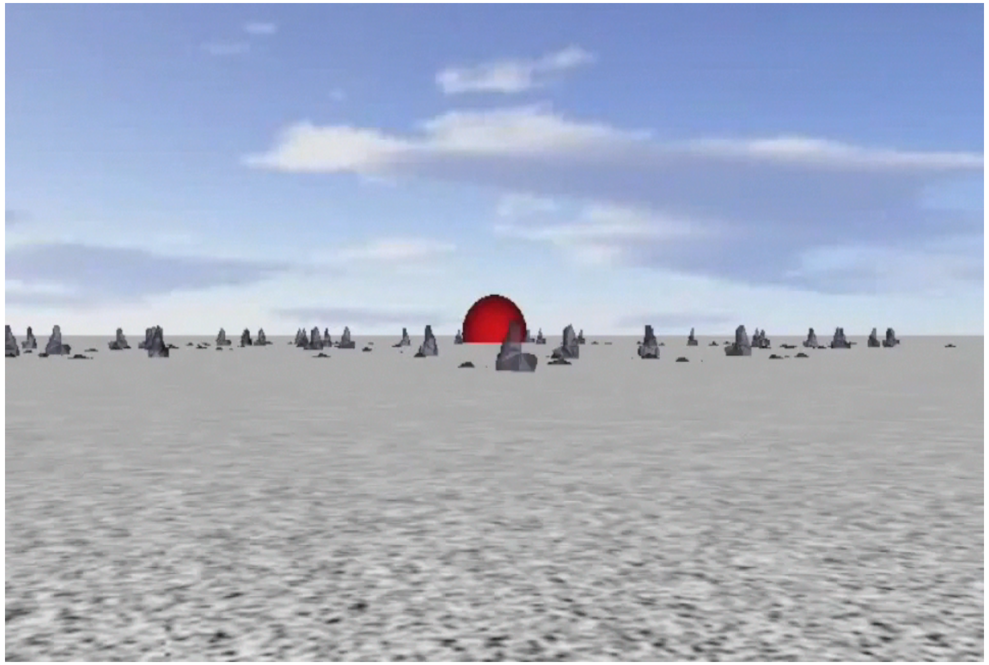
Bayesian Models for Perception

Dr. Frederike Petzschner

Translational Neuromodeling Unit

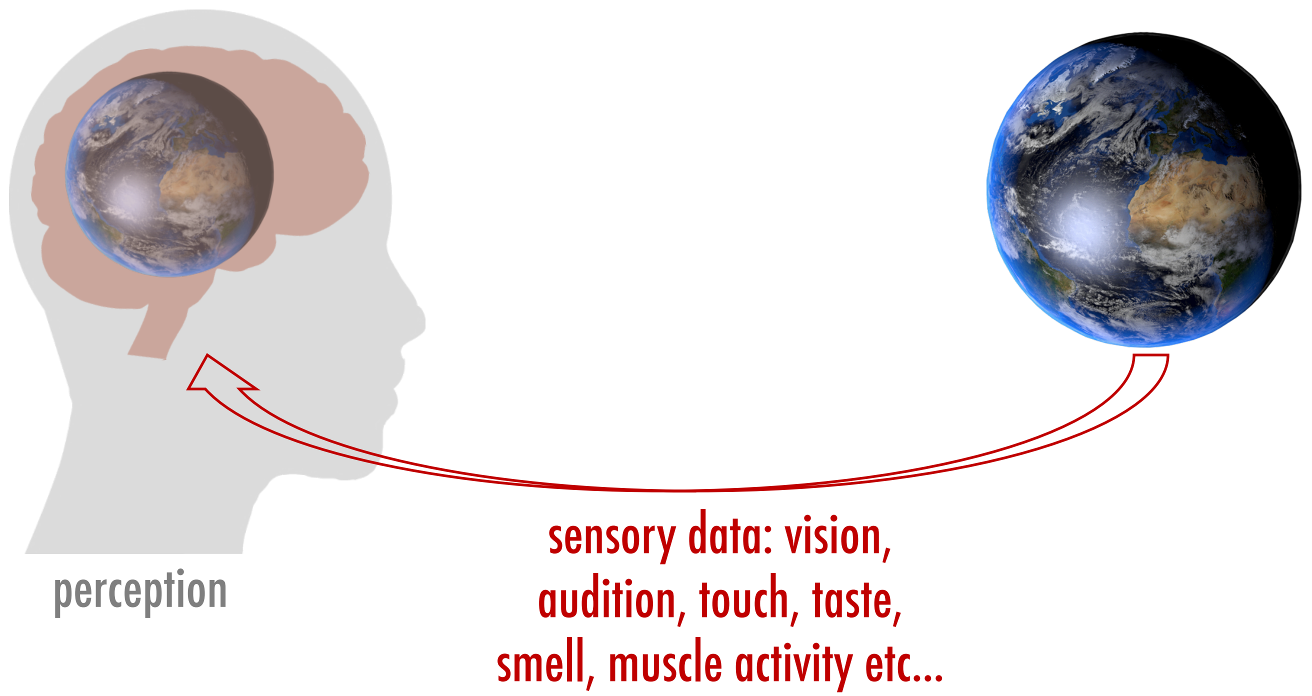
UZH & ETH



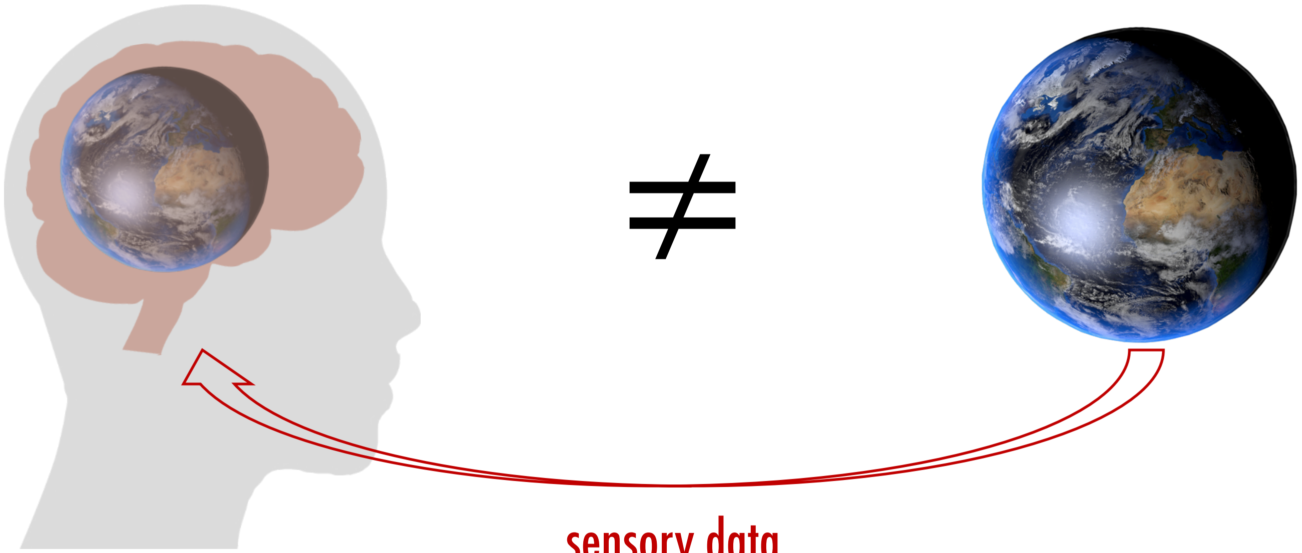
What is perception?
Why Bayes to model perception?
Do you we behave like a ‘Bayesian’?
Are humans Bayesian?

What is perception?

The brain’s dilemma:



The brain’s dilemma:

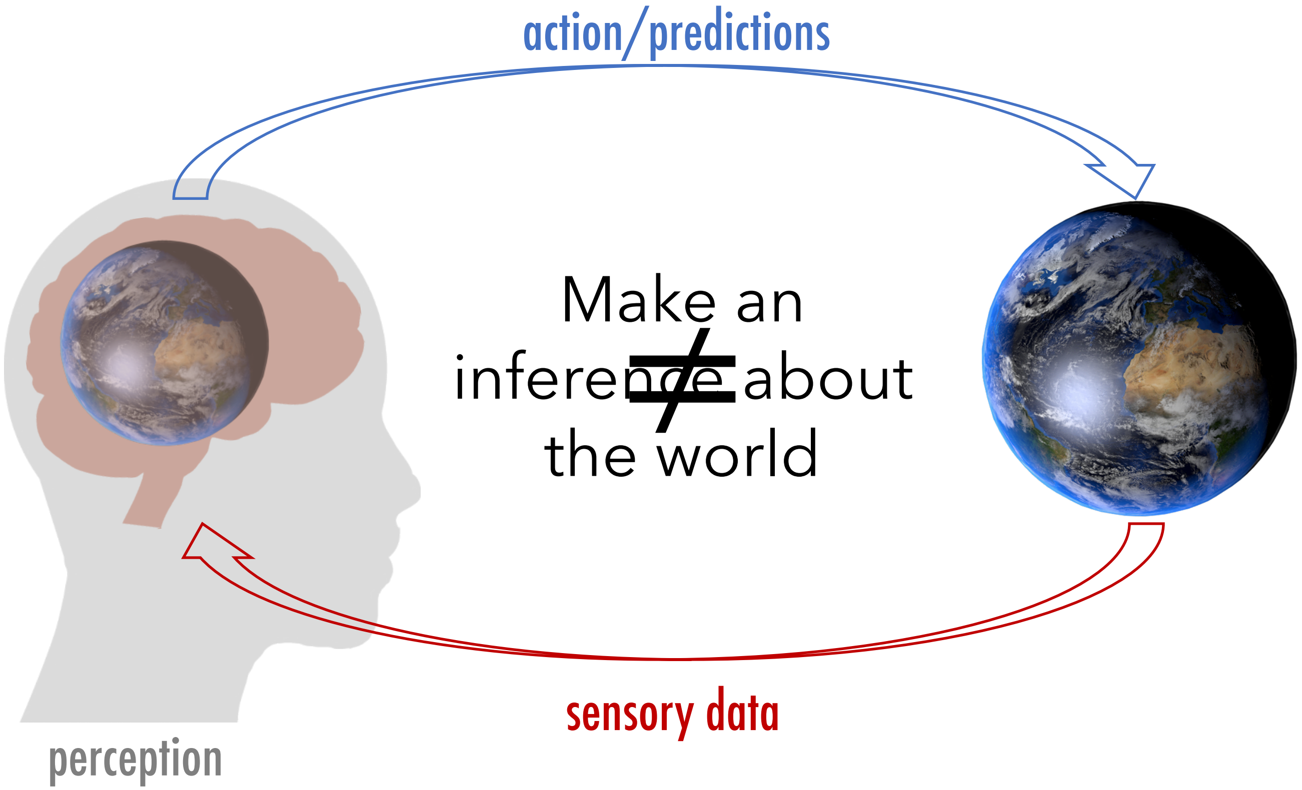


perception

sensory data

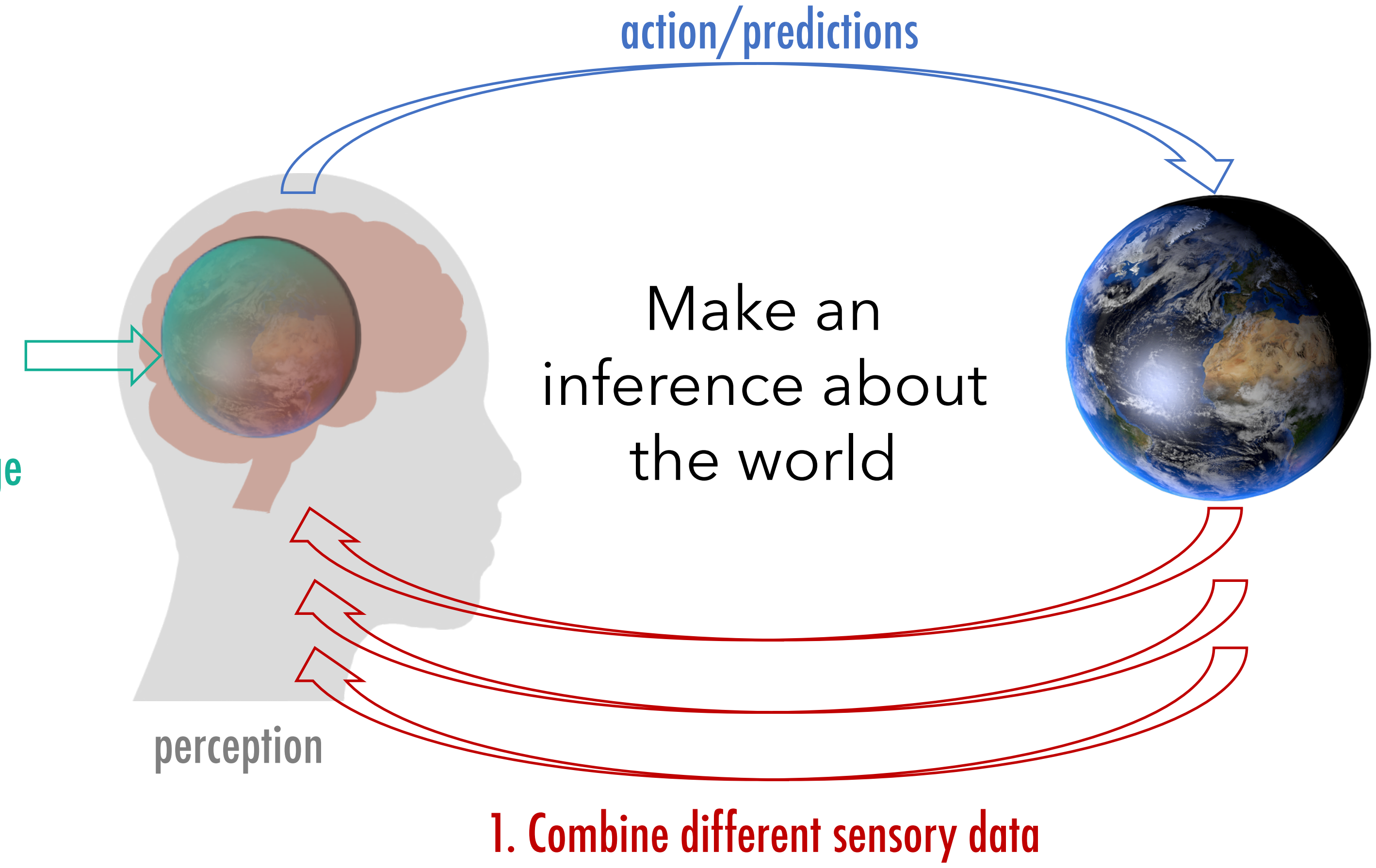
limited
noisy
ambiguous
hidden structure

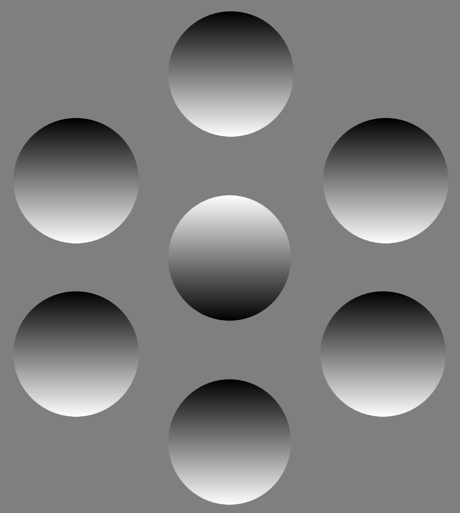
The brain’s dilemma:



The brain's solution:

Hermann
von Helmholtz
1821 - 1894

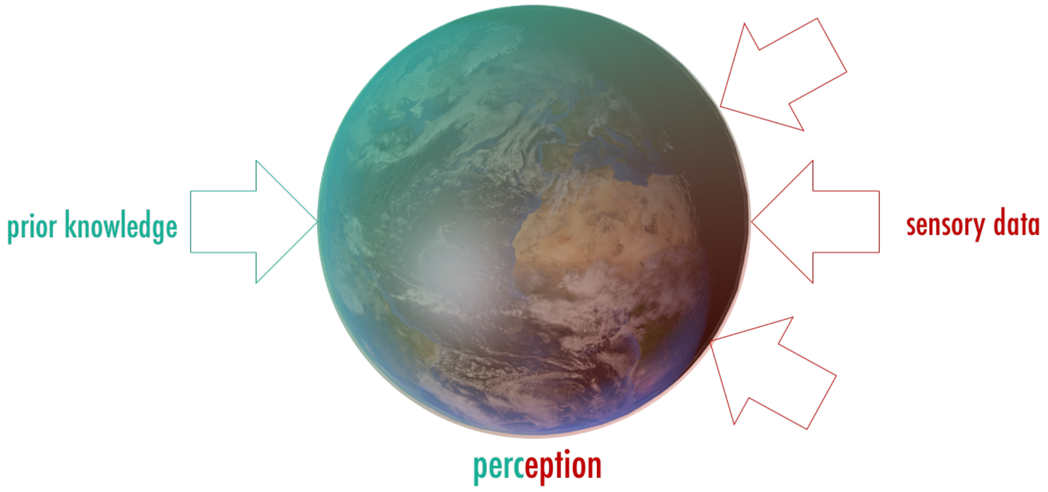




What is perception?

Perception is the result of a combination different types of noisy information:

- *Sensory data*
- *Prior knowledge*

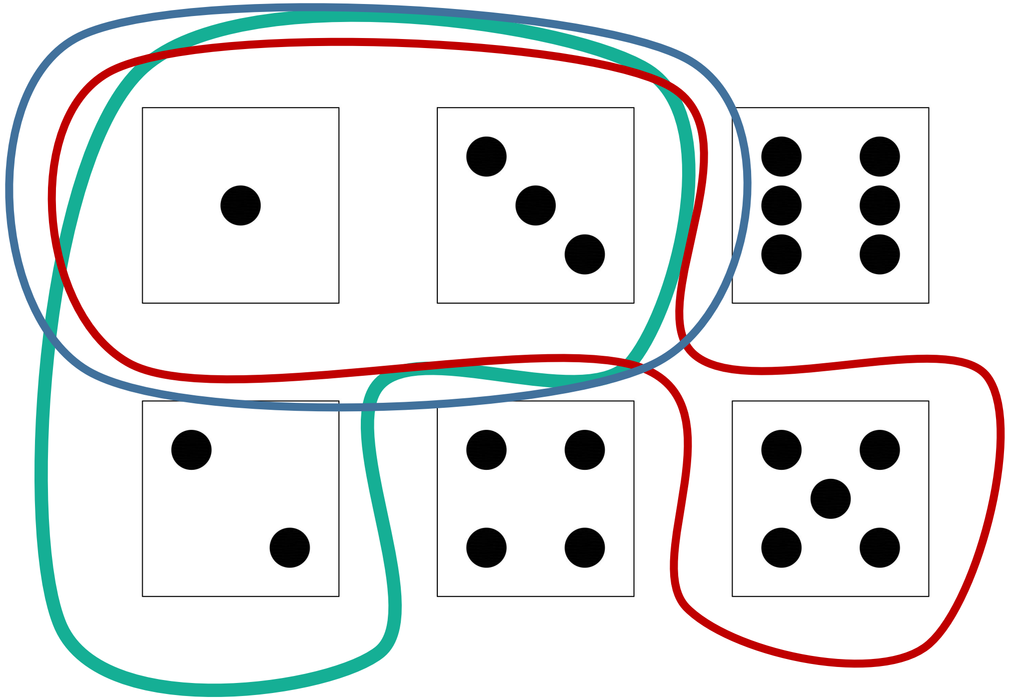


Why use Bayes to model perception?

What is Bayes’ Rule?

Statistics Tool for conditional probability distributions

P(A): Probability of rolling a dice and getting a number bellow 4
P(B): Probability of an odd number
P(A|B) = Probability of an number bellow 4 given that it is odd?



$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

Statistics Tool for conditional probability distribution

(1) $$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

(2) $$P(B|A) = \frac{P(B \cap A)}{P(A)} \quad \Rightarrow \quad P(B|A) \cdot P(A) = P(B \cap A) = P(A \cap B)$$

Equation (2) in (1):

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

BAYES' RULE

An example: Imagine you are a doctor and you want to find out if your patient has breast cancer.

You know: The probability of breast cancer is 1% for a woman at 40 who participates in a routine screening.

Now you get new information in form of a mammography.
You want to update your belief based on that new information.

What you also know, if a woman has breast cancer, the probability is 80% that she will have a positive mammography. If a woman does not have breast cancer, the probability is 9.6% that she will also have a positive mammography.

The woman you are testing as a positive mammography.
What are the chances of her having cancer?

A. greater than 90%
B. between 70% and 90% — 95 out of 100 doctors
C. between 50% and 70%
D. between 30% and 50%
E. between 10% and 30%
F. less than 10% — correct

Lets use Bayes' Rule:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

$$P(cancer|test +) = \frac{P(test + |cancer)\, P(cancer)}{P(test+)}$$

$$P(test +) = P(test + |cancer) \cdot P(cancer) + P(test + |no\ cancer) \cdot P(no\ cancer)$$

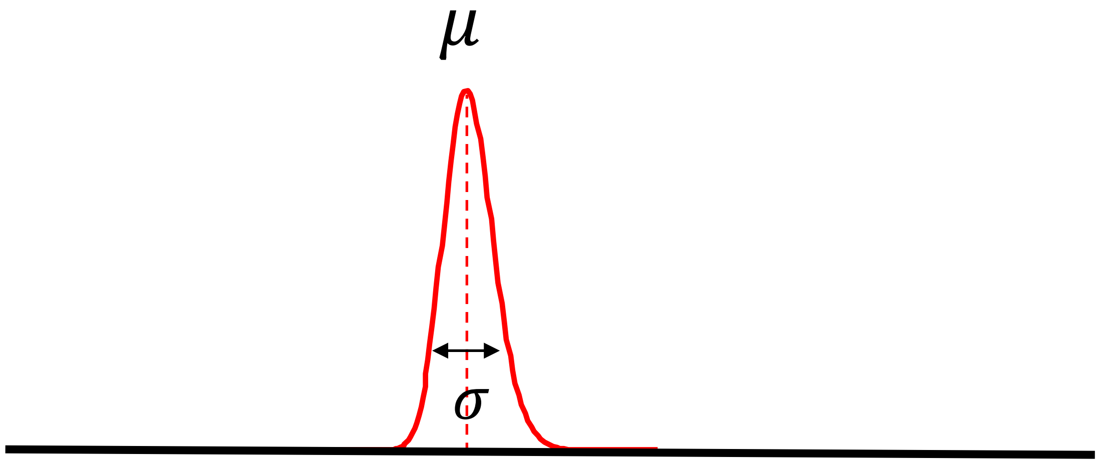
P(cancer): The probability of breast cancer is 1% for a woman
P(test+ given cancer):If a woman has breast cancer, the probability is 80% that she will have a positive mammography.
P(text+ given no cancer): If a woman does not have breast cancer, the probability is 9.6% that she will also have a positive mammography.

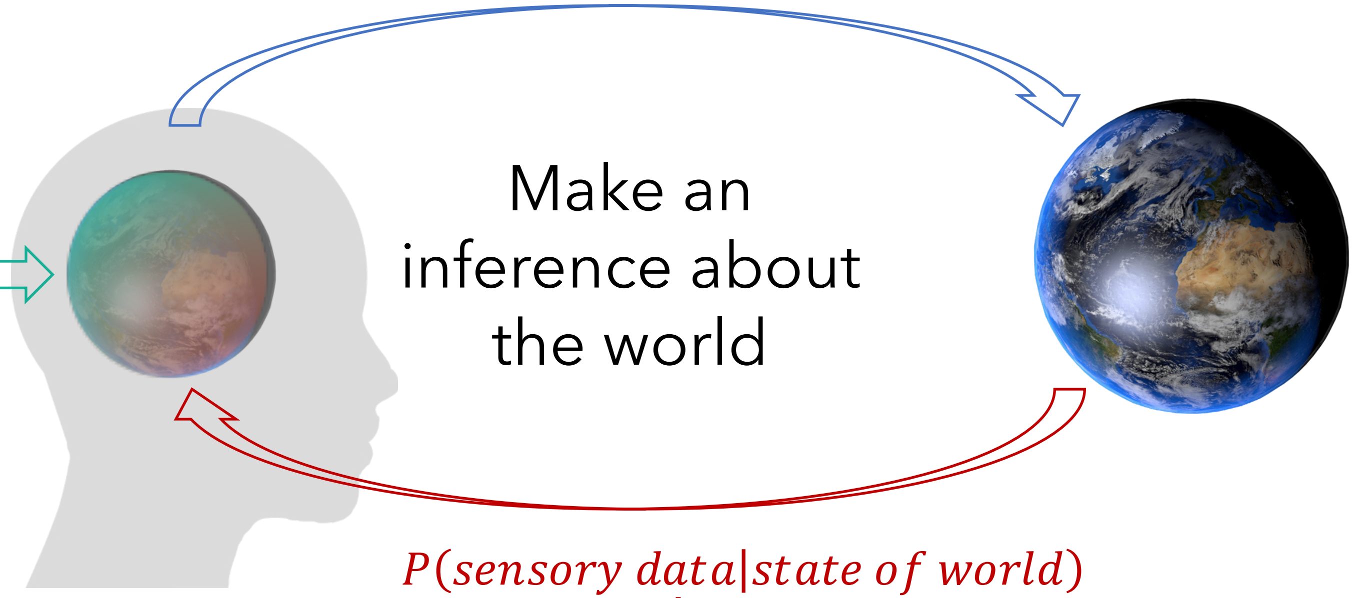
$$P(cancer|test +) = \frac{P(test +|cancer)\, P(cancer)}{P(test+)} = \frac{0.80 \cdot 0.01}{0.80 \cdot 0.01 + 0.96 \cdot 0.99} = 0.0457$$

Why use Bayes to model perception?

It's a trick...

- *Information can be described by probability distributions*
- *Sensory information and implicit assumptions (beliefs) may thus be formulates as (conditional) probability distributions*
- *The laws of probability can be used to calculate how these probability distributions can be combined in a statistical optimal manner*

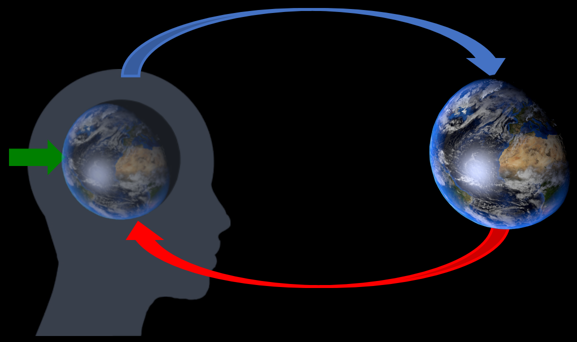


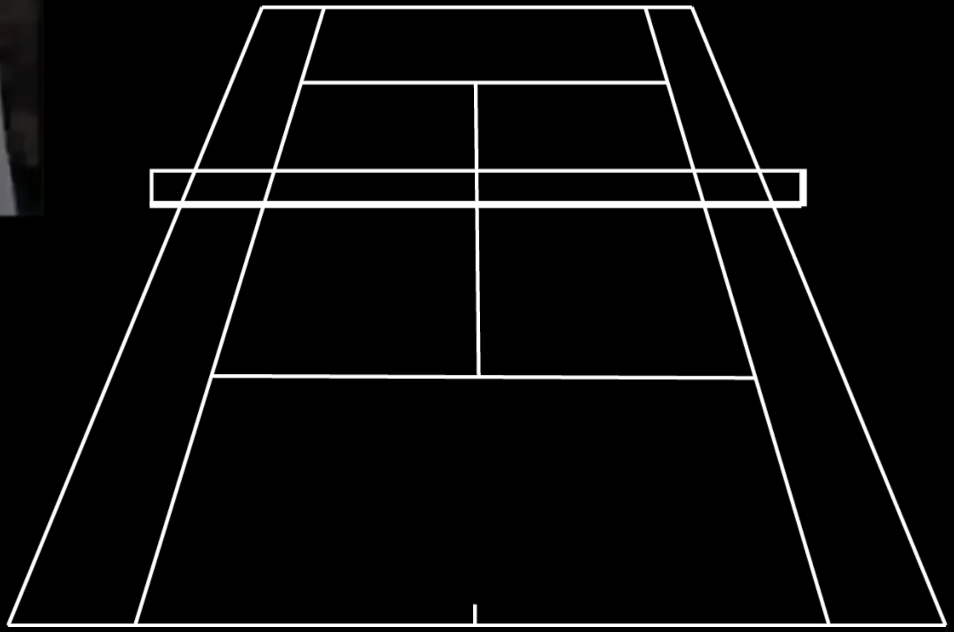


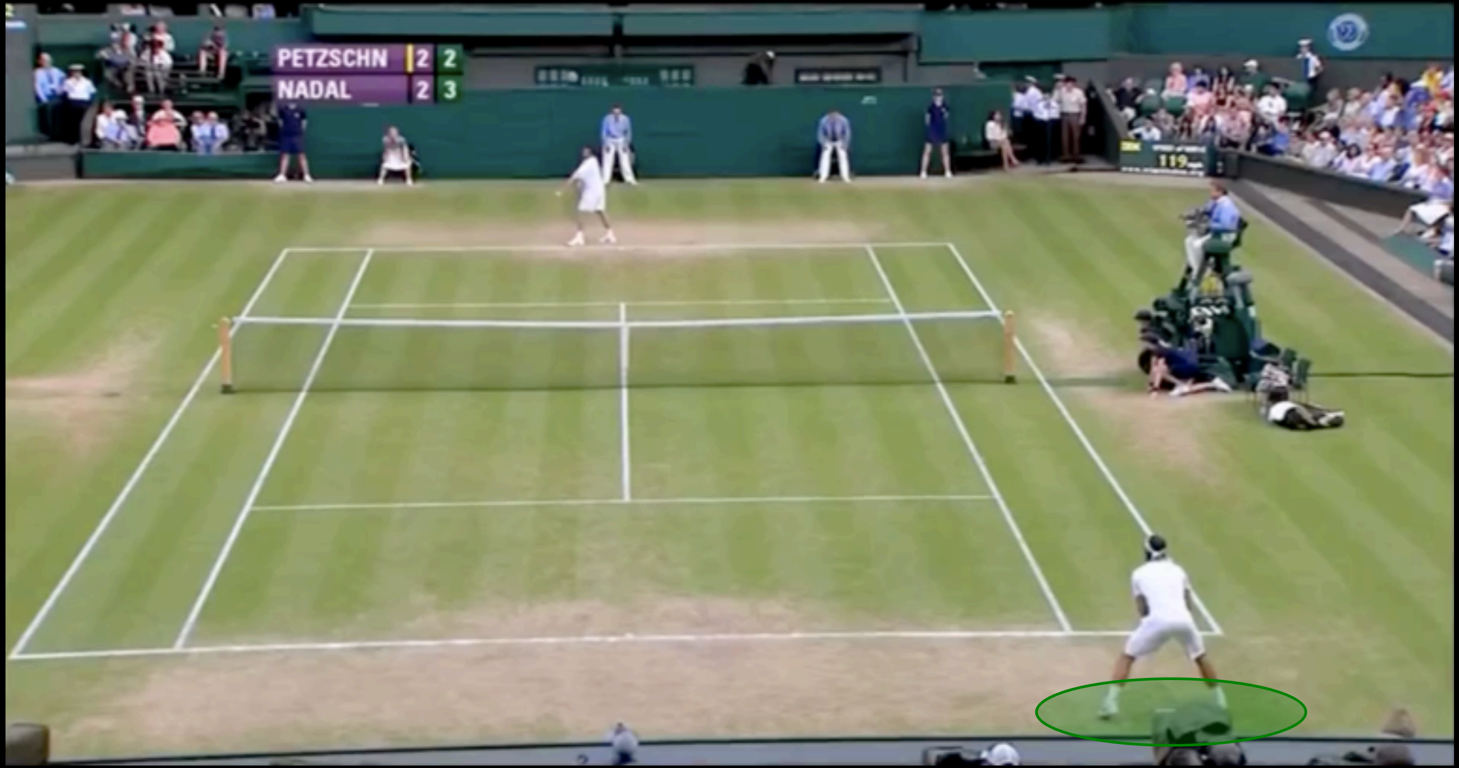
Bayes' Rule

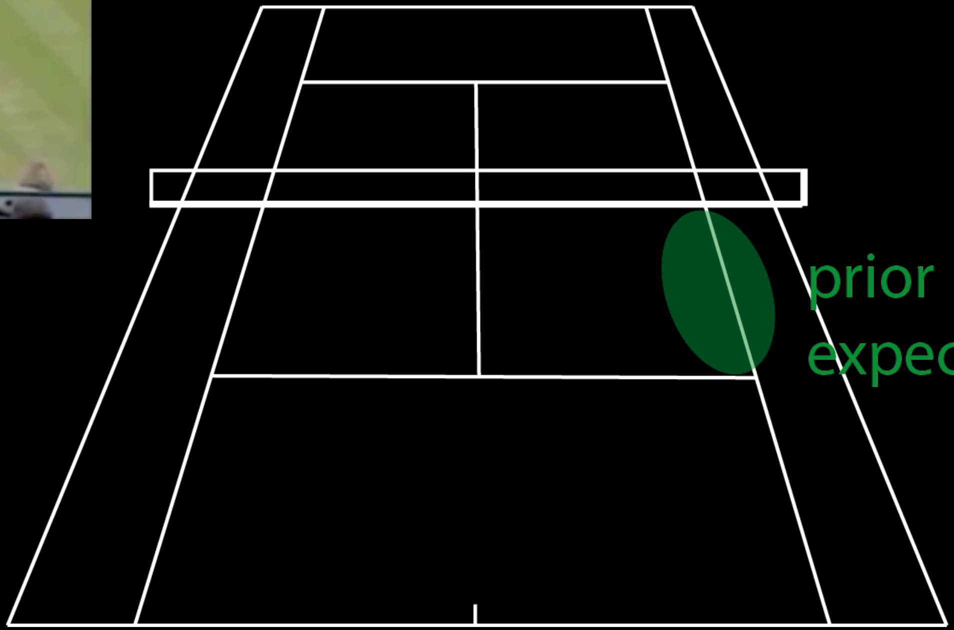
$$\underbrace{P(state\ of\ world|sensory\ data)}_{posterior} = \frac{\overbrace{P(sensory\ data|state\ of\ world)}^{likelihood}\,\overbrace{P(state\ of\ world)}^{prior}}{P(sensory\ data)}$$

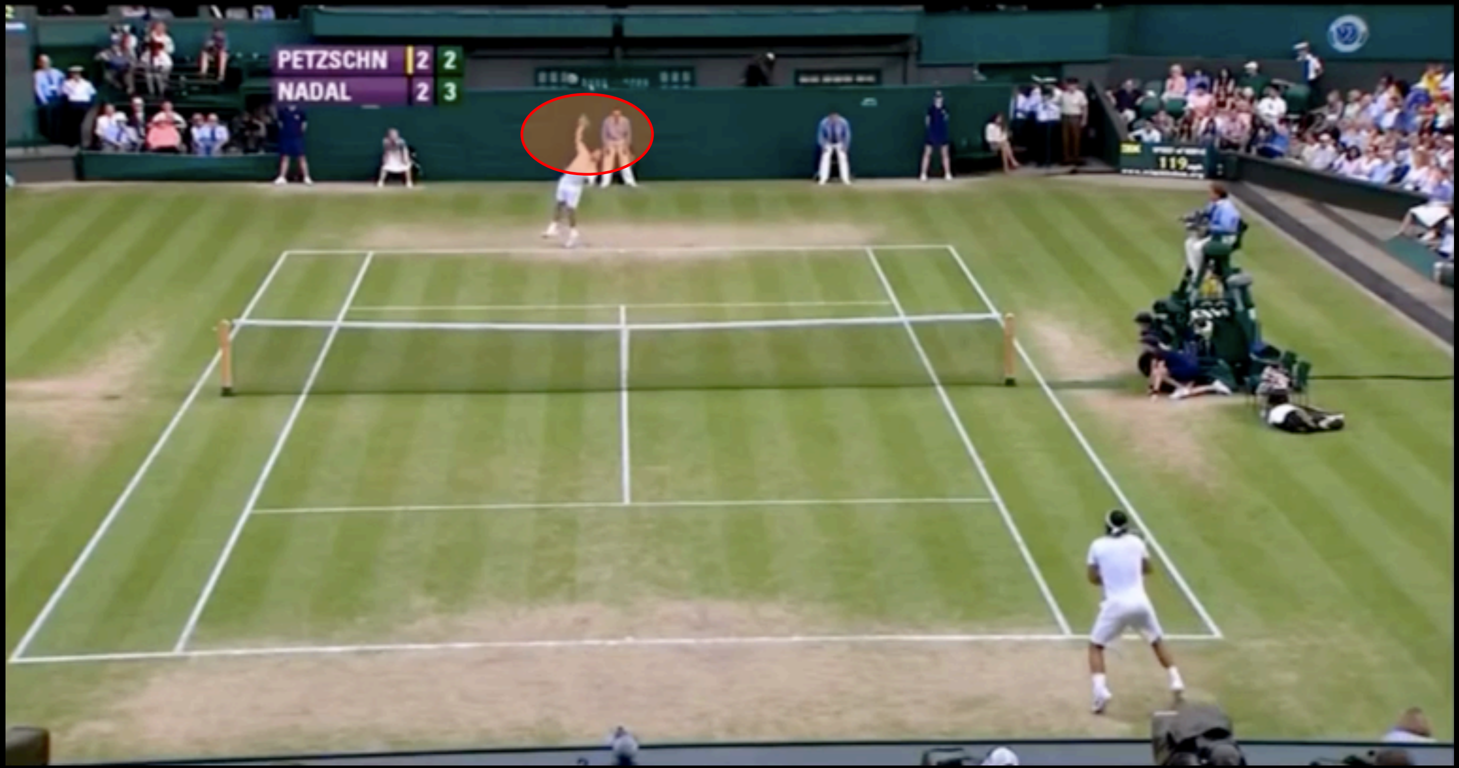
A Bayesian Model of Nadal

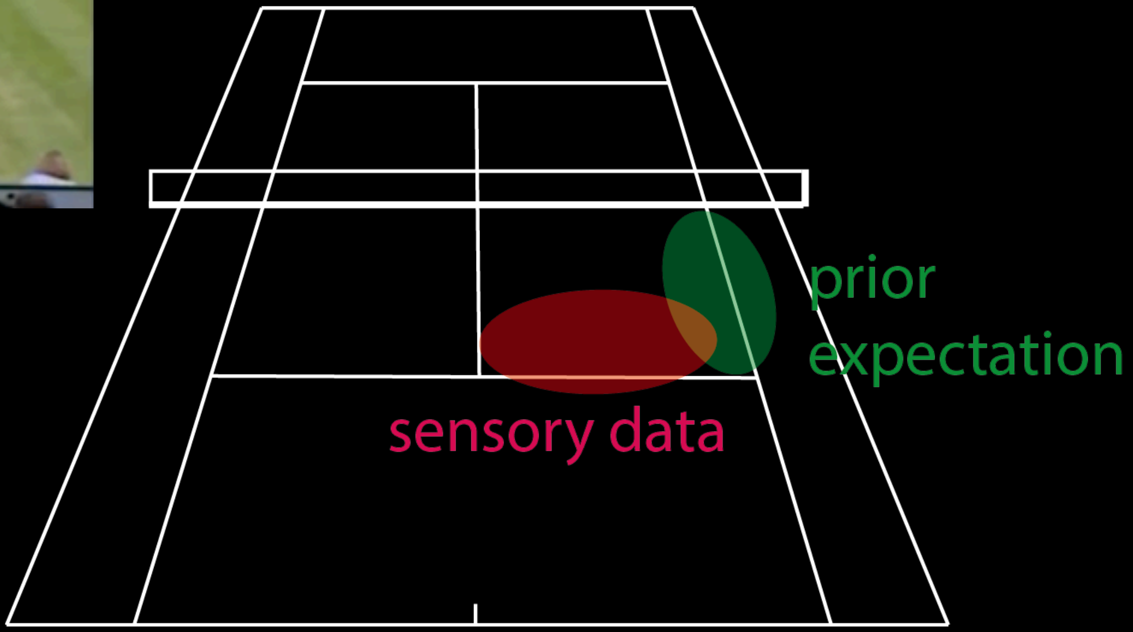


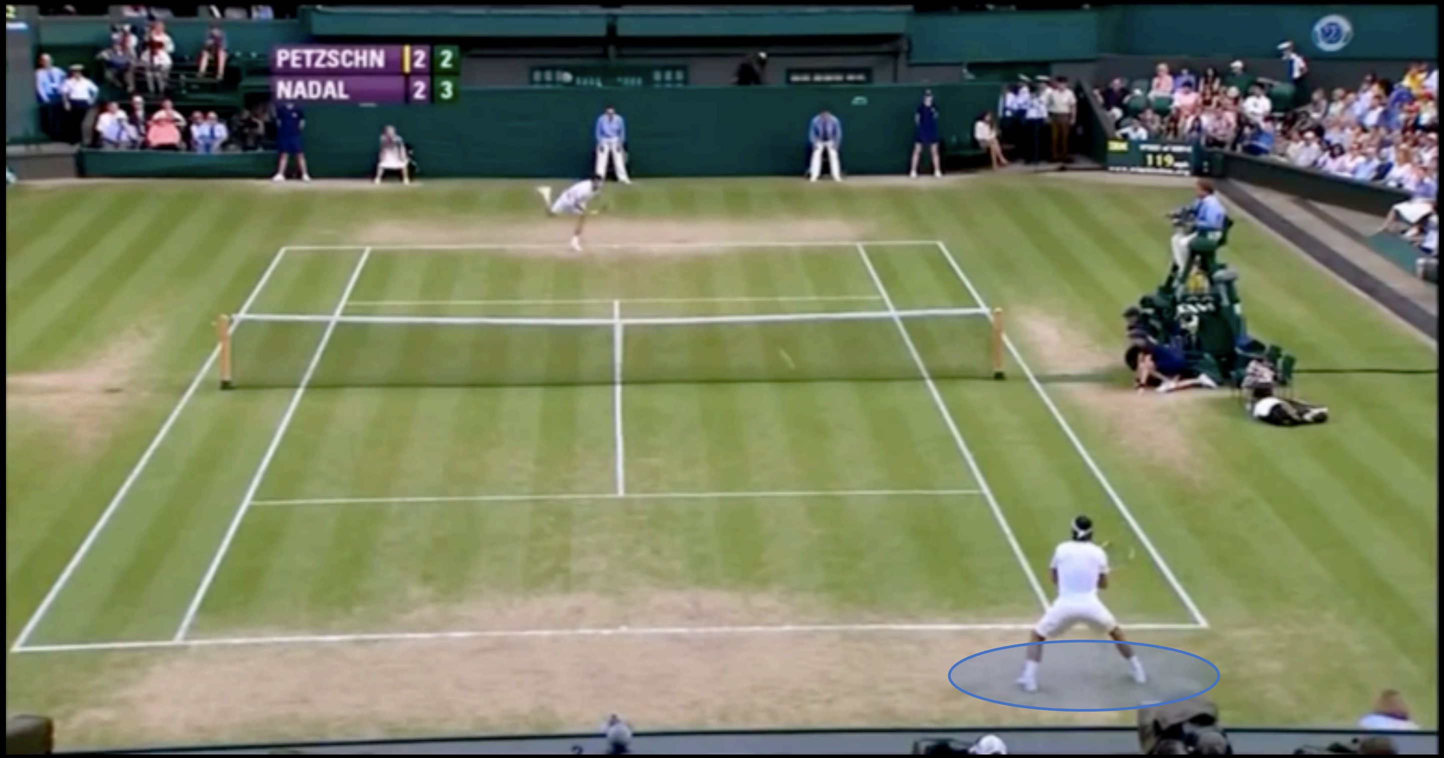


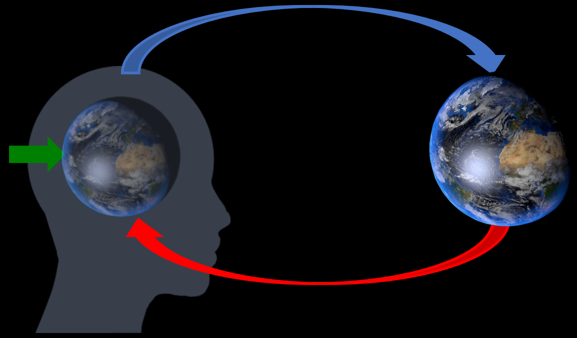

PETZSCHN 2 2
NADAL 2 3

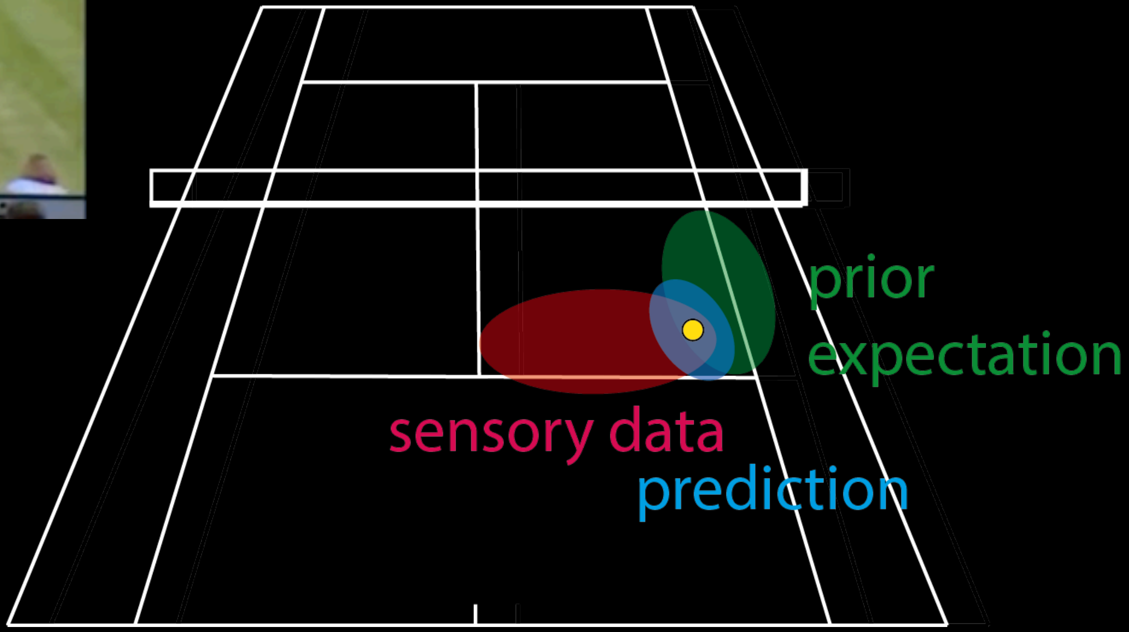

prior
expectation


PETZSCHN 2 2
NADAL 2 3

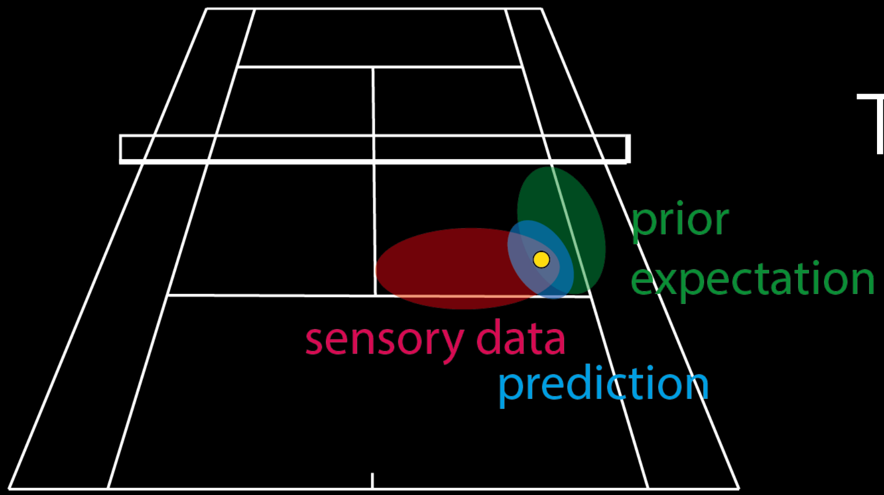

prior
expectation
sensory data


PETZSCHN 2 2
NADAL 2 3

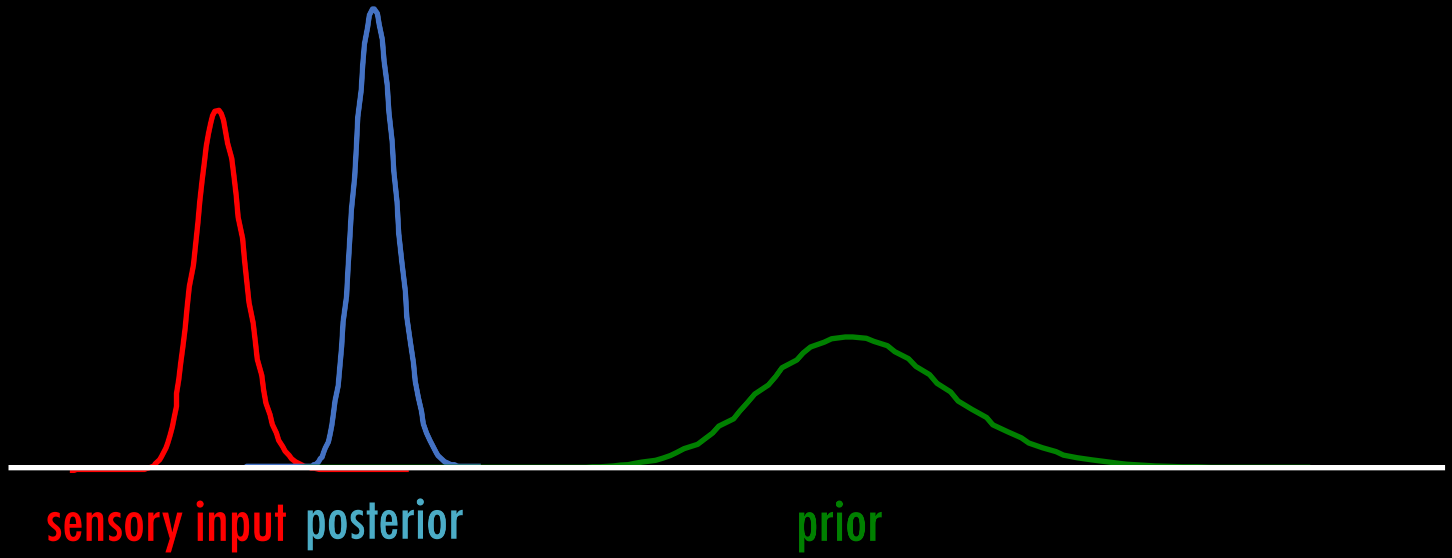



prior
expectation
sensory data
prediction

The ideal Nadal



$$P(world|sensory\ data) = \frac{P(sensory\ data|world)\ P(world)}{P(sensory\ data)}$$



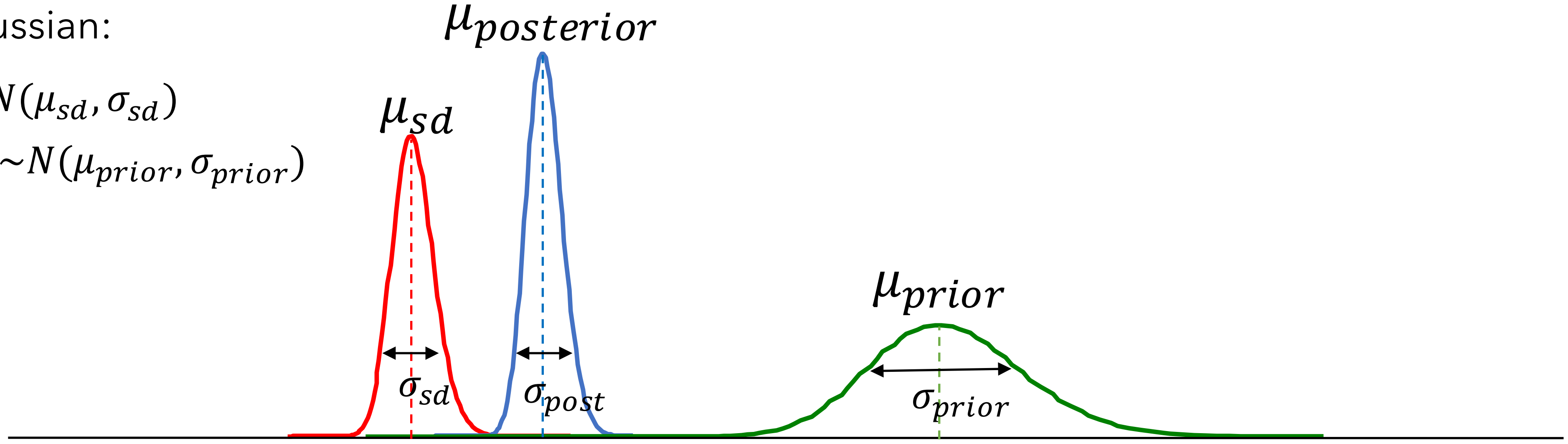
- The optimal combination takes uncertainty into account

$$P(world|sensory\ data) = \frac{P(sensory\ data|world)\ P(world)}{P(sensory\ data)}$$

If Gaussian:

$P_{sd} \sim N(\mu_{sd}, \sigma_{sd})$

$P_{prior} \sim N(\mu_{prior}, \sigma_{prior})$



$$\mu_{posterior} = w_{sd} \cdot \mu_{sd} + w_{prior} \cdot \mu_{prior}$$

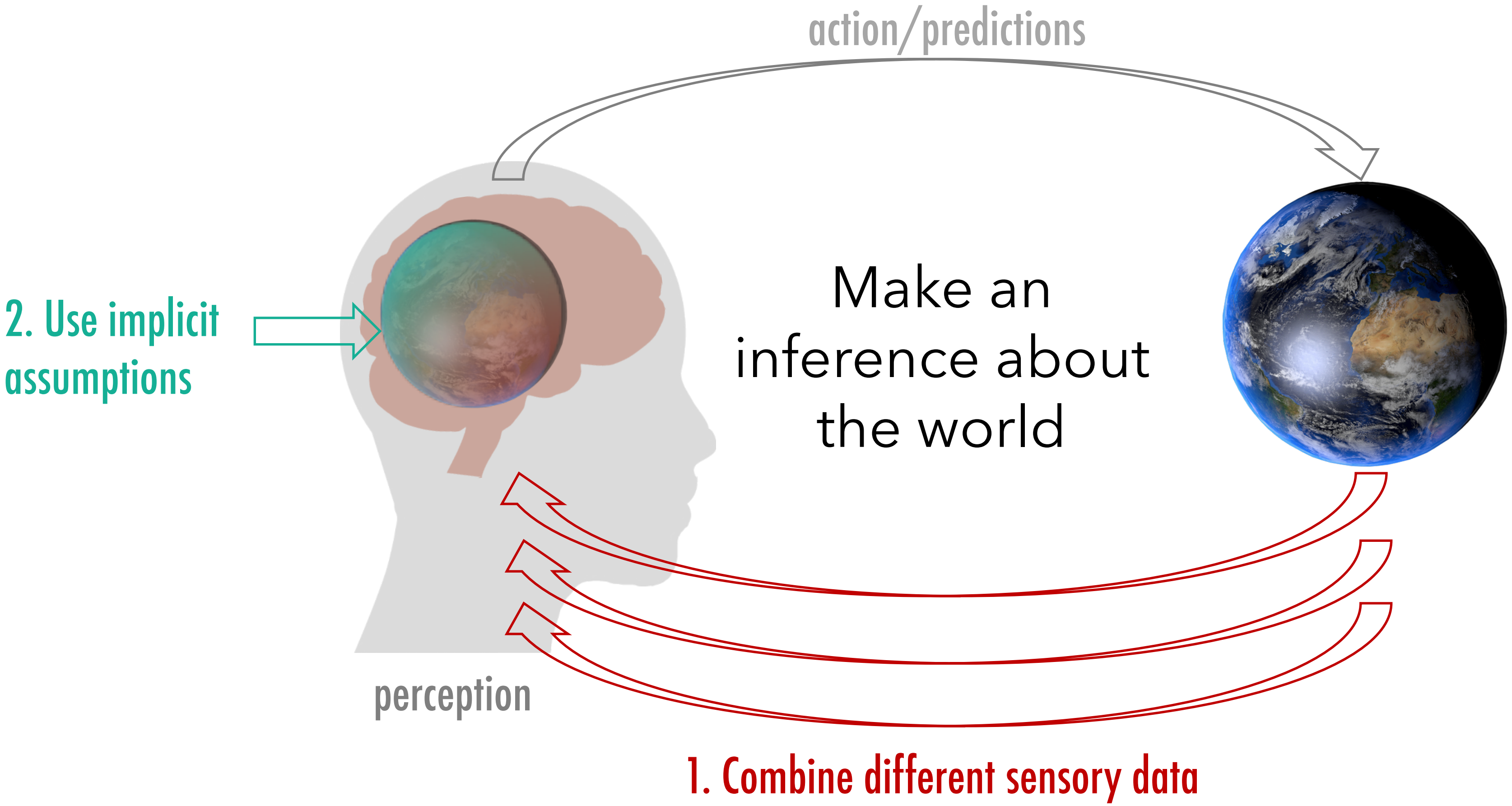
$$w_{sd} = \frac{\frac{1}{\sigma_{sd}^2}}{\frac{1}{\sigma_{sd}^2} + \frac{1}{\sigma_{prior}^2}}$$

$$\sigma_{posterior}^2 = \frac{\sigma_{sd}^2 \cdot \sigma_{prior}^2}{\sigma_{sd}^2 + \sigma_{prior}^2}$$

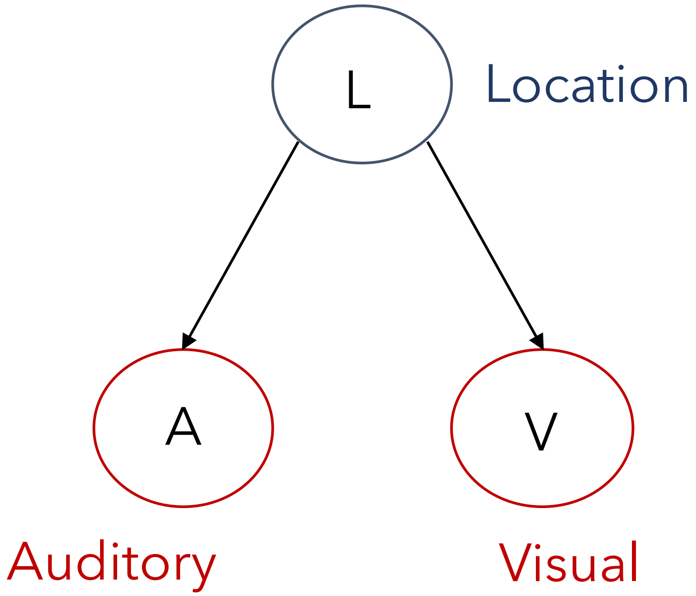
Why use Bayes to model perception?

It's a trick: Bayes' Rule is a statistical tool that describes how different types of information sources (different sensors and beliefs) can be combined in a statistical optimal manner.

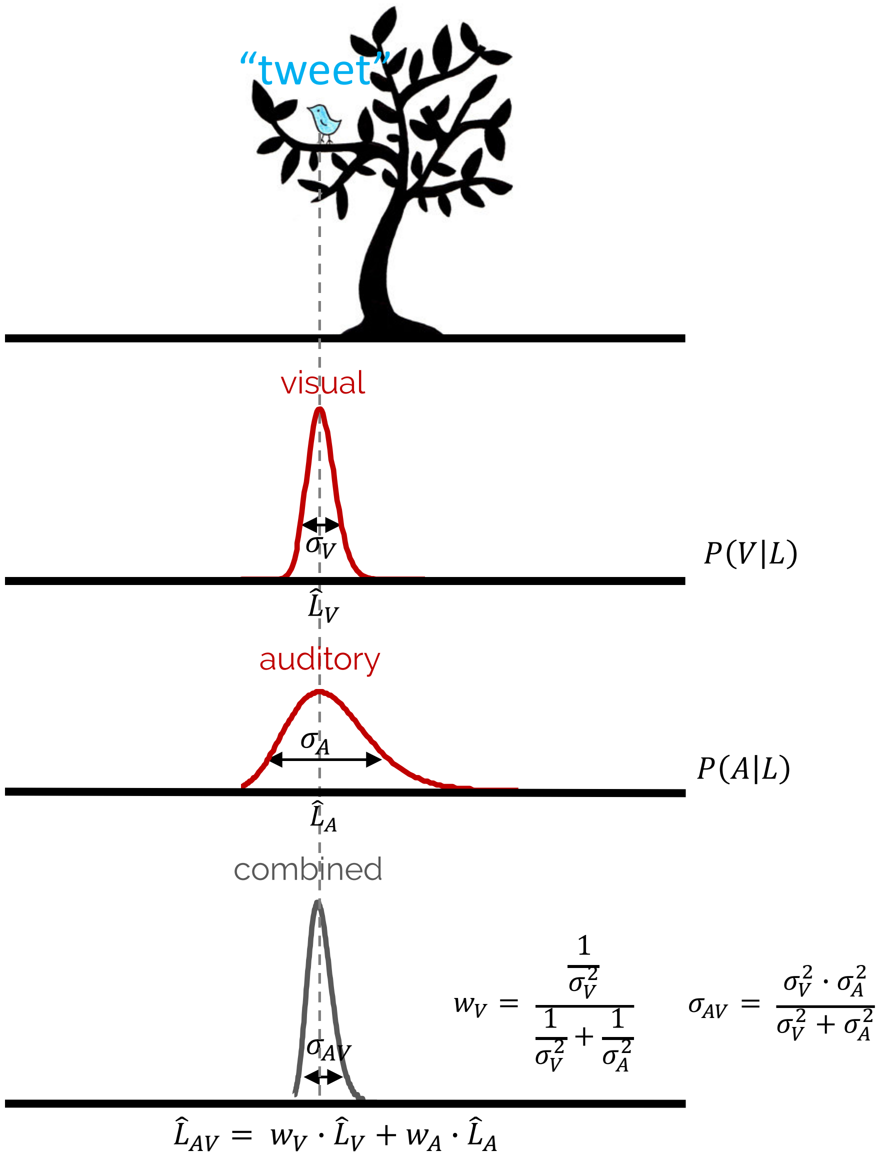
Do we behave like a ‘Bayesian’?

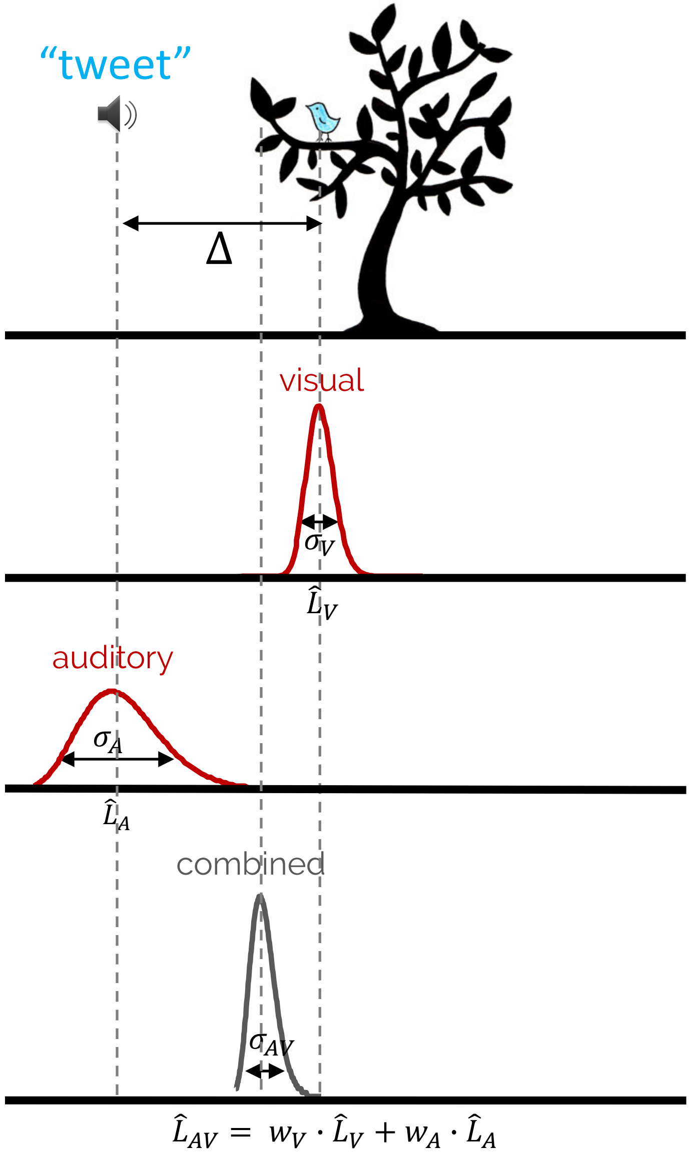

action/predictions
2. Use implicit assumptions
Make an inference about the world
perception
1. Combine different sensory data

Combination of visual and auditory information

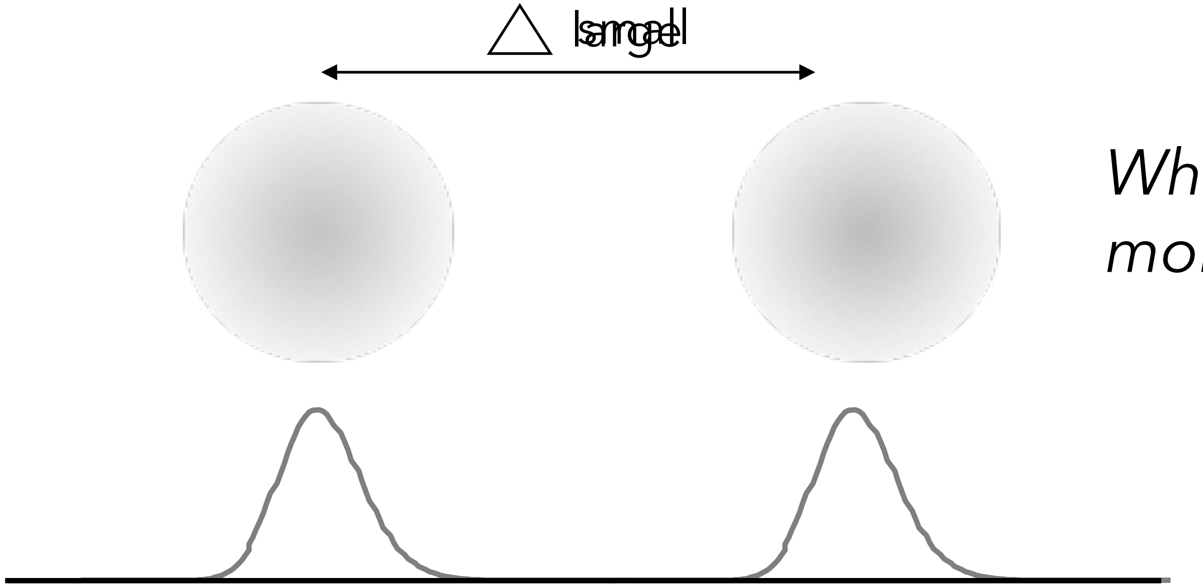


Combination of visual and auditory information: Ideal Observer

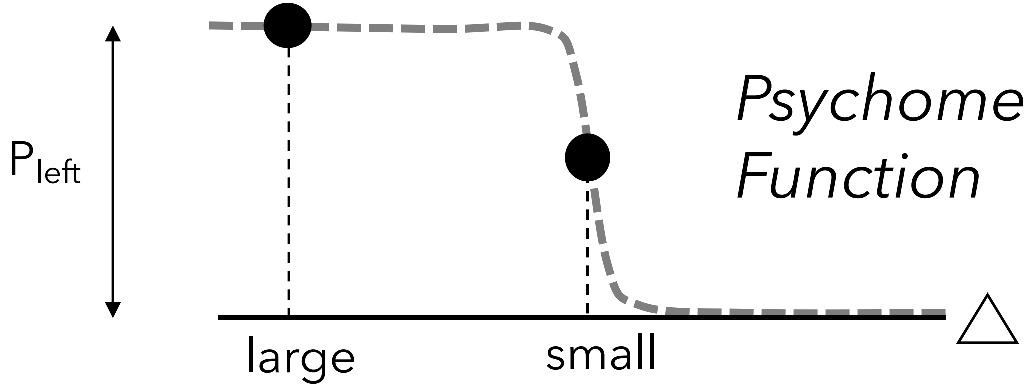




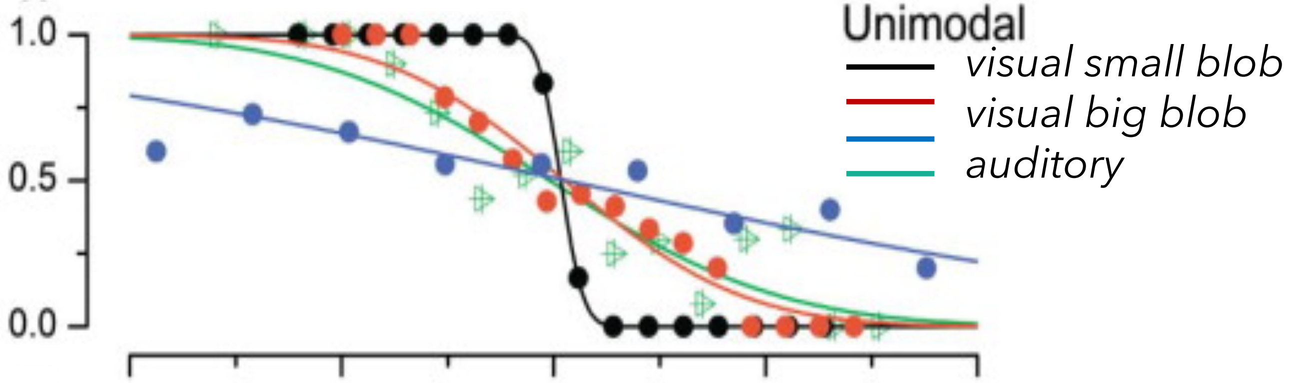
Measuring sensory uncertainty: The Psychometric Function



Which one was more left?

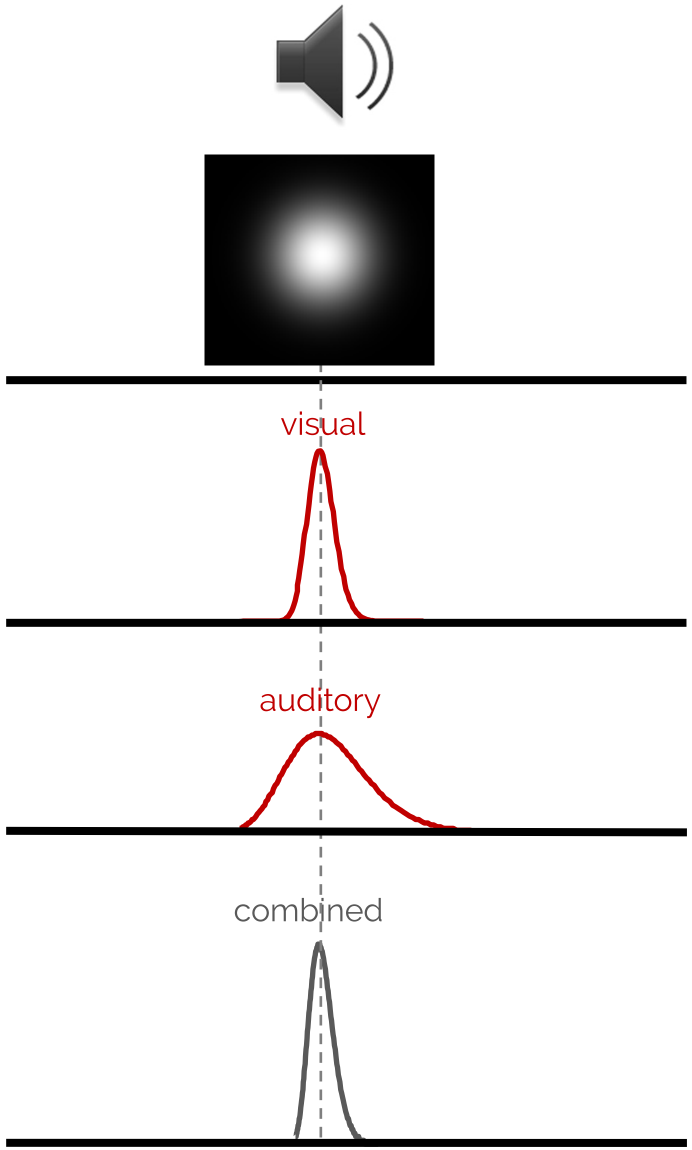


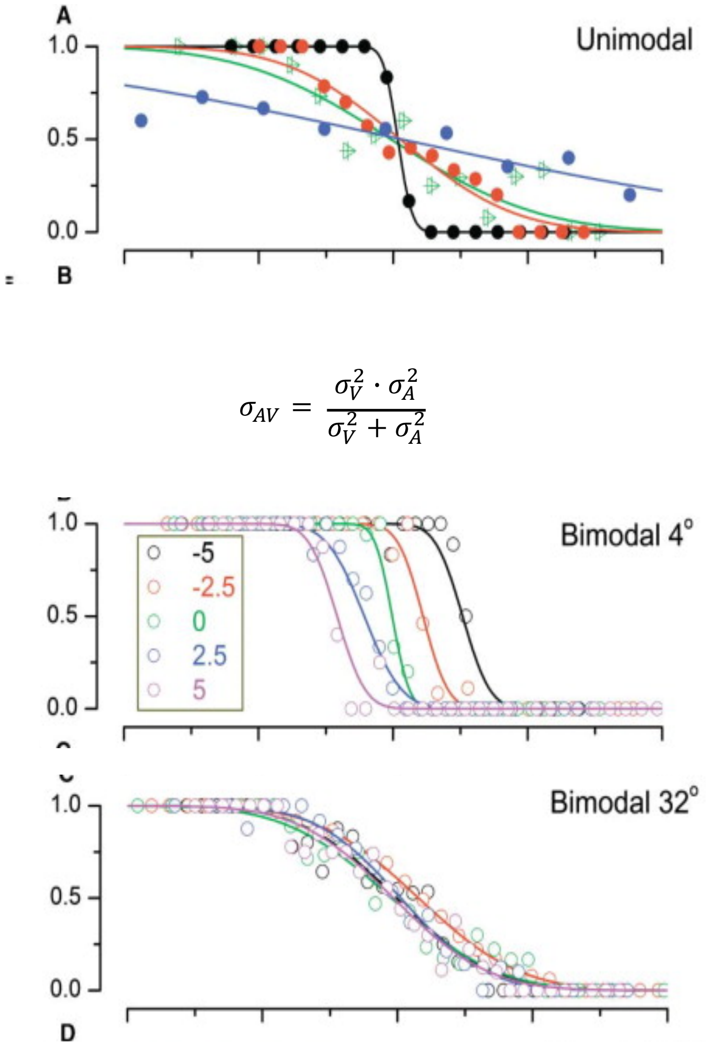
Psychometric Function



Alais & Burr, 2004

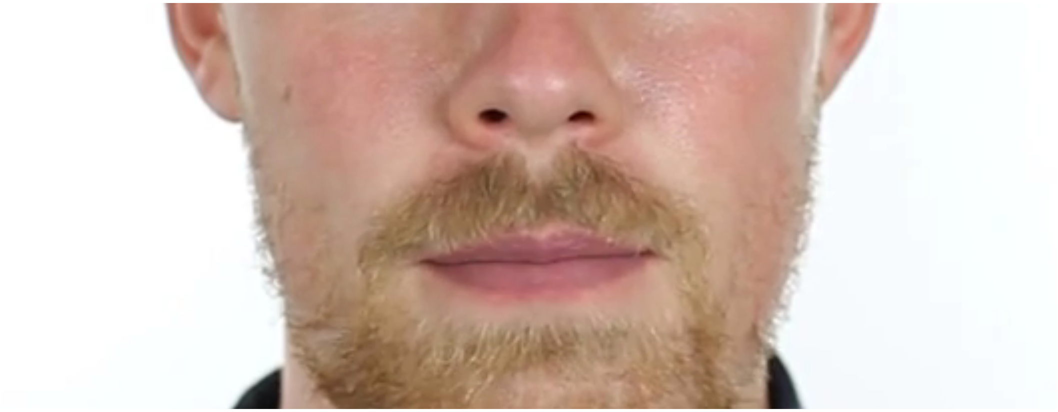
Combination of visual and auditory information: Human Observer





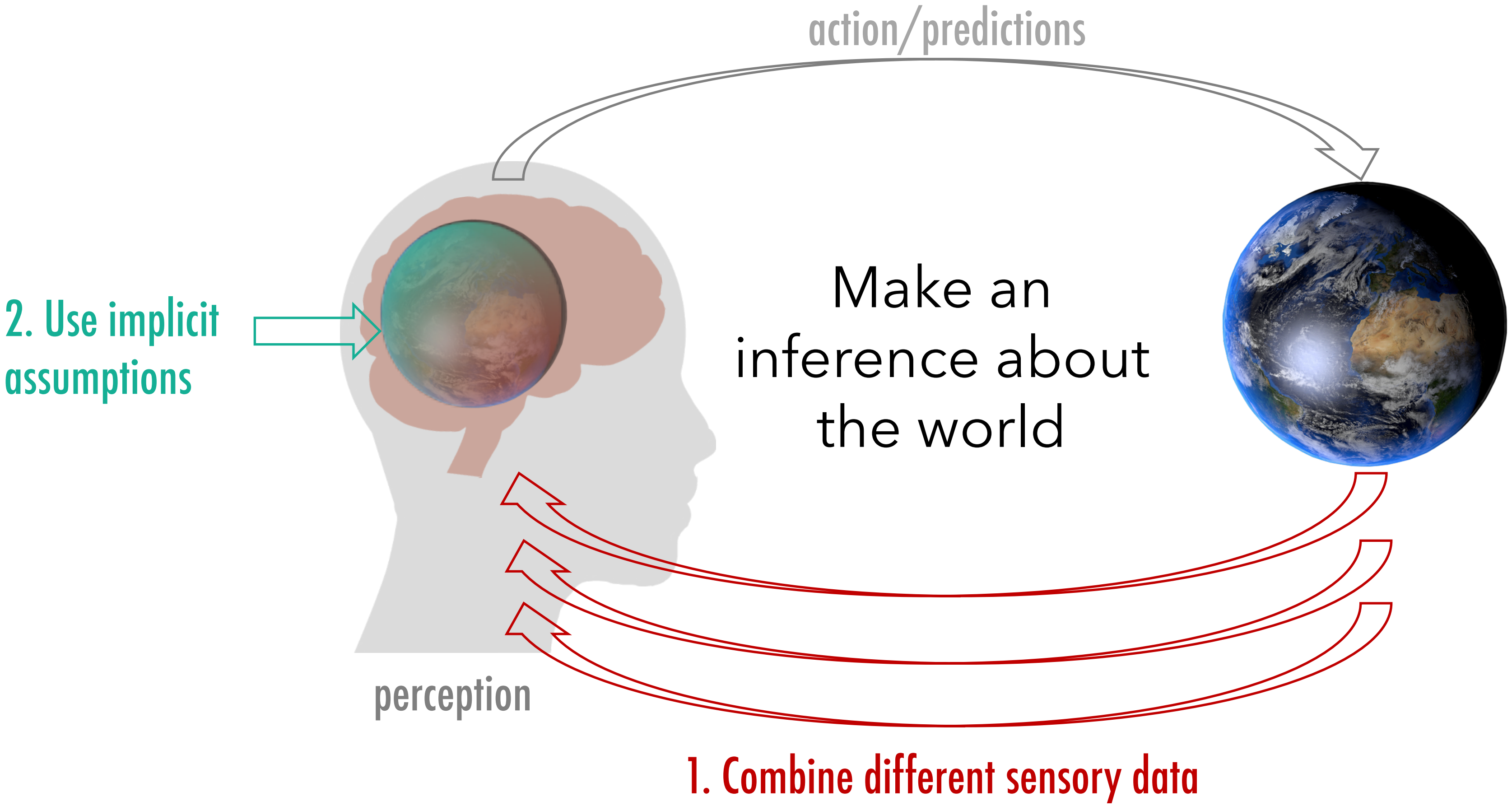
$$\sigma_{AV} = \frac{\sigma_V^2 \cdot \sigma_A^2}{\sigma_V^2 + \sigma_A^2}$$

Alais & Burr, 2004

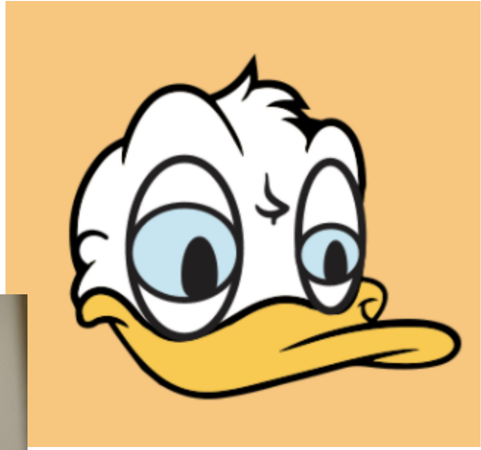


BAR

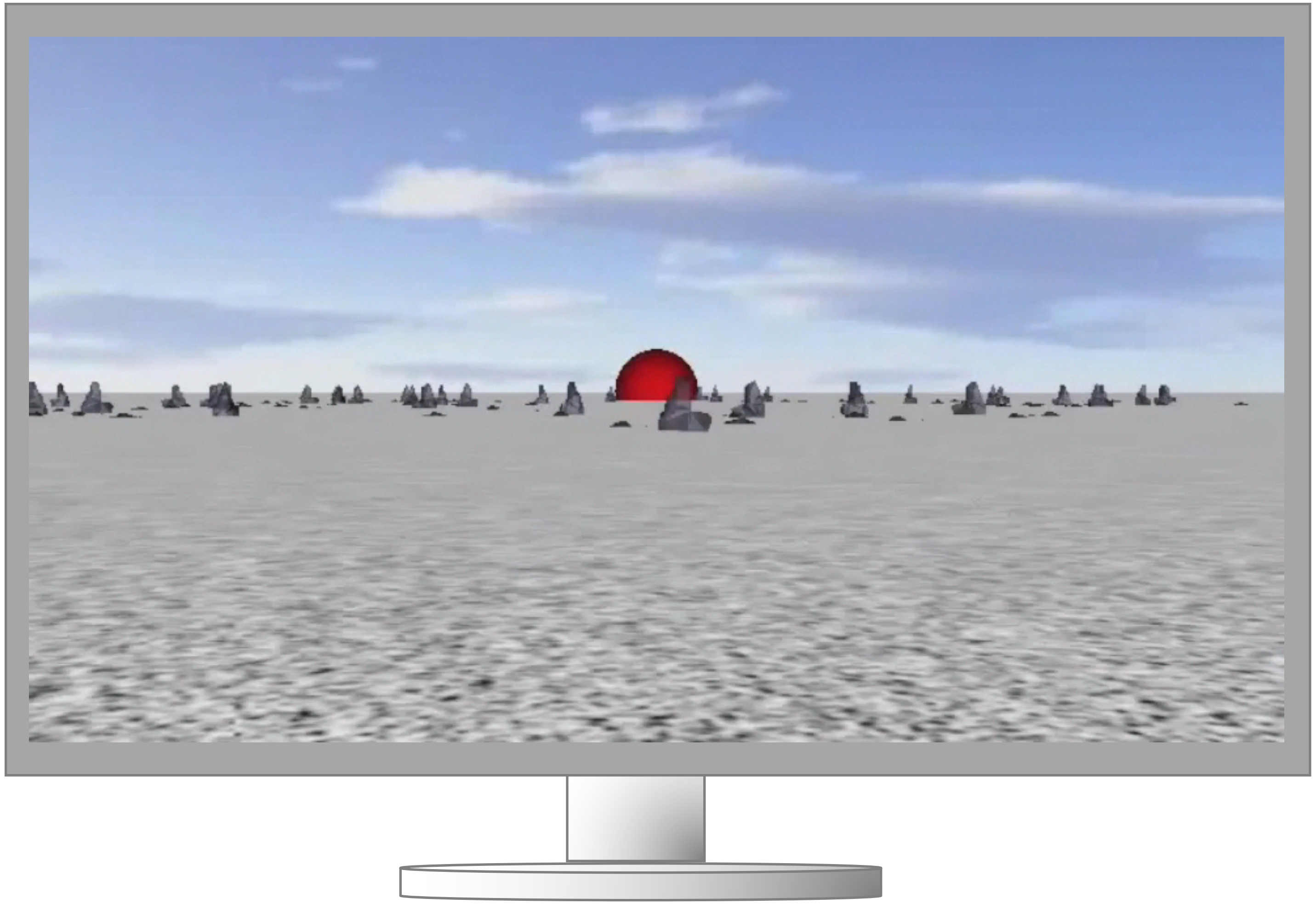
FAR


action/predictions
2. Use implicit assumptions
Make an inference about the world
perception
1. Combine different sensory data

Priors can be learned...

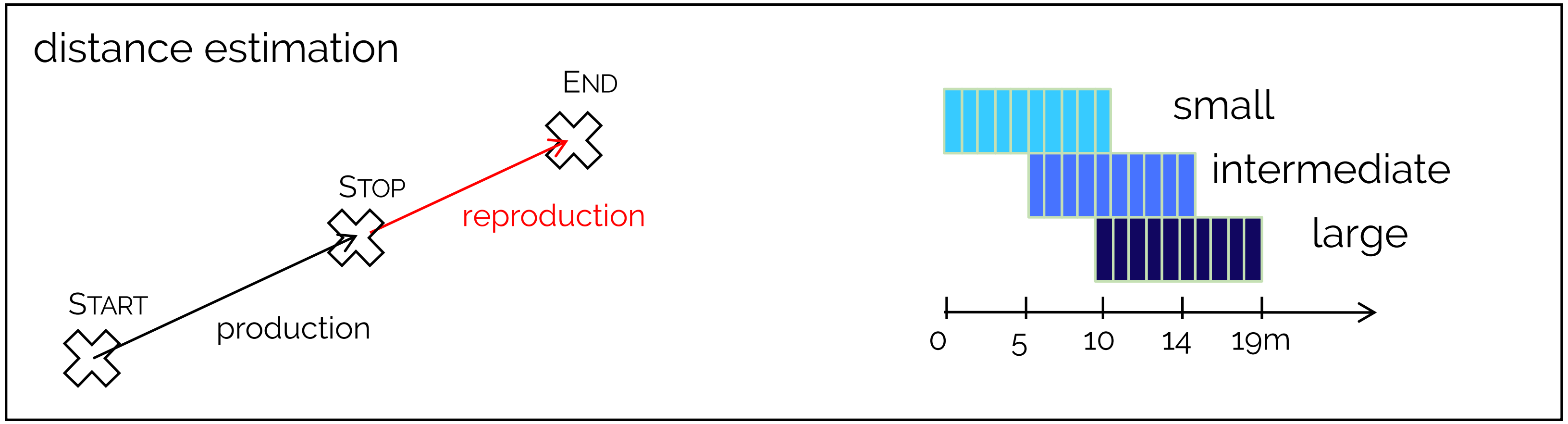


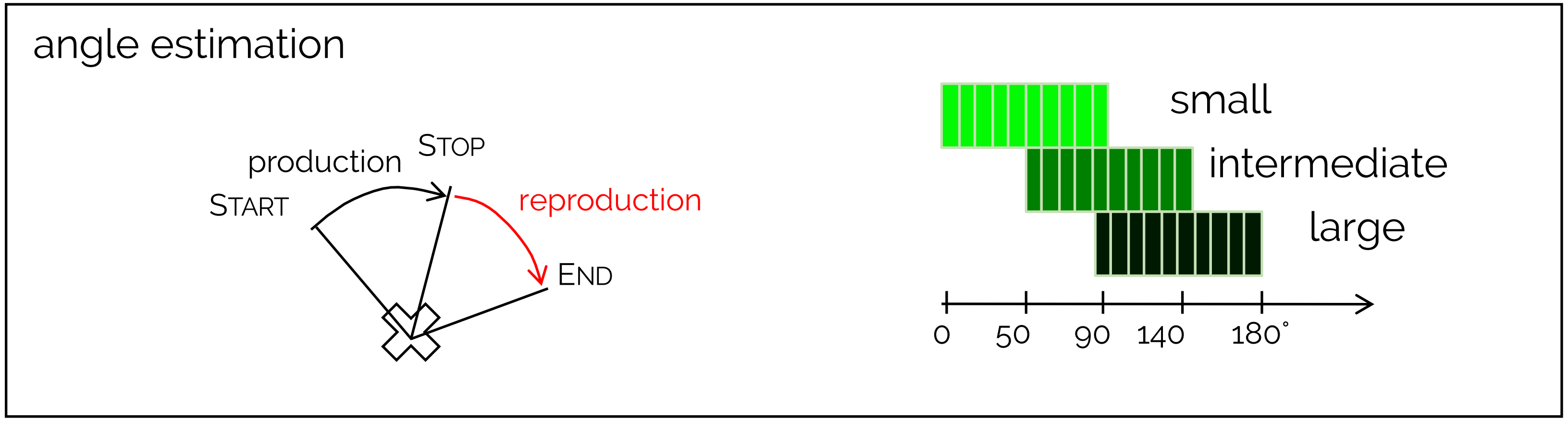
“optimal errors” in magnitude estimation



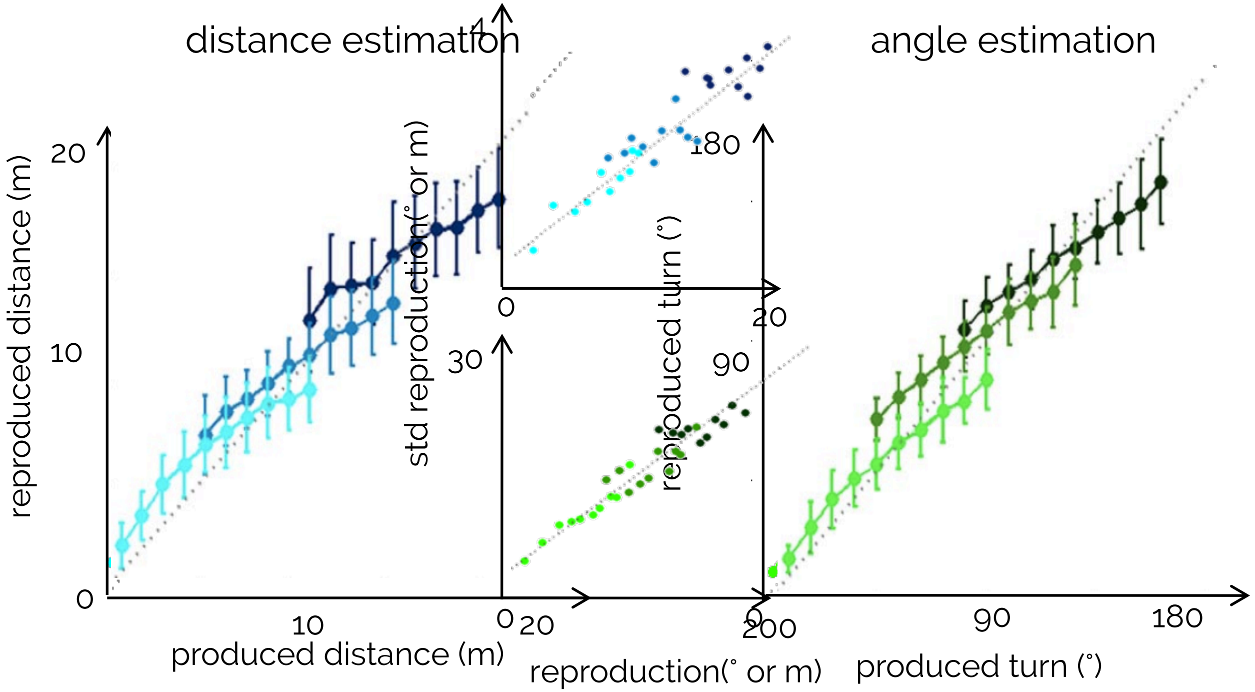
Petzschner & Glasauer, JoN, 2011

Varying the sample range

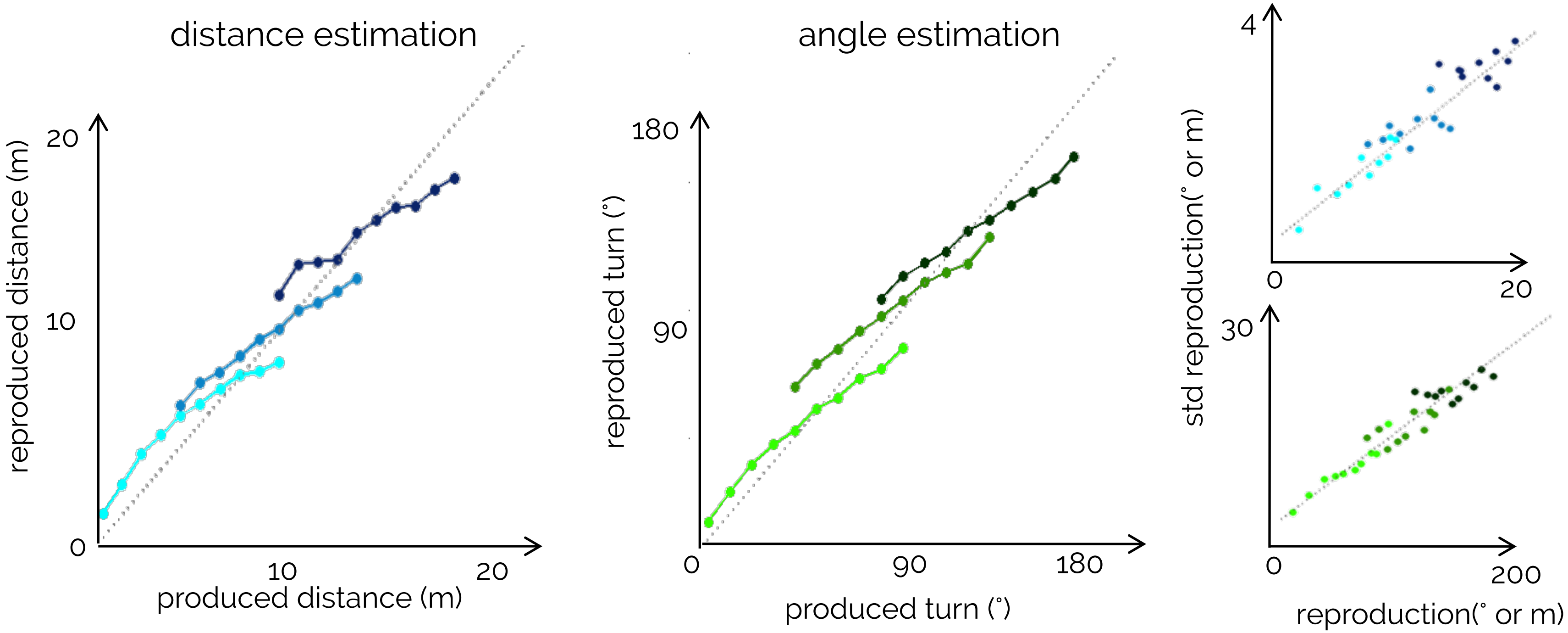




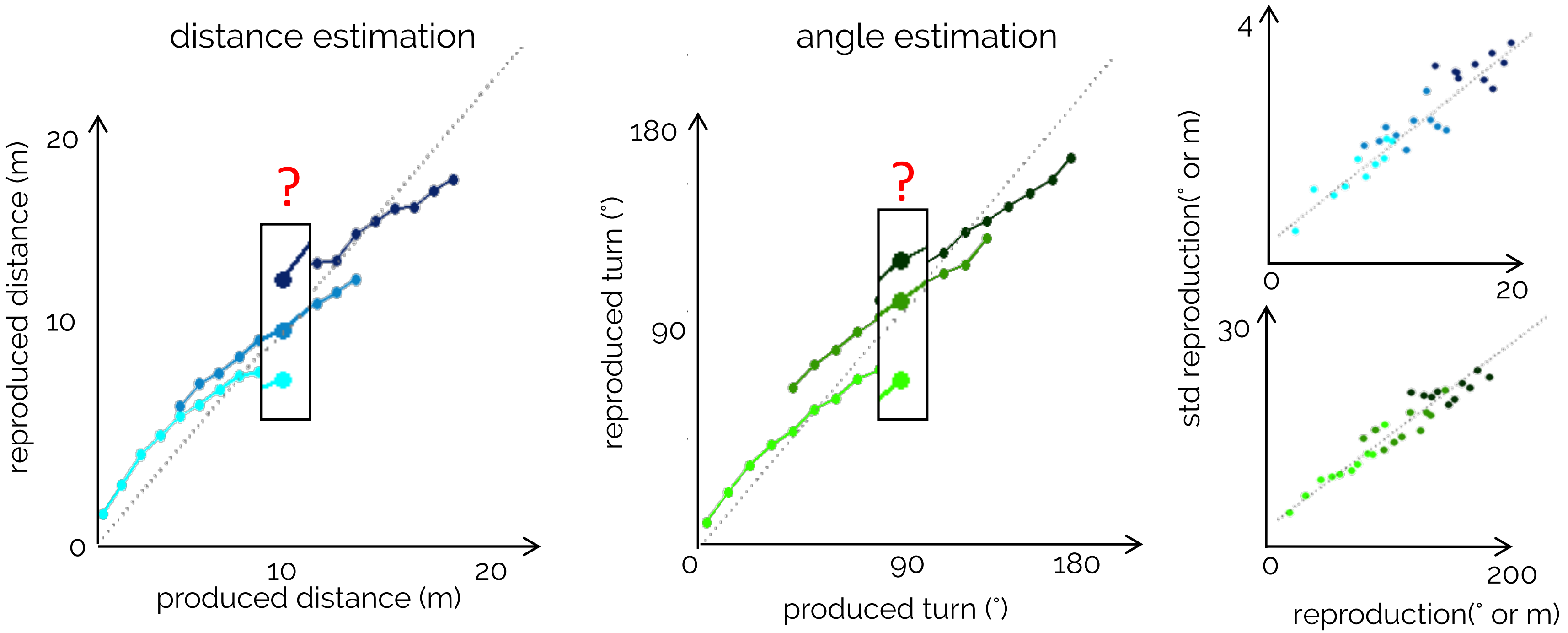
Petzschner & Glasauer, JoN, 2011



Petzschner & Glasauer, JoN, 2011



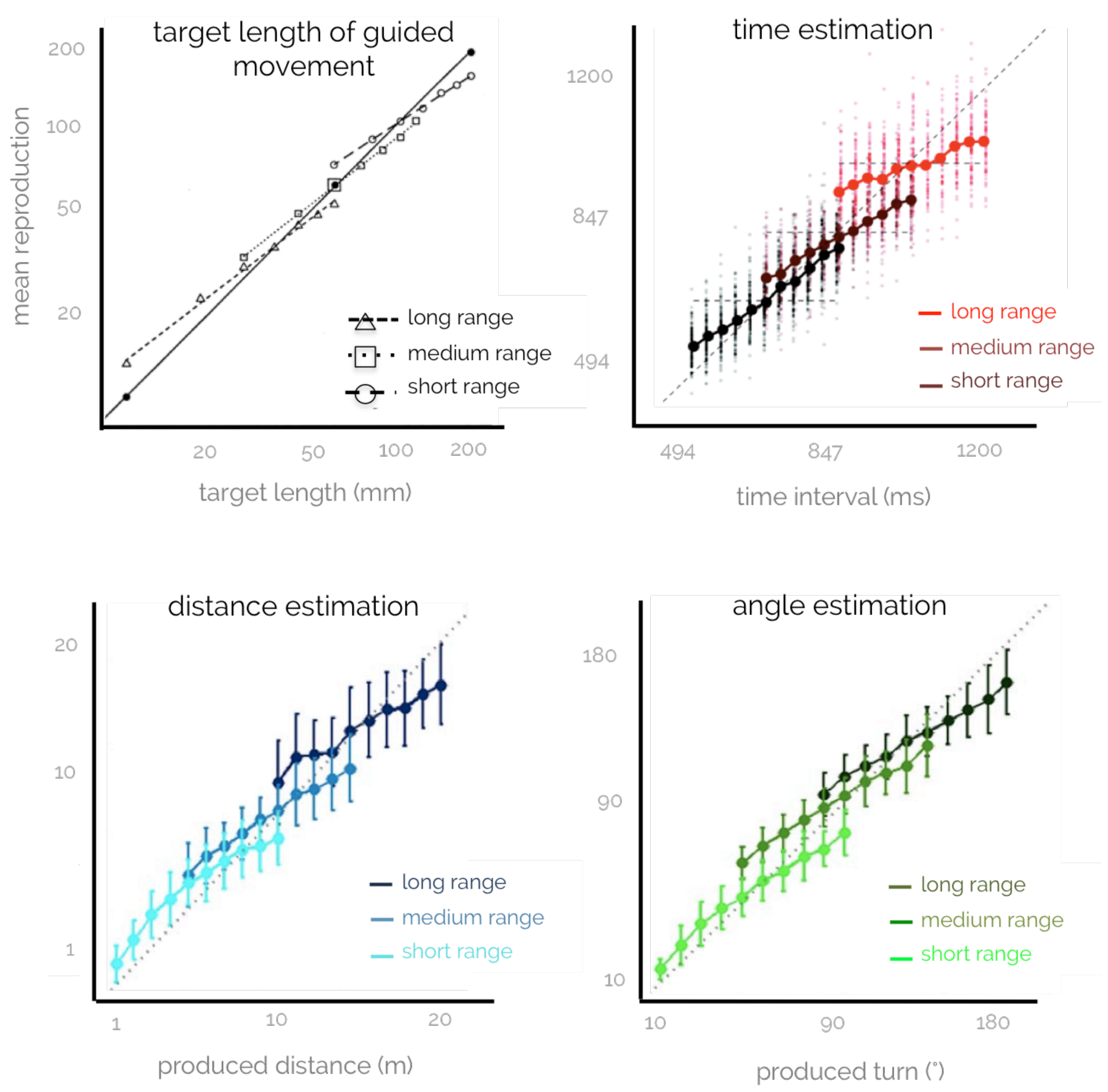
Petzschner & Glasauer, JoN, 2011



Prior knowledge: Experience

Petzschner & Glasauer, JoN, 2011

Let's take a look at the literature

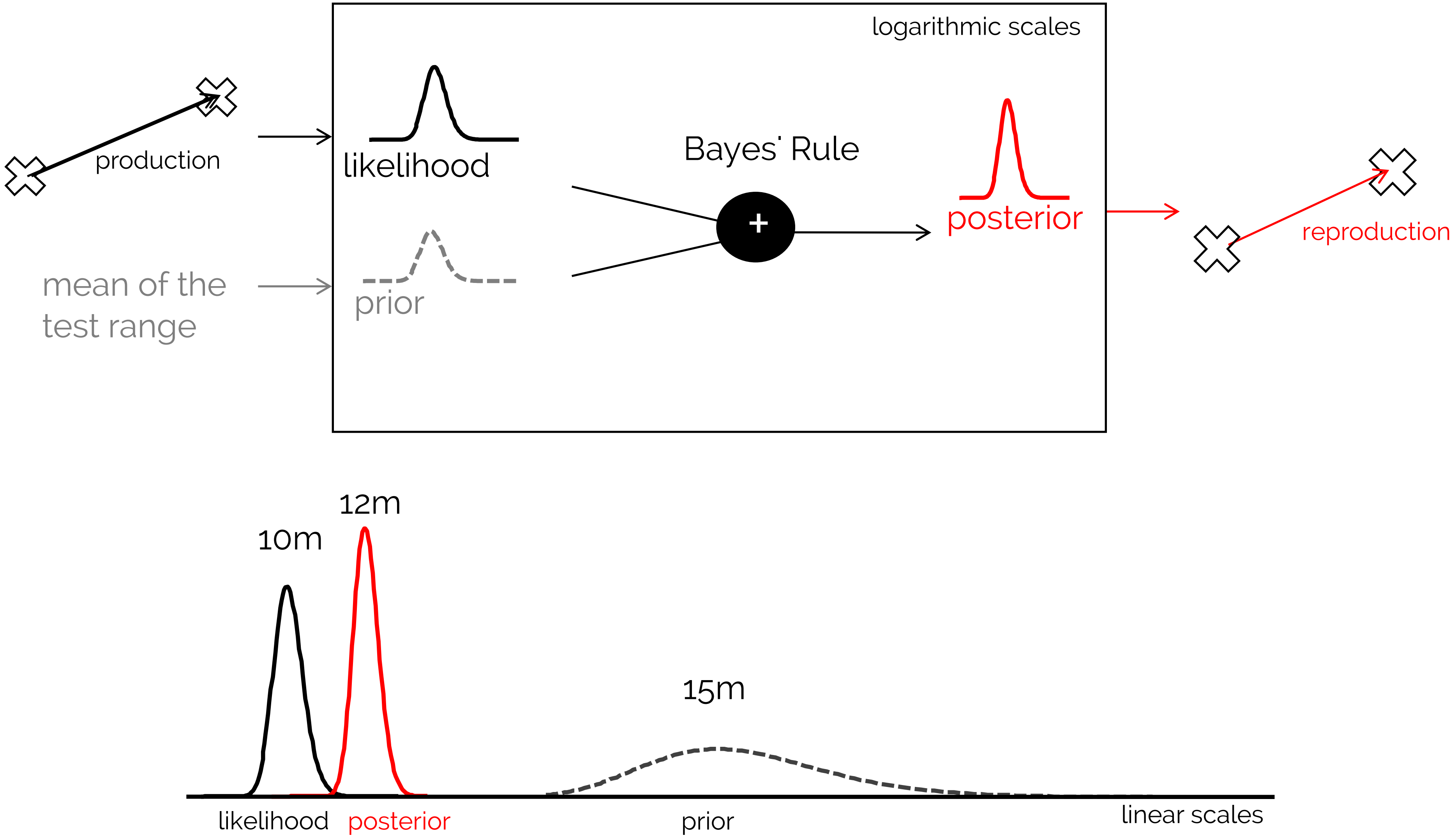


Petzschner et al, TiCS, 2015; Petzschner & Glasauer, JoN, 2011

A Bayesian Model for magnitude estimation

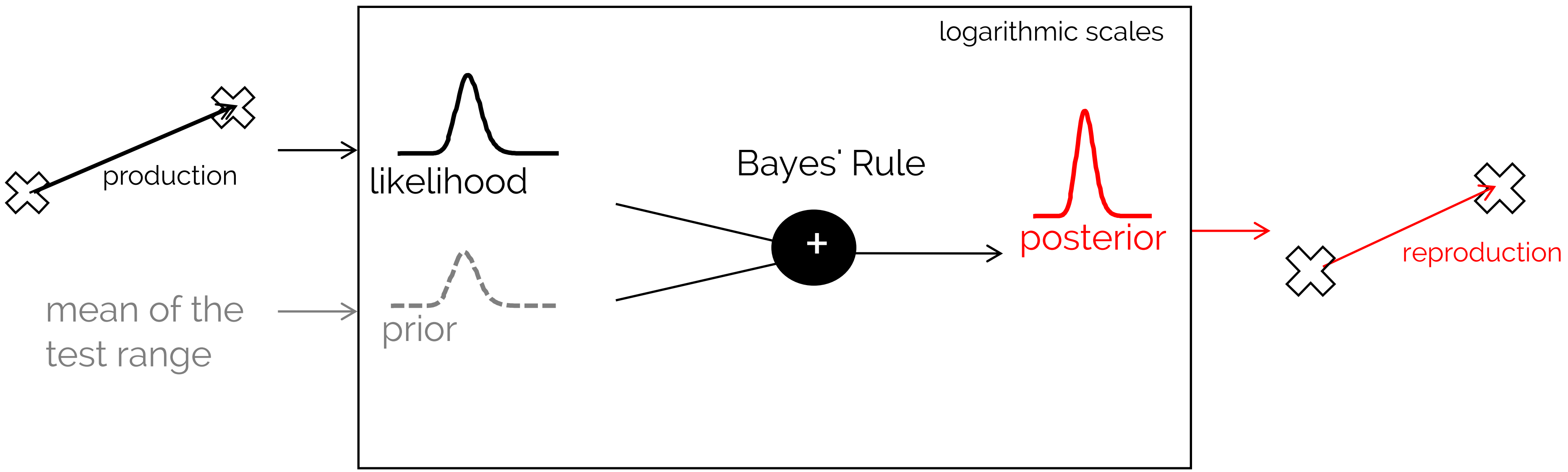
Petzschner & Glasauer, JoN, 2011; Petzschner et al, TiCS, 2015

A Bayesian Model for magnitude estimation



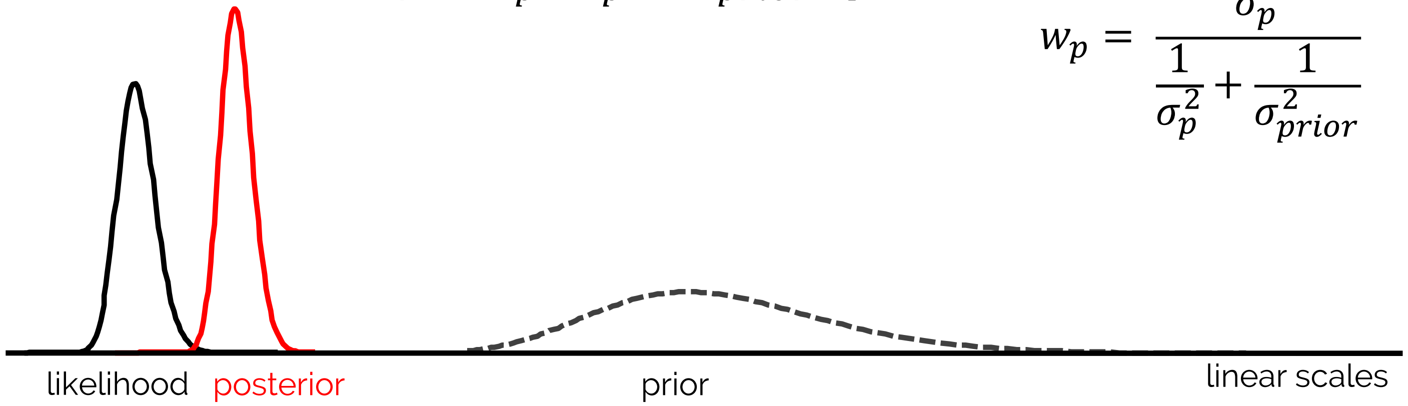
Petzschner & Glasauer, JoN, 2011; Petzschner et al, TiCS, 2015

A Bayesian Model for magnitude estimation



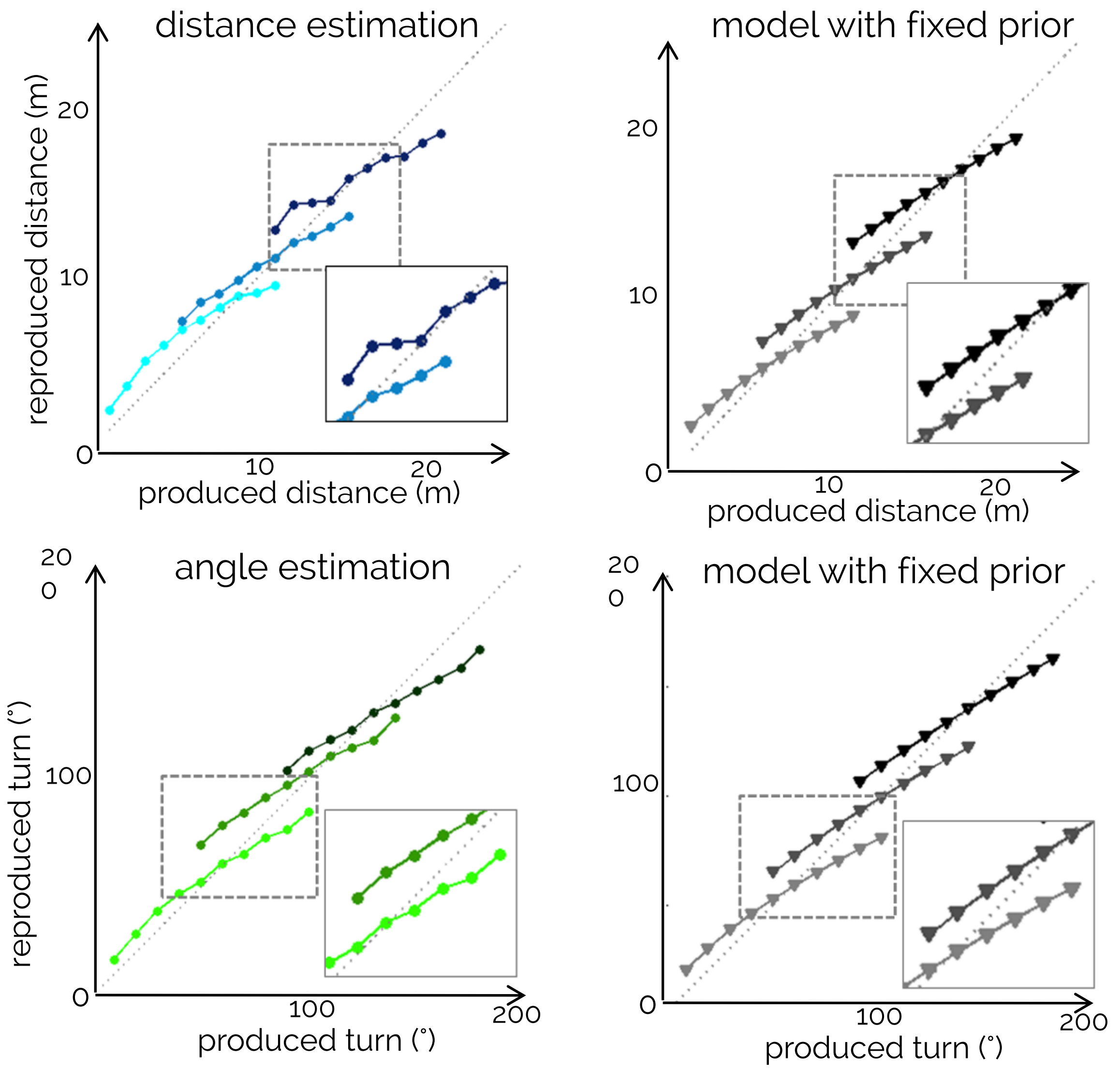
$$\widehat{d_r} = w_p \cdot d_p + w_{prior} \cdot prior$$

$$w_p = \frac{\frac{1}{\sigma_p^2}}{\frac{1}{\sigma_p^2} + \frac{1}{\sigma_{prior}^2}}$$



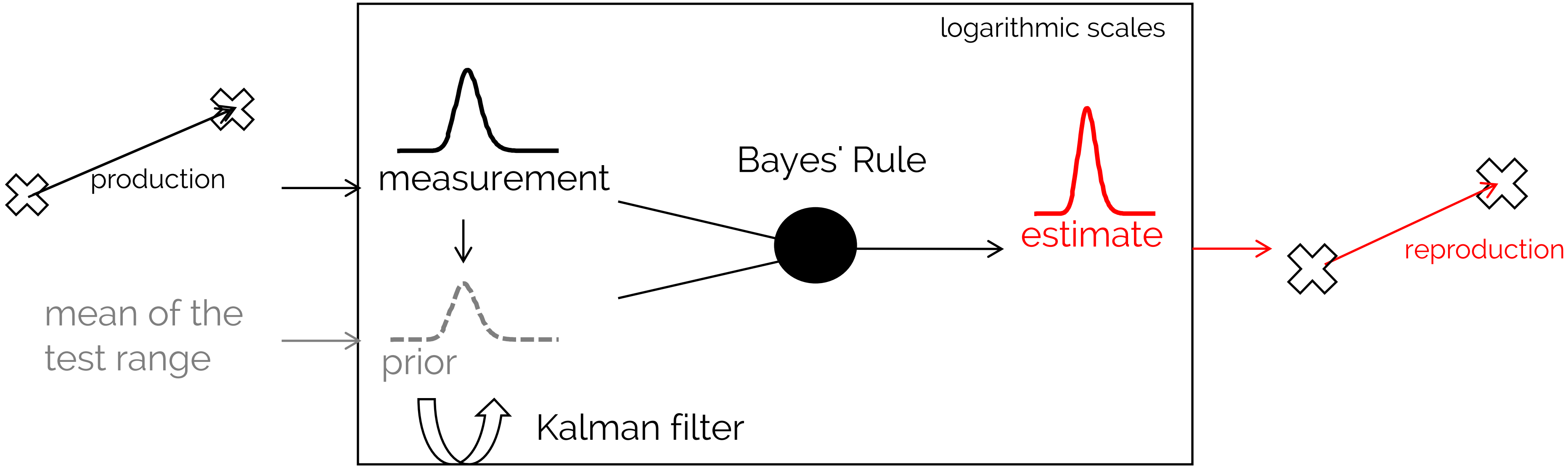
Petzschner & Glasauer, JoN, 2011; Petzschner et al, TiCS, 2015

Quantitative results



Petzschner & Glasauer, JoN, 2011; Petzschner et al, TiCS, 2015

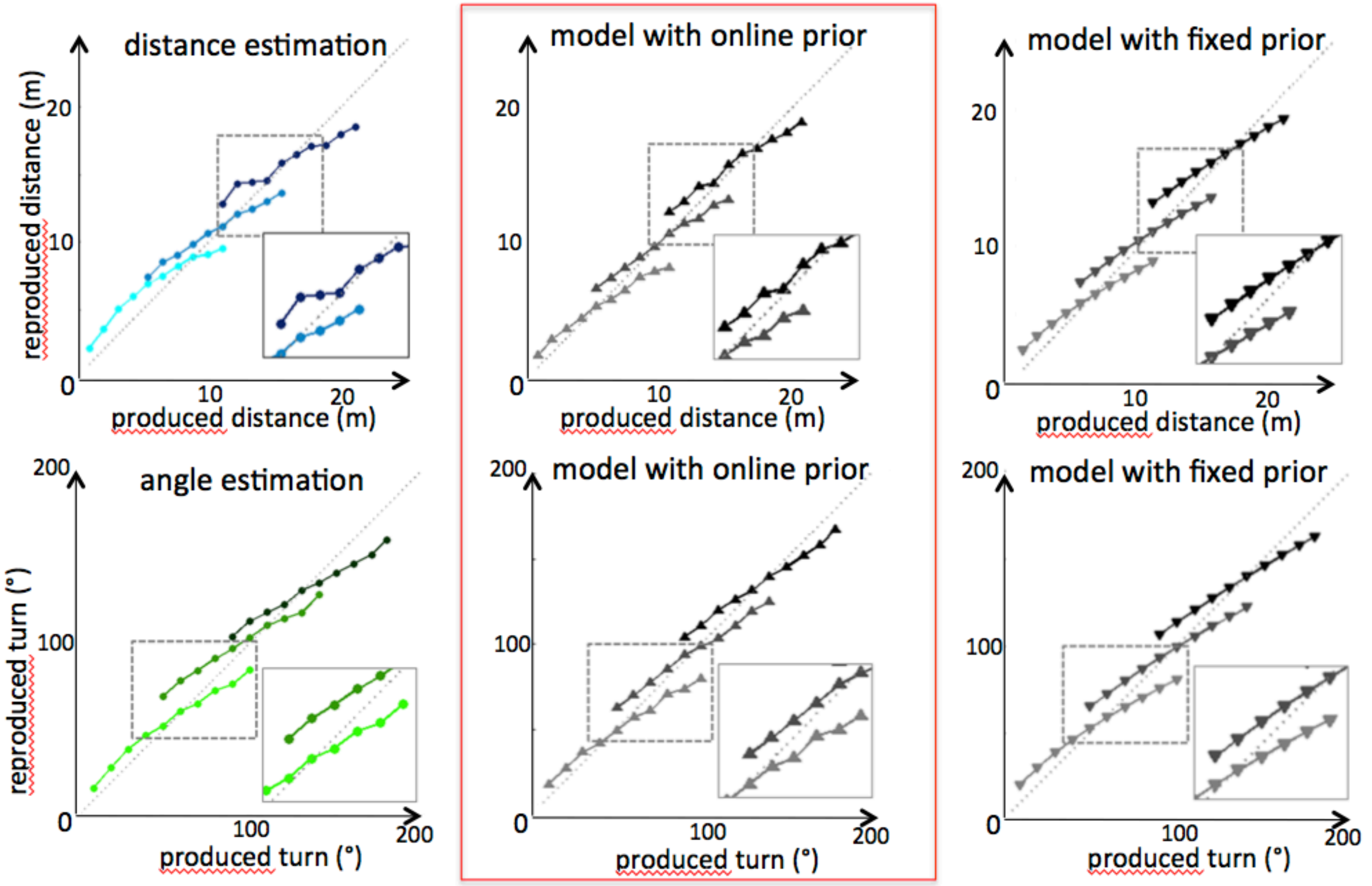
Bayesian Learning



"Todays posterior is tomorrows prior."

Petzschner & Glasauer, JoN, 2011; Petzschner et al, TiCS, 2015

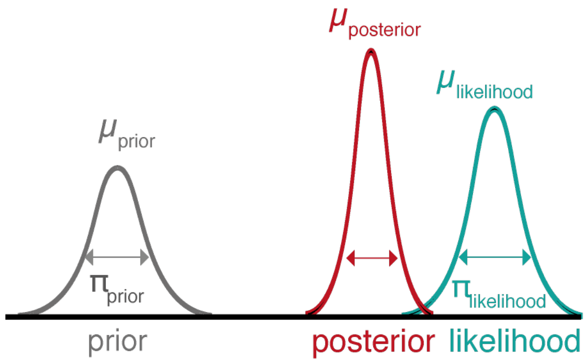
Kalman Filter



Petzschner & Glasauer, JoN, 2011; Petzschner et al, TiCS, 2015

Hierarchical Bayesian Learning

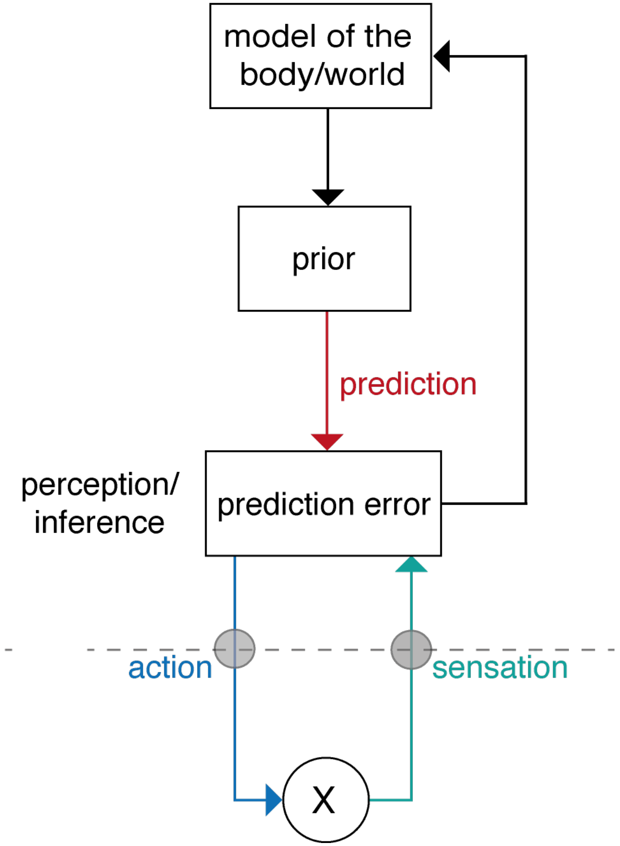
Hierarchical Gaussian Filter
Predictive Coding



Belief update

Δbelief ~precision · PE

$$\mu_{posterior} = \mu_{prior} + \frac{\pi_{likelihood}}{\pi_{likelihood} + \pi_{prior}} \cdot PE$$



→ *See the next two talks!*

Petzschner et al., Computational Psychosomatics and Computational Psychiatry: Towards a joint framework for differential diagnosis. Biol Psychiatry, 2017 (in press)

Applications of the Bayesian Models to Human Behavior

[Friston and Stephan, 2007; Knill and Pouget, 2004; Knill and Richards, 1996].

Magnitude Estimation [Shadlen, Kiani, Glasauer, Petzschner ...]

Visual perception [Weiss, Simoncelli, Adelson, Richards, Freeman, Feldman, Kersten, Knill, Maloney, Olshausen, Jacobs, Pouget, ...]

Language acquisition and processing [Brent, de Marken, Niyogi, Klein, Manning, Jurafsky, Keller, Levy, Hale, Johnson, Griffiths, Perfors, Tenenbaum, …]

Motor learning and motor control [Ghahramani, Jordan, Wolpert, Kording, Kawato, Doya, Todorov, Shadmehr, …]

Associative learning [Dayan, Daw, Kakade, Courville, Touretzky, Kruschke, …]

Memory [Anderson, Schooler, Shiffrin, Steyvers, Griffiths, McClelland, …]

Attention [Mozer, Huber, Torralba, Oliva, Geisler, Yu, Itti, Baldi, …]

Categorization and concept learning [Anderson, Nosfosky, Rehder, Navarro, Griffiths, Feldman, Tenenbaum, Rosseel, Goodman, Kemp, Mansinghka, …]

Reasoning [Chater, Oaksford, Sloman, McKenzie, Heit, Tenenbaum, Kemp, …]

Causal inference [Waldmann, Sloman, Steyvers, Griffiths, Tenenbaum, Yuille, …]

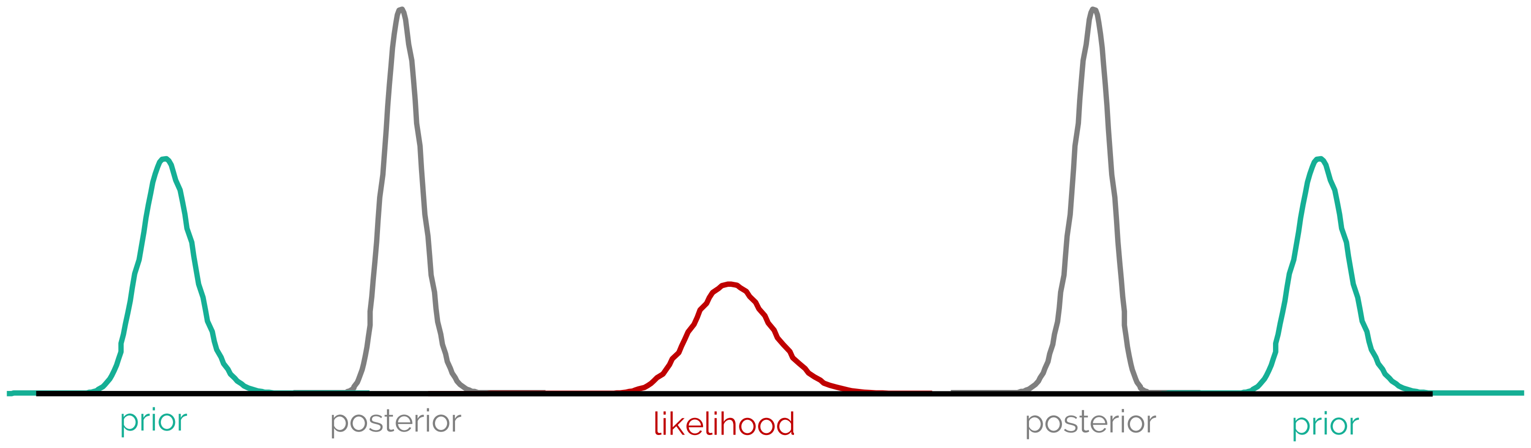
Decision making and theory of mind [Lee, Stankiewicz, Rao, Baker, Goodman, Tenenbaum, …]

Optimal motor control [Wolpert, Körding …]

Critique – Priors everywhere

[Bowers and Davis, 2012a,b; Griffiths et al., 2012, Colombo and Series, 2012; Jones and Love, 2011]

‘there are too many arbitrary ways that priors, likelihoods, utility functions, etc., can be altered in a Bayesian theory post hoc’.



Are humans 'Bayesian'?

These studies motivated conclusions that 'human perception is close to the Bayesian optimal suggesting the Bayesian process may be a fundamental element of sensory processing' [Körding and Wolpert, 2006] or, analogously, that there are myriad ways in which human observers behave as optimal Bayesian observers' [Knill and Pouget, 2004].

Marr's Three Levels of Analysis

[Bowers and Davis, 2012a,b; Griffiths et al., 2012, Colombo and Series, 2012; Jones and Love, 2011]

- Computation:

  "What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?"

- Algorithm:

  Cognitive psychology

- Implementation:

  Neurobiology

All models are wrong but some are useful!

Bayesian models can be used to reveal deviations in the way patients process different types of information

- **priors (bad experiences)**
- **likelihood (sensory data)**
- **precision**

All models are wrong but some are useful:

Bayesian models can be used to reveal deviations in the way patients process different types of information

When the world becomes 'too real': a Bayesian explanation of autistic perception

Elizabeth Pellicano[1,3] and David Burr[2,3]

Understanding why patients with schizophrenia do not perceive the hollow-mask illusion using dynamic causal modelling

Danai Dima[a,b,*], Jonathan P. Roiser[c], Detlef E. Dietrich[a,b], Catharina Bonnemann[a], Heinrich Lanfermann[d], Hinderk M. Emrich[a,b], Wolfgang Dillo[a]

No rapid audiovisual recalibration in adults on the autism spectrum

Marco Turi[1], Themelis Karaminis[2], Elizabeth Pellicano[2,4] & David Burr[3,4]

Horga G, Schatz KC, Abi-Dargham A, Peterson BS. Deficits in predictive coding underlie hallucinations in schizophrenia. J Neurosci. 2014 Jun 11;34(24):8072-82.

Computational Psychosomatics and Computational Psychiatry: Toward a Joint Framework for Differential Diagnosis

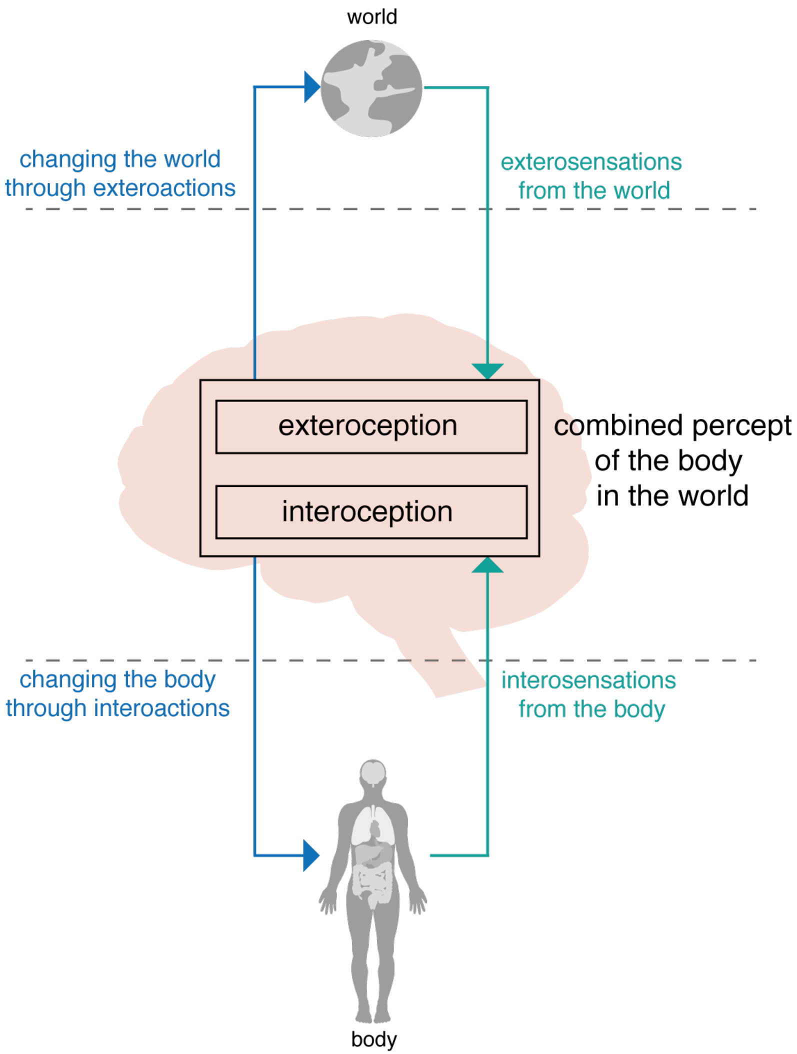
Frederike H. Petzschner, Lilian A.E. Weber, Tim Gard, and Klaas E. Stephan

Shift toward prior knowledge confers a perceptual advantage in early psychosis and psychosis-prone healthy individuals

Christoph Teufel[a,b,1], Naresh Subramaniam[b], Veronika Dobler[c,d], Jesus Perez[c,d], Johanna Finnemann[b,e], Puja R. Mehta[b], Ian M. Goodyer[c,d], and Paul C. Fletcher[b,d]

A. R. Powers, C. Mathys, P. R. Corlett. **Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors**. *Science*, August 2017

Perceiving your body



Petzschner et al., BiolPsych, 2017

What is perception?
Why Bayes to model perception?
Do you we behave like a ‘Bayesian’?
Are humans Bayesian?