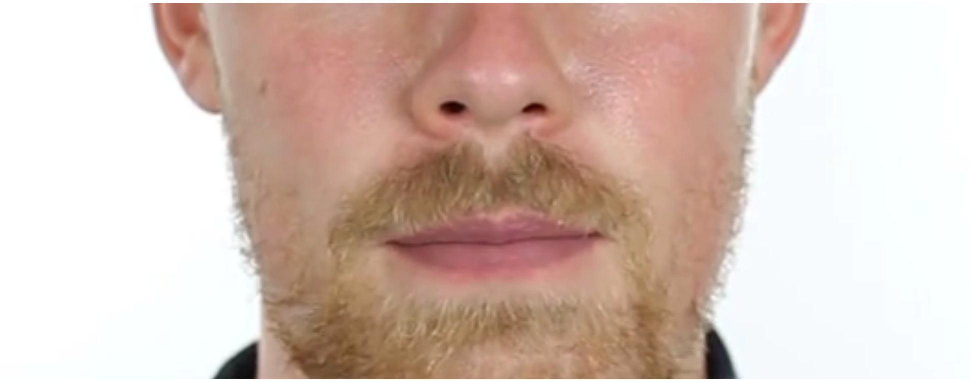
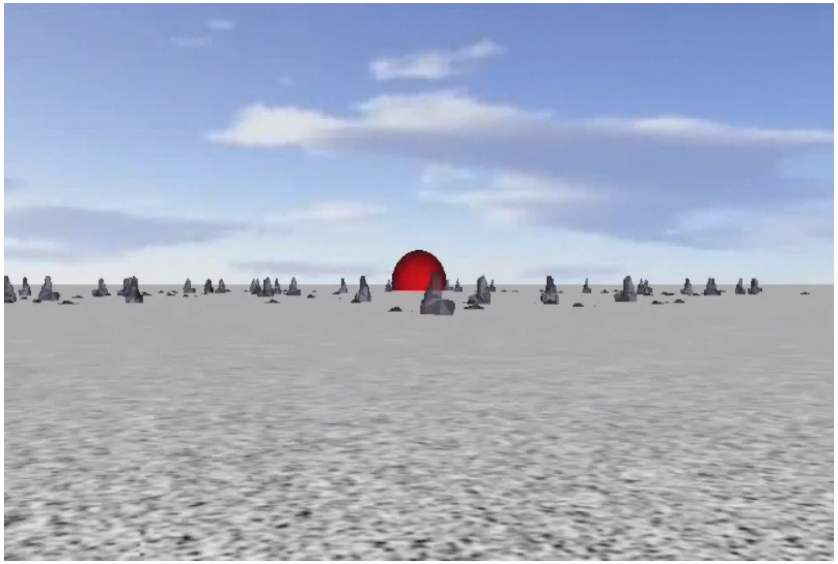
Bayesian Models for Perception

Dr. Frederike Petzschner

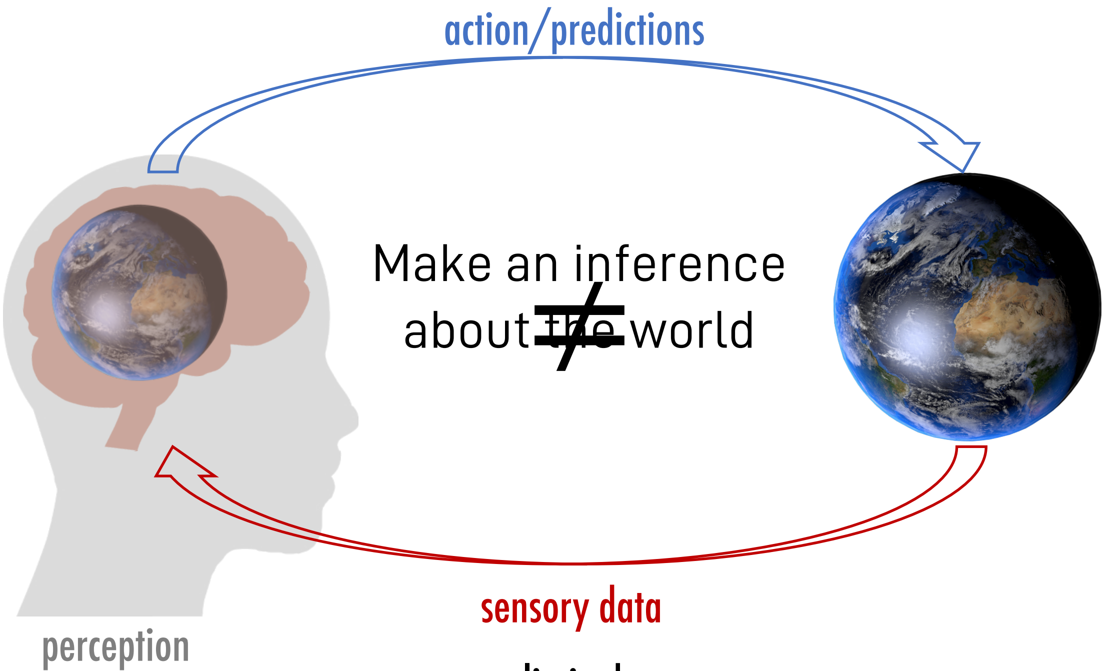




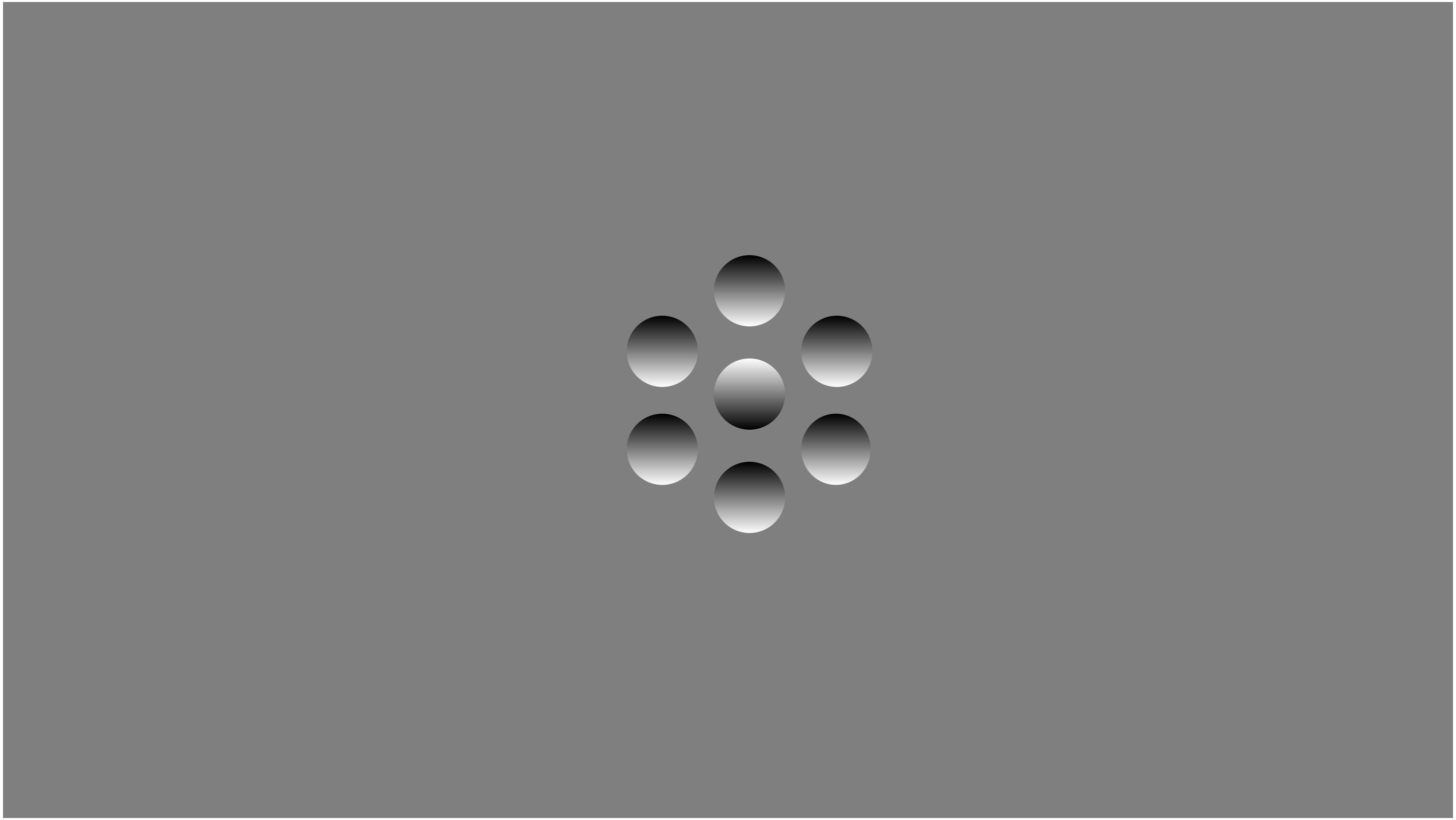
What is perception?
What is Bayes’ theorem?
Why Bayes to model perception?
Bayes Models for Behavior?
Bayes in the brain?

What is perception?

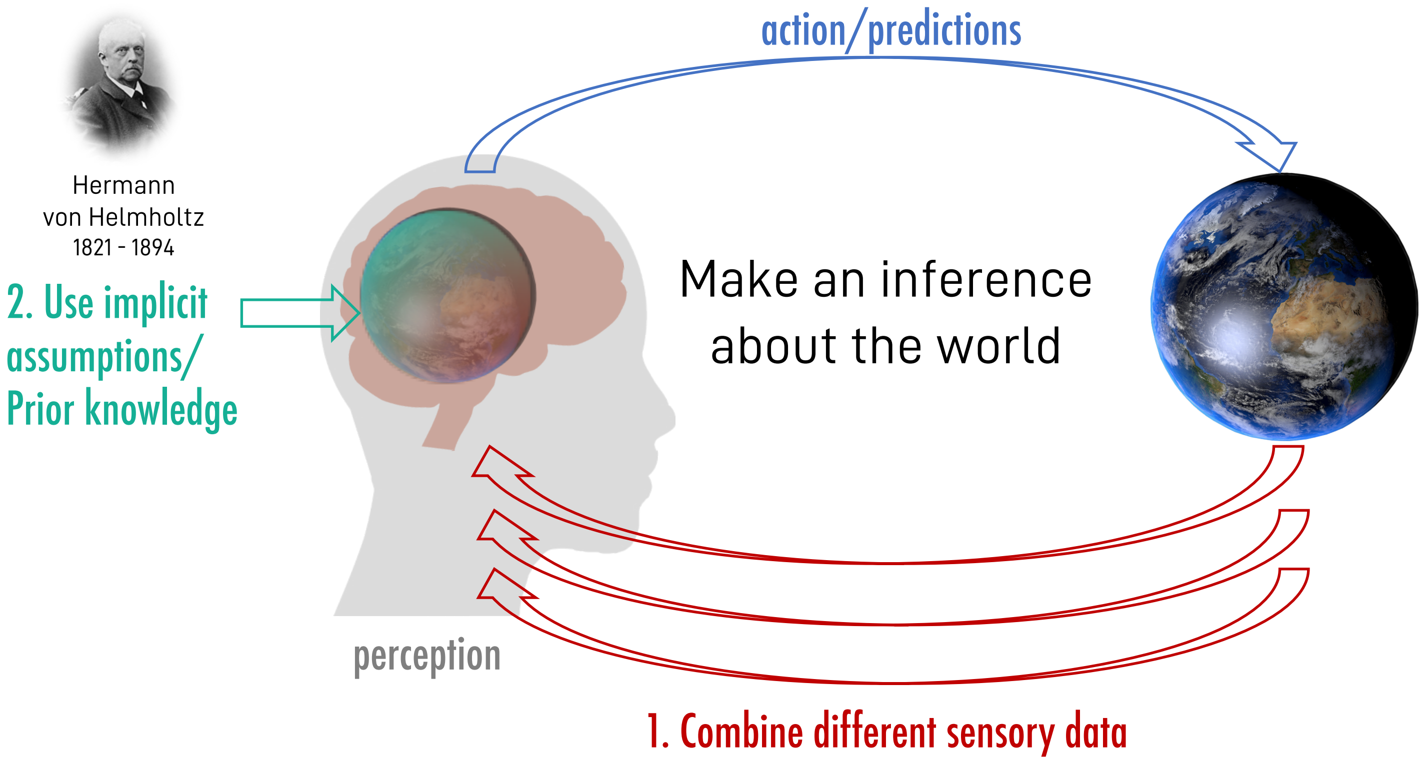
The brain’s dilemma:



limited
noisy
ambiguous
hidden structure



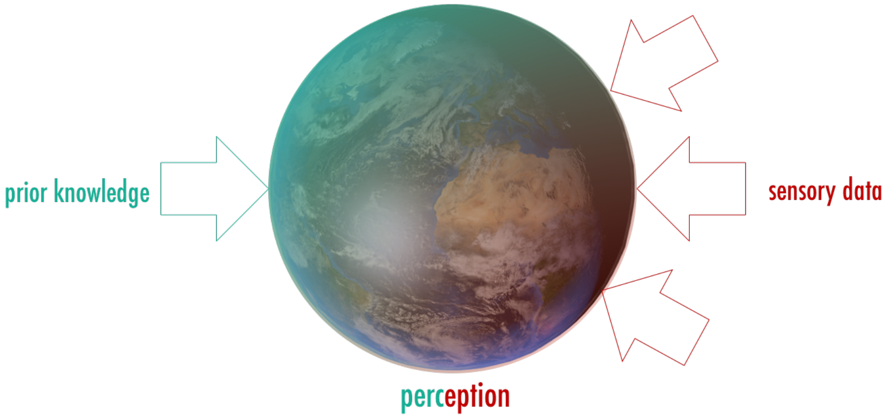
The brain's solution:



What is perception?

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

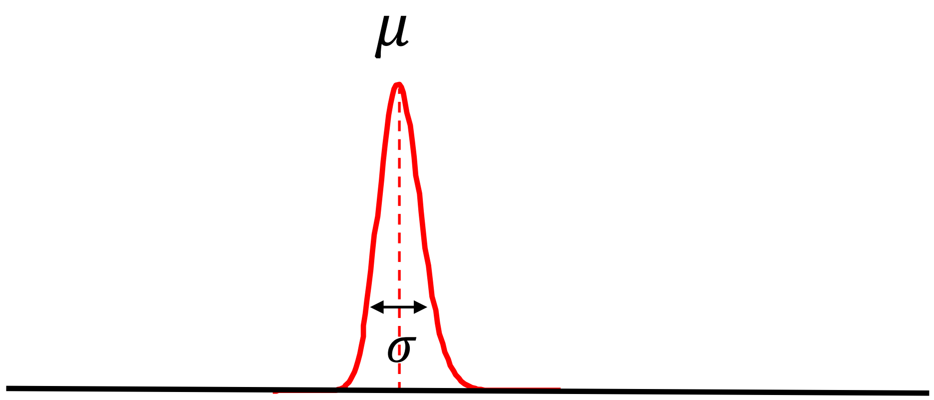
- Sensory data
- Prior knowledge



Why use Bayes to model perception?

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

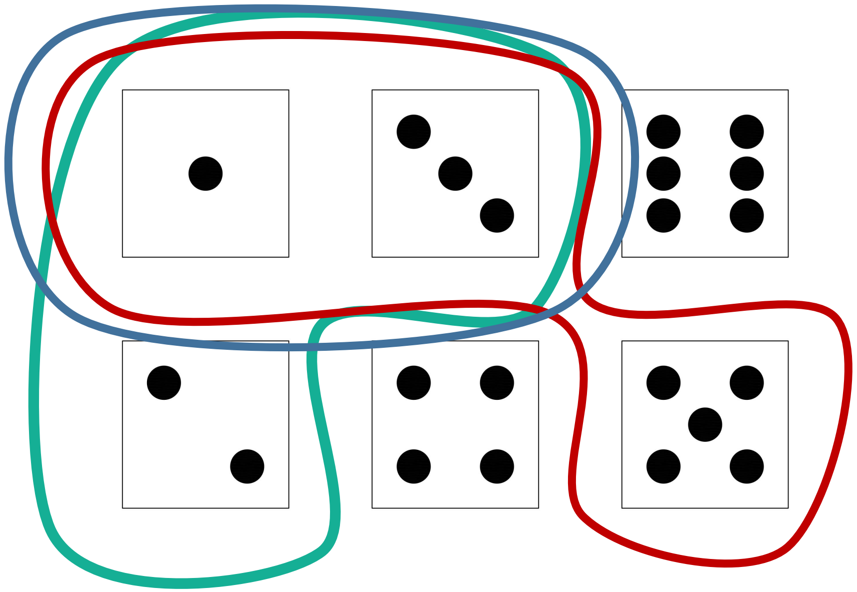


Statistics Tool for conditional probability distributions

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

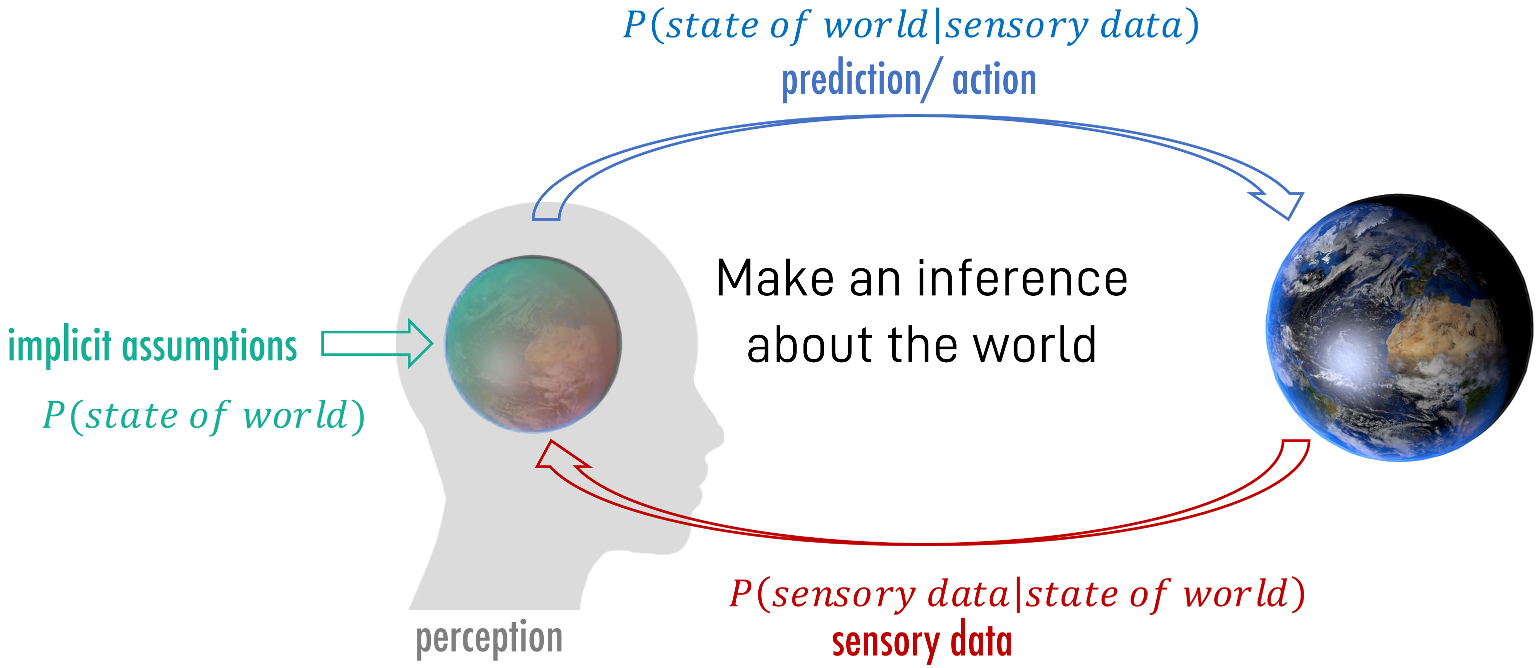
(1) $$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

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

Equation (2) in (1):

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

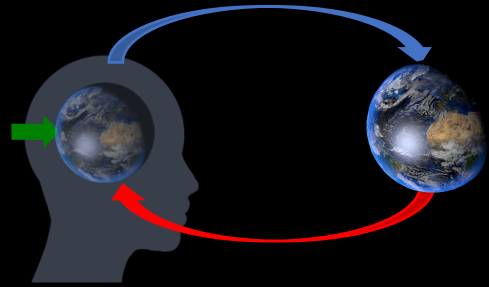
BAYES' RULE

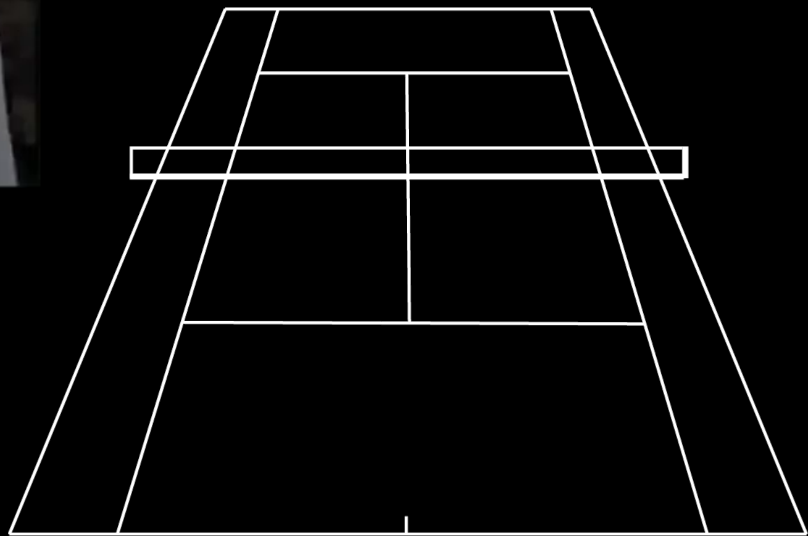


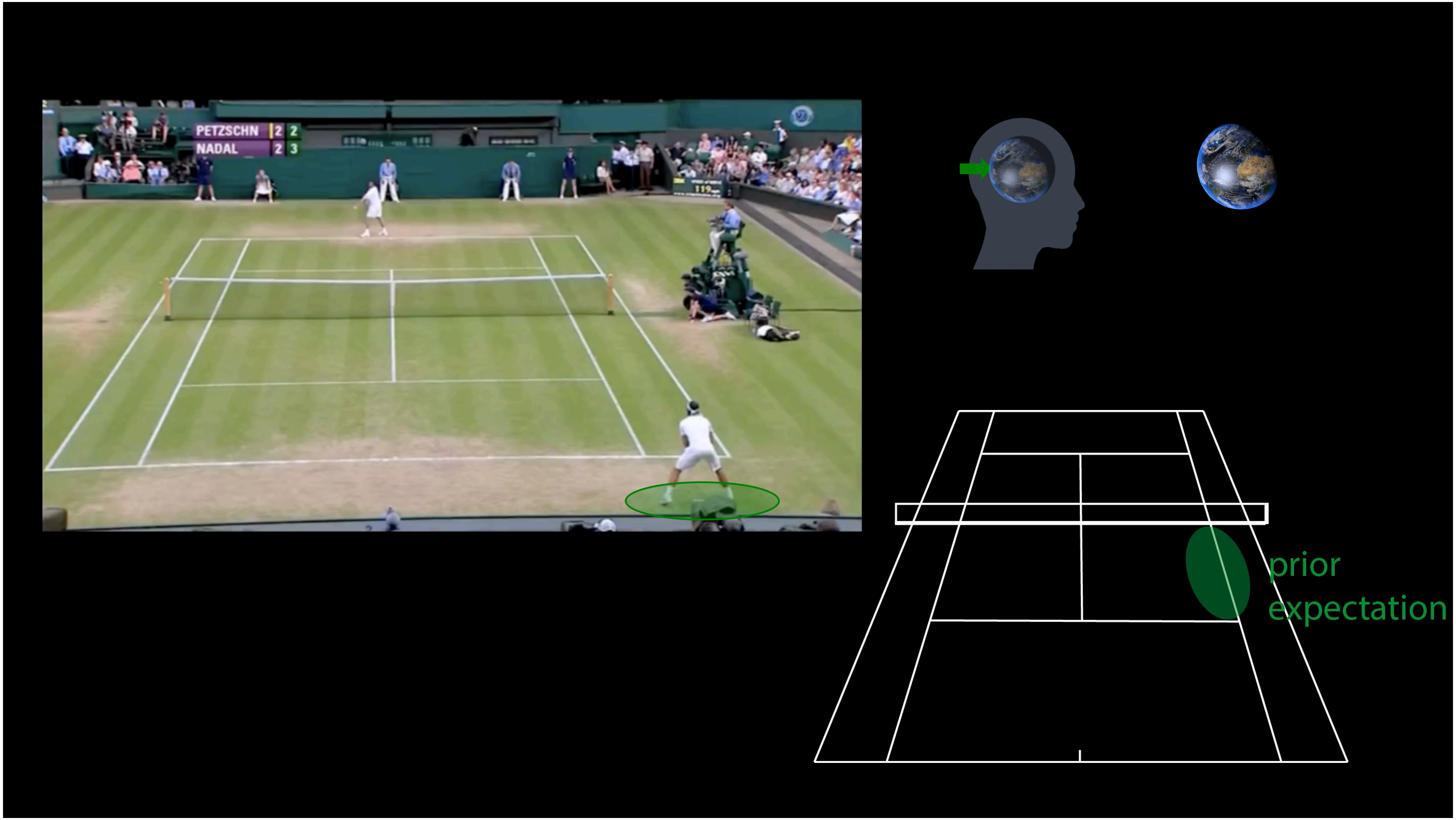
Bayes' Rule

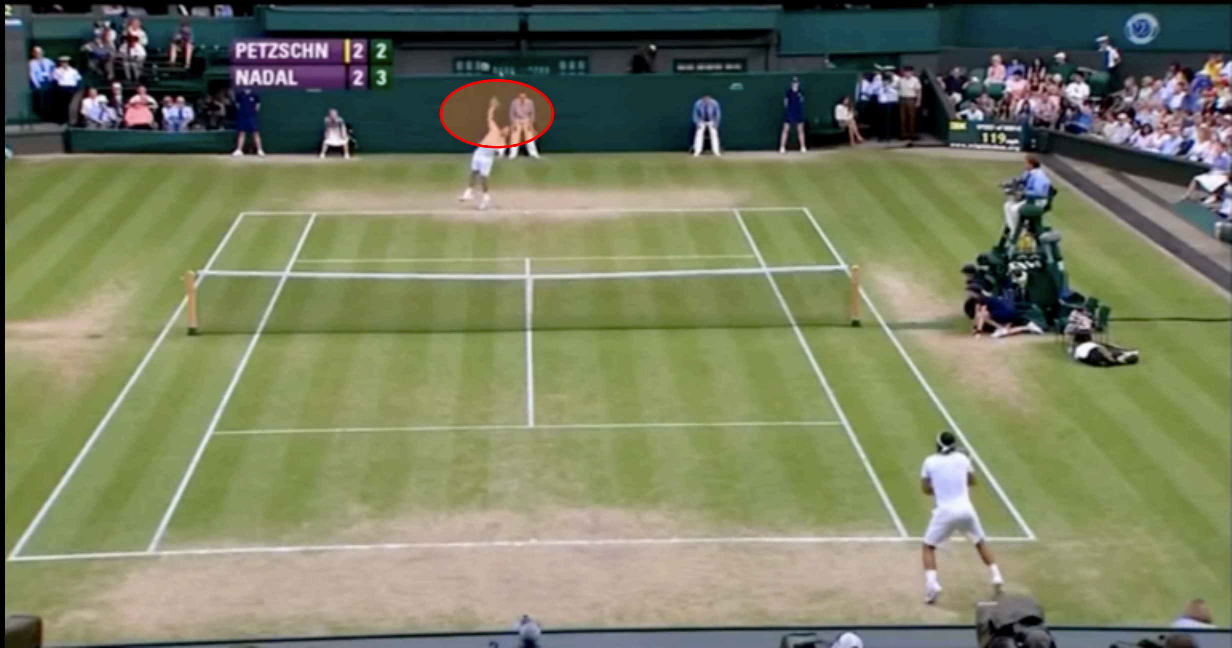
$$\underbrace{P(state\ of\ world|sensory\ data)}_{\text{posterior}} = \frac{\overbrace{P(sensory\ data|state\ of\ world)}^{\text{likelihood}}\overbrace{P(state\ of\ world)}^{\text{prior}}}{P(sensory\ data)}$$

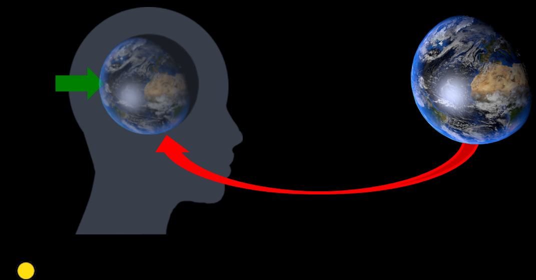
A Bayesian Model of Nadal

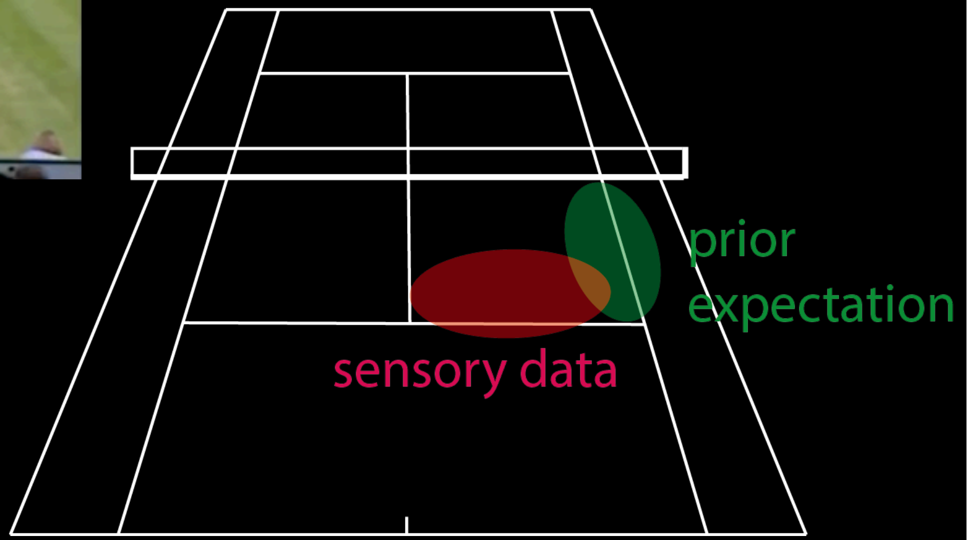


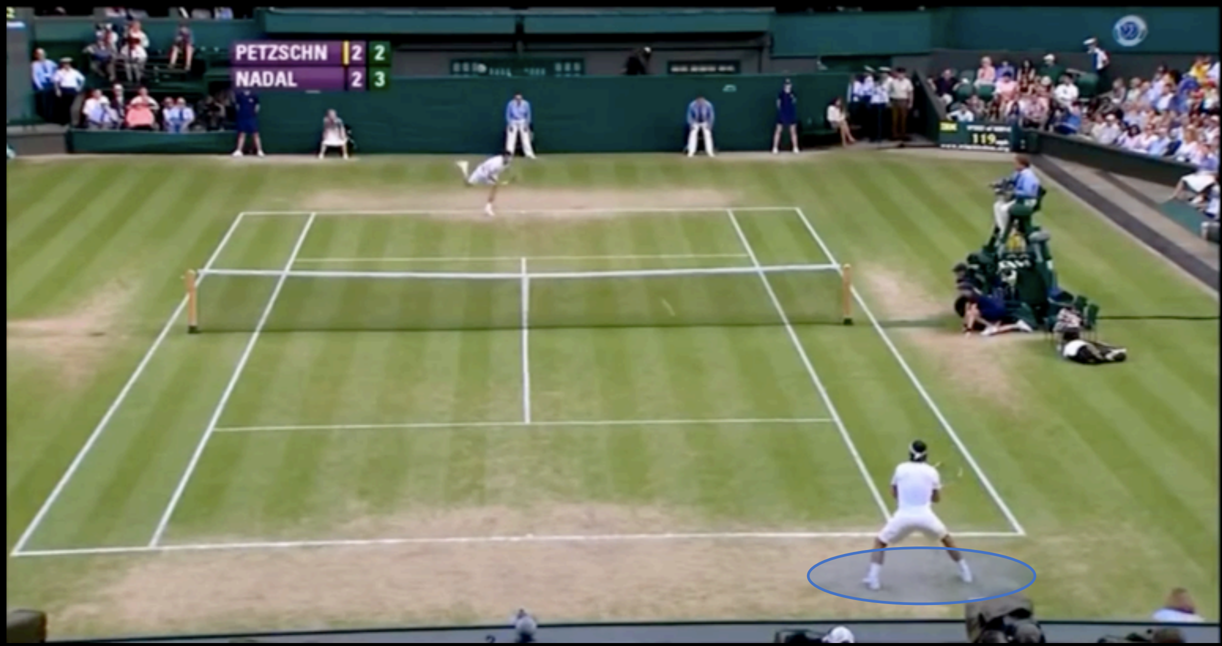


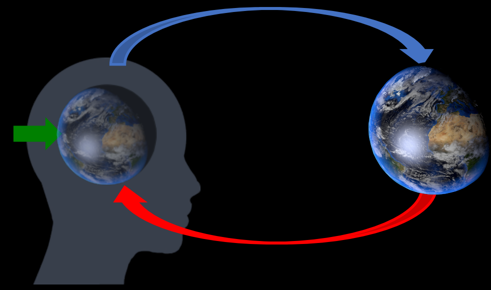

PETZSCHN 2 2
NADAL 2 3
prior
expectation

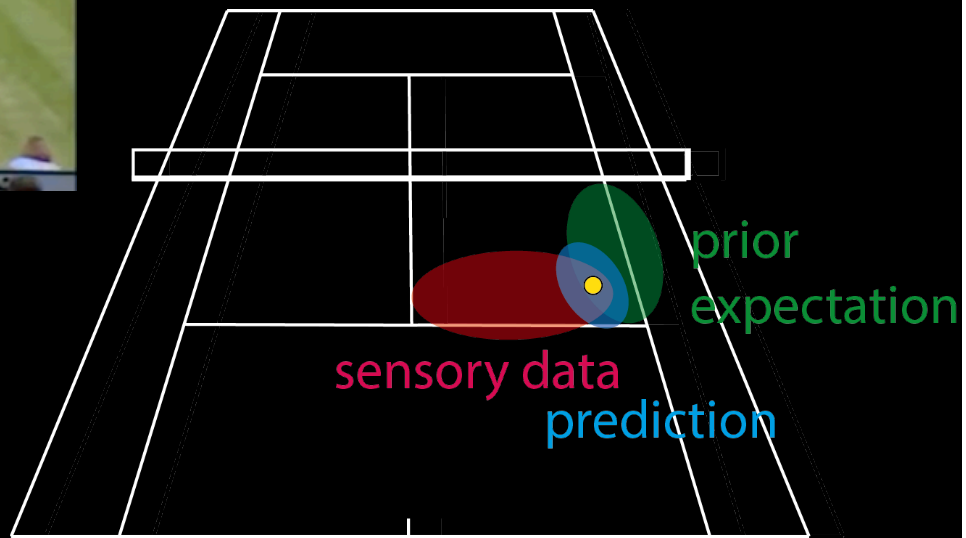

PETZSCHN 2 2
NADAL 2 3



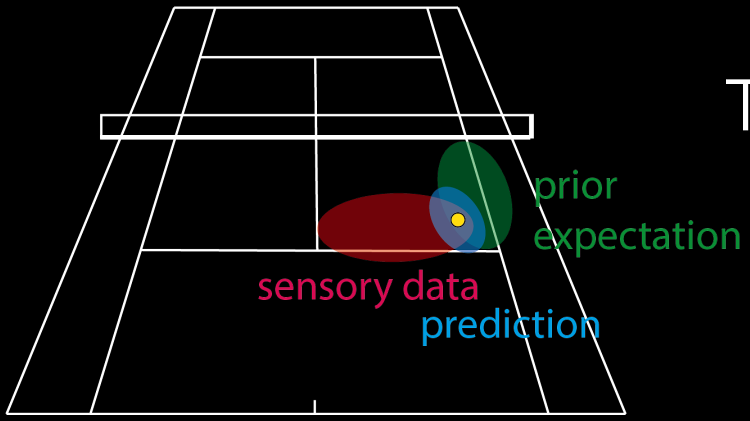

prior
expectation
sensory data


PETZSCHN 2 2
NADAL 2 3

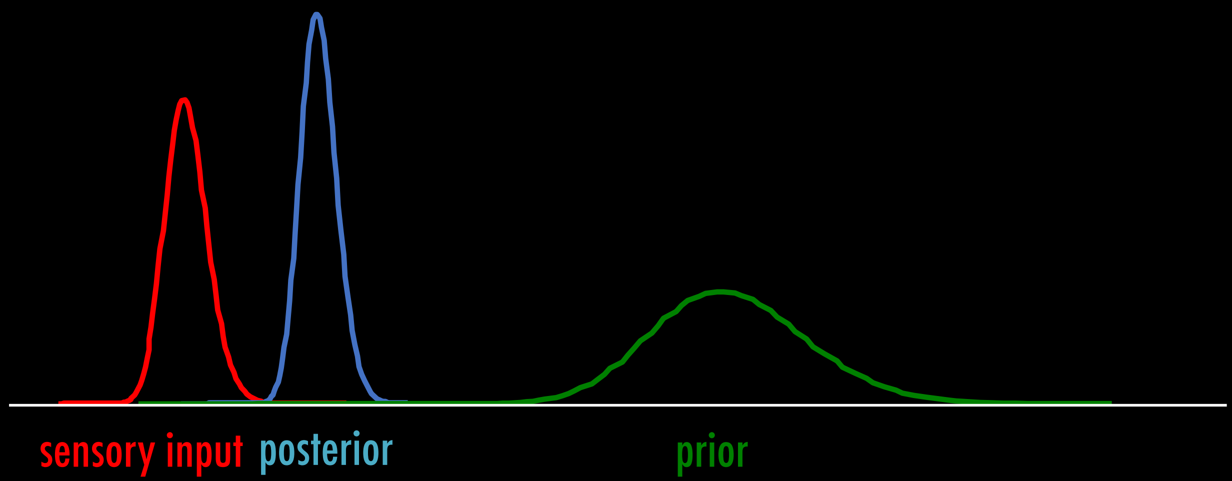



prior
expectation
sensory data
prediction

The ideal Nadal



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



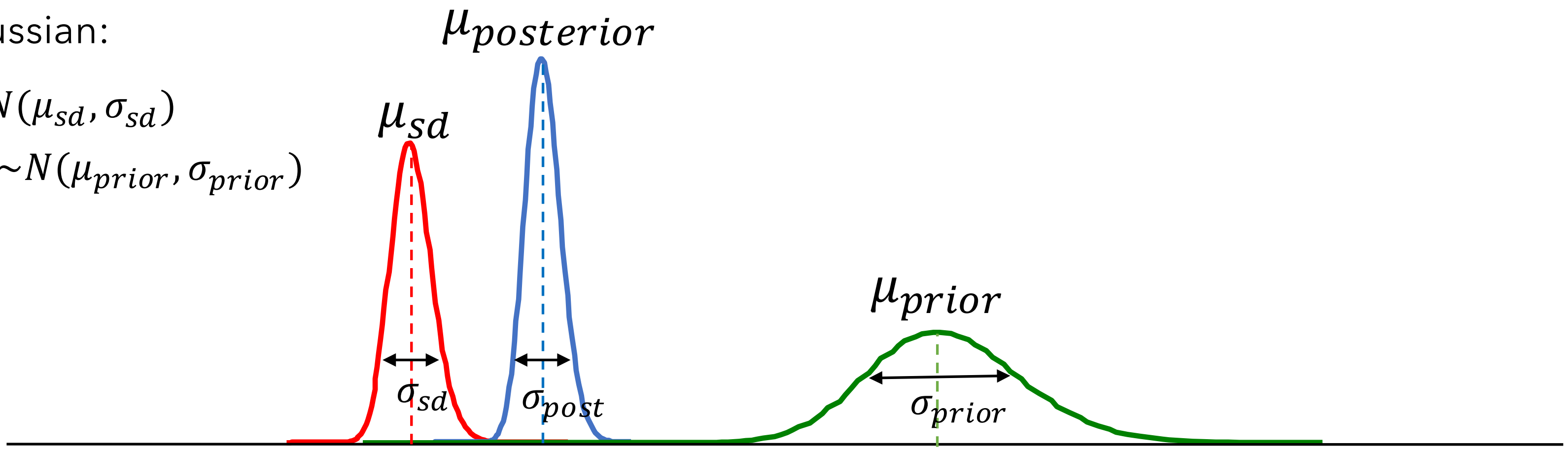
- The optimal combination takes uncertainty into account

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

If Gaussian:

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

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



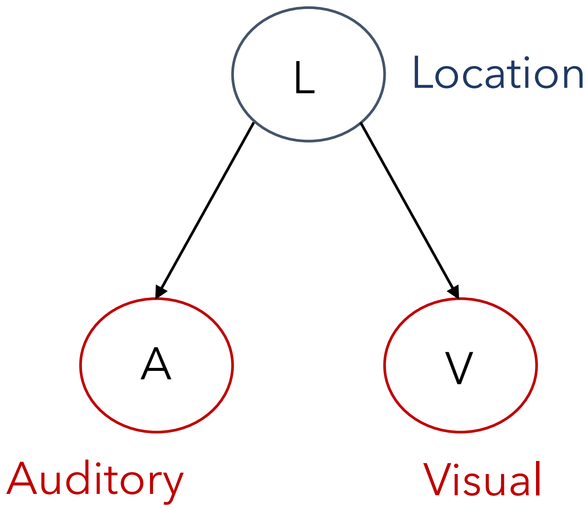
$$\mu_{posterior} = w_{sd} \cdot \mu_{sd} + w_{prior} \cdot \mu_{prior}$$

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

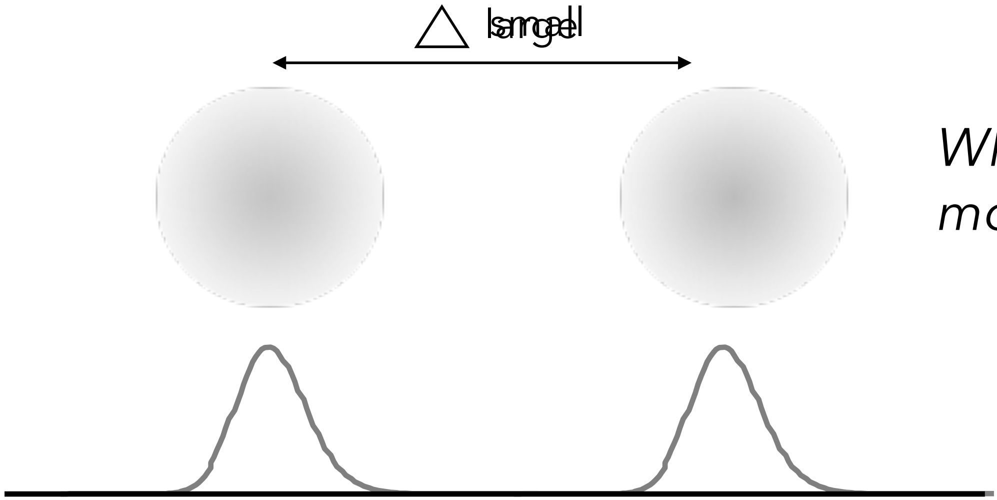
$$\sigma_{posterior}^2 = \frac{\sigma_{sd}^2 \cdot \sigma_{prior}^2}{\sigma_{sd}^2 + \sigma_{prior}^2}$$

Do we behave like a 'Bayesian'?

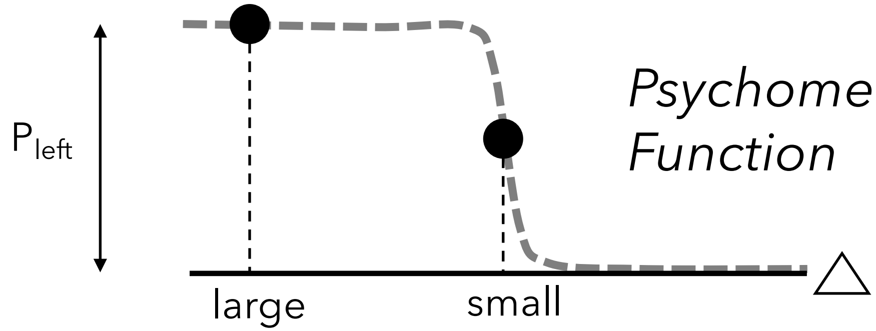
Combination of visual and auditory information



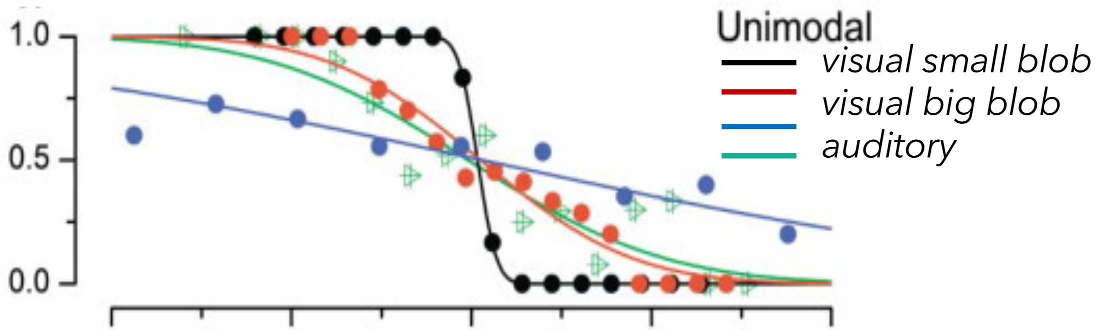
Measuring sensory uncertainty: The Psychometric Function



Which one was more left?

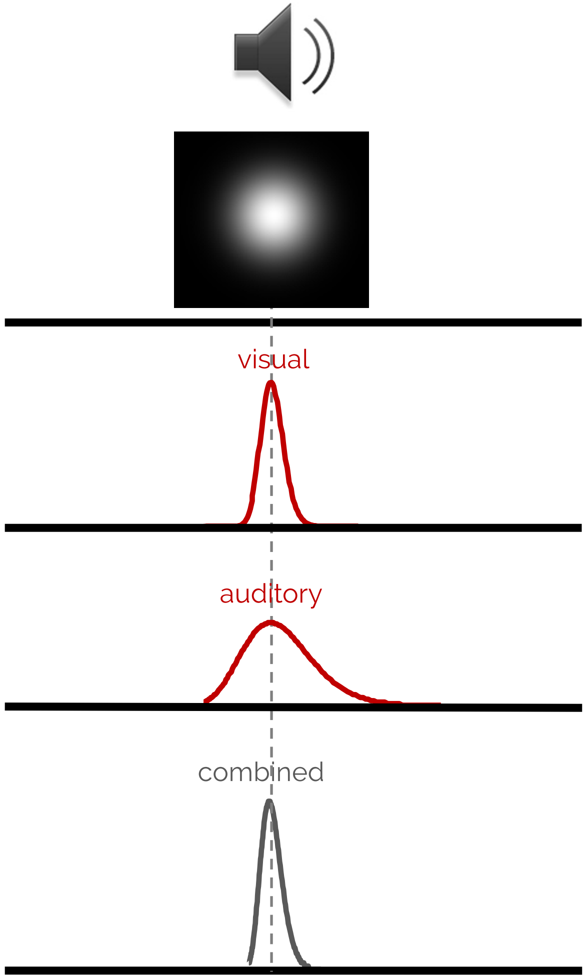


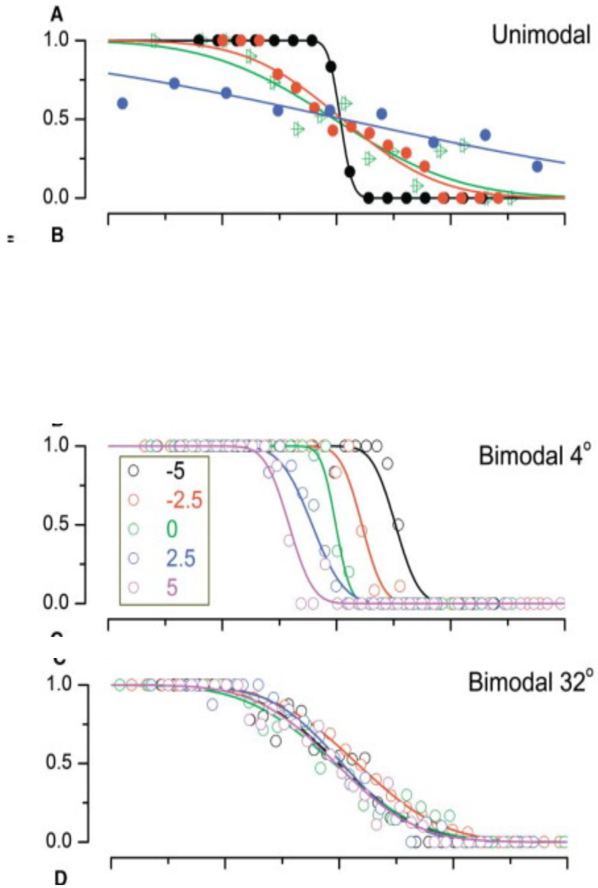
Psychometric Function



Alais & Burr, 2004

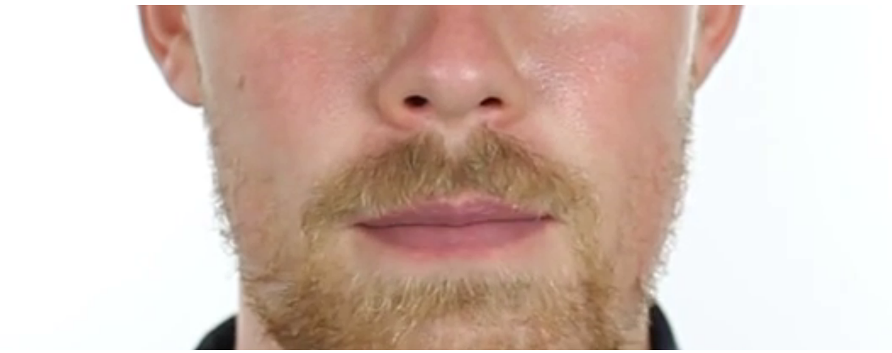
Combination of visual and auditory information: Human Observer



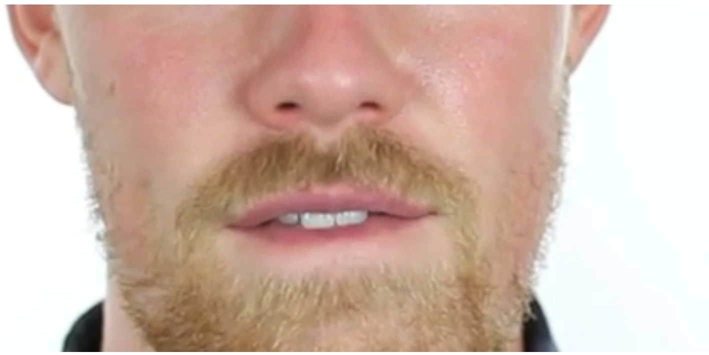


Alais & Burr, 2004

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

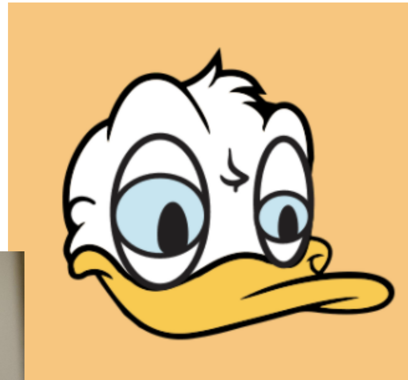


BAR

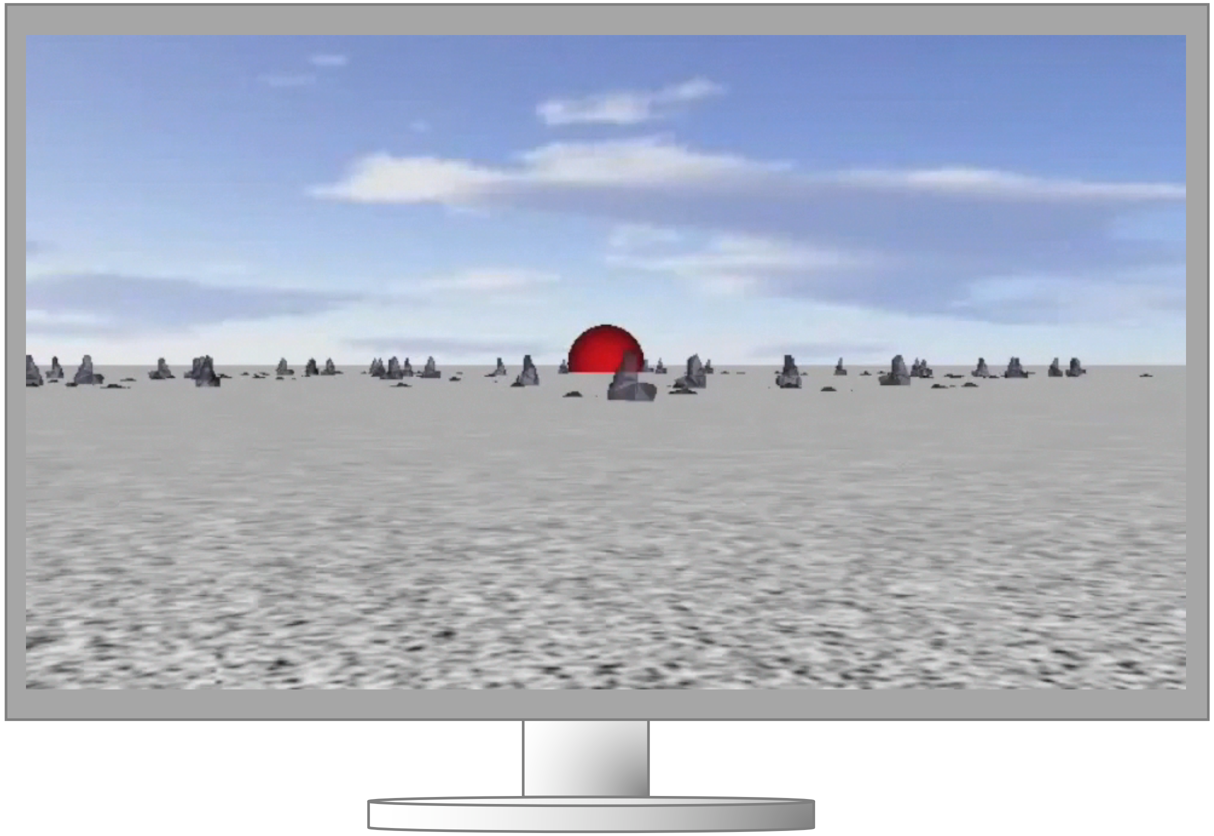


FAR

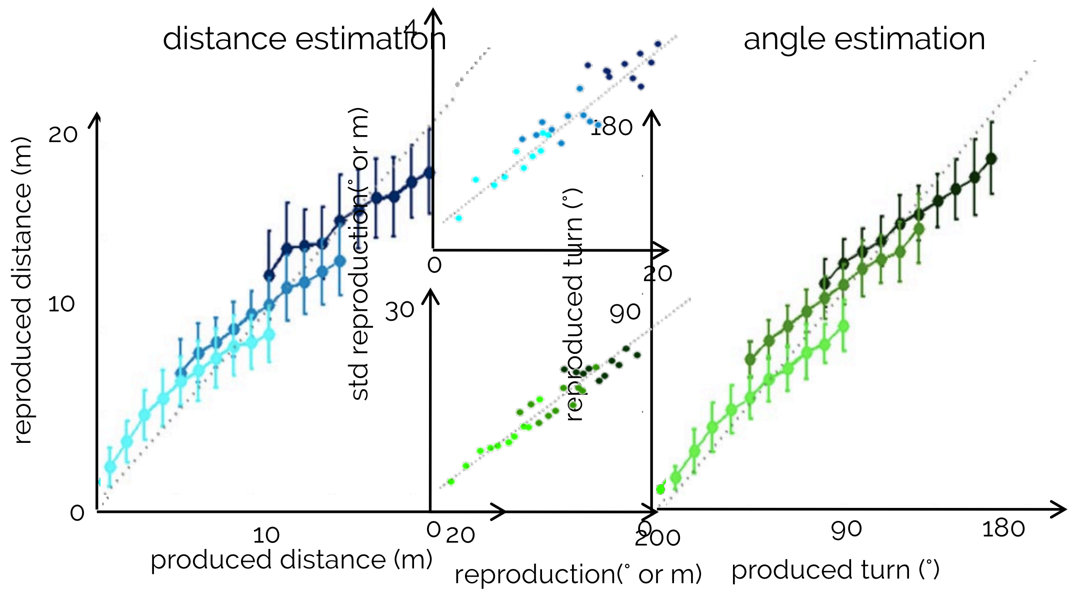
Priors can be learned…



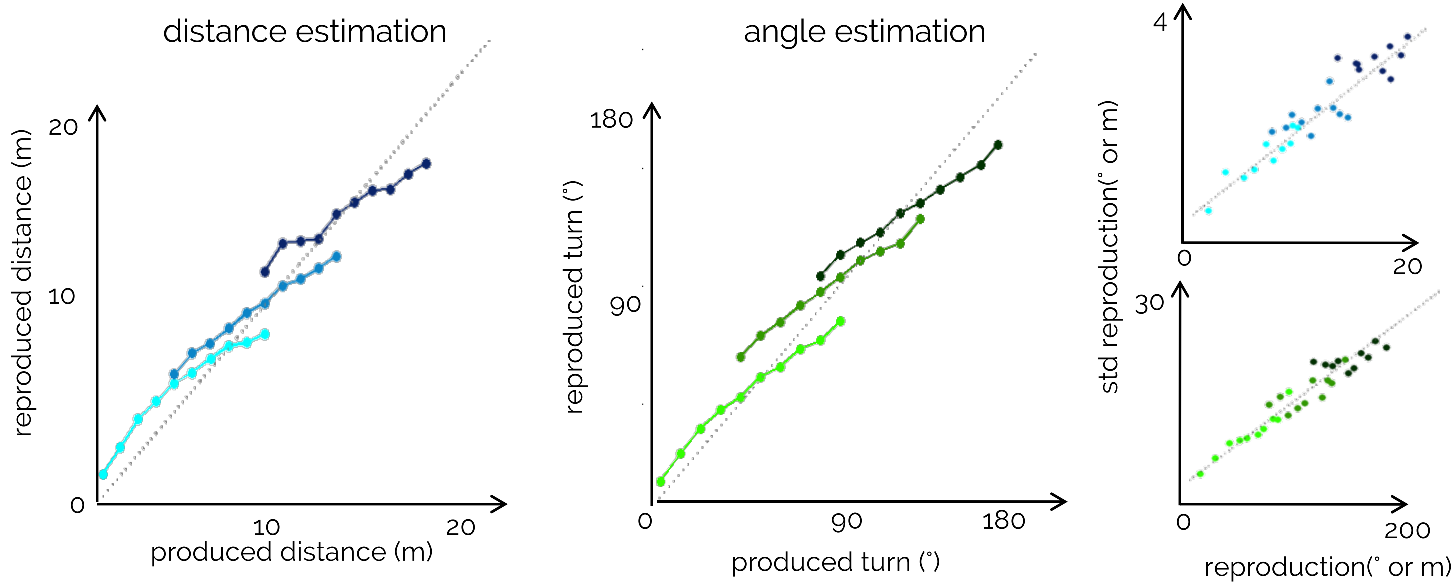
“optimal errors” in magnitude estimation



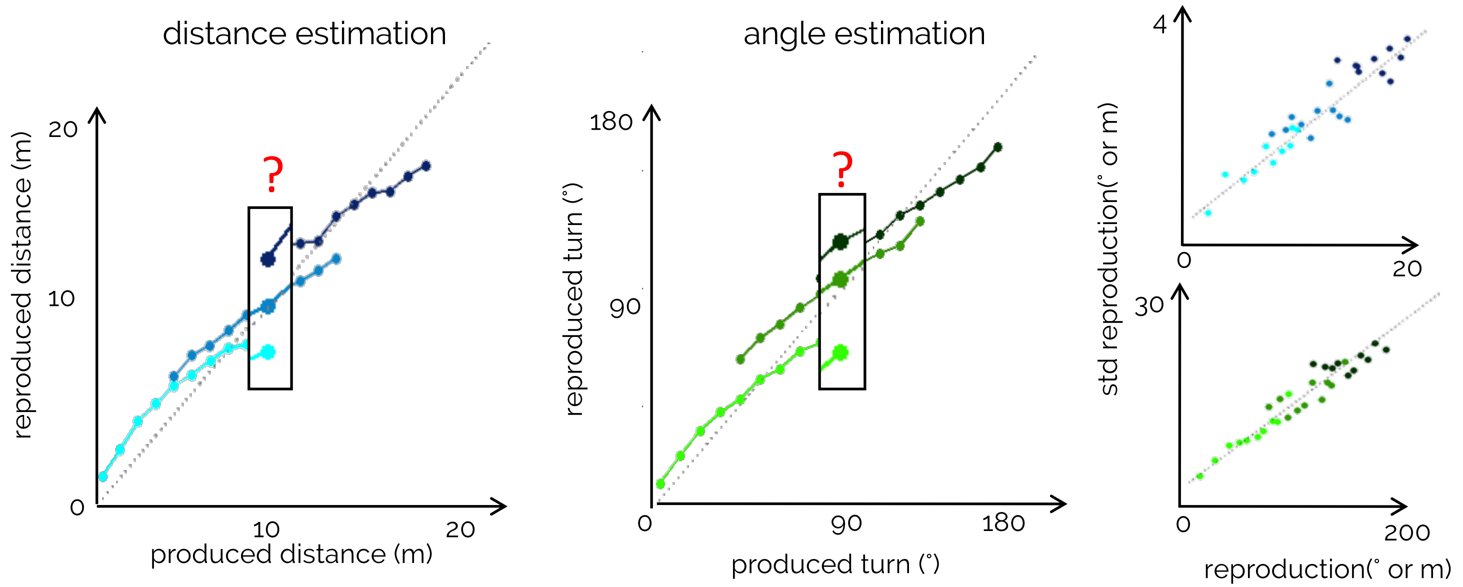
Petzschner & Glasauer, JoN, 2011



Petzschner & Glasauer, JoN, 2011



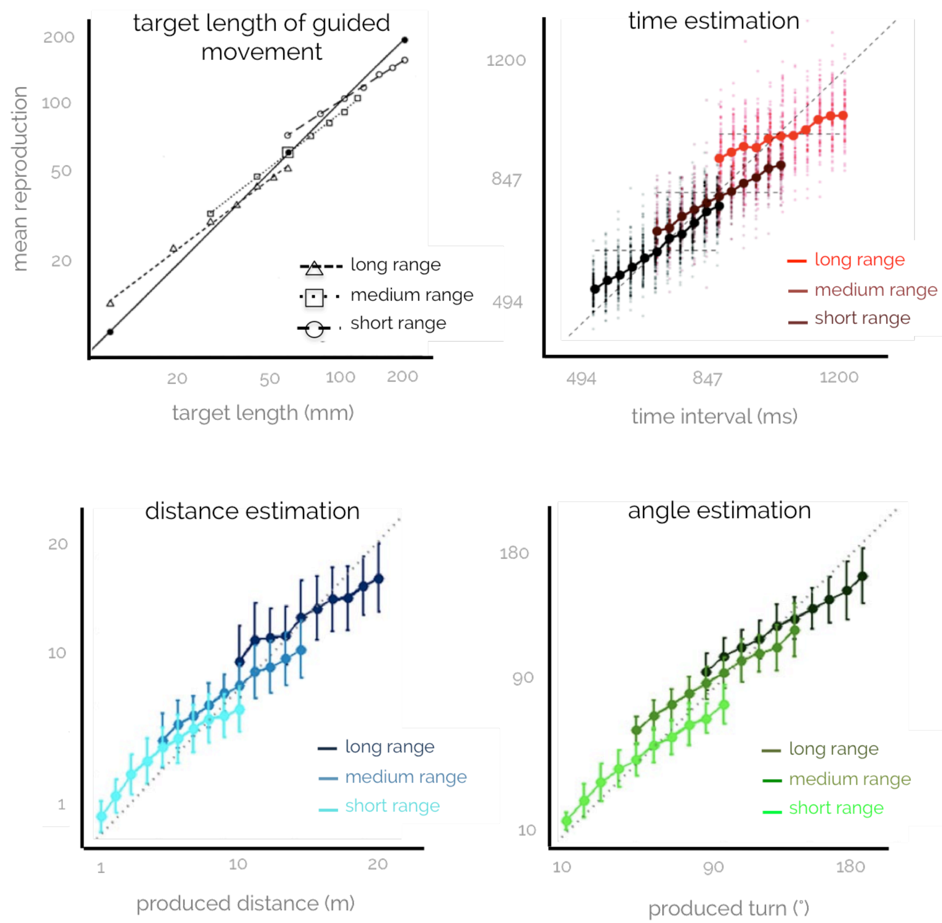
Petzschner & Glasauer, JoN, 2011



Prior knowledge: Experience

Petzschner & Glasauer, JoN, 2011

Let's take a look at the literature

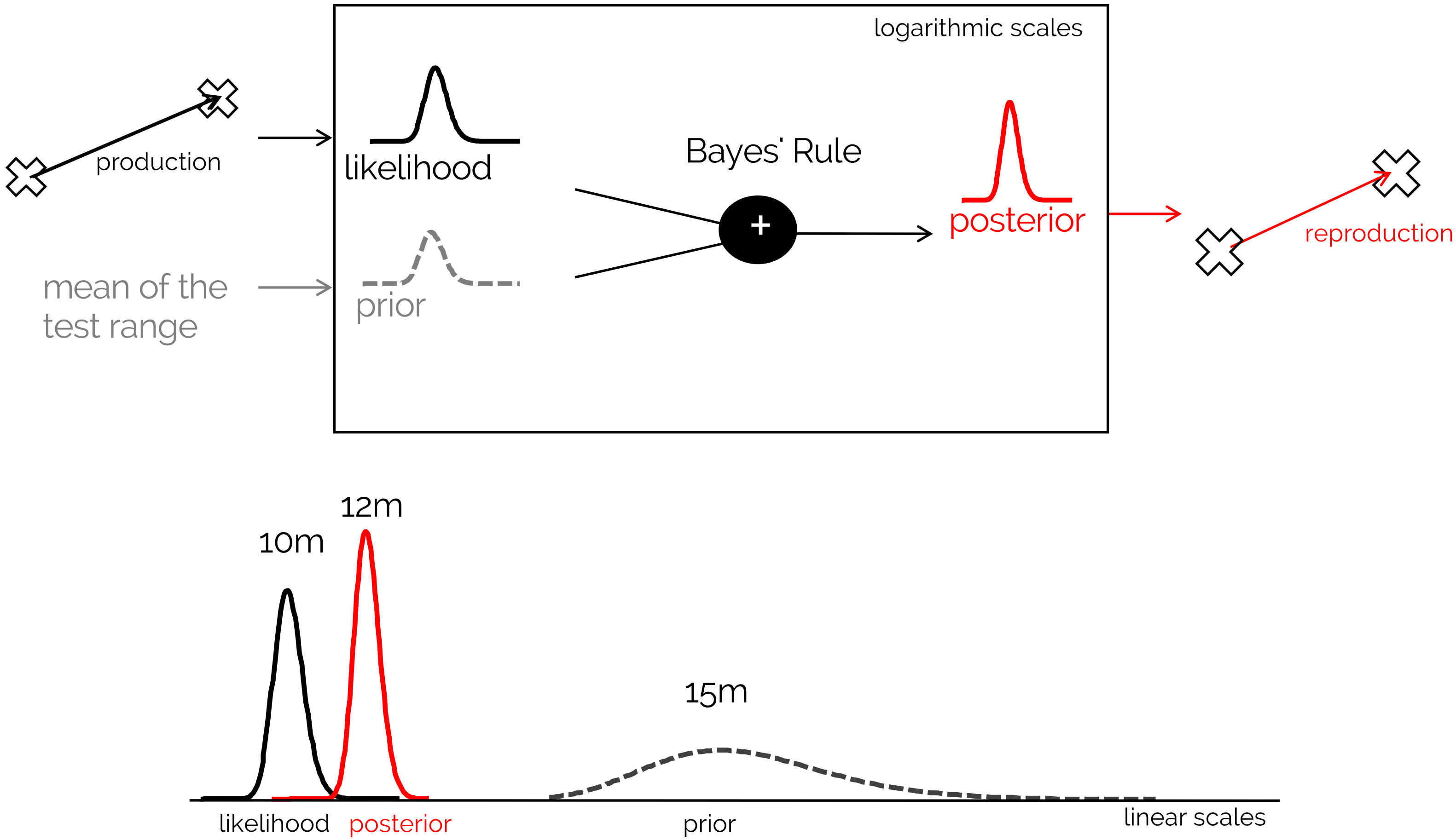


Petzschner et al, TiCS, 2015; Petzschner & Glasauer, JoN, 2011

A Bayesian Model for magnitude estimation

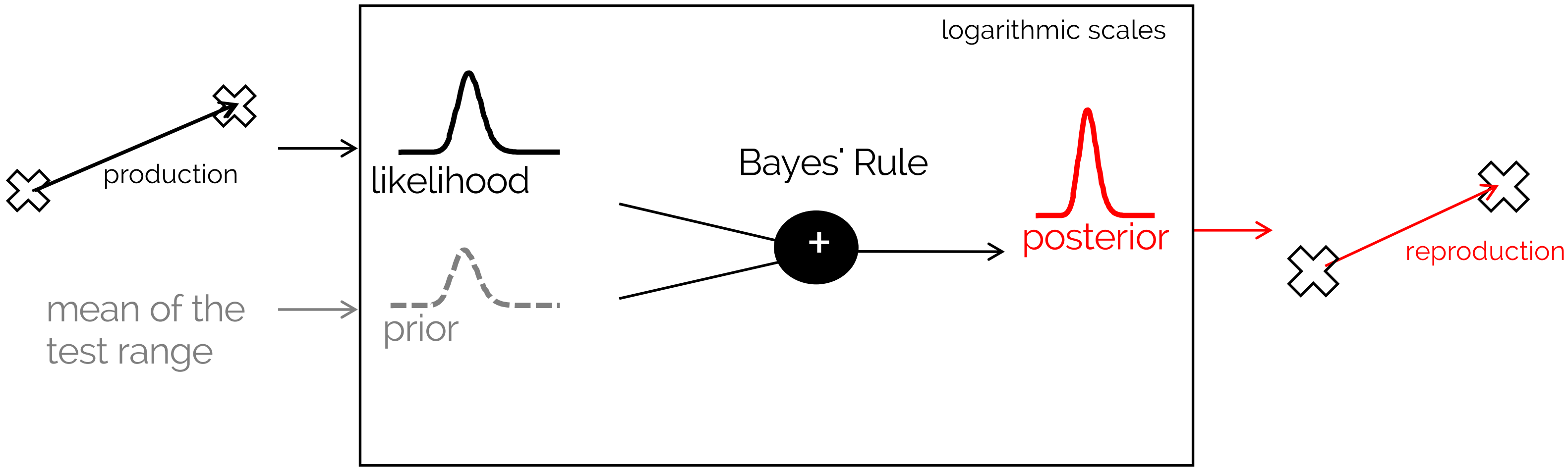
Petzschner & Glasauer, JoN, 2011; Petzschner et al, TiCS, 2015

A Bayesian Model for magnitude estimation



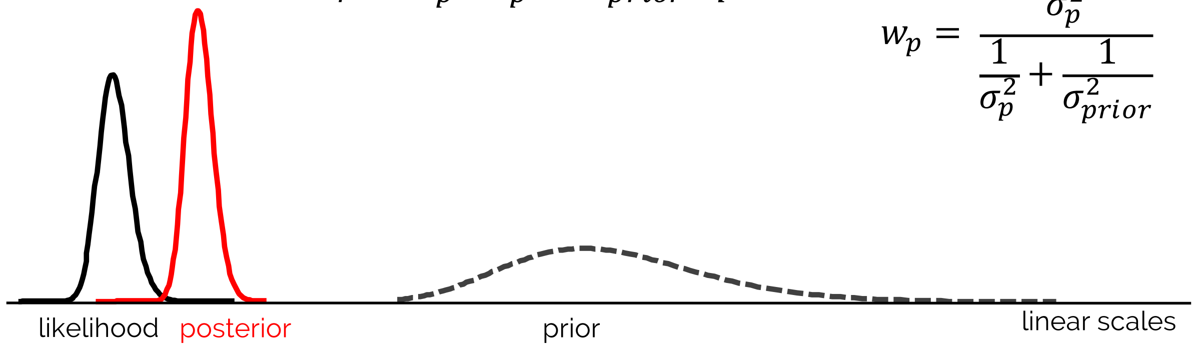
Petzschner & Glasauer, JoN, 2011; Petzschner et al, TiCS, 2015

A Bayesian Model for magnitude estimation



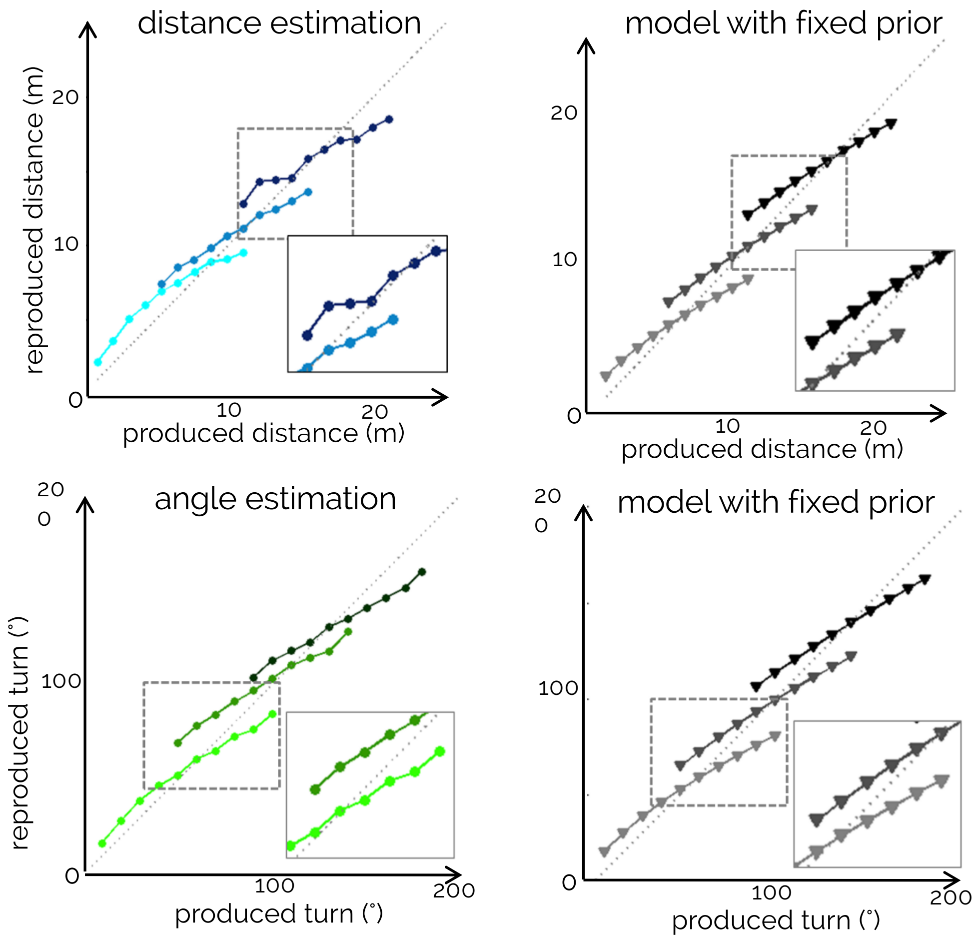
$$\widehat{d_r} = w_p \cdot d_p + w_{prior} \cdot prior$$

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



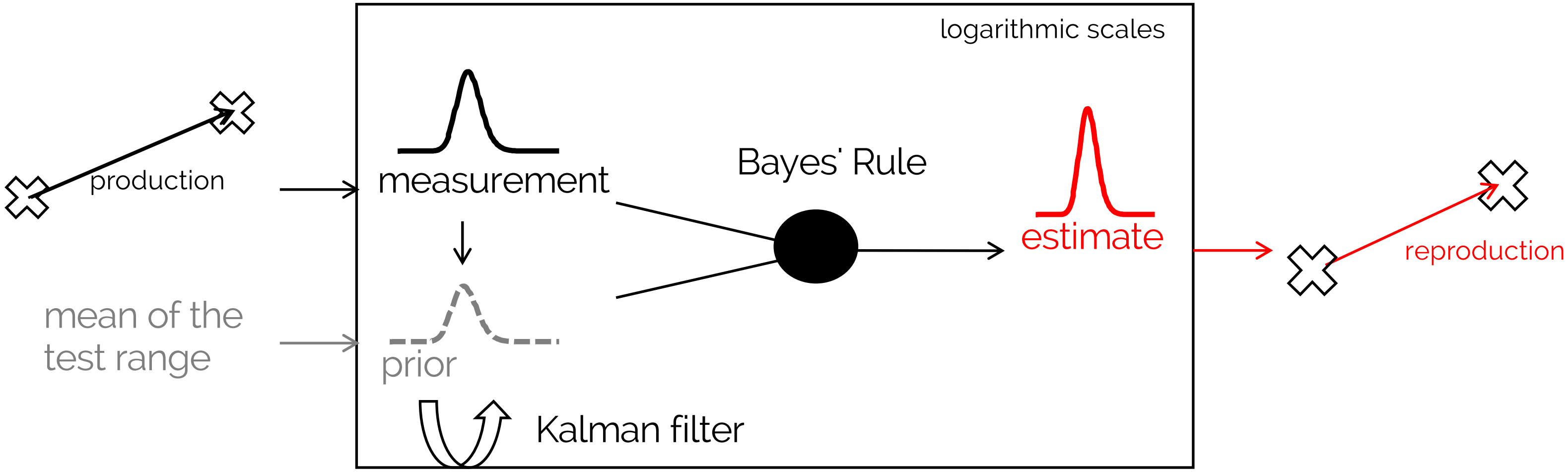
Petzschner & Glasauer, JoN, 2011; Petzschner et al, TiCS, 2015

Quantitative results



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

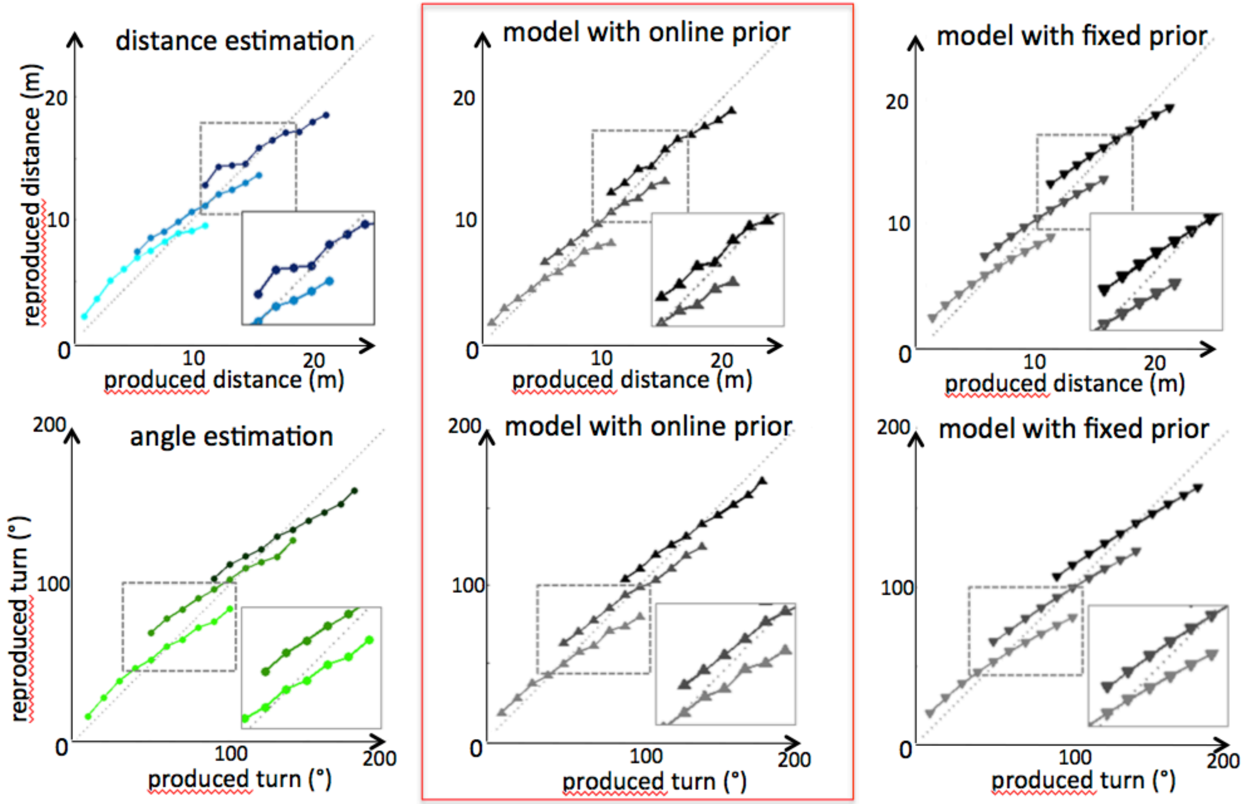
Bayesian Learning



"Todays posterior is tomorrows prior."

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

Bayesian Learning

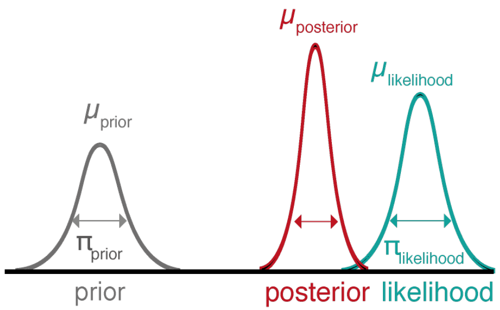


→ Talk by Christoph Mathys on the HGF

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

Hierarchical Bayesian Learning

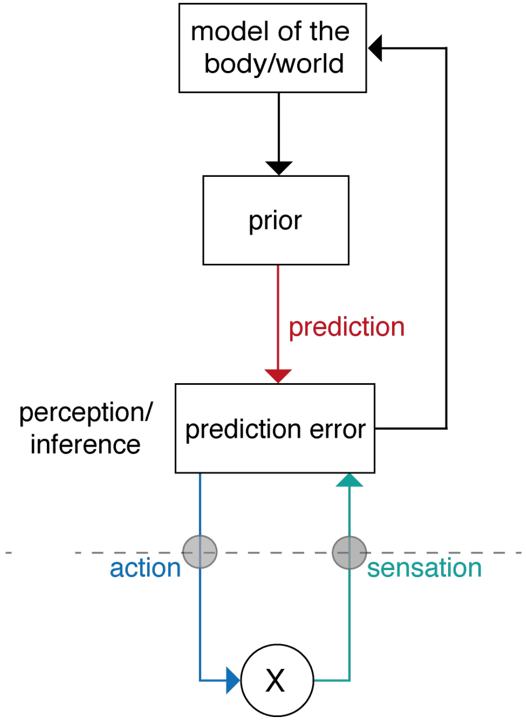
Hierarchical Gaussian Filter
Predictive Coding



Belief update

Δbelief ~precision · PE

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



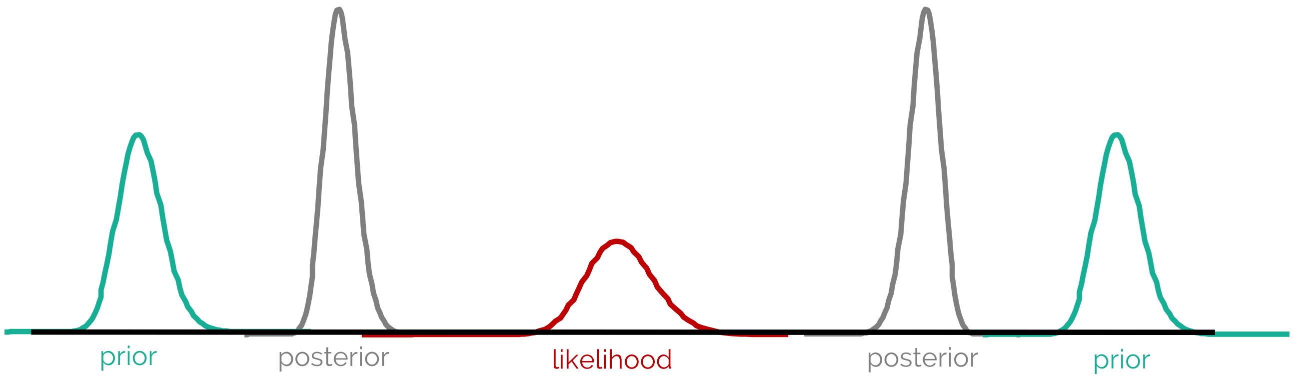
→ *See the next two talks!*
- → *HGF*
- → *Predictive Coding*

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

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

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 …]

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

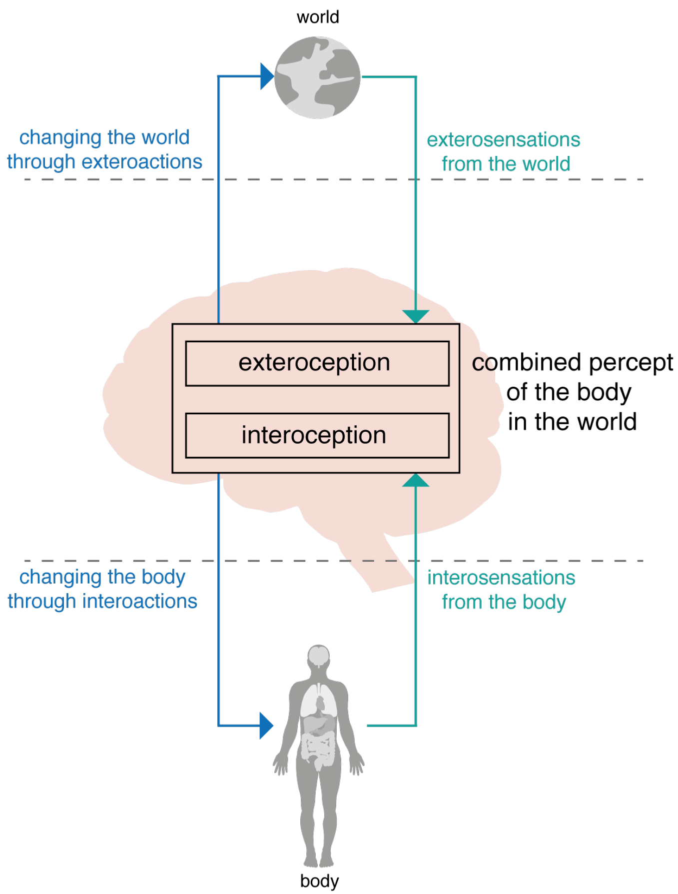
Frederike H. Petzschner, Lilian A.E. Weber, Tim Gard, and Klaas E. Stephan

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

Dr. Frederike Petzschner
Translational Neuromodeling Unit
University of Zurich & ETH
@rikepetzschner