Testing the Bayesian Brain:
A Statistical Model for Sampling-Based Probability Estimates

Henrik Singmann

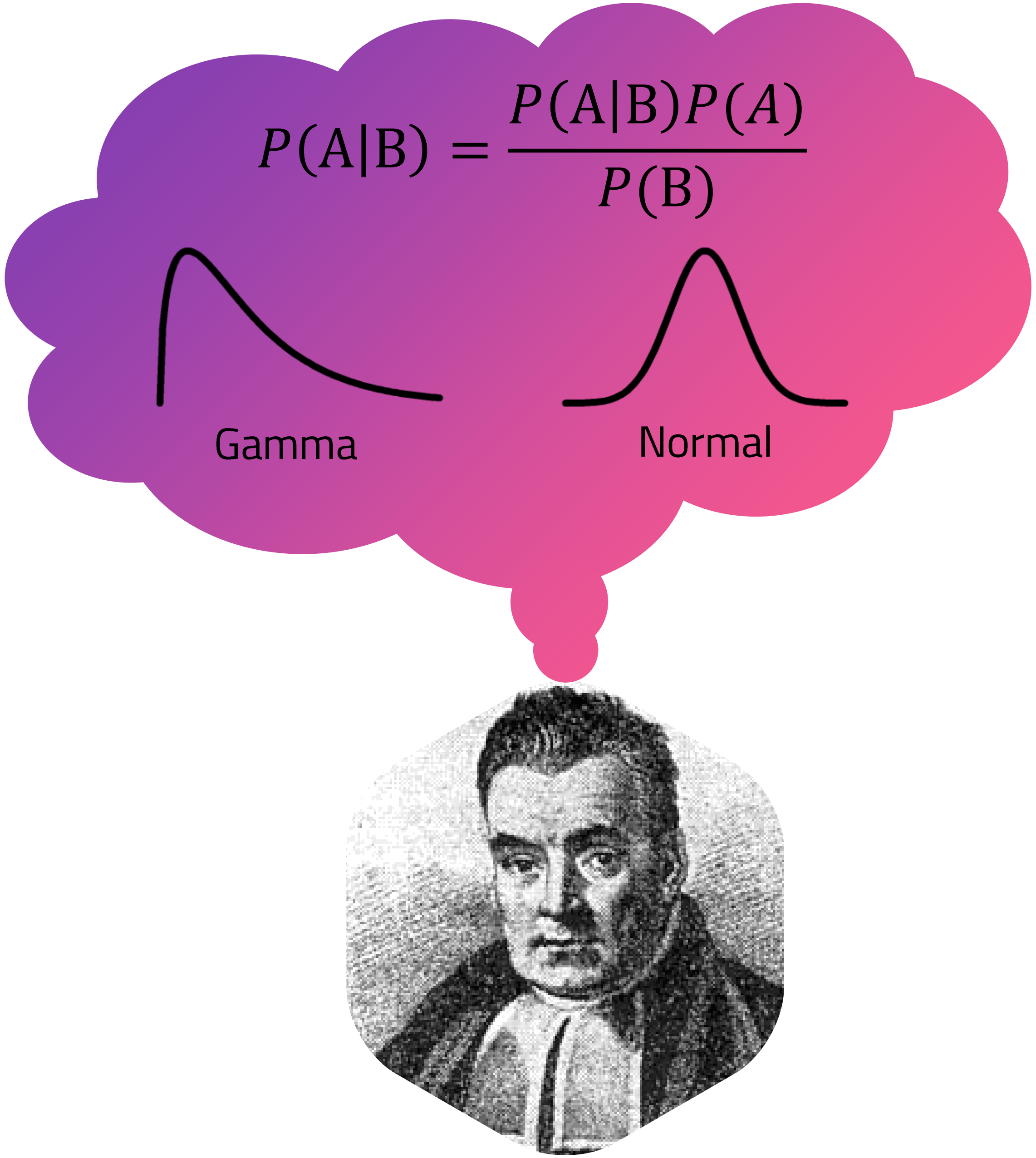
University of Zurich (from October: University of Warwick)

http://singmann.org

@HenrikSingmann

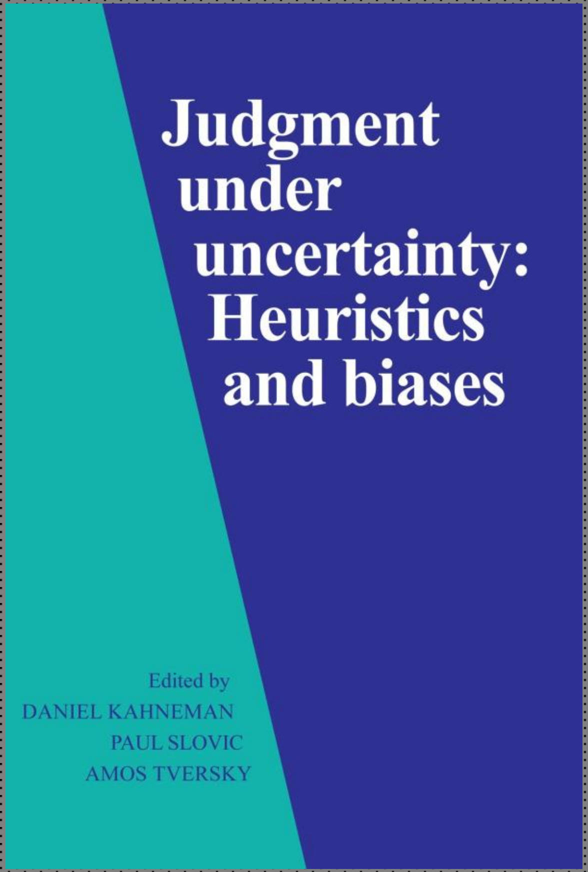
David Kellen

Syracuse University



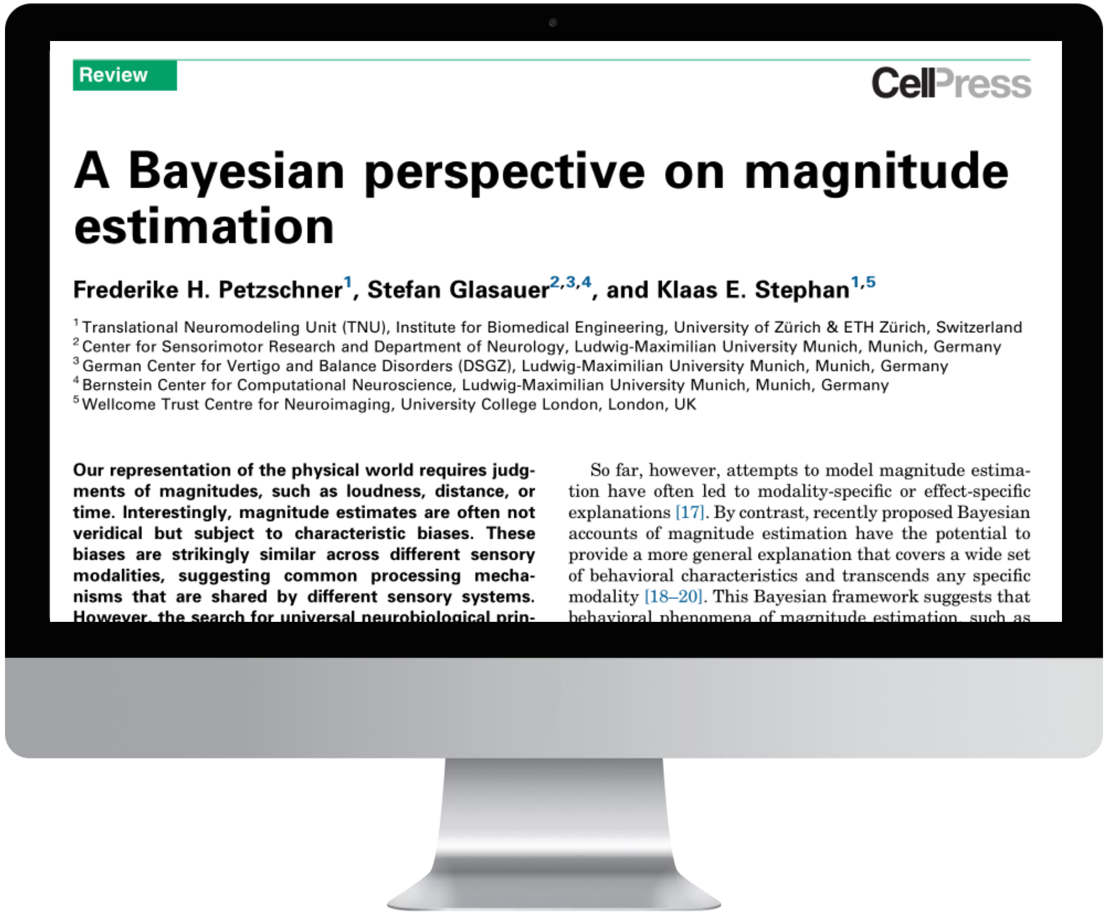
Probabilities in the head

(Thomas Bayes)



People are not Bayesian!
(at least in higher-level tasks)

Resurgence of Bayesian Models since ca. 2000


Review

CellPress

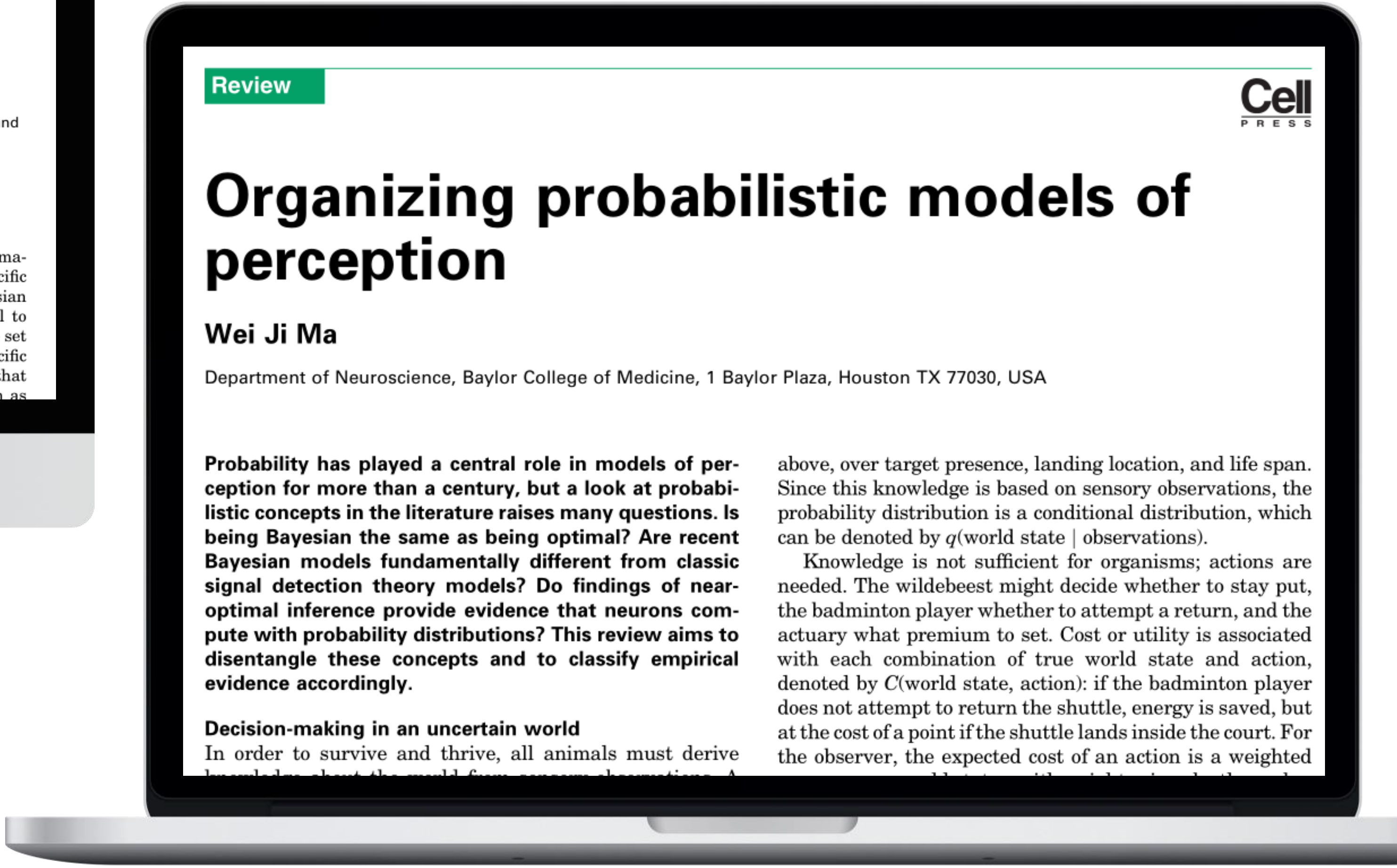
A Bayesian perspective on magnitude estimation

Frederike H. Petzschner[1], Stefan Glasauer[2,3,4], and Klaas E. Stephan[1,5]

[1] Translational Neuromodeling Unit (TNU), Institute for Biomedical Engineering, University of Zürich & ETH Zürich, Switzerland
[2] Center for Sensorimotor Research and Department of Neurology, Ludwig-Maximilian University Munich, Munich, Germany
[3] German Center for Vertigo and Balance Disorders (DSGZ), Ludwig-Maximilian University Munich, Munich, Germany
[4] Bernstein Center for Computational Neuroscience, Ludwig-Maximilian University Munich, Munich, Germany
[5] Wellcome Trust Centre for Neuroimaging, University College London, London, UK

Our representation of the physical world requires judgments of magnitudes, such as loudness, distance, or time. Interestingly, magnitude estimates are often not veridical but subject to characteristic biases. These biases are strikingly similar across different sensory modalities, suggesting common processing mechanisms that are shared by different sensory systems. However, the search for universal neurobiological prin-

So far, however, attempts to model magnitude estimation have often led to modality-specific or effect-specific explanations [17]. By contrast, recently proposed Bayesian accounts of magnitude estimation have the potential to provide a more general explanation that covers a wide set of behavioral characteristics and transcends any specific modality [18–20]. This Bayesian framework suggests that


Review

Cell PRESS

Organizing probabilistic models of perception

Wei Ji Ma

Department of Neuroscience, Baylor College of Medicine, 1 Baylor Plaza, Houston TX 77030, USA

Probability has played a central role in models of perception for more than a century, but a look at probabilistic concepts in the literature raises many questions. Is being Bayesian the same as being optimal? Are recent Bayesian models fundamentally different from classic signal detection theory models? Do findings of near-optimal inference provide evidence that neurons compute with probability distributions? This review aims to disentangle these concepts and to classify empirical evidence accordingly.

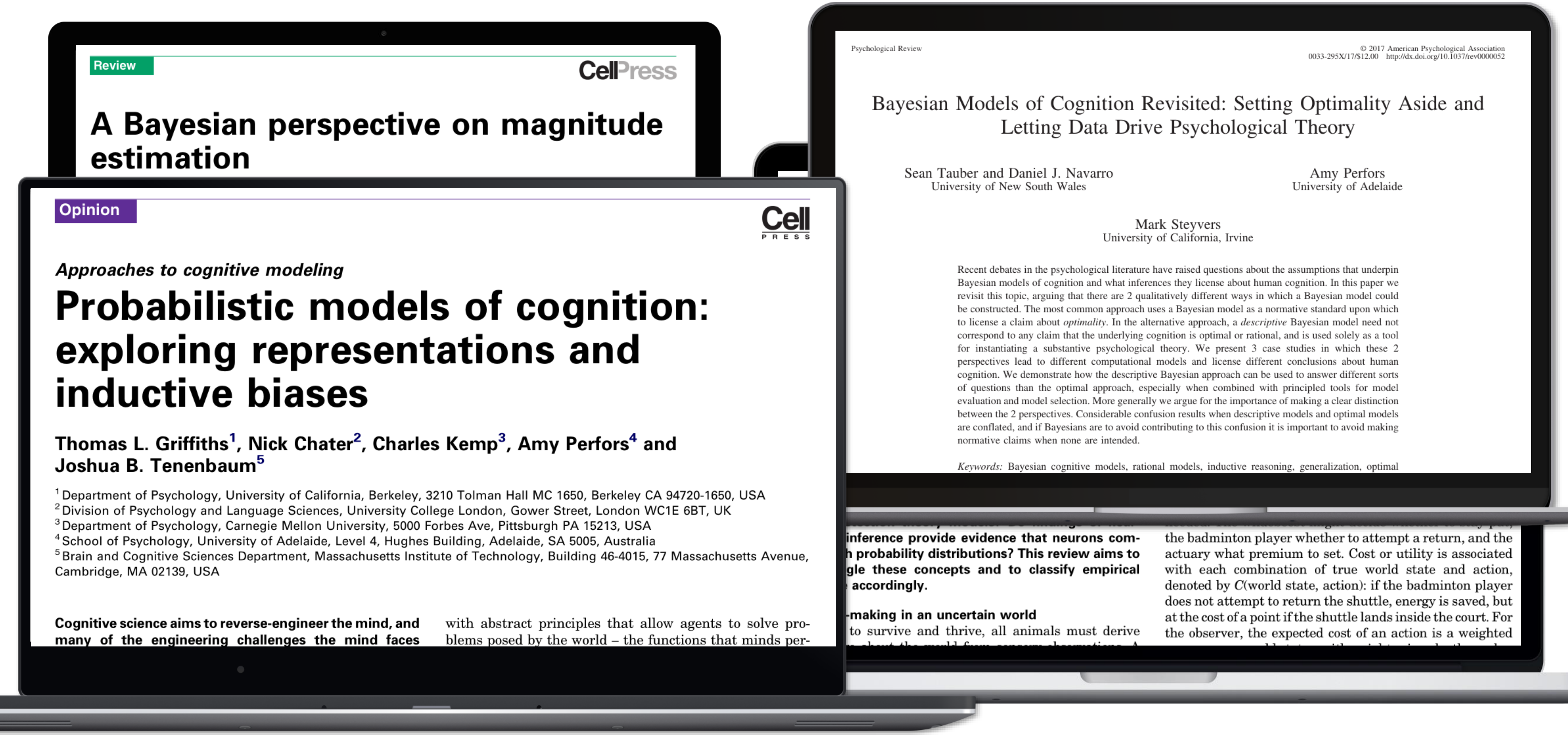
Decision-making in an uncertain world

In order to survive and thrive, all animals must derive

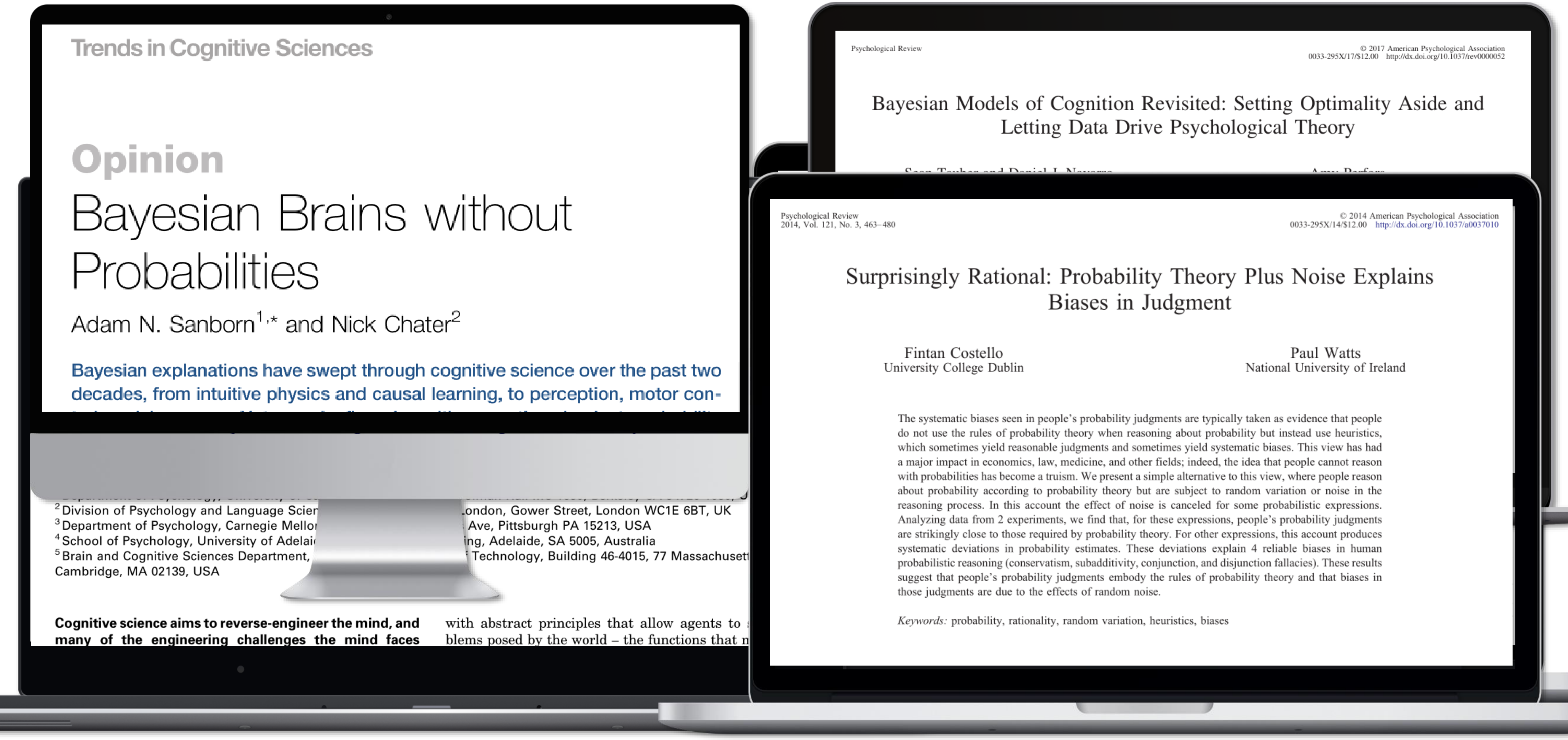
above, over target presence, landing location, and life span. Since this knowledge is based on sensory observations, the probability distribution is a conditional distribution, which can be denoted by q(world state | observations).

Knowledge is not sufficient for organisms; actions are needed. The wildebeest might decide whether to stay put, the badminton player whether to attempt a return, and the actuary what premium to set. Cost or utility is associated with each combination of true world state and action, denoted by C(world state, action): if the badminton player does not attempt to return the shuttle, energy is saved, but at the cost of a point if the shuttle lands inside the court. For the observer, the expected cost of an action is a weighted

Resurgence of Bayesian Models since ca. 2000



Resurgence of Bayesian Models since ca. 2000



Surprisingly rational: Probability theory plus noise explains biases in judgment.

Costello, Fintan,Watts, Paul

Psychological Review, Vol 121(3), Jul 2014, 463-480

The Mind is a **Bayesian Sampler**

Trends in Cognitive Sciences
2016

Opinion

Bayesian Brains without Probabilities

Adam N. Sanborn[1,*] and Nick Chater[2]

Surprisingly rational: Probability theory plus noise explains biases in judgment.

Costello, Fintan,Watts, Paul

Psychological Review, Vol 121(3), Jul 2014, 463-480

Brain does not represent probability distributions

The Mind is a **Bayesian Sampler**

Trends in Cognitive Sciences
2016

Opinion

Bayesian Brains without Probabilities

Adam N. Sanborn[1,*] and Nick Chater[2]

Surprisingly rational: Probability theory plus noise explains biases in judgment.

Costello, Fintan,Watts, Paul

Psychological Review, Vol 121(3), Jul 2014, 463-480

Brain does not represent probability distributions

Probability distributions approximated via sampling from memory

The Mind is a **Bayesian Sampler**

Trends in Cognitive Sciences
2016

Opinion

Bayesian Brains without Probabilities

Adam N. Sanborn[1,*] and Nick Chater[2]

Surprisingly rational: Probability theory plus noise explains biases in judgment.

Costello, Fintan,Watts, Paul

Psychological Review, Vol 121(3), Jul 2014, 463-480

The Mind is a **Bayesian Sampler**

Trends in Cognitive Sciences
2016

Opinion

Bayesian Brains without Probabilities

Adam N. Sanborn[1,*] and Nick Chater[2]

Brain does not represent probability distributions

Probability distributions approximated via sampling from memory

Test Bayesian brain:
Can sampling based Bayesian model describe participants' behavior?

Surprisingly rational: Probability theory plus noise explains biases in judgment.

Costello, Fintan,Watts, Paul

Psychological Review, Vol 121(3), Jul 2014, 463-480

The Mind is a **Bayesian Sampler**

Trends in Cognitive Sciences
2016

Opinion

Bayesian Brains without Probabilities

Adam N. Sanborn[1,*] and Nick Chater[2]

Brain does not represent probability distributions

Probability distributions approximated via sampling from memory

Test Bayesian brain:
Can sampling based Bayesian model describe participants' behavior?

Noisy sampler:
Model for probability estimation, combines
- Bayesian brain
- Sampling
- Noise

Surprisingly rational: Probability theory plus noise explains biases in judgment.

Costello, Fintan,Watts, Paul

Psychological Review, Vol 121(3), Jul 2014, 463-480

The Mind is a **Bayesian Sampler**

Trends in Cognitive Sciences
2016

Opinion

Bayesian Brains without Probabilities

Adam N. Sanborn[1,*] and Nick Chater[2]

Brain does not represent probability distributions

Probability distributions approximated via sampling from memory

Test Bayesian brain:
Can sampling based Bayesian model describe participants' behavior?

Noisy sampler:
Model for probability estimation, combines
- Bayesian brain
- Sampling
- Noise

Kind of.

Surprisingly rational: Probability theory plus noise explains biases in judgment.

Costello, Fintan, Watts, Paul

Psychological Review, Vol 121(3), Jul 2014, 463-480

The Mind is a **Bayesian Sampler**

Trends in Cognitive Sciences
2016

Opinion

Bayesian Brains without Probabilities

Adam N. Sanborn[1,*] and Nick Chater[2]

Surprisingly rational: Probability theory plus noise explains biases in judgment.

Costello, Fintan, Watts, Paul

Psychological Review, Vol 121(3), Jul 2014, 463-480

What is the probability that the weather will be *cloudy* on a randomly selected day in Ireland?

The Mind is a **Bayesian Sampler**

Trends in Cognitive Sciences
2016

Opinion

Bayesian Brains without Probabilities

Adam N. Sanborn[1,*] and Nick Chater[2]

Surprisingly rational: Probability theory plus noise explains biases in judgment.

Costello, Fintan,Watts, Paul

Psychological Review, Vol 121(3), Jul 2014, 463-480

The Mind is a **Bayesian Sampler**

Trends in Cognitive Sciences

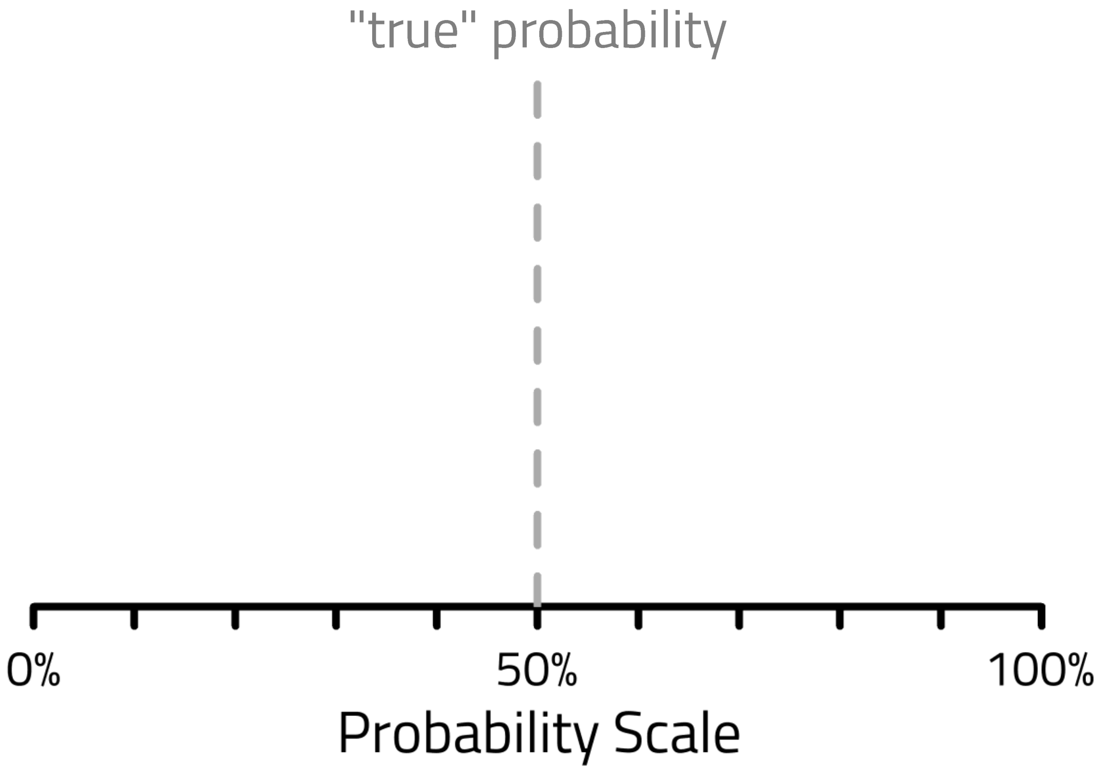
2016

Opinion

Bayesian Brains without Probabilities

Adam N. Sanborn[1,*] and Nick Chater[2]

What is the probability that the weather will be *cloudy* on a randomly selected day in Ireland?



Surprisingly rational: Probability theory plus noise explains biases in judgment.

Costello, Fintan,Watts, Paul

Psychological Review, Vol 121(3), Jul 2014, 463-480

The Mind is a **Bayesian Sampler**

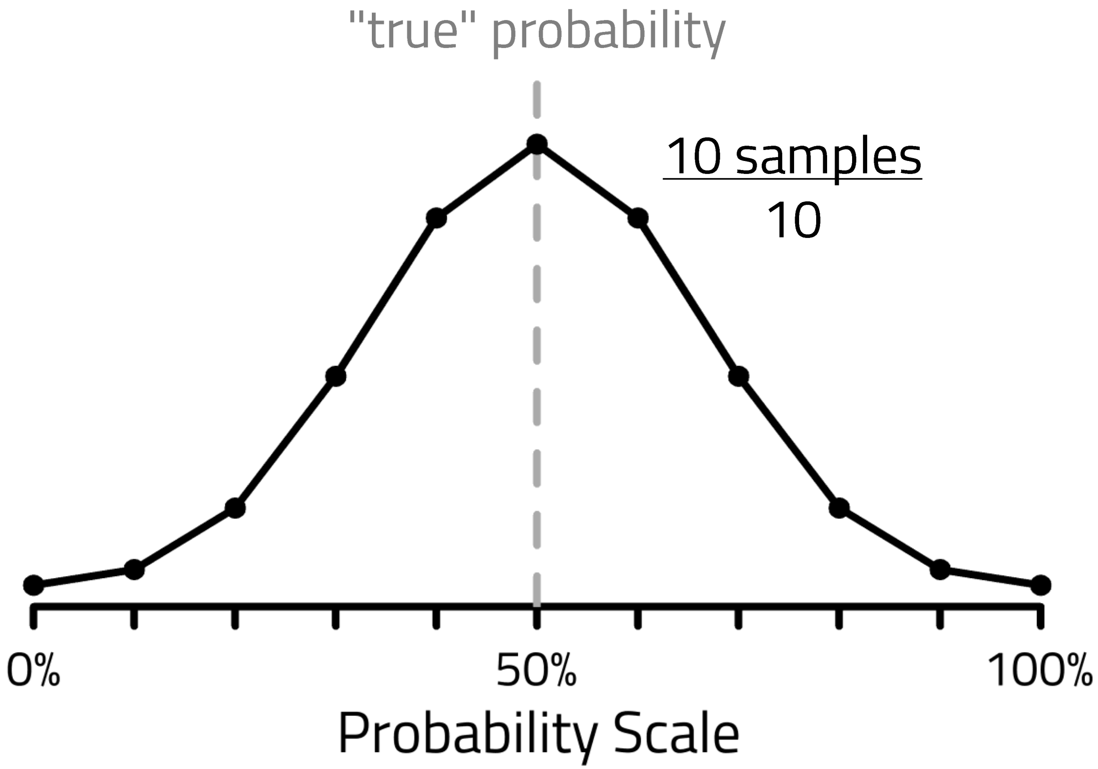
Trends in Cognitive Sciences
2016

Opinion

Bayesian Brains without Probabilities

Adam N. Sanborn[1,*] and Nick Chater[2]

What is the probability that the weather will be *cloudy* on a randomly selected day in Ireland?



Surprisingly rational: Probability theory plus noise explains biases in judgment.

Costello, Fintan,Watts, Paul

Psychological Review, Vol 121(3), Jul 2014, 463-480

The Mind is a **Bayesian Sampler**

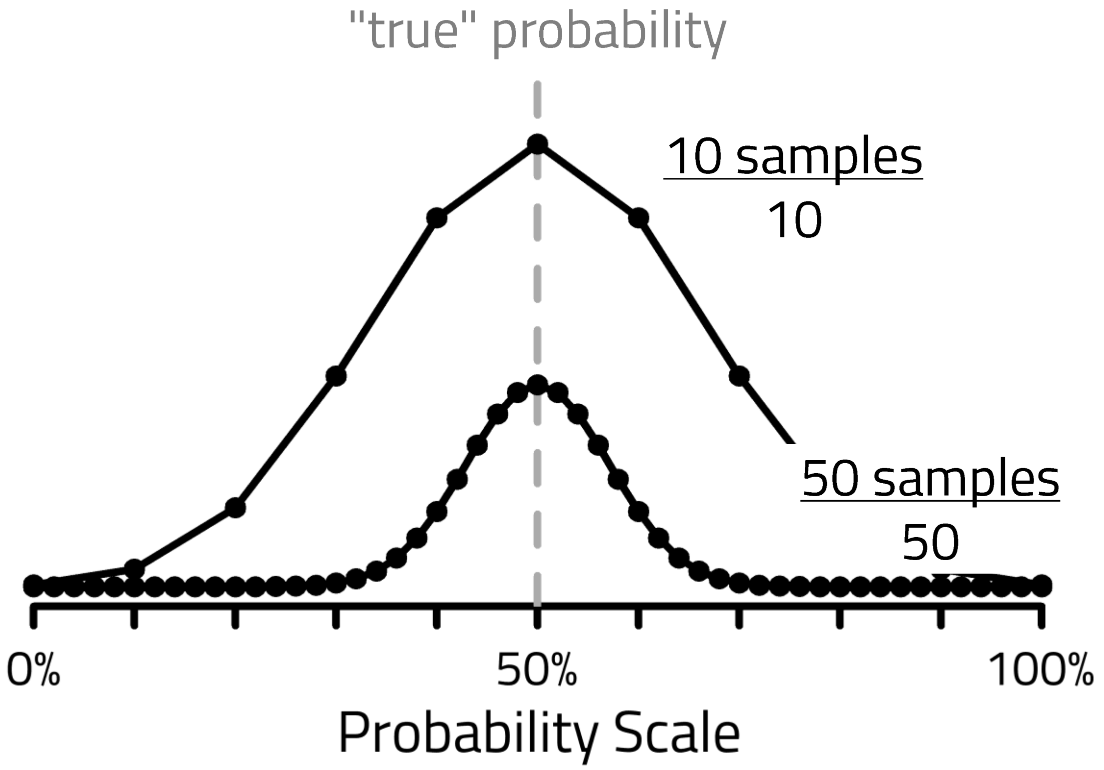
Trends in Cognitive Sciences
2016

Opinion

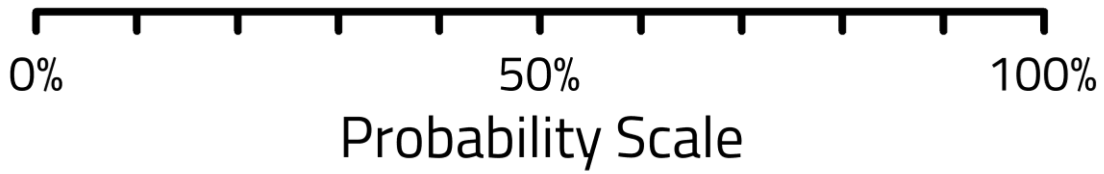
Bayesian Brains without Probabilities

Adam N. Sanborn[1,*] and Nick Chater[2]

What is the probability that the weather will be *cloudy* on a randomly selected day in Ireland?



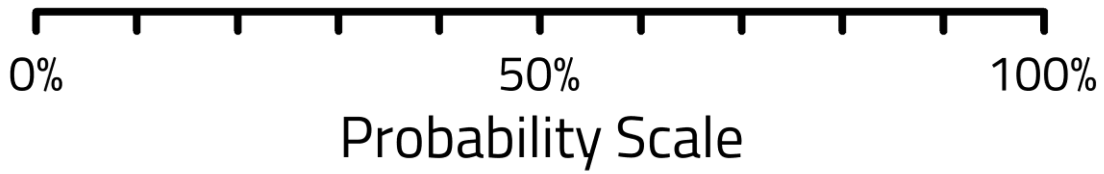
What is the probability that the weather will be … on a randomly selected day in Ireland?



01 | *cloudy?*

Costello & Watts (2014)

What is the probability that the weather will be … on a randomly selected day in Ireland?

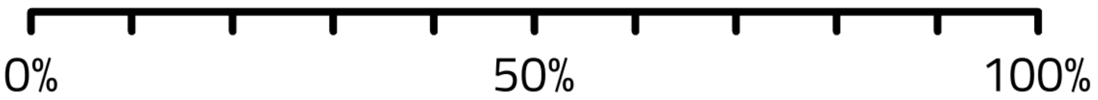


01 | *cloudy?*

02 | *cold?*

Costello & Watts (2014)

What is the probability that the weather will be ... on a randomly selected day in Ireland?



Probability Scale

01 | *cloudy?*

02 | *cold?*

03 | *cloudy or cold?*

Costello & Watts (2014)

What is the probability that the weather will be … on a randomly selected day in Ireland?

0% 50% 100%

Probability Scale

01 | *cloudy?*

02 | *cold?*

03 | *cloudy or cold?*

04 | *cloudy and cold?*

Costello & Watts (2014)

What is the probability that the weather will be ... on a randomly selected day in Ireland?

0% 50% 100%

Probability Scale

01 | *cloudy?*

02 | *cold?*

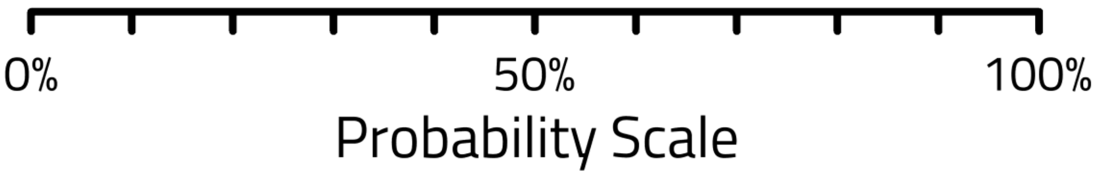
03 | *cloudy or cold?*

04 | *cloudy and cold?*

05 | *cloudy and not cold?*

Costello & Watts (2014)

What is the probability that the weather will be … on a randomly selected day in Ireland?



01 | *cloudy?*

02 | *cold?*

03 | *cloudy or cold?*

04 | *cloudy and cold?*

05 | *cloudy and not cold?*

06 | *cold and not cloudy?*

Costello & Watts (2014)

What is the probability that the weather will be ... on a randomly selected day in Ireland?

0% 50% 100%

Probability Scale

01 | *cloudy?*

02 | *cold?*

03 | *cloudy or cold?*

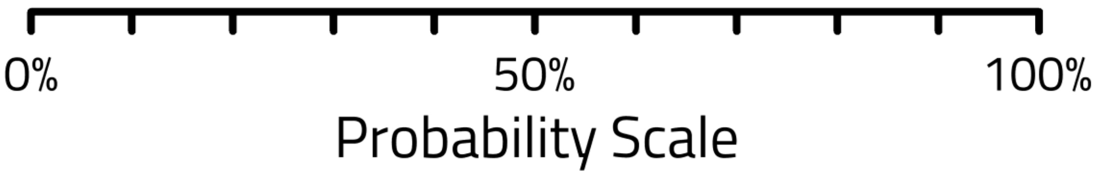
04 | *cloudy and cold?*

05 | *cloudy and not cold?*

06 | *cold and not cloudy?*

Costello & Watts (2014)

What is the probability that the weather will be … on a randomly selected day in Ireland?



01 | *cloudy?*

02 | *cold?*

03 | *cloudy or cold?*

04 | *cloudy and cold?*

05 | *cloudy and not cold?*

06 | *cold and not cloudy?*

Costello & Watts (2014)

What is the probability that the weather will be ... on a randomly selected day in Ireland?

0% 50% 100%

Probability Scale

01 | *cloudy?*

02 | *cold?*

03 | *cloudy or cold?*

joint probability distribution over *cold* & *cloudy*?

04 | *cloudy and cold?*

05 | *cloudy and not cold?*

06 | *cold and not cloudy?*

e.g., Oaksford & Chater (1994)

Noisy Sampler Generative Model for Probability Estimation

Joint Probability Distribution

Provides "true" probability

(e.g., Oakford & Chater, 2014)

Bayesian Sampling

Probabilistic transformation of

"true" probability

Memory & Response Noise

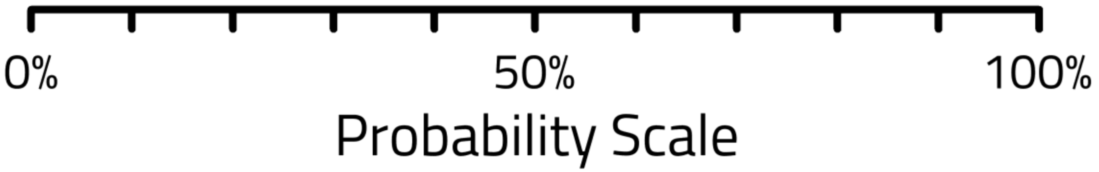
Sometimes sampling is erroneous

(Costello & Watts, 2014) *and*

sometimes response is noisy

Noisy Sampler

Discrete Probability Distribution on

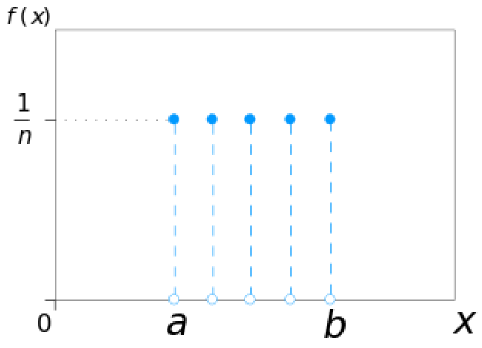


Mixture Model

$(1 - \phi)$

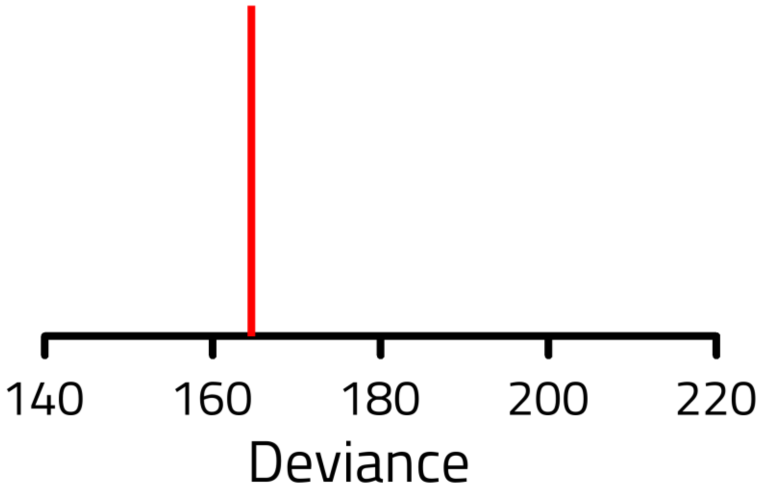
$$\hat{P}(A) = \frac{\text{Noisy number of } A \text{ instances}}{\text{Total number of samples}}$$

ϕ



Discrete uniform distribution to adjacent points

Noisy Sampler Generative Model for Probability Estimation



Joint Probability Distribution

Provides "true" probability

(e.g., Oakford & Chater, 2014)

Bayesian Sampling

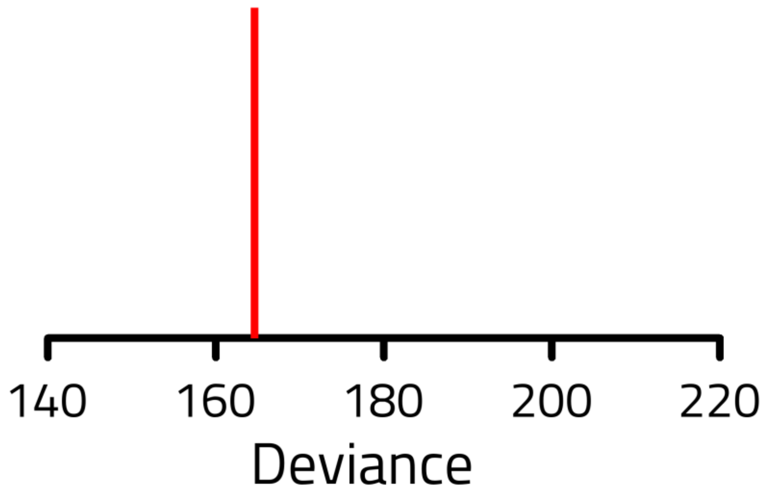
Probabilistic transformation of "true" probability

Memory & Response Noise

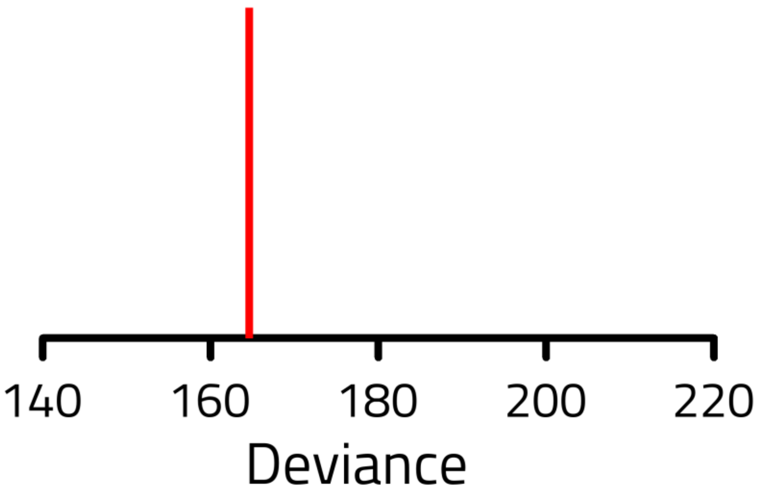
Sometimes sampling is erroneous (Costello & Watts, 2014) *and* sometimes response is noisy

Derive distribution of deviance under null hypothesis (H_0)

Noisy Sampler
Generative Model for Probability Estimation



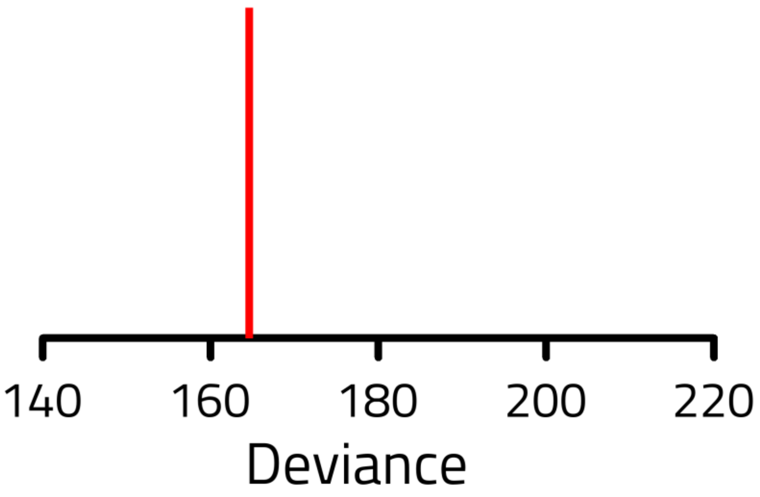
Noisy Sampler
Generative Model for Probability Estimation



Derive distribution of deviance under null hypothesis (H_0)

01 Generate synthetic data from best fitting parameter estimates

Noisy Sampler Generative Model for Probability Estimation

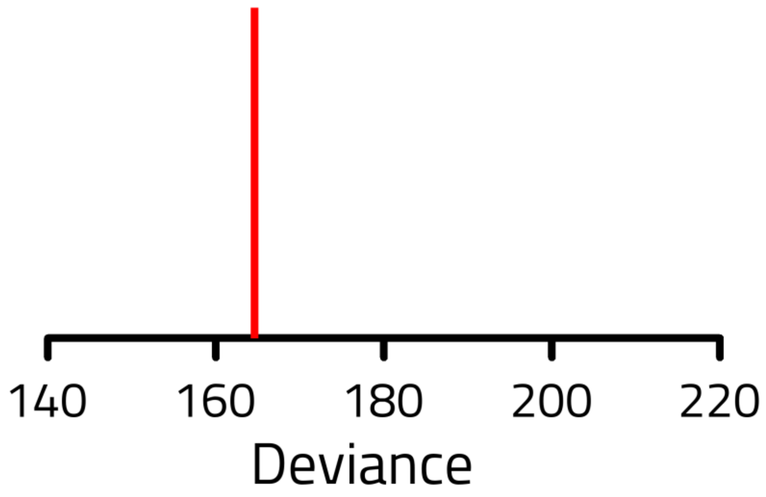


Derive distribution of deviance under null hypothesis (H_0)

01 Generate synthetic data from best fitting parameter estimates

02 Fit *noisy sampler* to synthetic data

Noisy Sampler
Generative Model for Probability Estimation



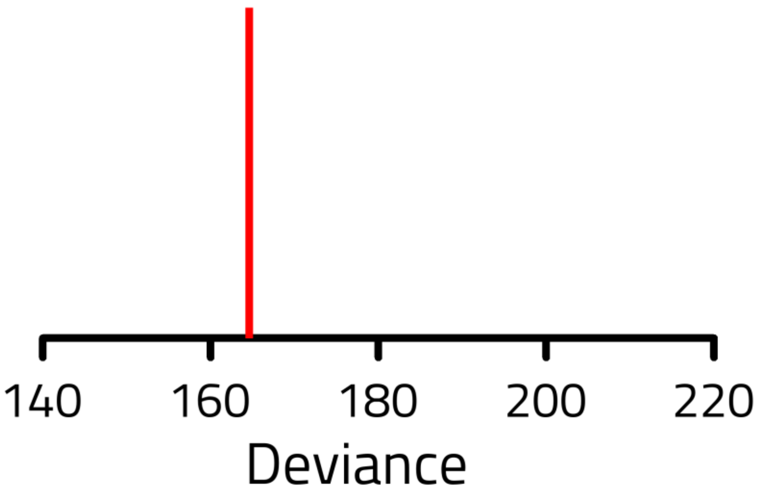
Derive distribution of deviance under null hypothesis (H_0)

01 Generate synthetic data from best fitting parameter estimates

02 Fit *noisy sampler* to synthetic data

03 Record deviance of fit to synthetic data

Noisy Sampler
Generative Model for Probability Estimation



Derive distribution of deviance under null hypothesis (H_0)

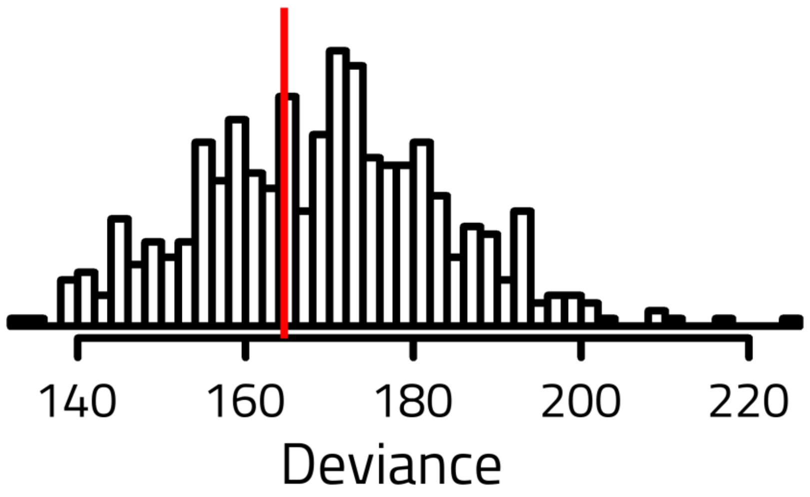
01 Generate synthetic data from best fitting parameter estimates

02 Fit *noisy sampler* to synthetic data

03 Record deviance of fit to synthetic data

repeat 500 times

Noisy Sampler
Generative Model
for Probability
Estimation



Derive distribution of deviance under null hypothesis (H_0)

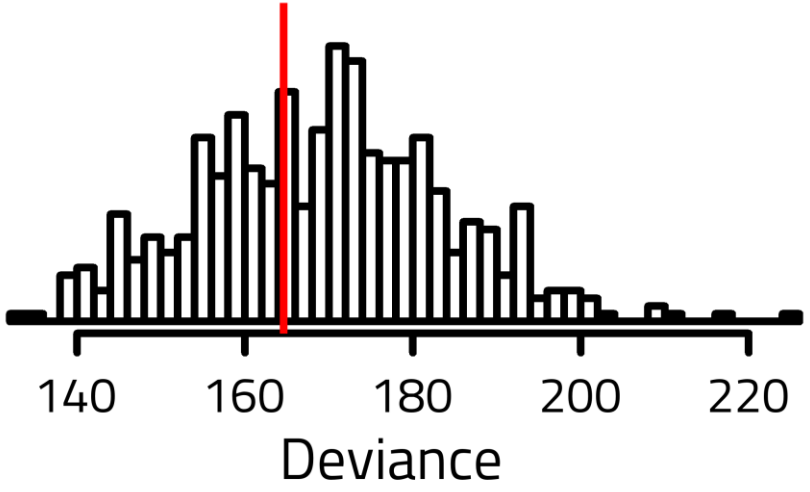
01 Generate synthetic data from best fitting parameter estimates

02 Fit *noisy sampler* to synthetic data

03 Record deviance of fit to synthetic data

repeat 500 times

Noisy Sampler Generative Model for Probability Estimation



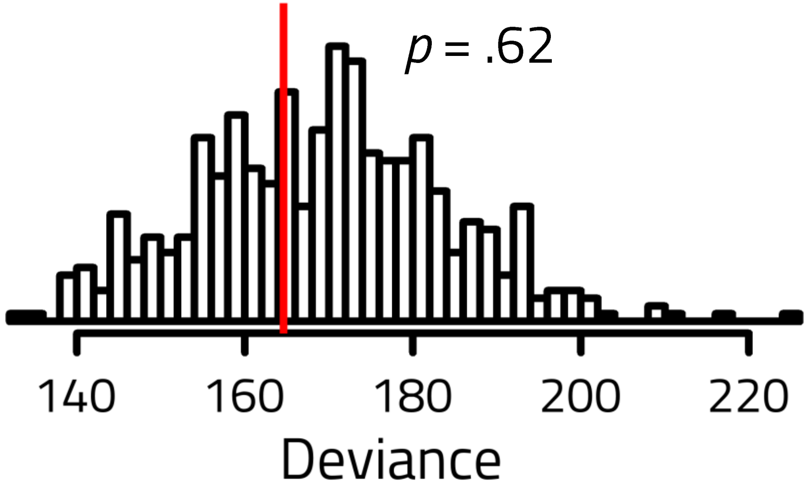
Derive distribution of deviance under null hypothesis (H_0)

- 01 Generate synthetic data from best fitting parameter estimates
- 02 Fit *noisy sampler* to synthetic data
- 03 Record deviance of fit to synthetic data

repeat 500 times

p-value: probability of sampling test statistic at least as extreme as observed under H_0

Noisy Sampler Generative Model for Probability Estimation



Derive distribution of deviance under null hypothesis (H_0)

01 Generate synthetic data from best fitting parameter estimates

02 Fit *noisy sampler* to synthetic data

03 Record deviance of fit to synthetic data

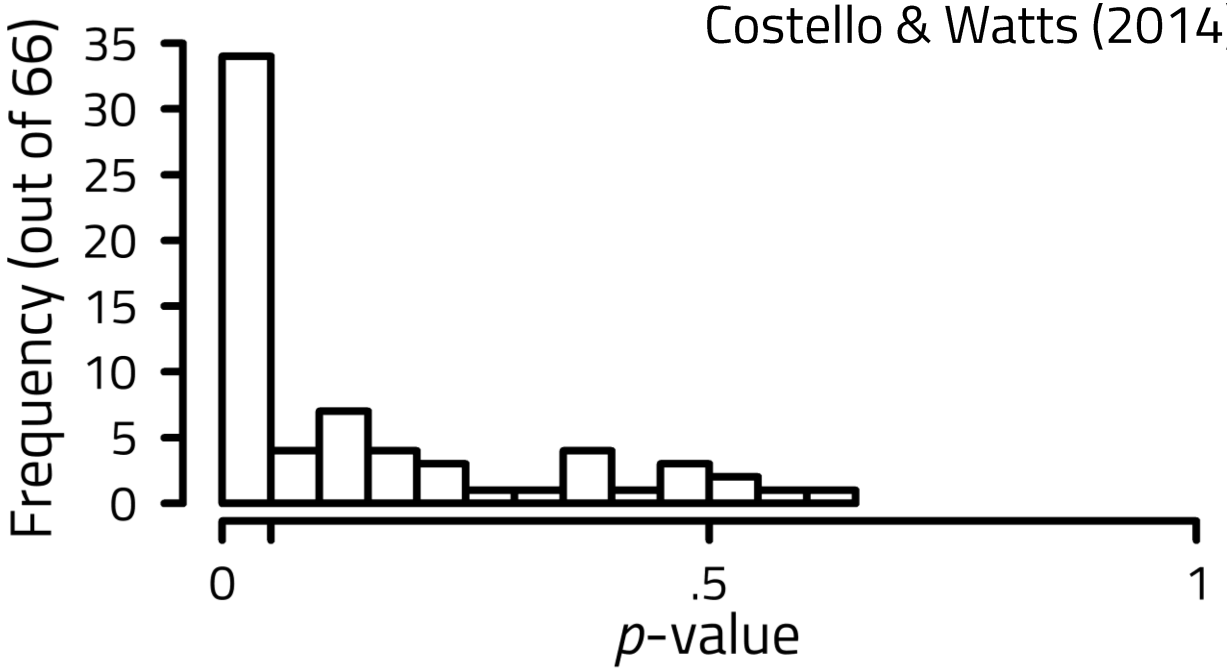
repeat 500 times

p-value: probability of sampling test statistic at least as extreme as observed under H_0

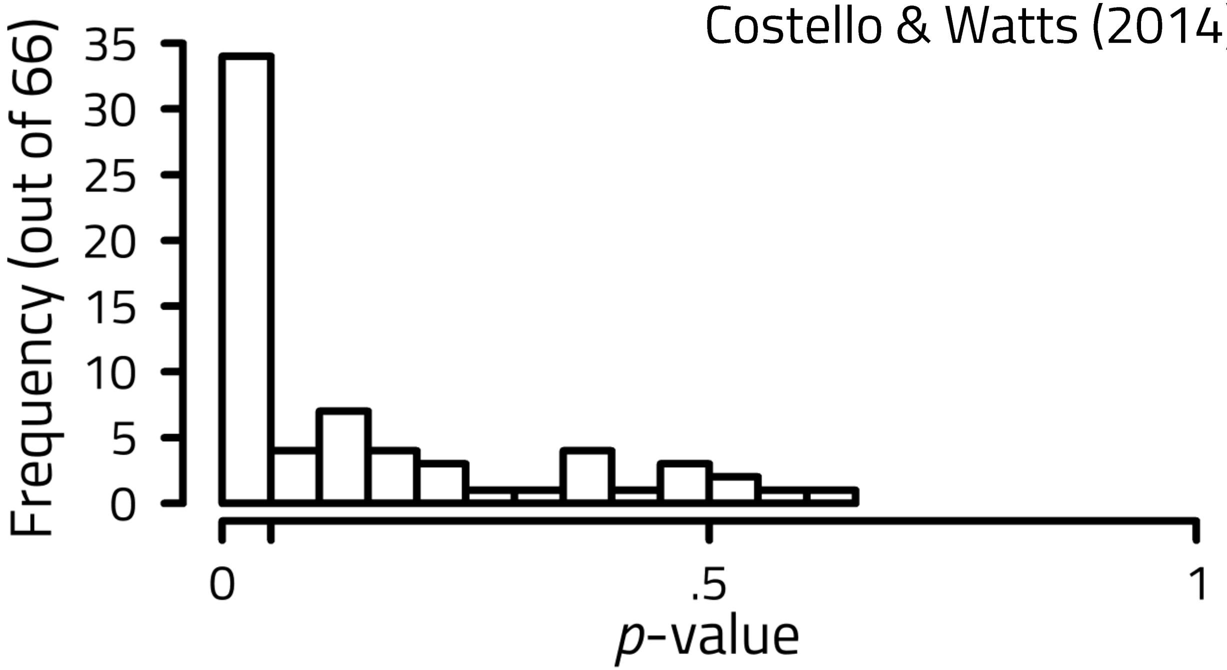
1.

Testing the Bayesian brain

Costello & Watts (2014)



Costello & Watts (2014)



1.

Testing the Bayesian brain

50% of participants Bayesian according to *noisy sampler*.

1.

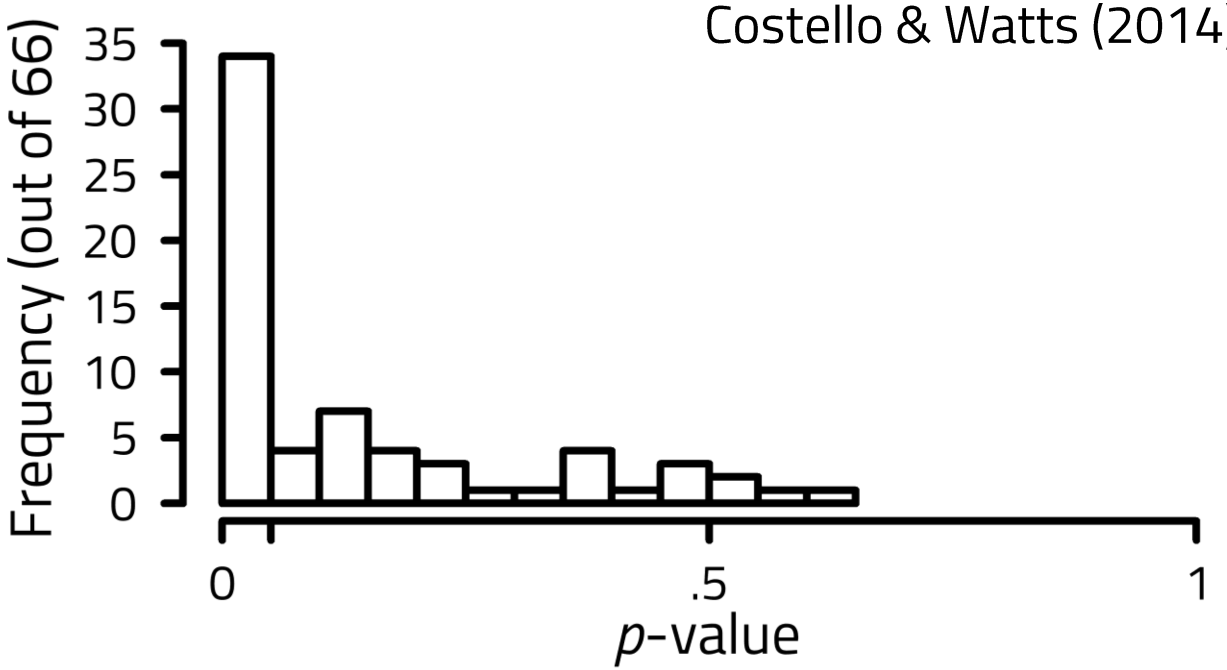
Testing the Bayesian brain

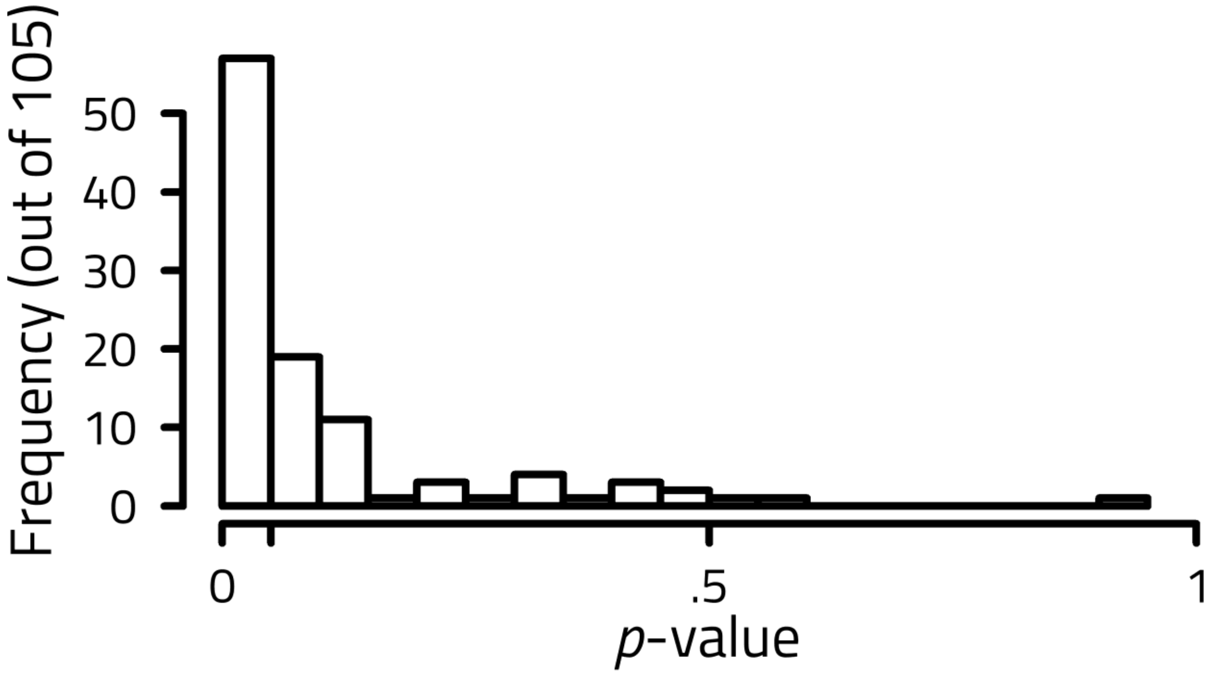
50% of participants Bayesian according to *noisy sampler*.

2.

MTurk online replication

46% of participants Bayesian according to *noisy sampler*.

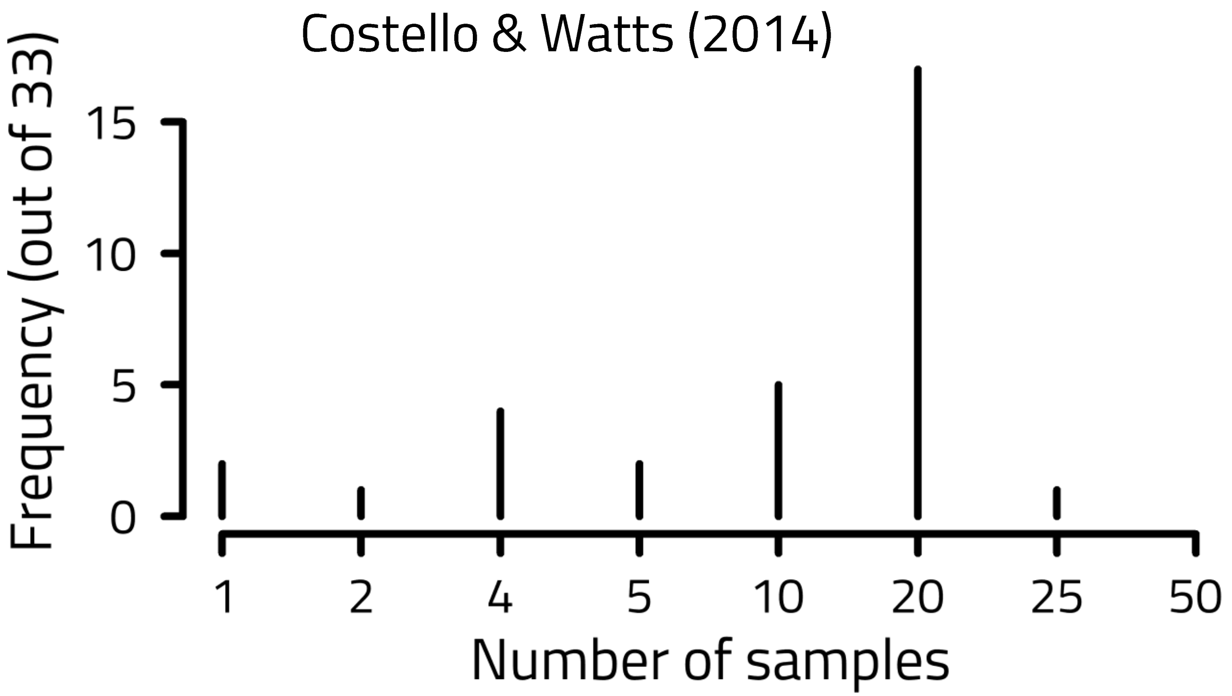




1.

How many samples do individuals take?

Restricted to *non-rejected* participants.



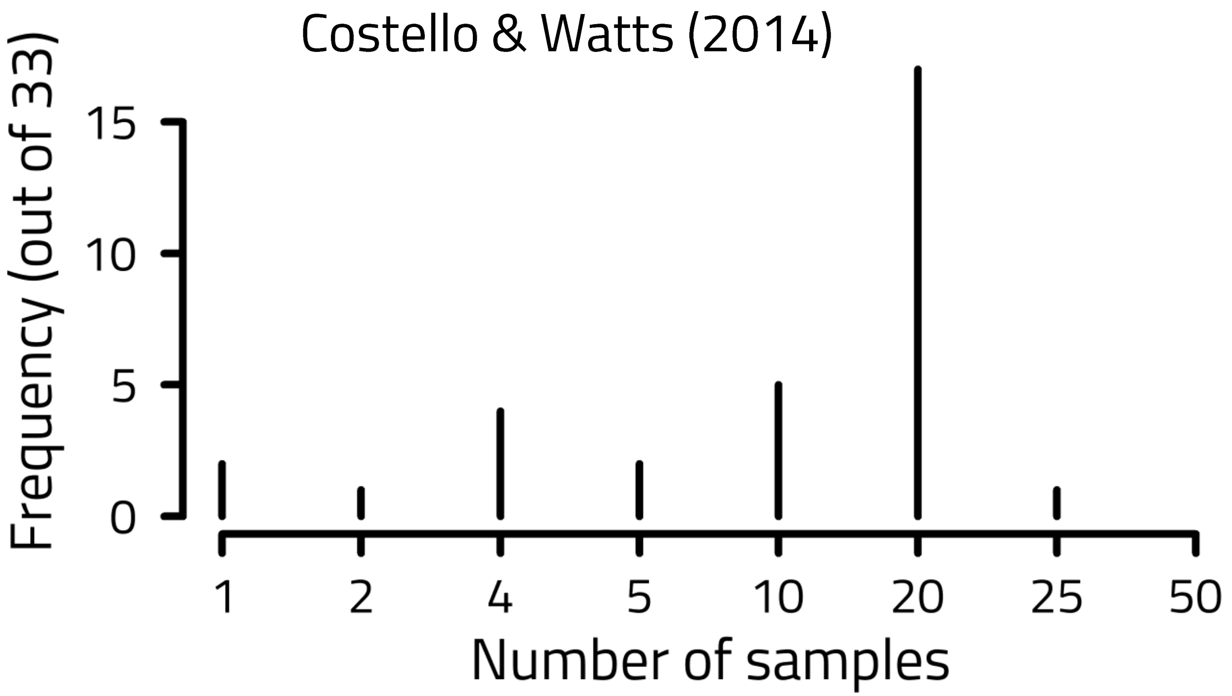
1.

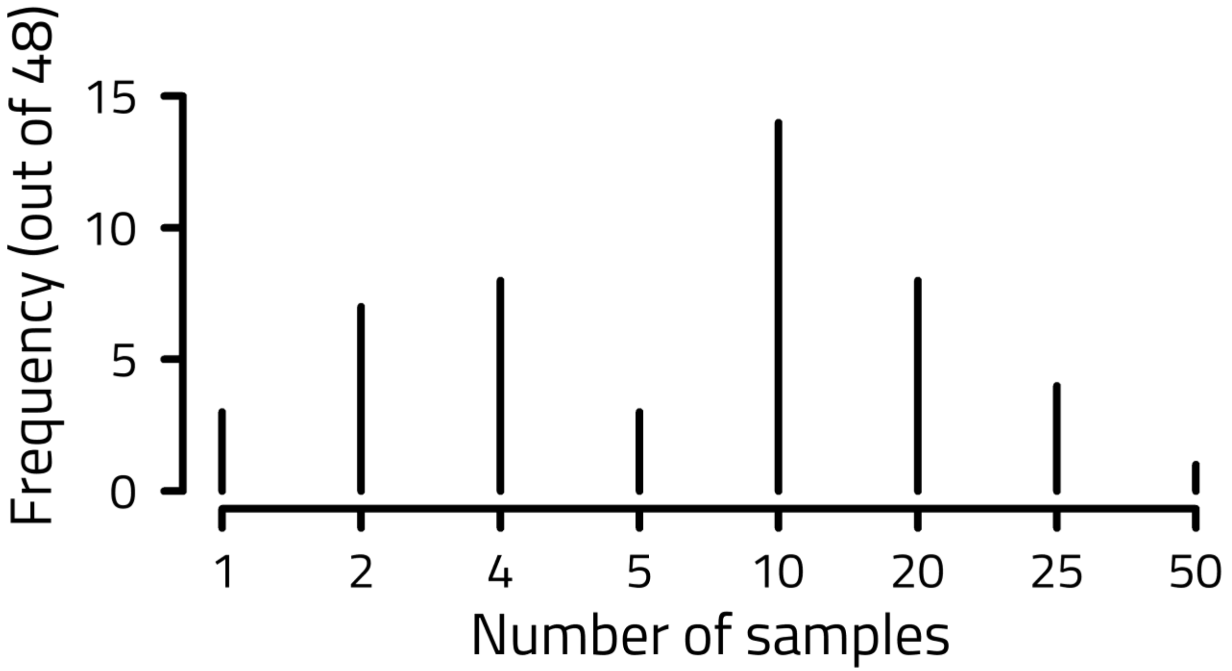
How many samples do individuals take?

Restricted to *non-rejected* participants.

2.

MTurk online replication





Does sampling save
the Bayesian brain?

THANK YOU FOR YOUR ATTENTION!

Questions?

cold – cloudy

cold – sunny

windy – cold

cloudy – rainy

cloudy – icy

sunny – icy

sunny – rainy

icy – windy

rainy – windy

= 42 independent data points per participant

Costello & Watts (2014)

MODELING

19 for joint probability distributions
1 mixture weight
1 noise parameter
1 number of samples

cold – cloudy

cold – sunny

windy – cold

cloudy – rainy

cloudy – icy

sunny – icy

sunny – rainy

icy – windy

rainy – windy

= 42 independent data points per participant

Costello & Watts (2014)