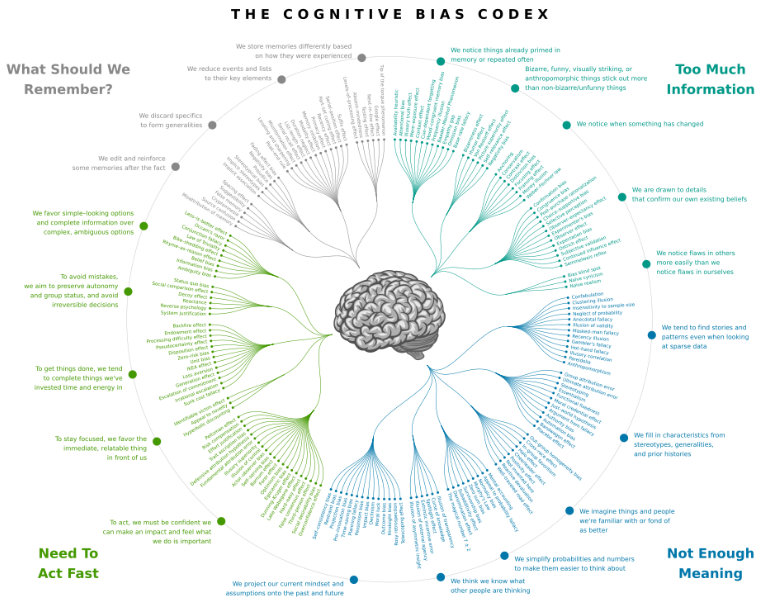


Fig. 2. People's predictions for various everyday phenomena. The top row of plots shows the empirical distributions of the total duration or extent, t_{total}, for each of these phenomena. The first two distributions are approximately Gaussian, the third and fourth are approximately power-law, and the fifth and sixth are approximately Erlang. The bottom row shows participants' predicted values of t_{total} for a single observed sample t of a duration or extent for each phenomenon. Black dots show the participants' median predictions of t_{total}. Error bars indicate 68% confidence intervals (estimated by a 1,000-sample bootstrap). Solid lines show the optimal Bayesian predictions based on the empirical prior distributions shown above. Dashed lines show predictions made by estimating a subjective prior, for the pharaohs and waiting-times stimuli, as explained in the main text. Dotted lines show predictions based on a fixed uninformative prior (Gott, 1993).

Griffiths et Tenenbaum, 2006. *Psychological Science*

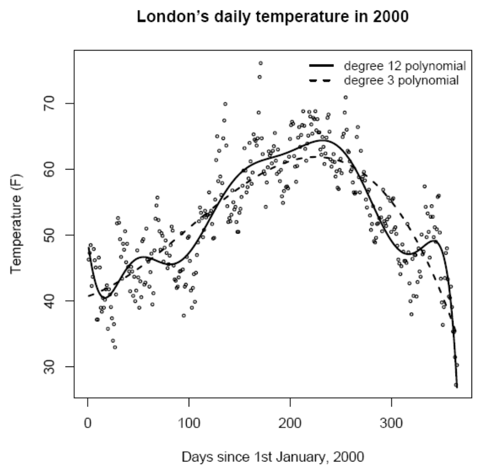


Tversy and Kahneman, 1970's

When biases lead to better inferences
(vs. complex reasoning processes)

When information is scarce, degraded, uncertain, complex, noisy
When the environment is sufficiently predictable

- **Predictive superiority**
  Generalizes better
- **Robustness to uncertainty**
  Ignoring information can make predictions less sensitive to noise and small samples
- **Cognitive efficiency**
  Reduces cost while maintaining sufficient performance (Martignon et al., 2008)



Gigerenzer & Brighton, 2009

- **By simplifying**, heuristics introduce a bias:

→ This bias reduces the instability of predictions (variance)
→ Improves robustness and generalizability to similar situations, especially under uncertainty

- **By complexifying**, models reduce bias but become more sensitive to noise :

→ This increases the variance of predictions
→ Reduces ability to generalize to new situations

Prediction error in predictive models:

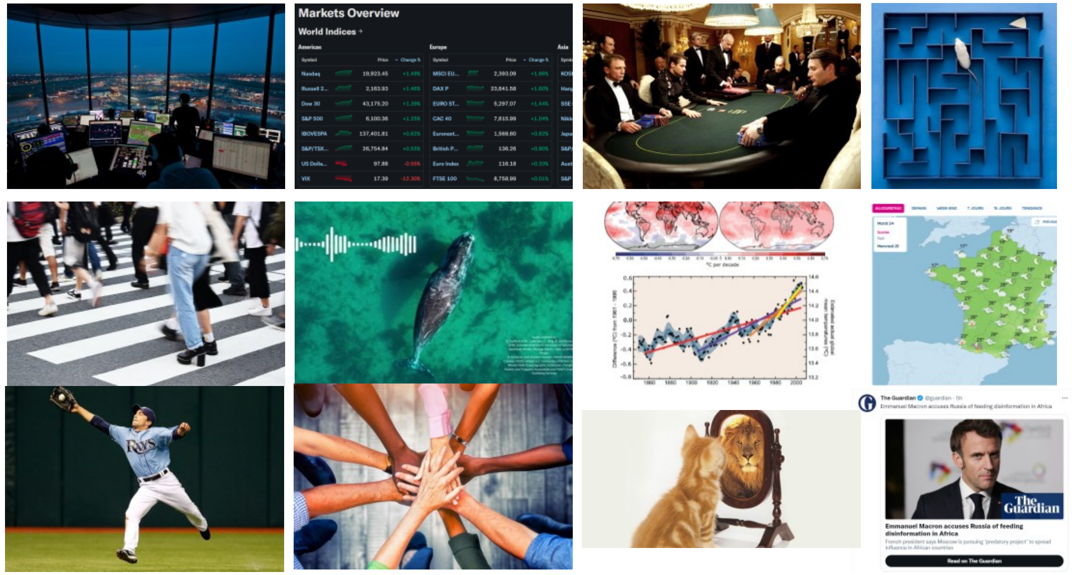
+ noise (ε)

When biases prevent better inferences
(vs. complex reasoning processes)

When information is abundant, clear, certain

When the environment is insufficiently predictable

And/or when we have the opportunity to mobilise lots of resources (time, computation, energy)



Where does the Bayesian brain fit in the picture?

Bottemanne, 2025. *Neuroscience*

The predictive brain

- Predictive coding (Hierarchical top-down predictions and bottom-up prediction errors)
- Bayesian brain (Brain maintains and updates posterior beliefs over latent causes)
- Free energy principle (Minimizing surprise to maintaining adaptive states)
- Active inference (Agents update beliefs and select actions to minimize expected free energy, trading exploitation-exploration)

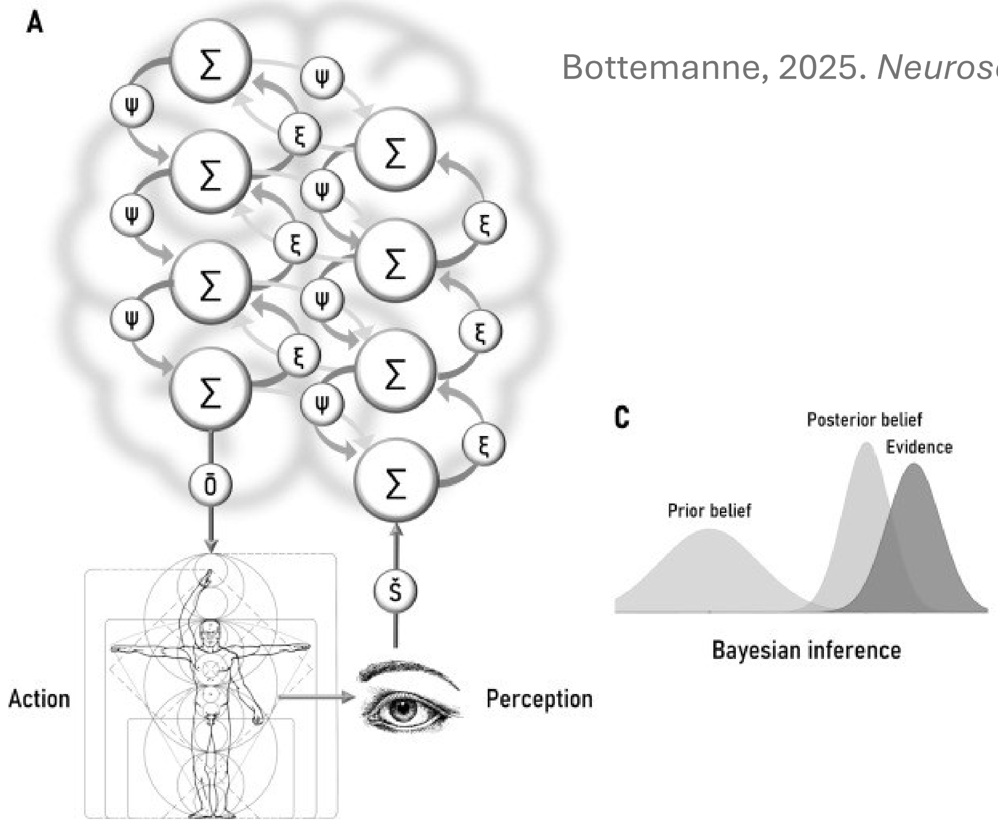


Fig. 1. Hierarchical predictive processing in Bayesian brain theory. (A) Bayesian brain theory supposes that the brain encodes a hierarchy of beliefs (∑) that are used to generate goal-directed actions (O) and predictions about sensory signals (Ψ) and sensory signal processing (S). These predictions are updated by prediction errors (ξ) resulting from actions and changes in the environment.

Can the Bayesian brain, and should it, explain the best and worst inferences?

Identifying the correct architecture

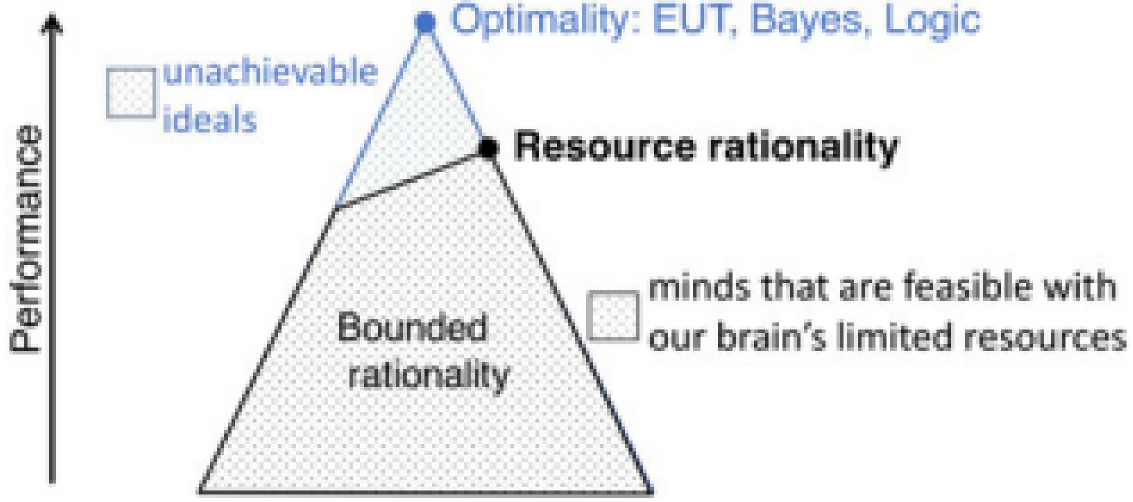


Figure 1. Resource rationality and its relationship to optimality and Tversky and Kahneman's concept of bounded rationality. The horizontal dimension corresponds to alternative cognitive mechanisms that achieve the same level of performance. Each dot represents a possible mind. The gray dots are minds with bounded cognitive resources and the blue dots are minds with unlimited computational resources. The thick black line symbolizes the bounds entailed by people's limited cognitive resources. Resource limitations reflect anatomical, physiological, and metabolic con-

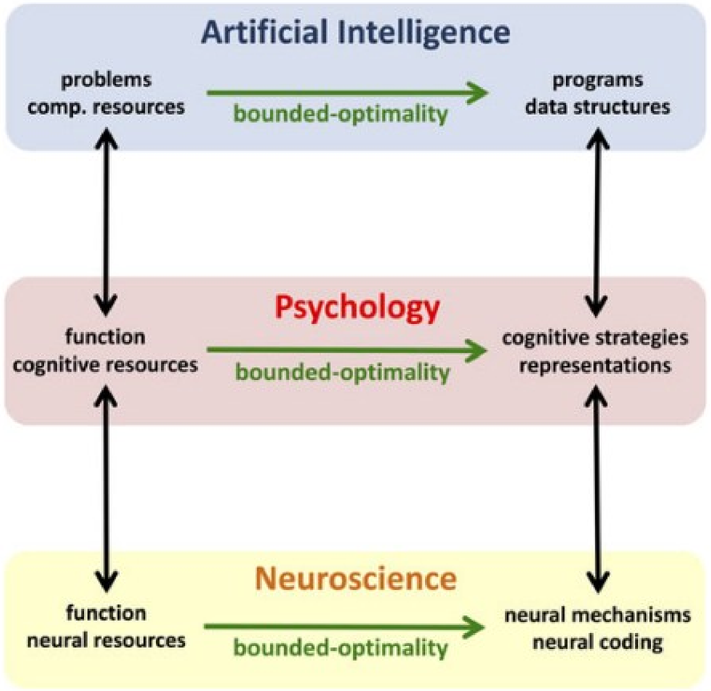
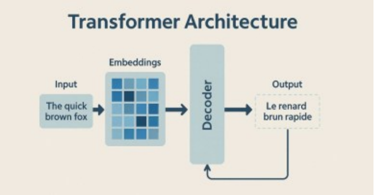
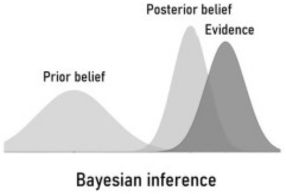


Figure 4. Resource-rational analysis connects levels of analysis.





We operate in a space where we need to recover complex latent structures despite scarce resources, have interpretability despite little data and provide posteriors on the right hypotheses/structures