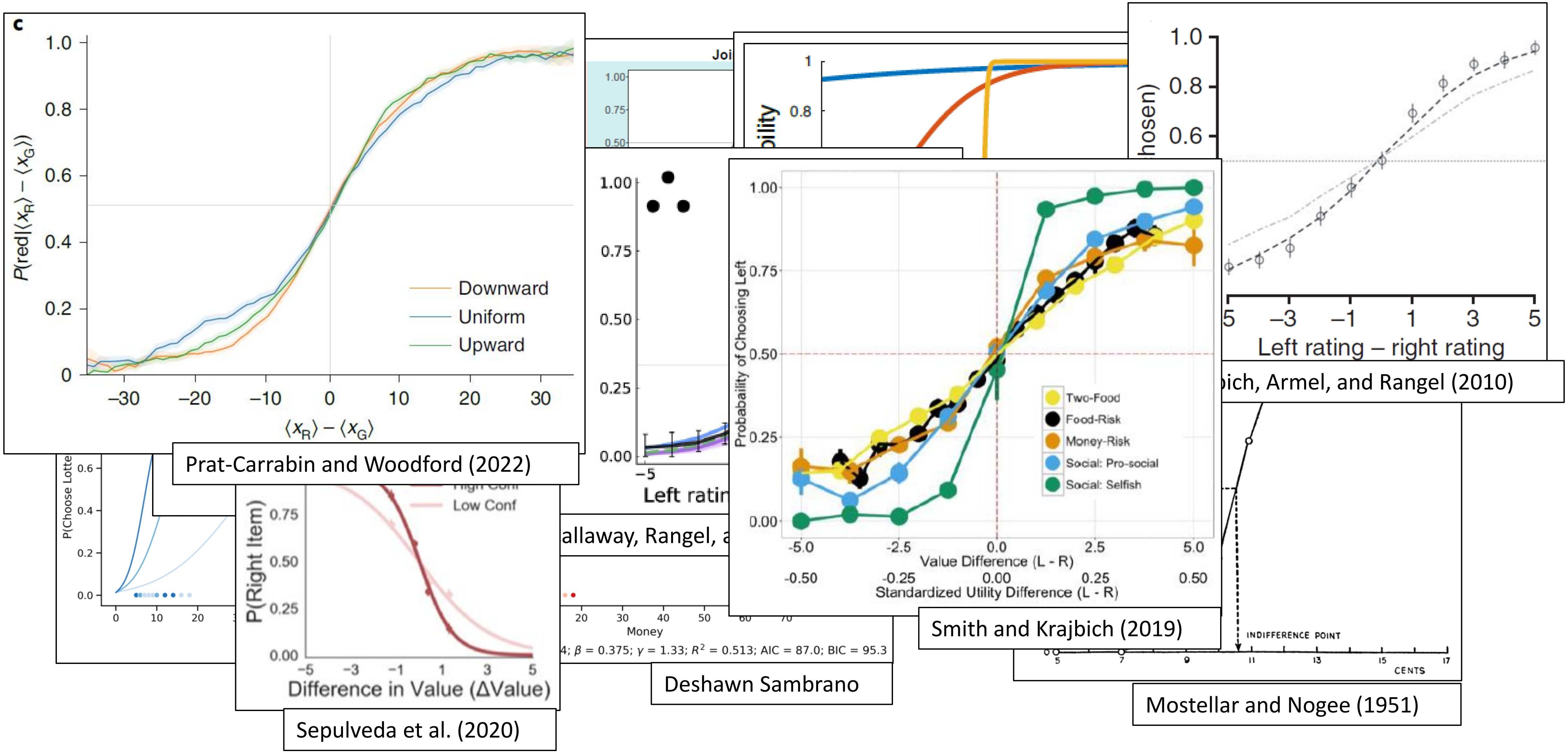
Causal manipulation of early noise biases numerosity judgements towards adaptive prior means

Brenden Eum *slides
Antonio Rangel
Michael Woodford

RNL Lab Meeting - SDN Proseminar

Choices exhibit randomness



Prat-Carrabin and Woodford (2022)

…allaway, Rangel, a…

Smith and Krajbich (2019)

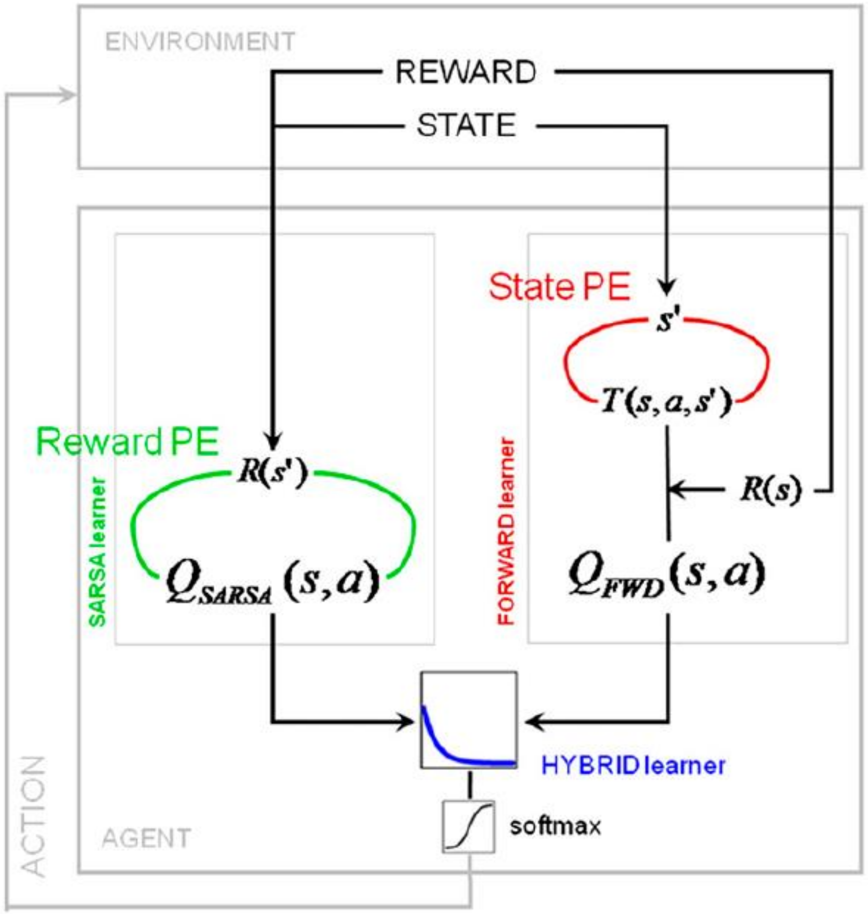
…bich, Armel, and Rangel (2010)

Sepulveda et al. (2020)

Deshawn Sambrano

Mostellar and Nogee (1951)

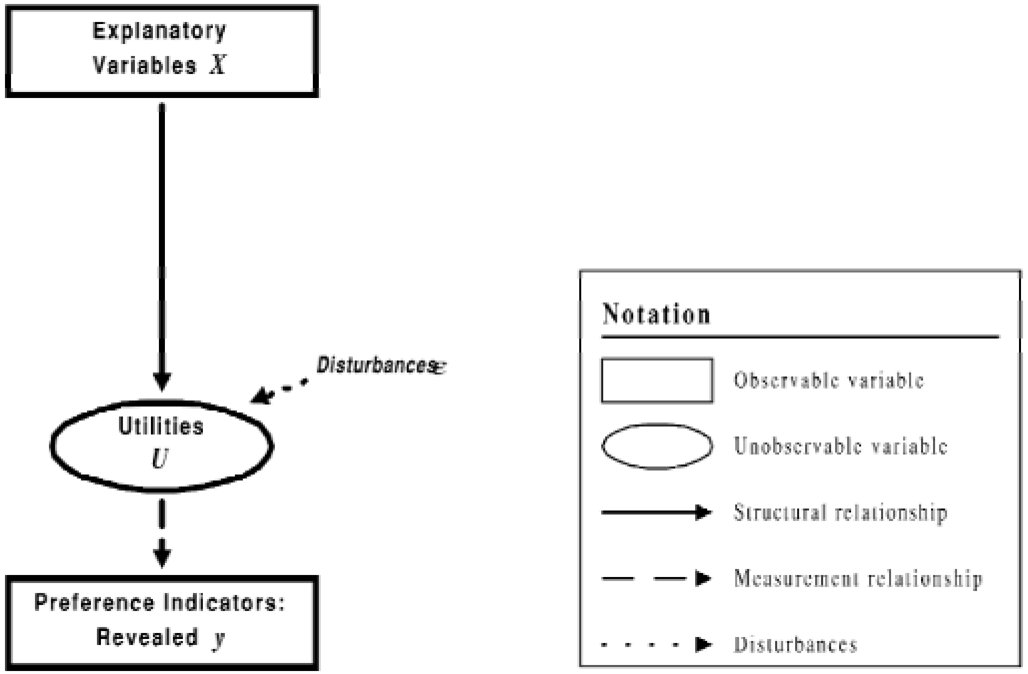
Modeling randomness: Softmax



Glascher, Daw, Dayan, and O'Doherty (2010)

Modeling randomness: Random utility models

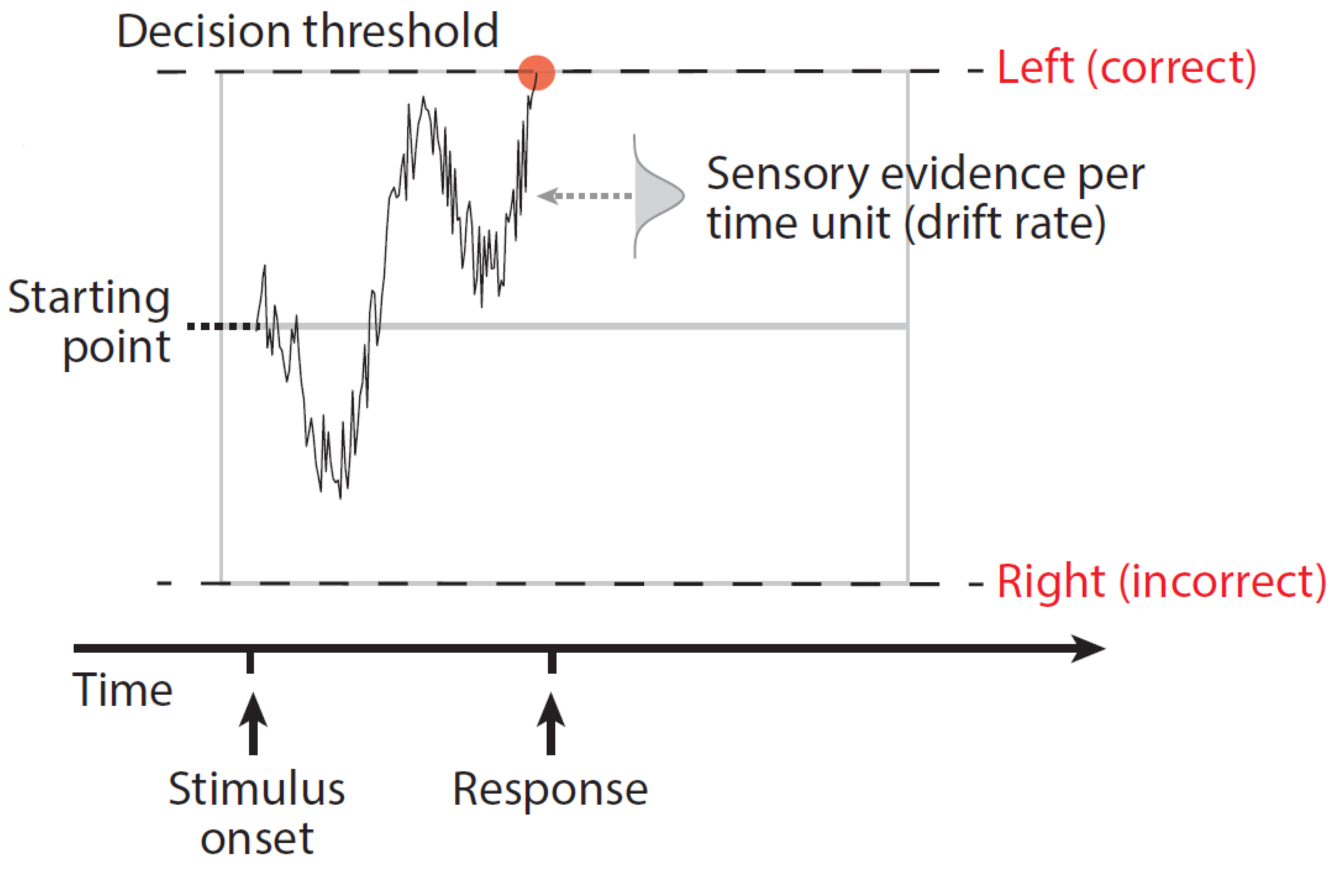
$$U = V(X;\beta) + \epsilon$$



Walker and Ben-Akiva (2002)

Modeling randomness: Drift-Diffusion-Model

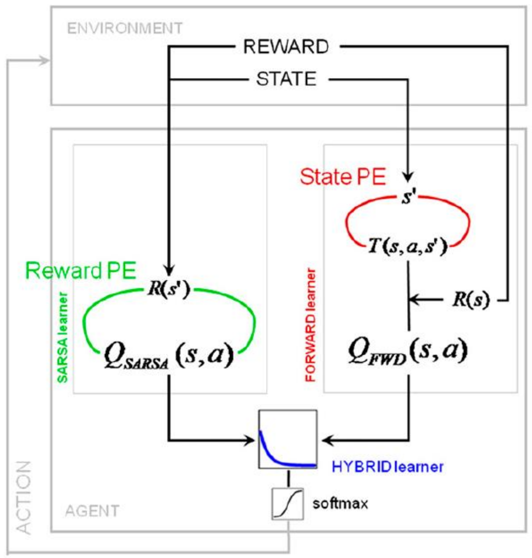
$$E_t = E_{t-1} + \mu + \epsilon$$

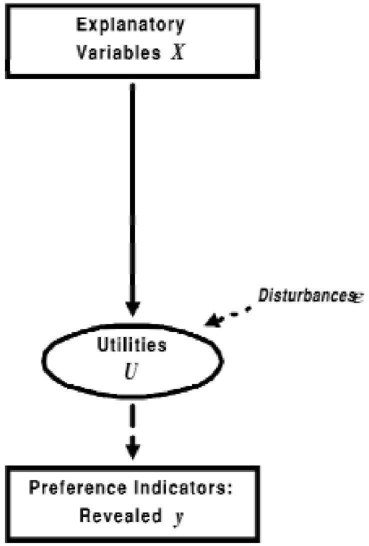


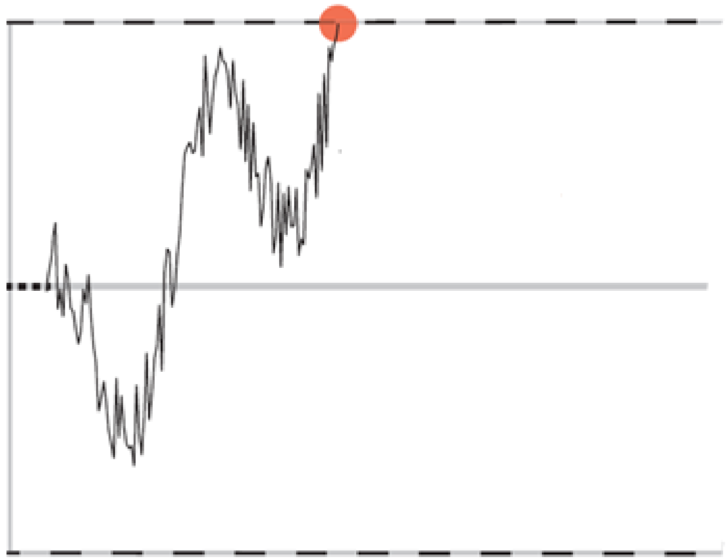
Forstmann, Ratcliff, and Wagenmakers (2016)

“Late” noise

Noise is added at the time of decision making.

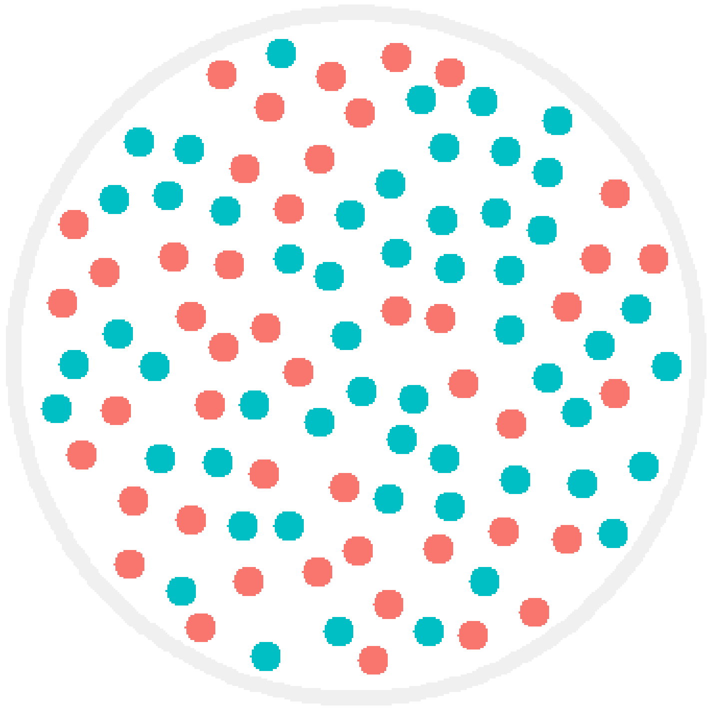


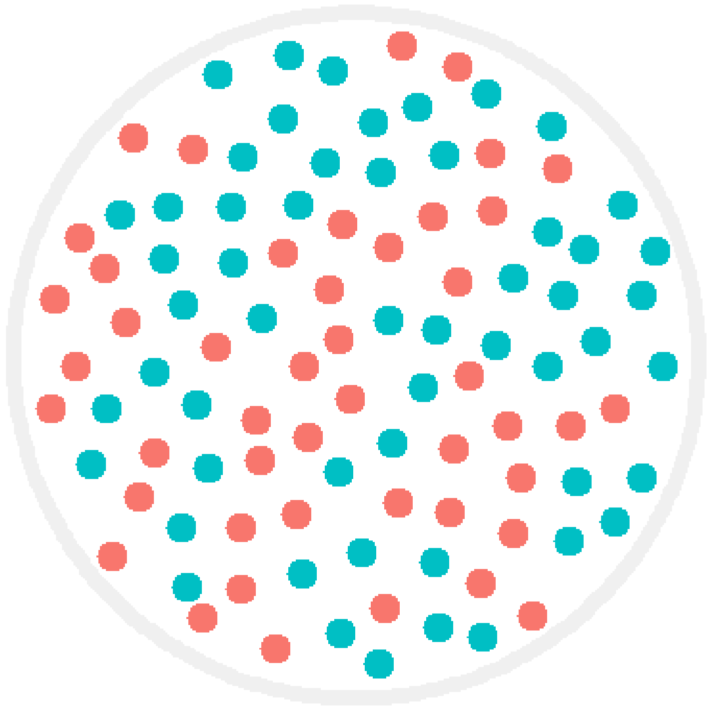


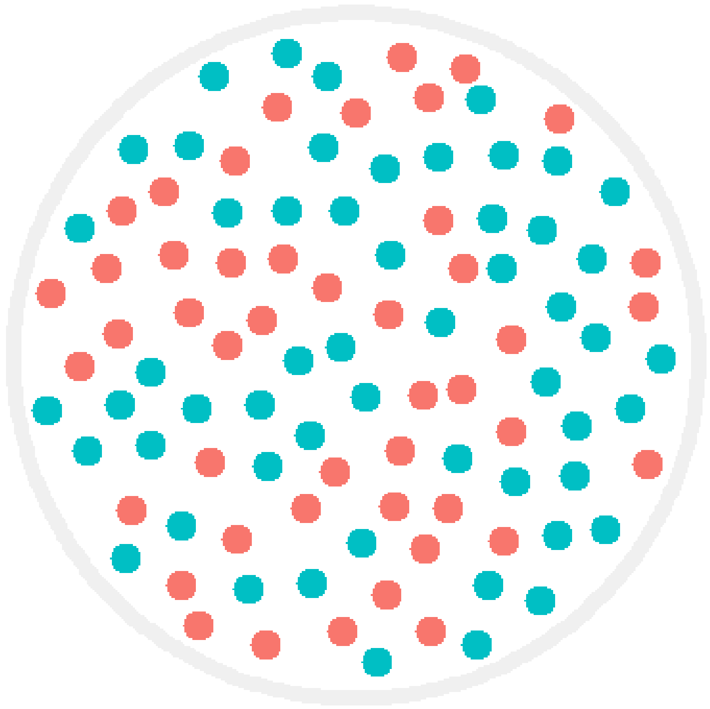


+

Guess how many blue dots there are on each of the following screens.







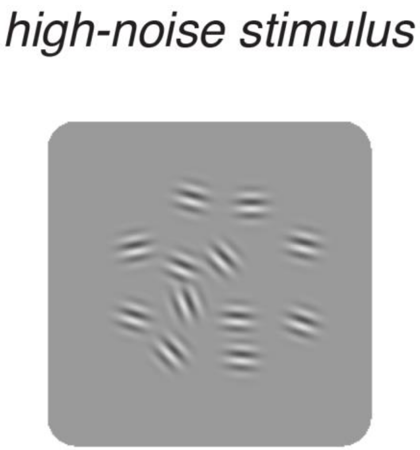
54

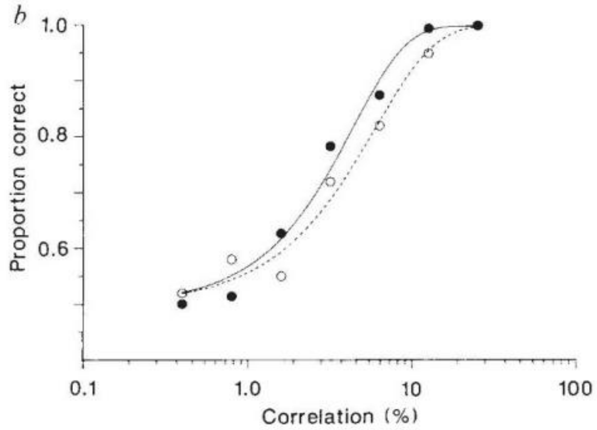
Is noise entering into the decision-making process solely at the late stage?

Where else might it be entering at?

“Early” noise

Noise is added at the time of stimulus encoding.





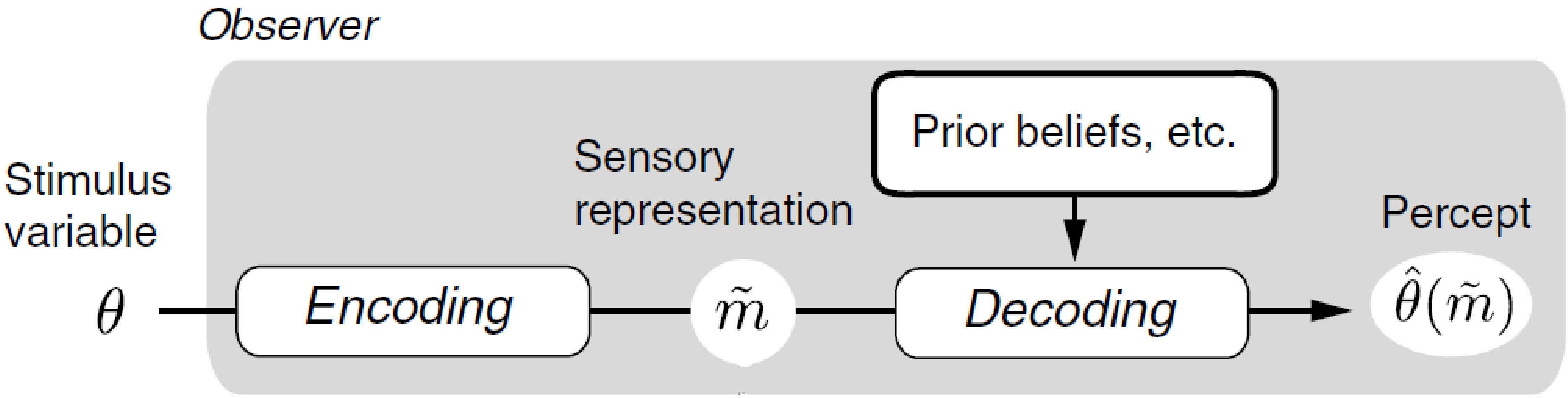
> “Can you do Addition?” the White Queen asked. “What’s one and one and one and one and one and one and one and one and one and one?”
>
> “I don’t know,” said Alice. “I lost count.”
>
> “She can’t do Addition,” the Red Queen interrupted.

Wei and Stocker (2012)

Newsome, Britten, and Movshon (1989)

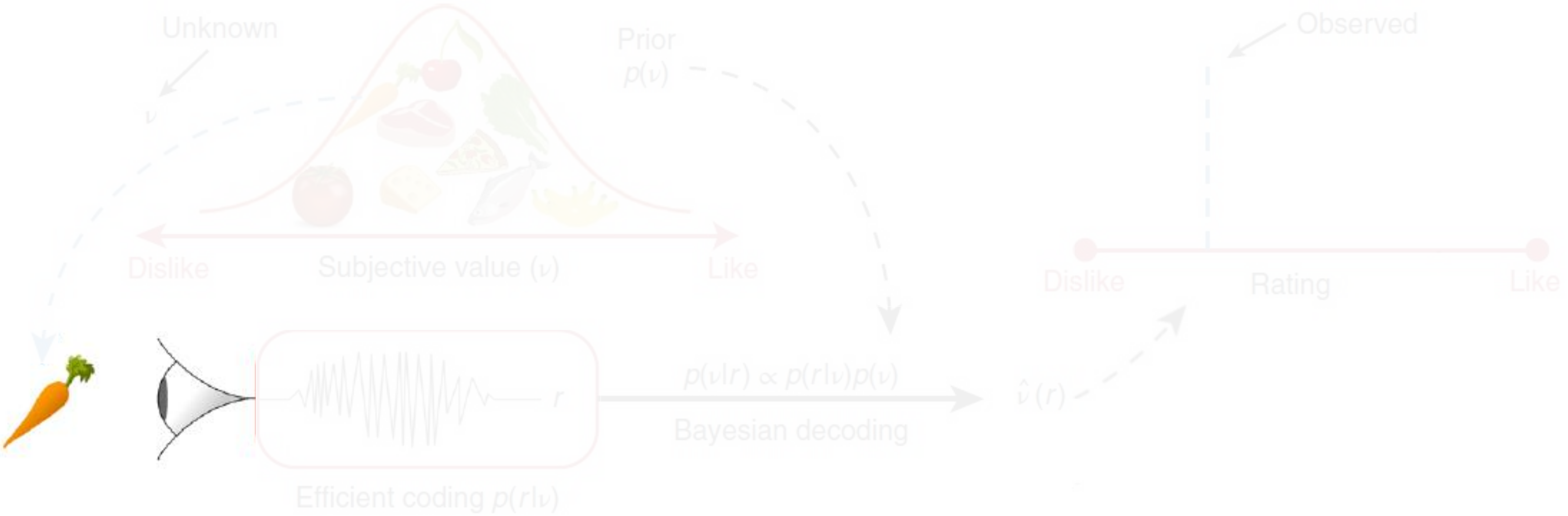
Quote from *Through the Looking Glass* by Lewis Carroll
Example from *The Number Sense: How the Mind Creates Mathematics* by Stanislas Dehaene

Encoding-decoding models



Wei and Stocker (2017)

Encoding-decoding models

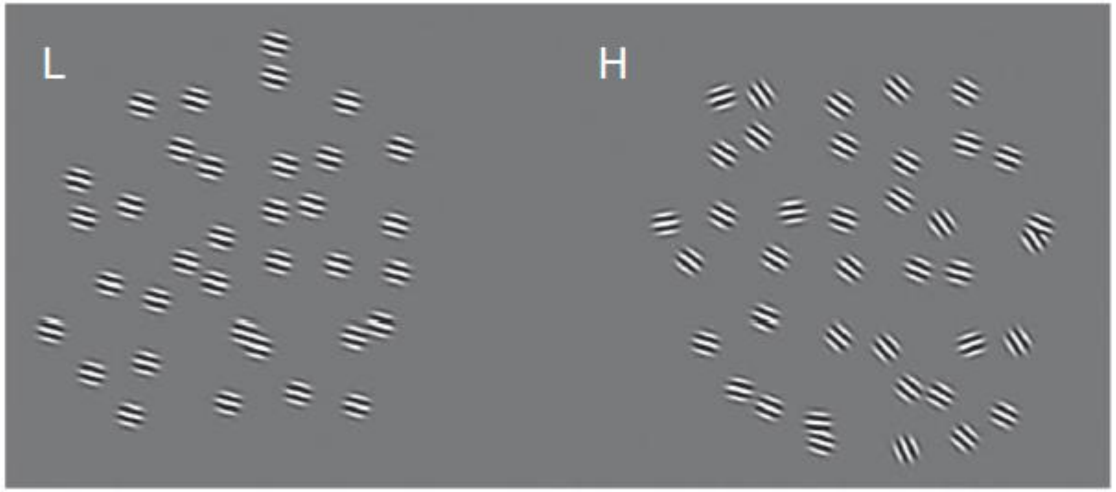


Polania, Woodford, and Ruff (2019), with edits

Model Class 1: Stable naturalistic priors

In perception: Olshausen and Field (1996); Simoncelli and Olshausen (2001); Olshausen and Field (2004); Geisler (2011); Girshick, Landy, and Simoncelli (2011)

In numerosity perception: Dehaene (1997); Cheyette and Piantadosi (2020)



Model Class 1: Stable naturalistic priors

In perception: Olshausen and Field (1996); Simoncelli and Olshausen (2001); Olshausen and Field (2004); Geisler (2011); Girshick, Landy, and Simoncelli (2011)

In numerosity perception: Dehaene (1997); **Cheyette and Piantadosi (2020) [CP (2020)]**

How often is numerosity x encountered and represented follows a power law: $P(x) \propto \frac{1}{x^2}$ (Piantadosi and Cantlon 2017)

CP (2020): Efficient encoding

$Q(r|x)$ is the probability that observed numerosity x is represented internally with quantity r.

Optimize: $Q(r|x)$ is chosen to minimize expected squared error between r and x.

$$E[(x-r)^2] = \sum_x P(x) \sum_r Q(r|x)(x-r)^2$$

Cost: KL divergence between natural prior and likelihood.

$$D_{KL}[P(x)||Q(r|x)] \leq \min(B, Rt)$$

CP (2020): Decoding

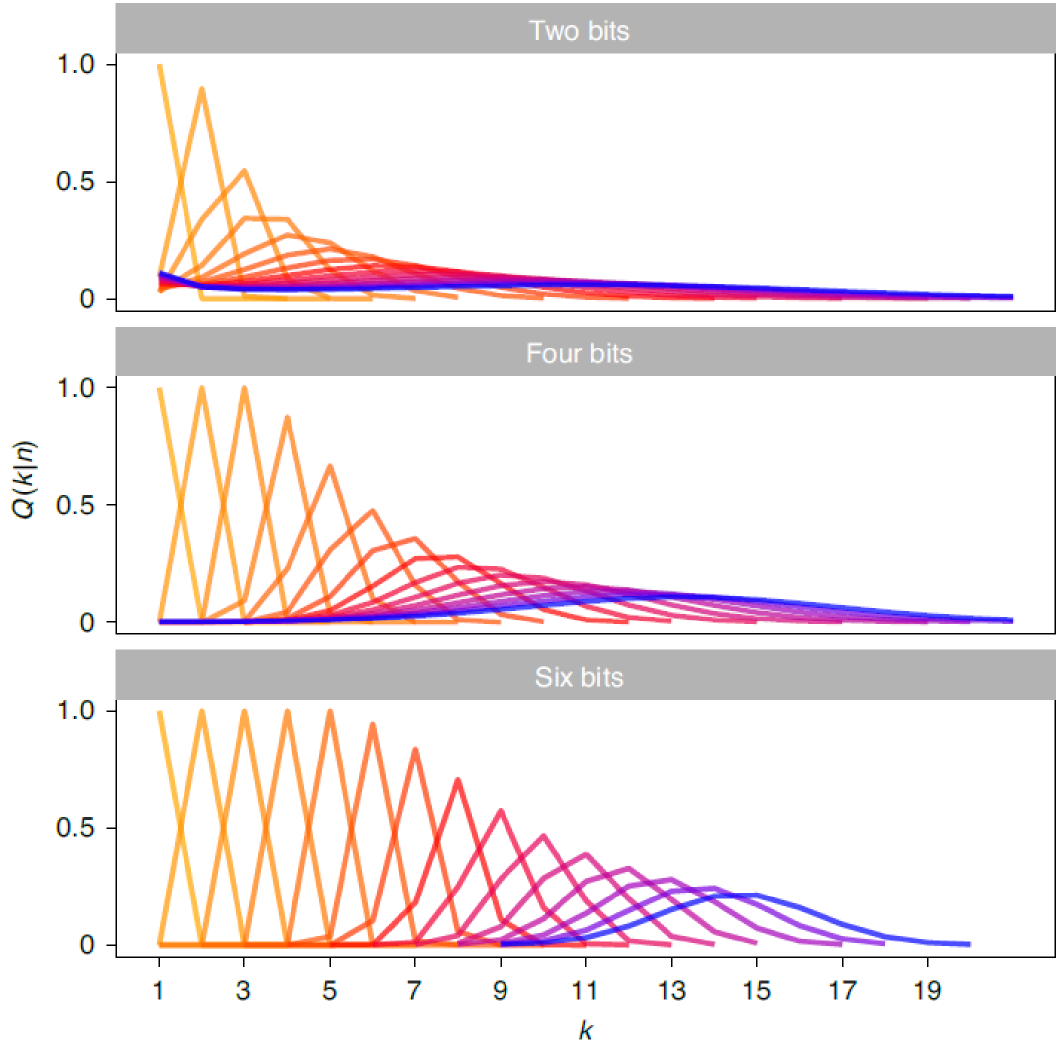
Analytical solution:

$$Q(r|x) \propto P(r) \exp\left(-\left(\frac{P(x)}{\lambda_x}\right)(x-r)^2\right)$$

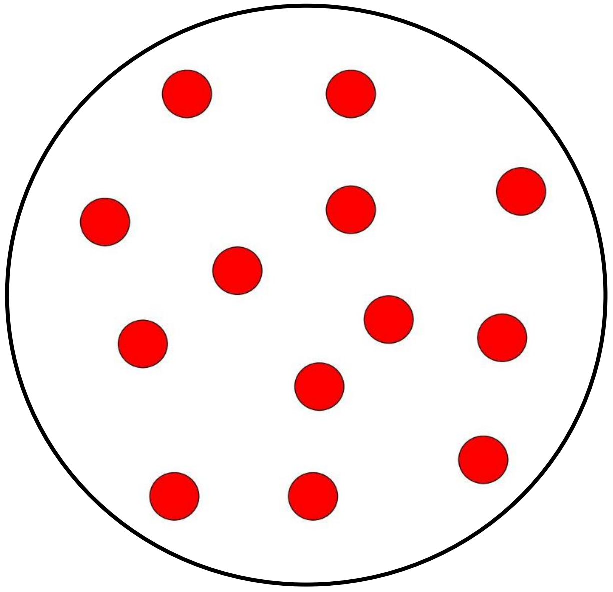
which ends up **looking like a weighted Gaussian** with variance $\frac{\lambda_x}{2P(x)}$.

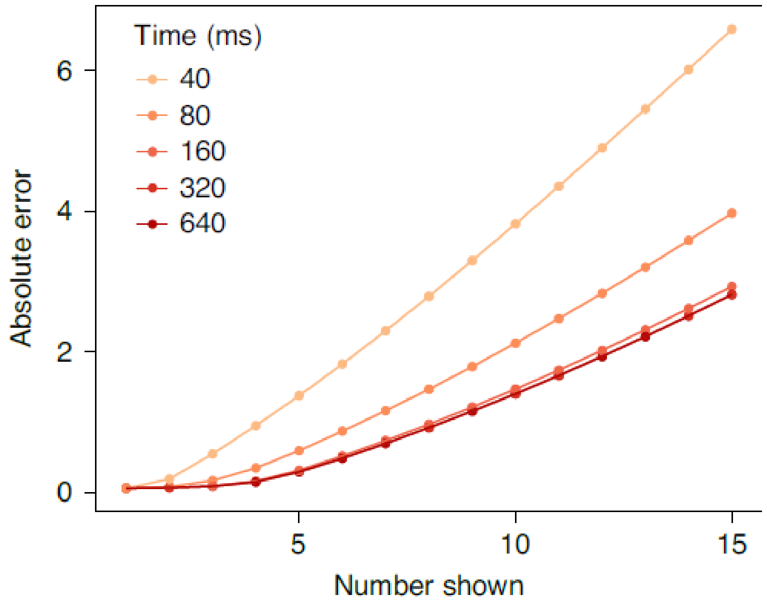
* Because $P(x)$ follows a decreasing power law, this variance grows in x.

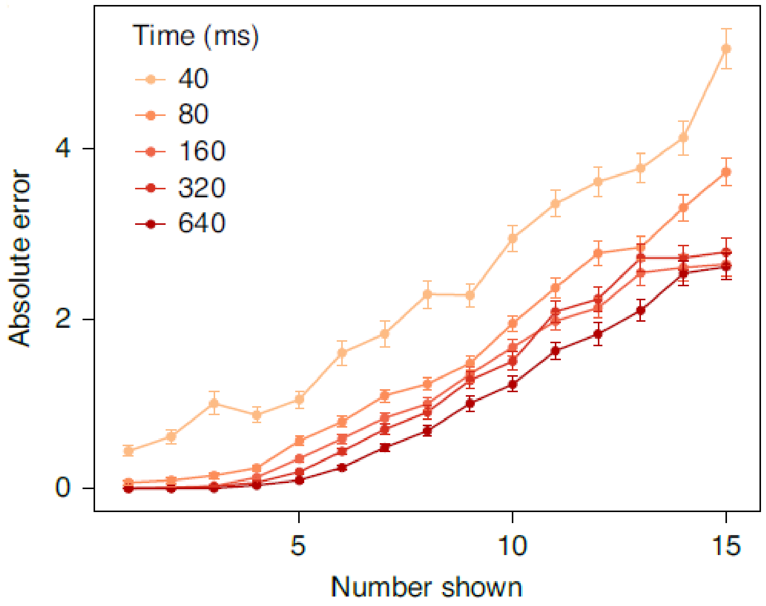
CP (2020)



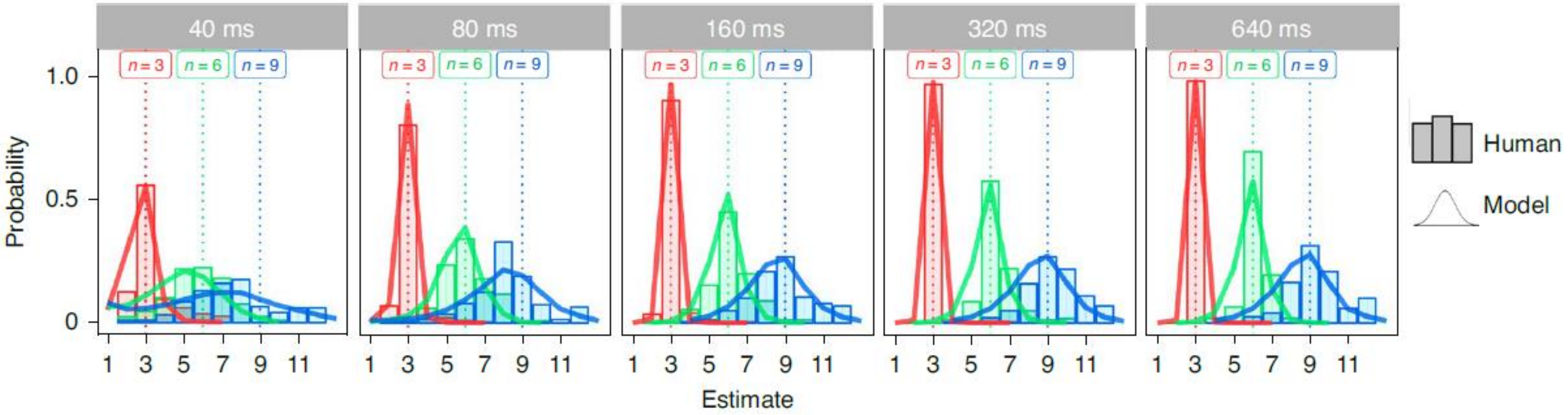
CP (2020)




Time (ms)
40
80
160
320
640
Absolute error
0
2
4
6
5
10
15
Number shown


Time (ms)
40
80
160
320
640
Absolute error
0
2
4
5
10
15
Number shown

CP (2020)

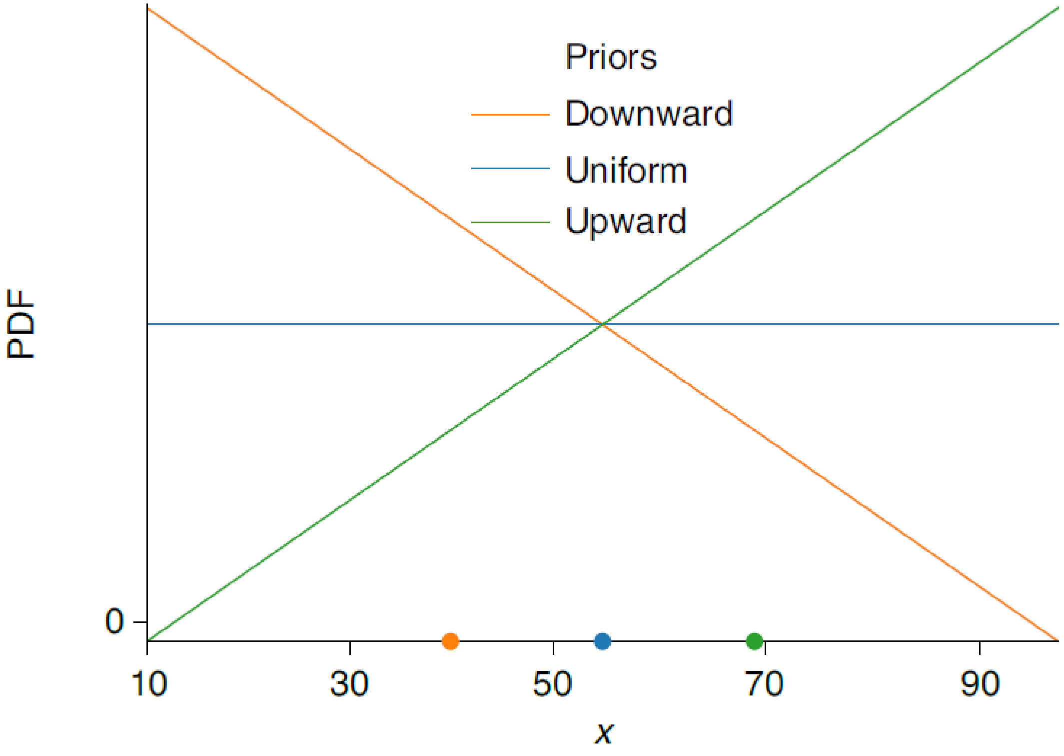


40 ms
80 ms
160 ms
320 ms
640 ms
n = 3
n = 6
n = 9
1.0
0.5
0
Probability
1 3 5 7 9 11
Estimate
Human
Model

Model Class 2: Efficient coding, Bayesian decoding (ECBD)

Why can't the prior adapt to different contexts?

[Prat-Carrabin et al. (2021), **Prat-Carrabin and Woodford (2022)**]



ECBD: Efficient coding

Remember x is the numerosity, r is its representation.

$r = (r_1, \ldots, r_n)$ where $\forall i, r_i \sim p(r_i|x)$.

Optimize the amount of information your representation carries about the true magnitude by minimizing expected error…

$$\min \int \hat{P}(x) I(r|x)^{-\frac{\rho}{2}} dx$$

Subject to a bound on informativeness…

$$\int \sqrt{I(r|x)} dx \leq B$$

ECBD: Bayesian decoding

The optimal Fisher information:

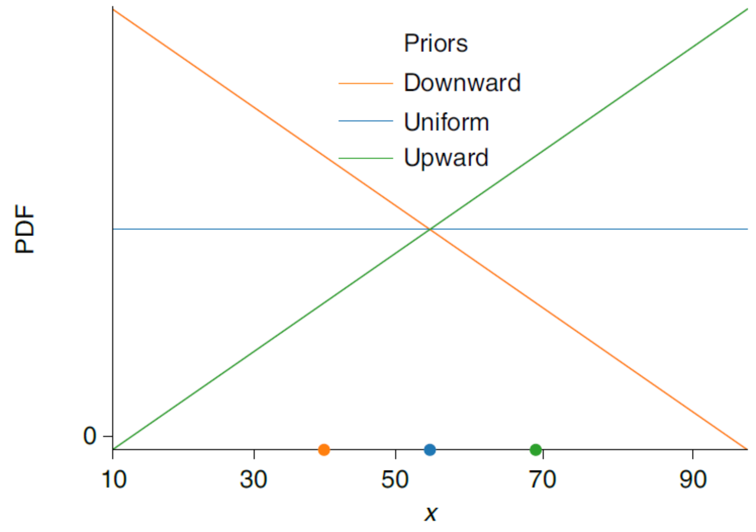
$$I^*(r|x) \propto \hat{P}(x)^{\frac{2}{\rho+1}}$$

If $\frac{d}{dx}\hat{P}(x) > 0$, then $\frac{d}{dx} bias(x) < 0$.

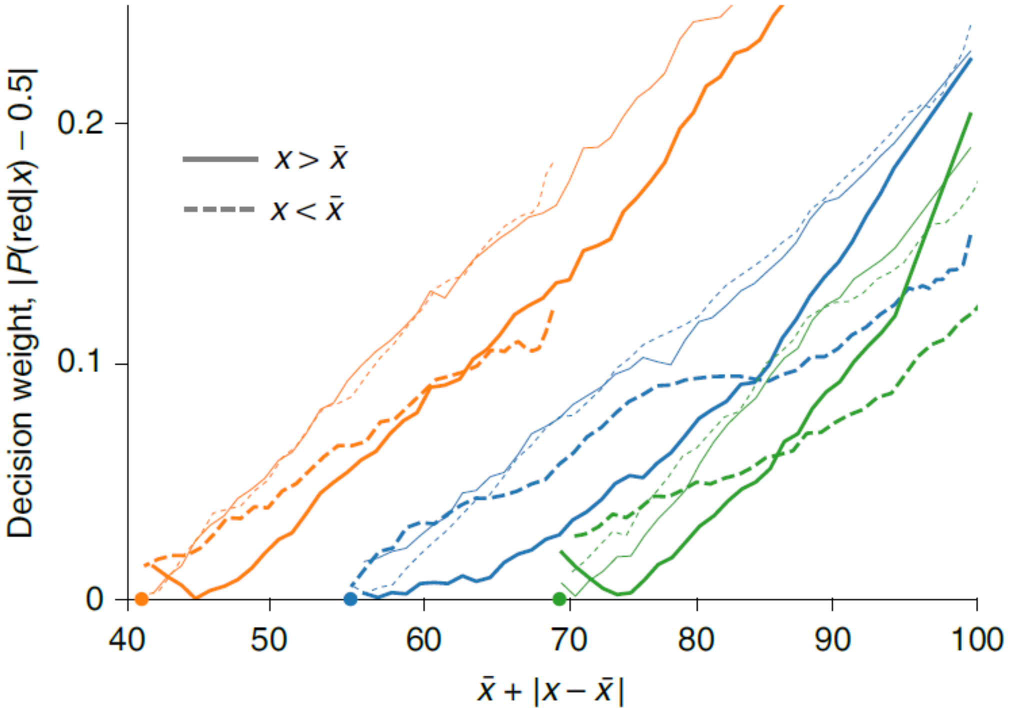
If $\frac{d}{dx}\hat{P}(x) < 0$, then $\frac{d}{dx} bias(x) > 0$.

The approximate distribution of estimates:

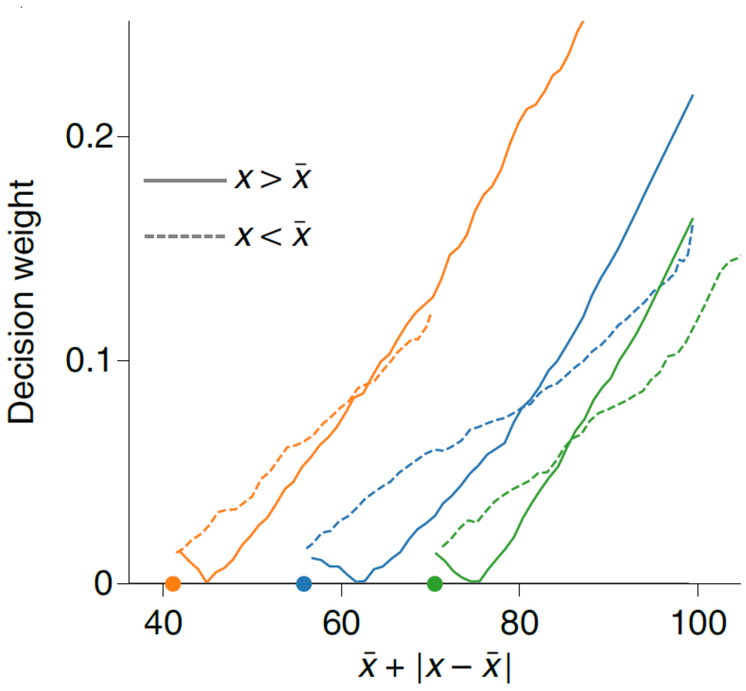
$$\hat{x}(r) \sim N\left(x + \underbrace{f\left(\frac{d}{dx}\frac{1}{I^*(r|x)}\right)}_{\text{bias}}, \underbrace{\frac{1}{I^*(r|x)}}_{\text{noise}}\right)$$



ECBD

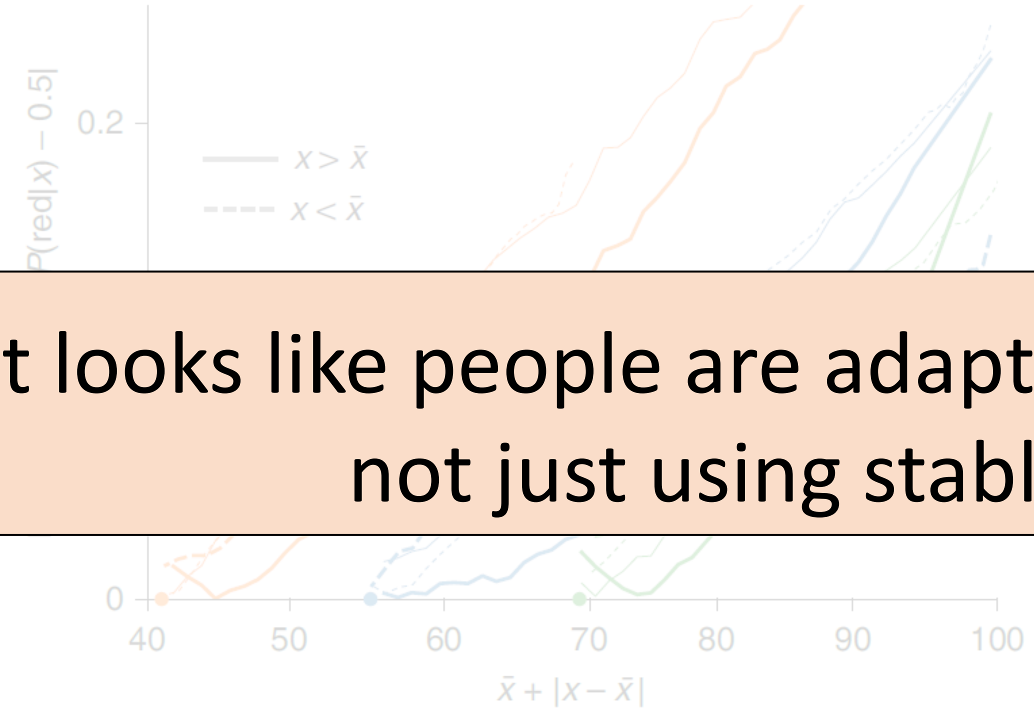


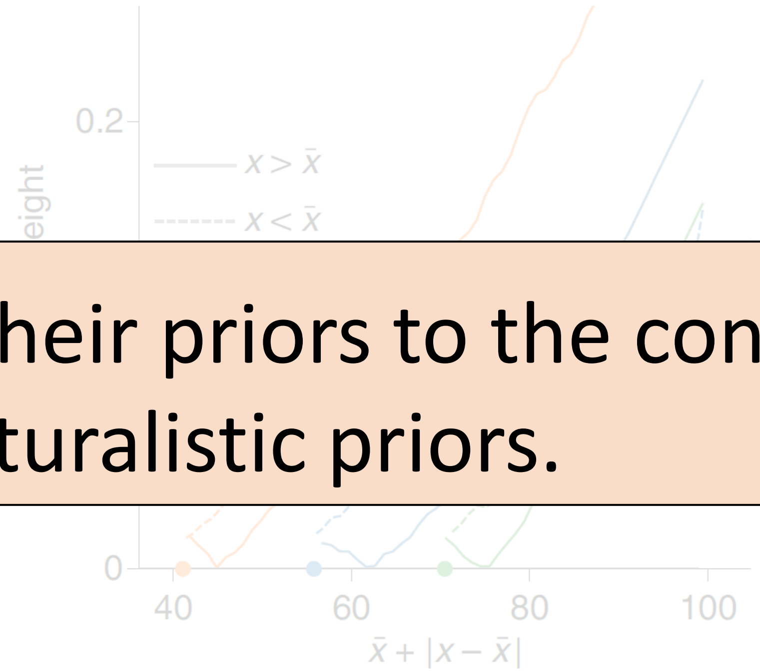
$\approx$



Prat-Carrabin and Woodford (2022)

ECBD





It looks like people are adapting their priors to the context, not just using stable naturalistic priors.

$\approx$

Prat-Carrabin and Woodford (2022)

ECBD: (Over)simplified Bayesian decoding

If…

Prior: $x \sim N\left(\mu, \frac{1}{p_x}\right)$ Representation: $r \sim N\left(x, \frac{1}{p_r}\right)$

Then the posterior is…

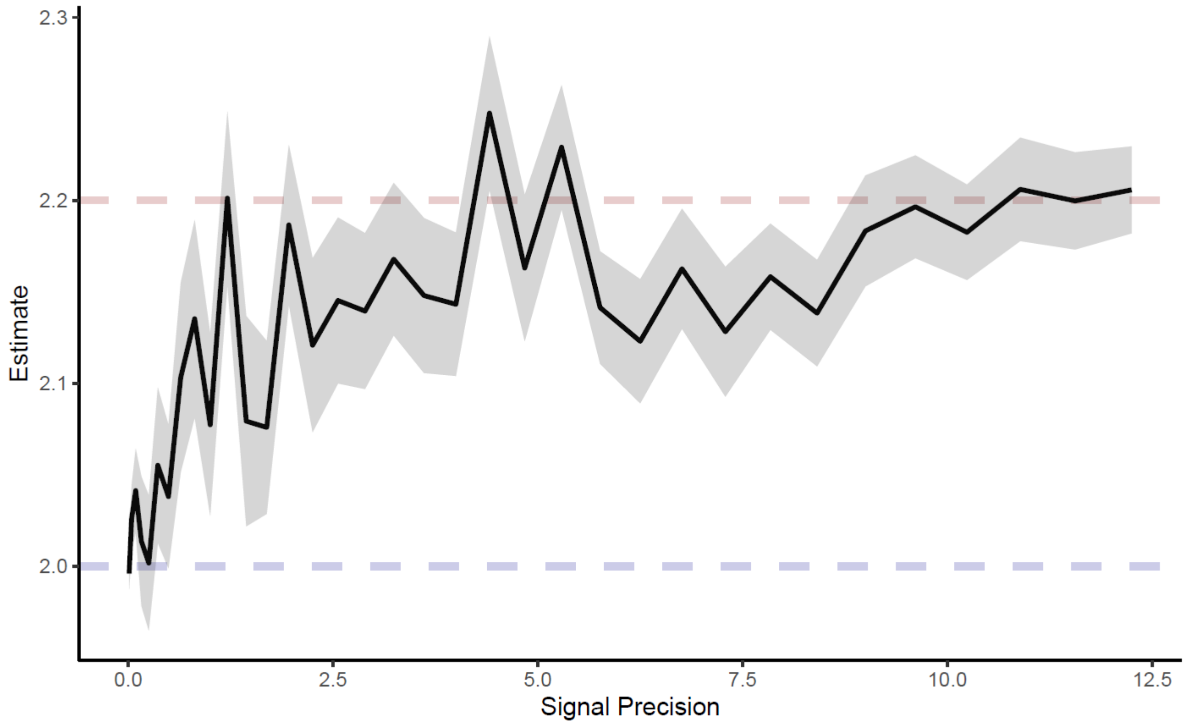
$$\hat{x}(r) \sim N\left(\mu + \beta(r - \mu), \frac{\beta^2}{p_r}\right)$$

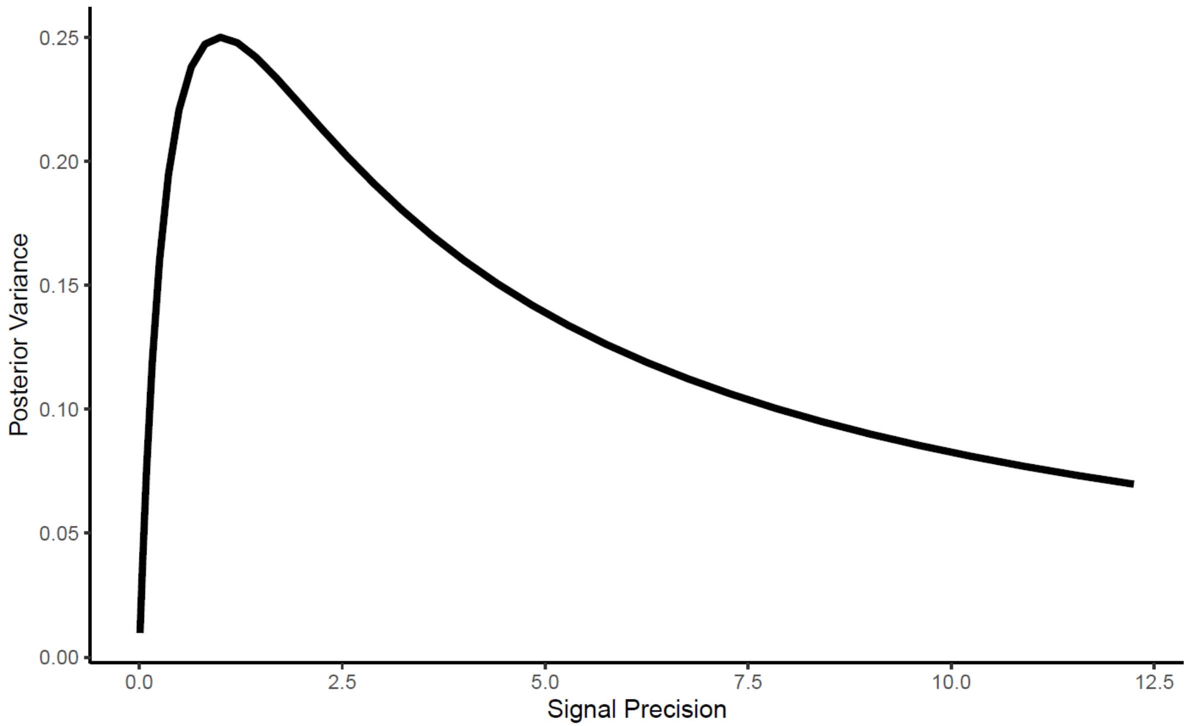
where $\beta = \frac{\mathrm{p_r}}{\mathrm{p_x} + \mathrm{p_r}}$

see Woodford (2020) for detailed explanation

Simple ECBD

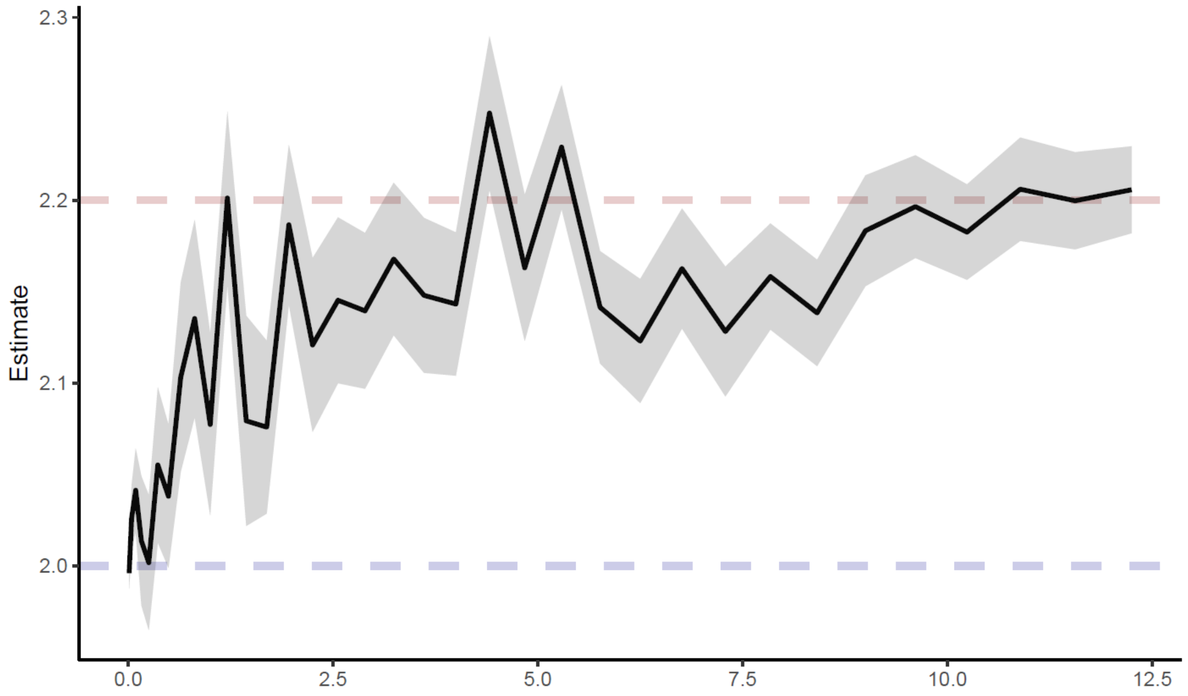
Fix $x = 2.2, \mu = 2, \mathrm{p_x} = 1.$



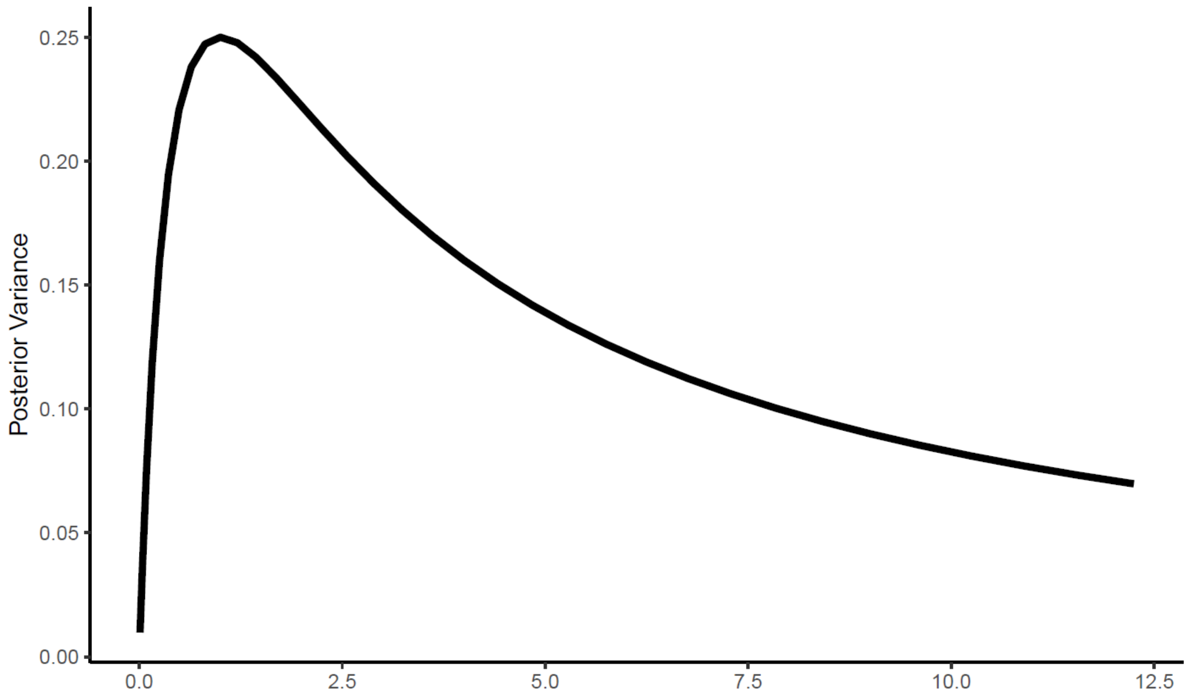


Simple ECBD

Fix $x = 2.2, \mu = 2, \mathrm{p_x} = 1.$

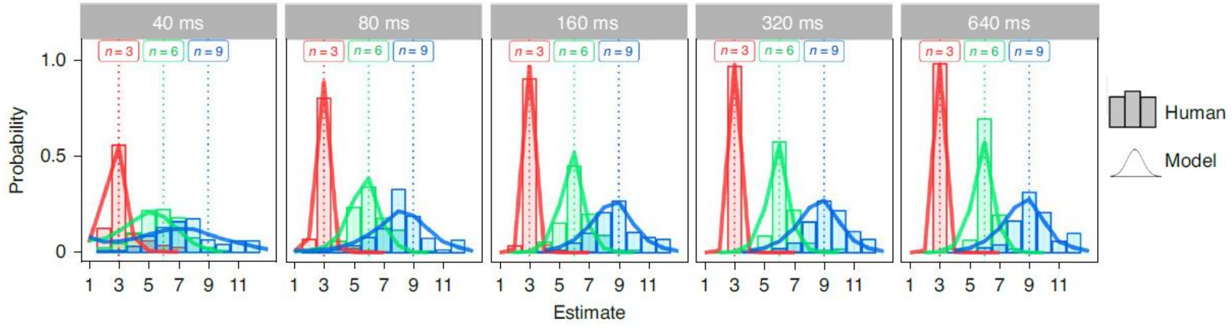


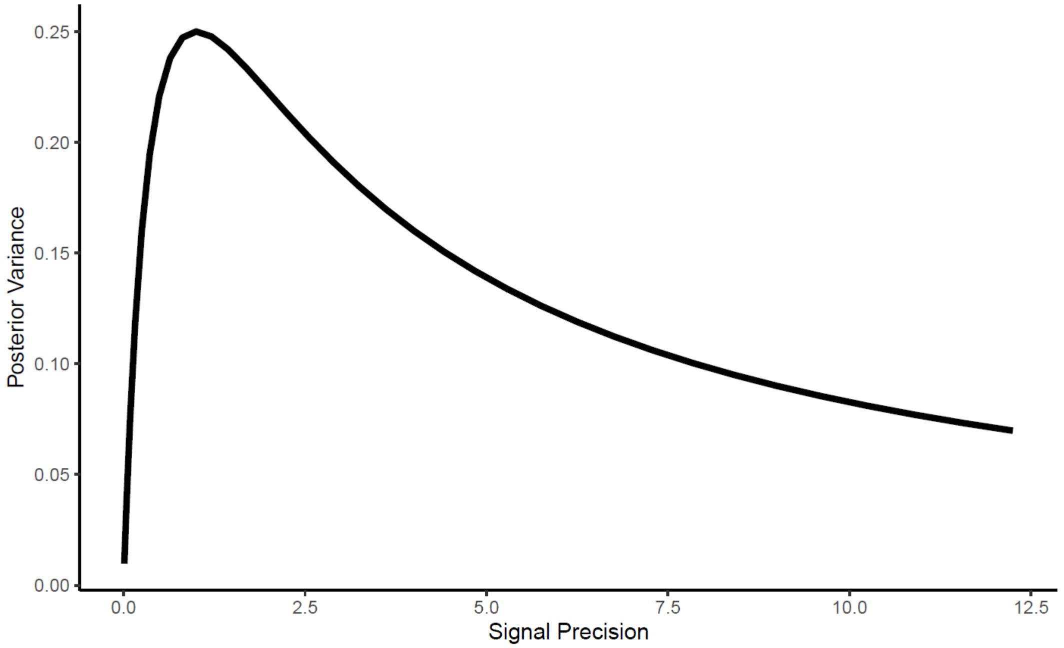
Duration of stimulus presentation (arbitrary)

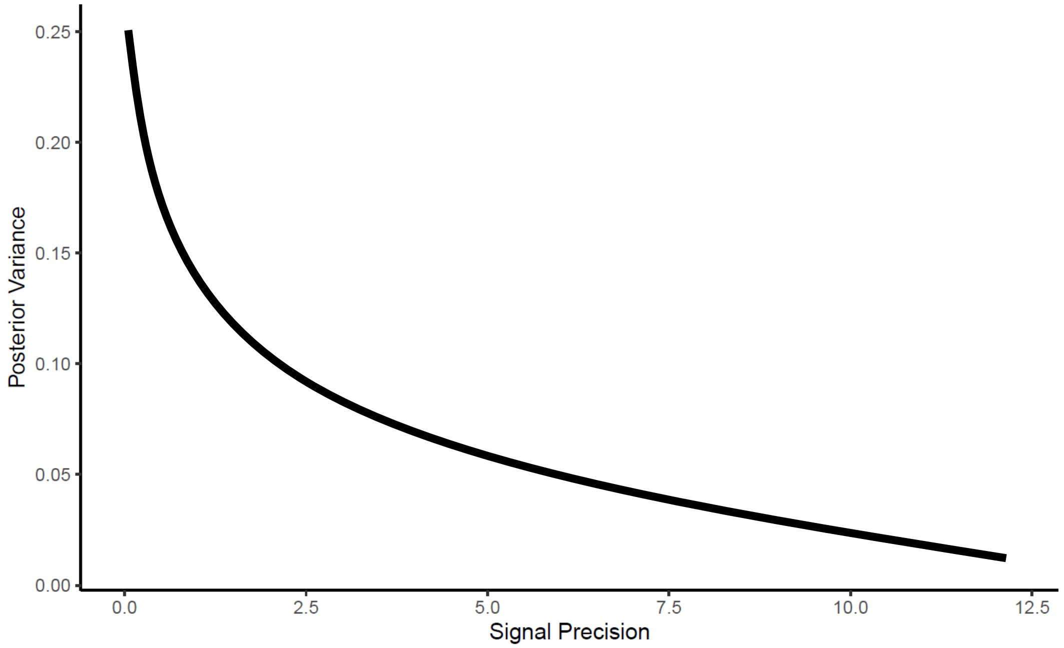


Duration of stimulus presentation (arbitrary)

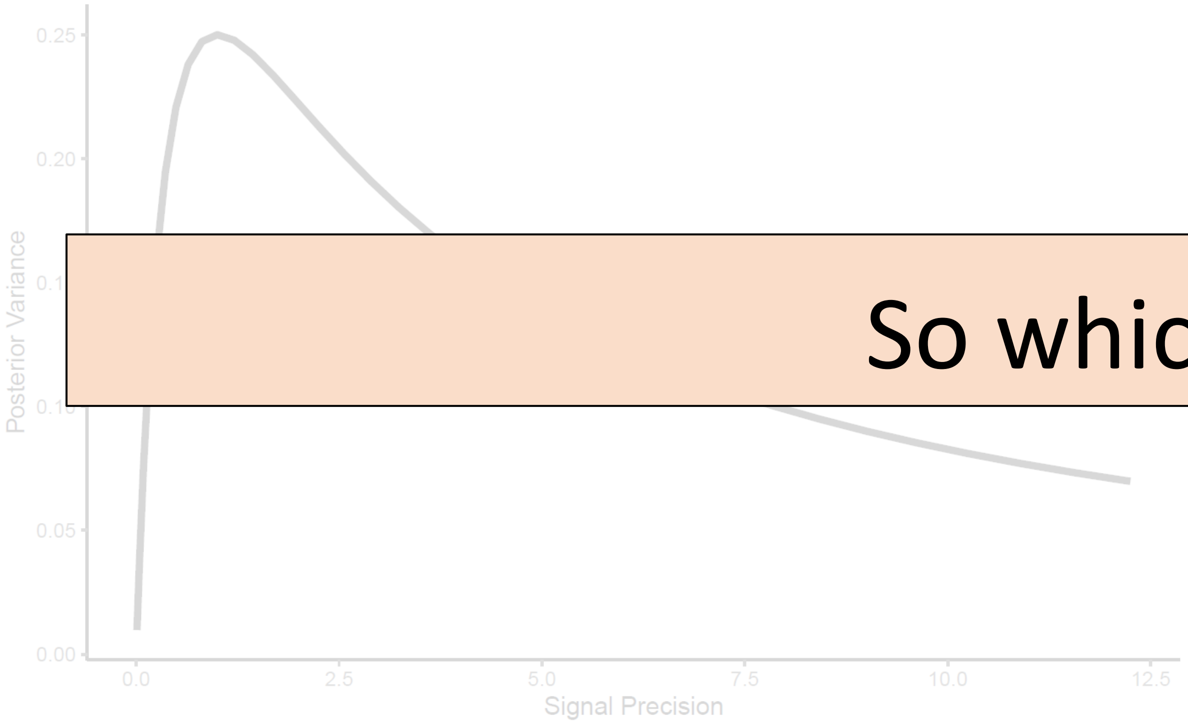
ECBD vs. Stable naturalistic priors







ECBD vs. Stable naturalistic priors



So which is it?

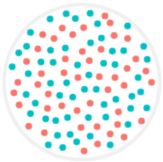
Task

100 trials — 100 trials — 100 trials — 100 trials

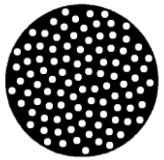
B **R** **B** **R**

+

Fixation
1500 ms



Stimulus
t ms



Mask
100 ms

Response
$\leq$2000 ms

$t \in \{100, 300, 500, 750, 1000, 3000\}$

90 trials: 54-46

10 trials: 80-20

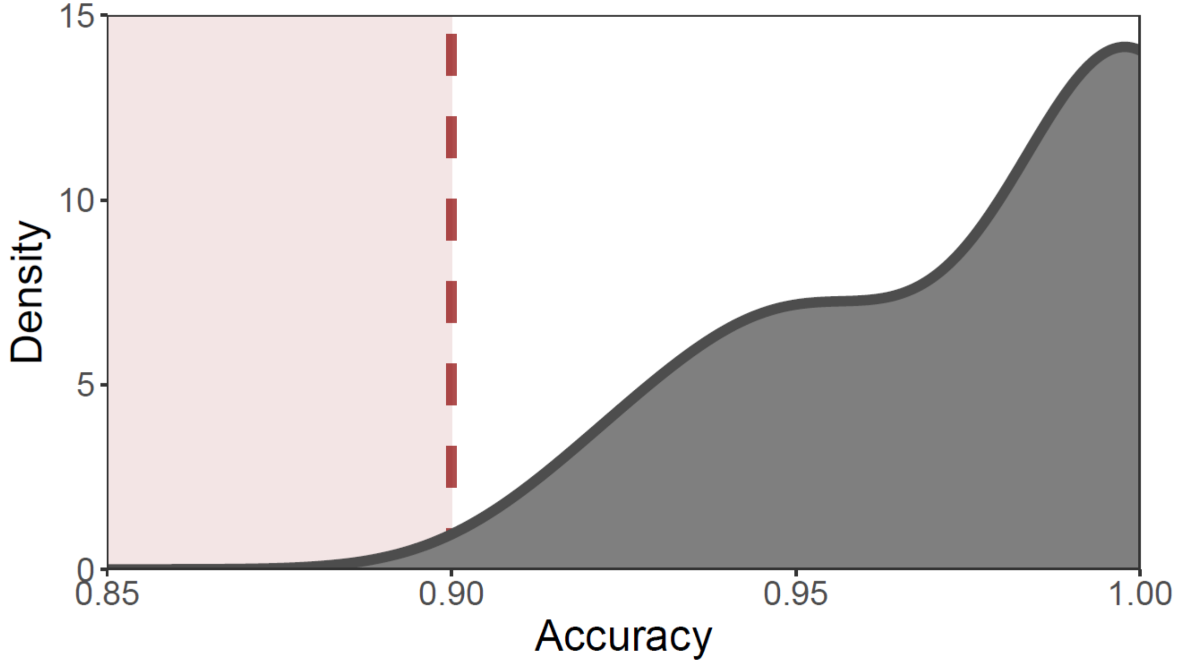
Precommitment to high quality data

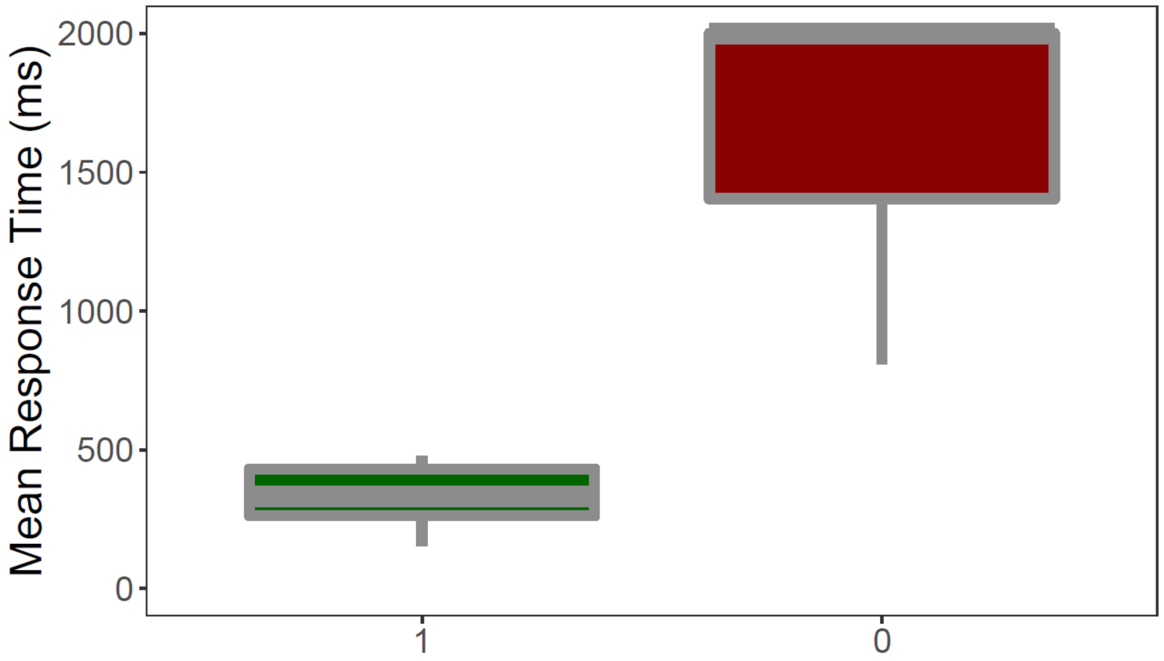
Drop if…

- Accuracy on “sanity check” trials is <90%.

Results

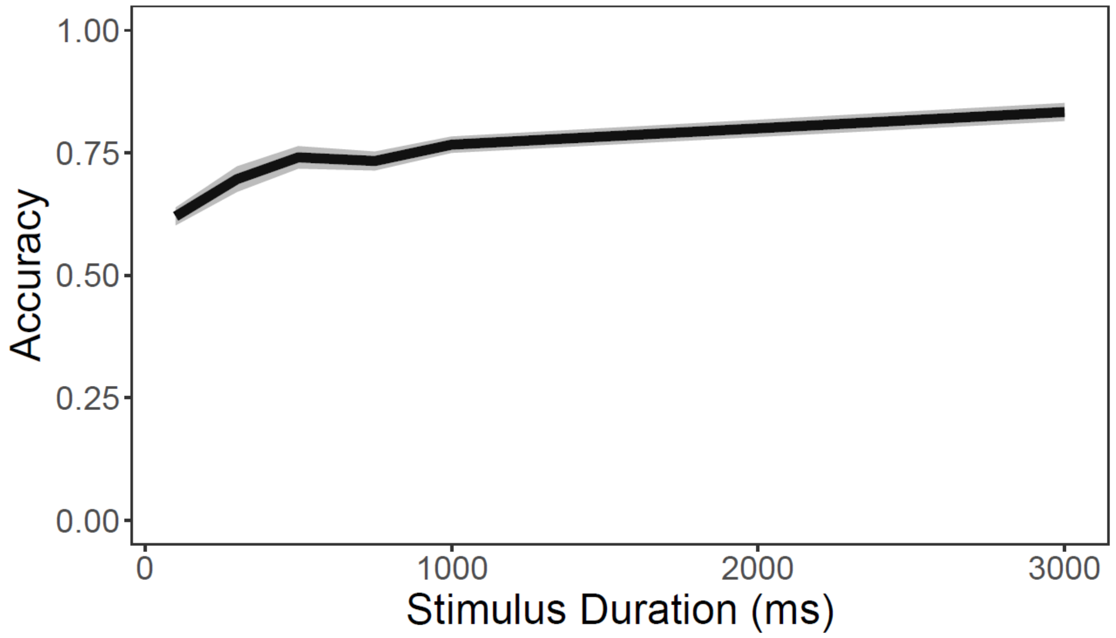
Quality check

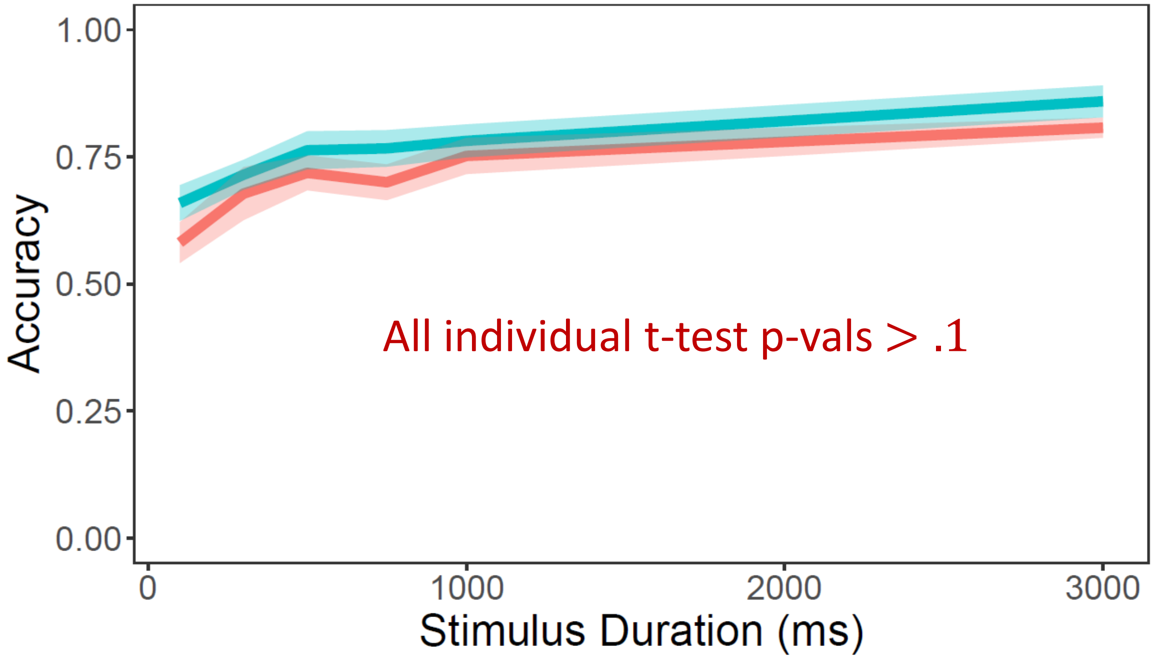




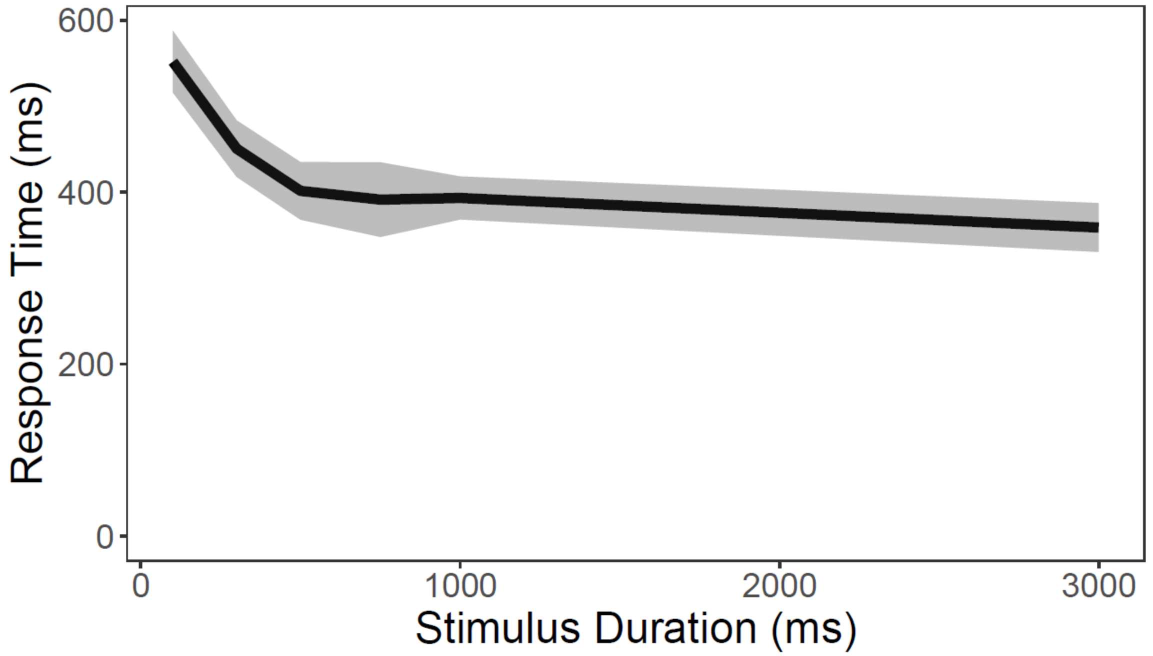
Dropped 1 out of 10 total subjects.

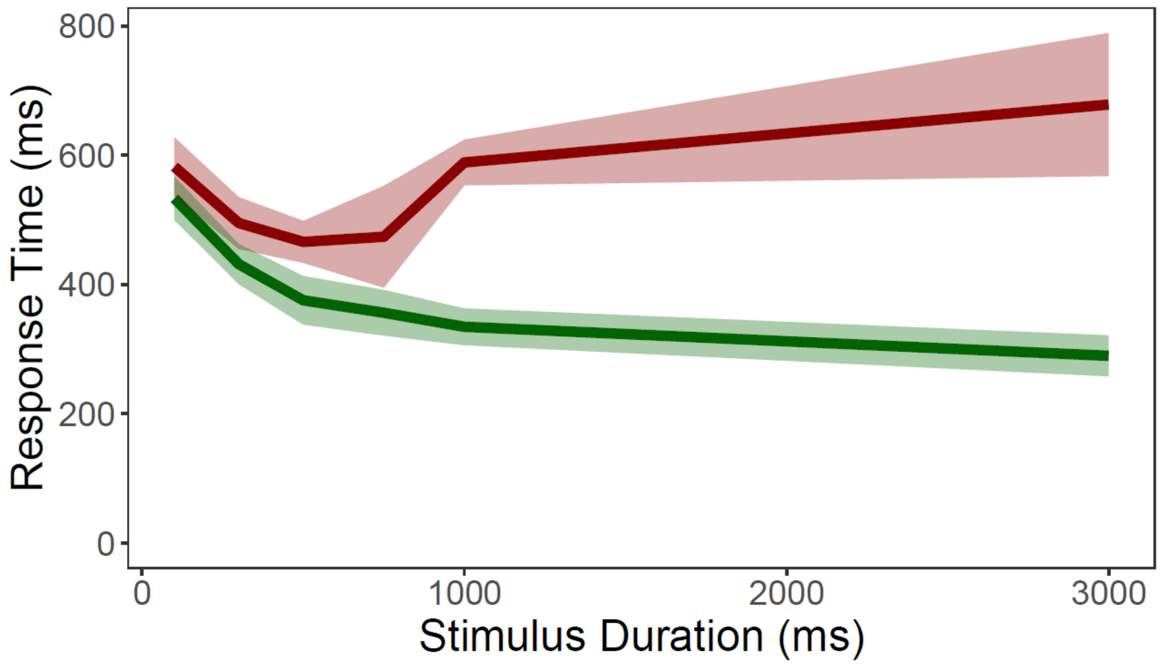
Basic psychometrics



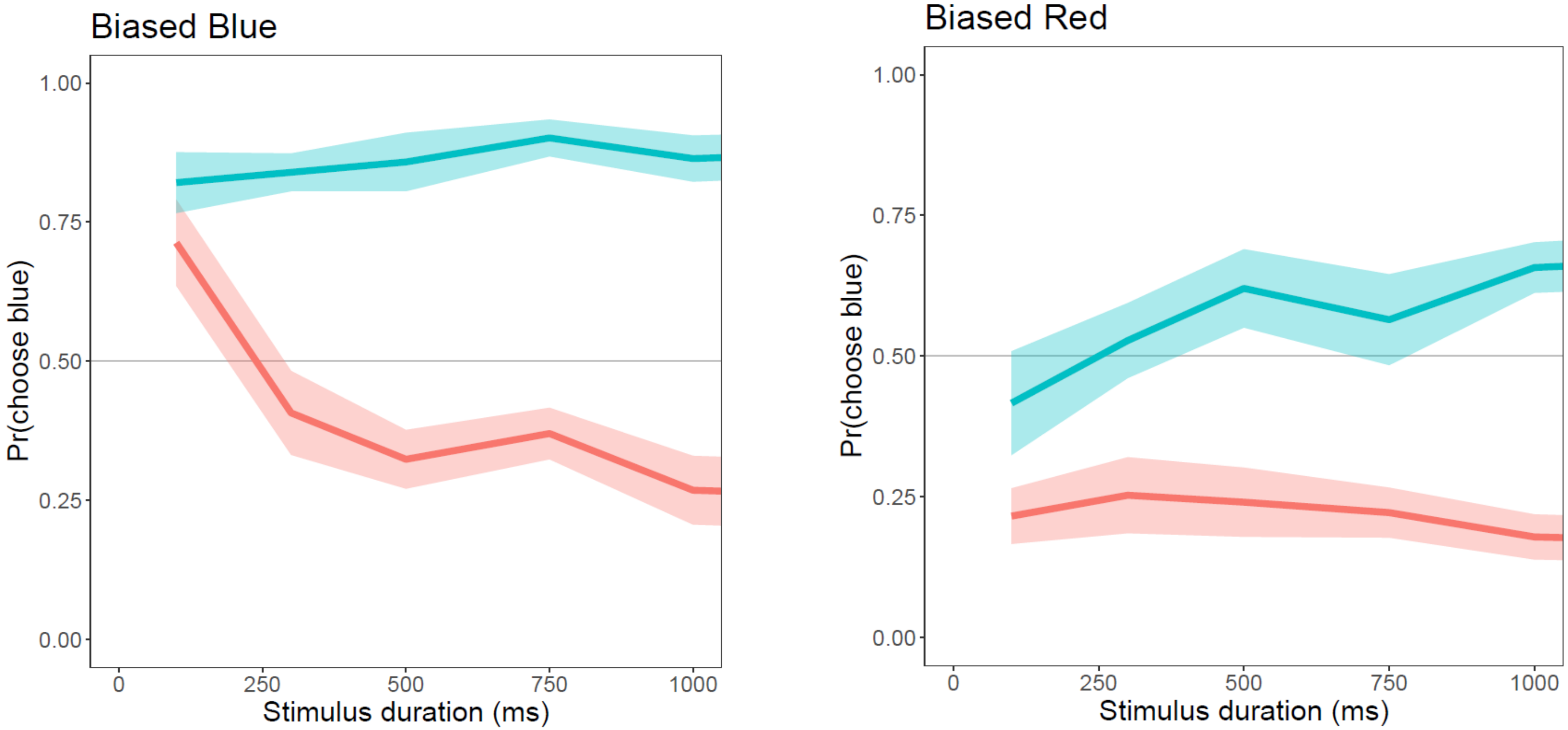


Response time

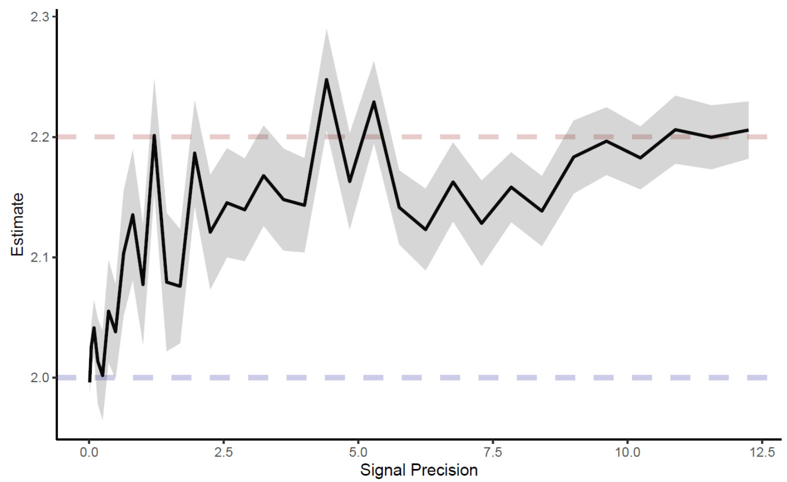


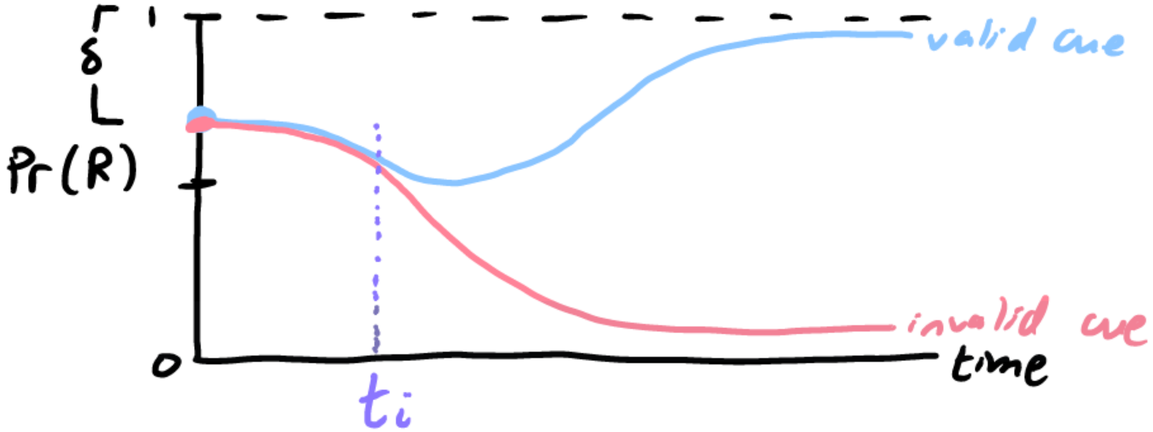


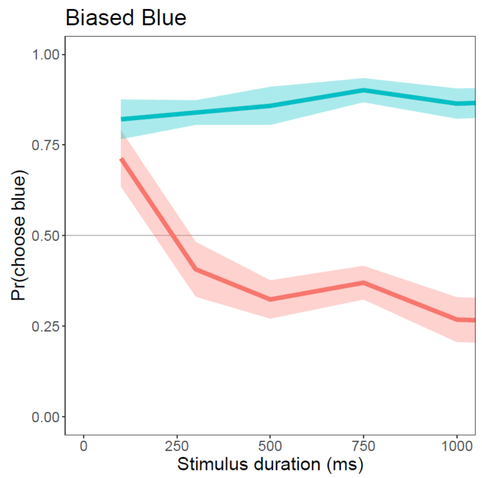
Effects of signal precision on estimates

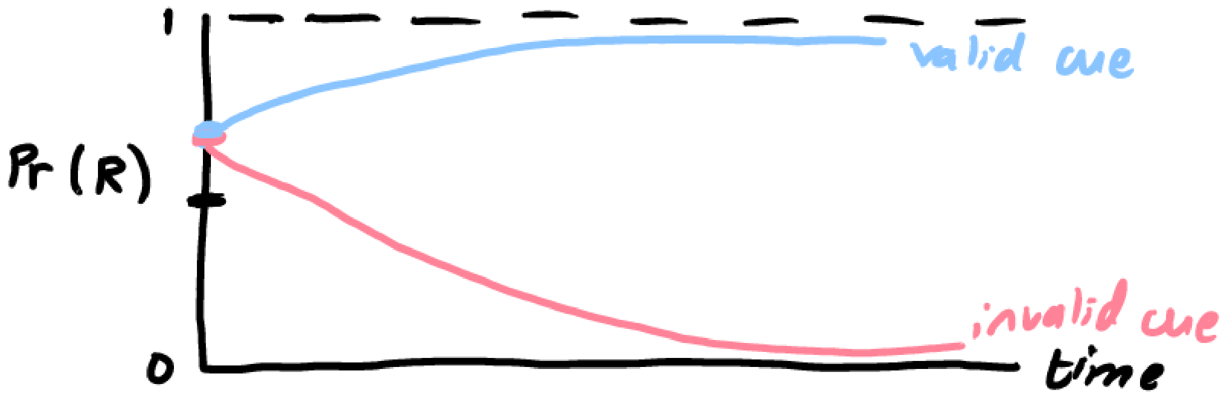


Quick note (**TAKE THIS OUT FOR ACTUAL TALKS)

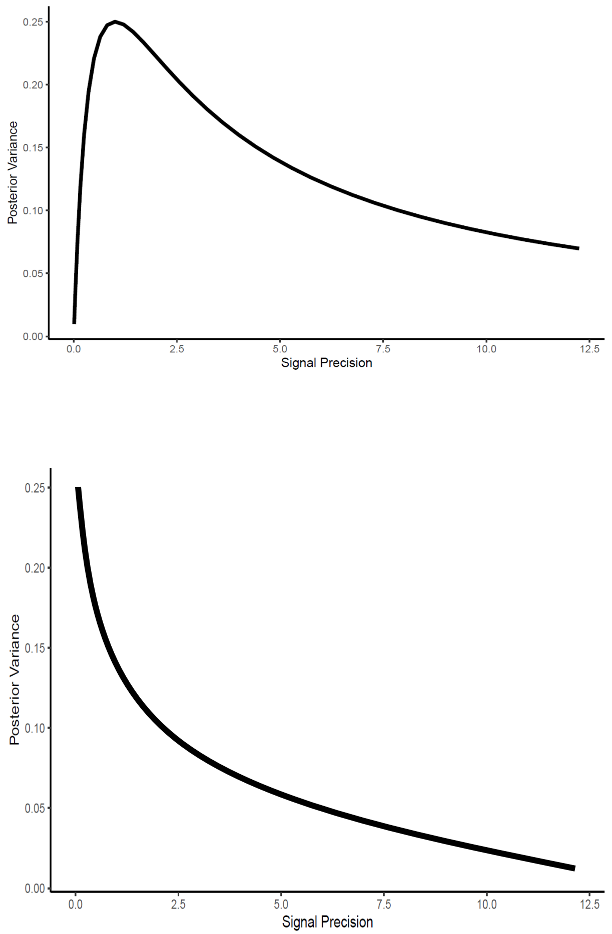


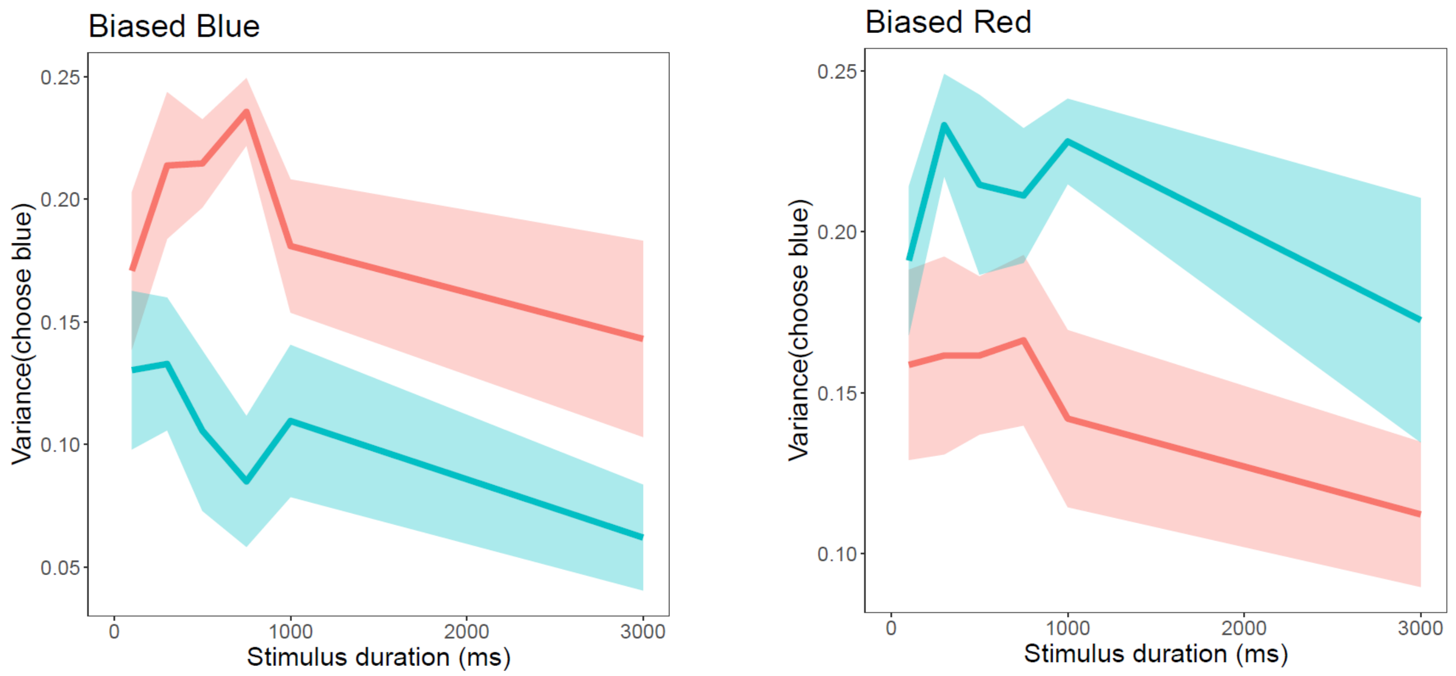






Quick note (**TAKE THIS OUT FOR ACTUAL TALKS)



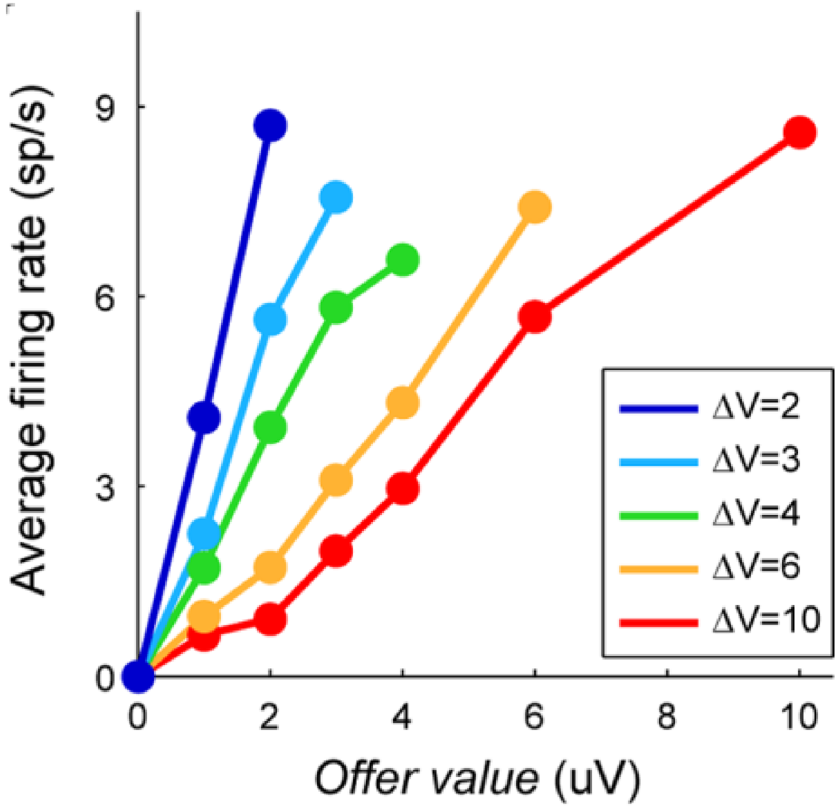


Take-aways from pilot study

1. We should be focused on looking at average variance over stimulus duration as our test between competing classes of models, not average estimate.
2. We should be focusing on shorter stimulus durations, perhaps based on the times from CP (2020) (40, 80, 160, 320, 640, 1280 ms).
3. How can we check if 54-46 dot splits are a reasonable amount of difficulty? Average estimates do look like they're properly separating in stimulus duration.

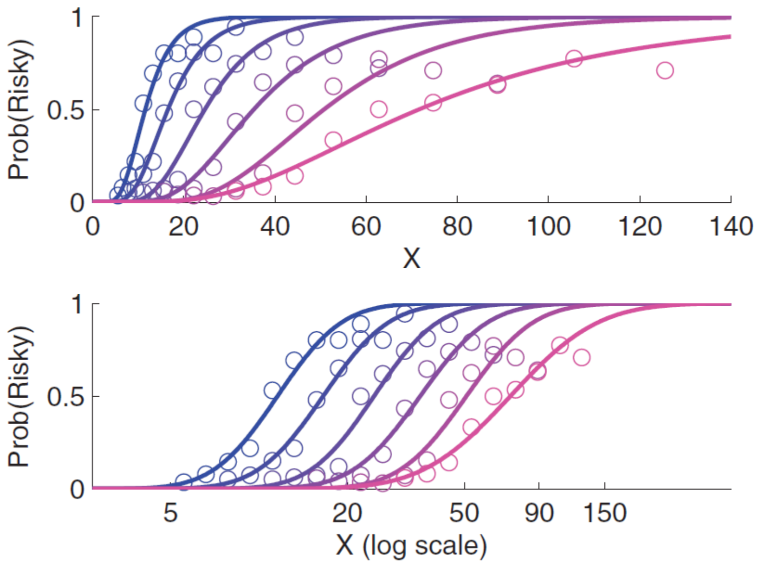
Why does any of this matter?

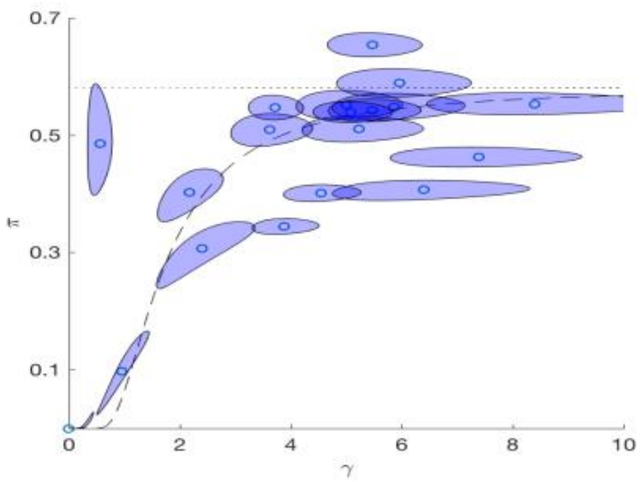
Range normalization for value neurons



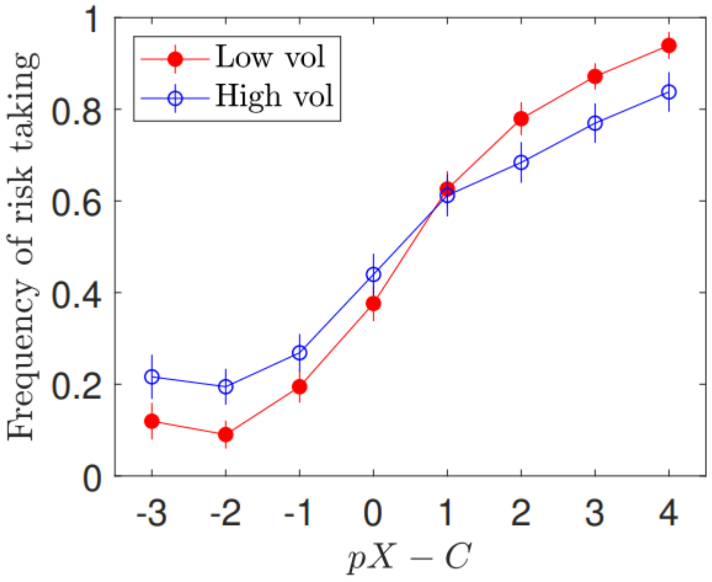
Padoa-Schioppa (2009)

Risk aversion results from perceptual bias



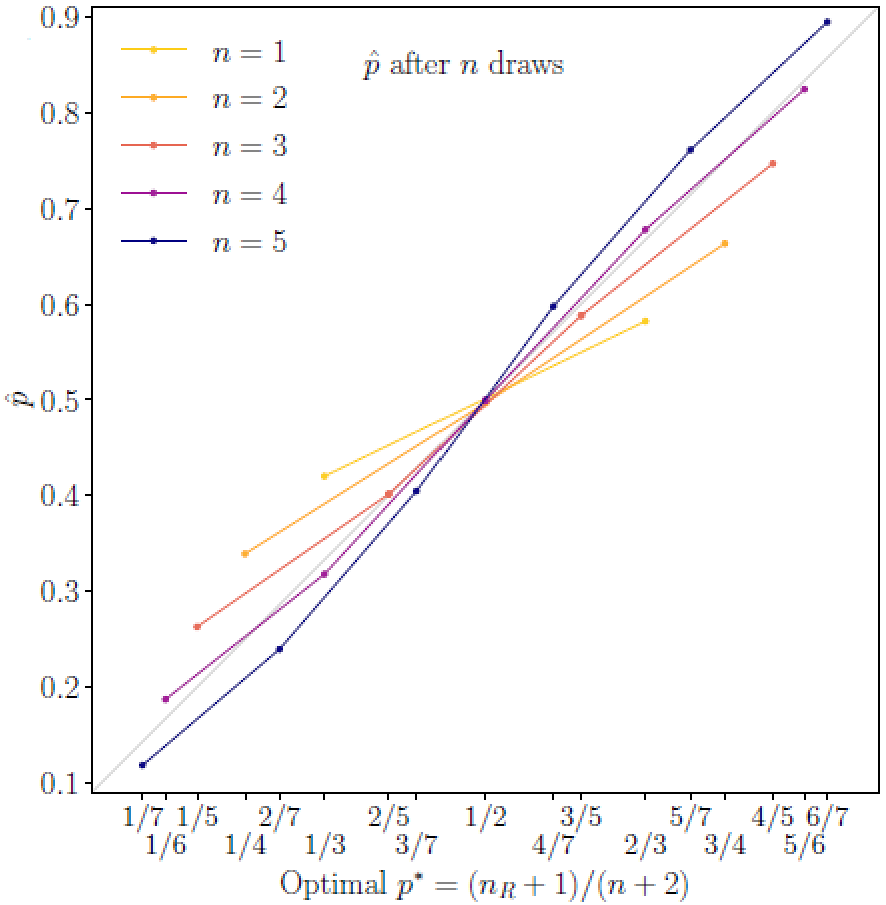


Khaw, Li, and Woodford (2021)



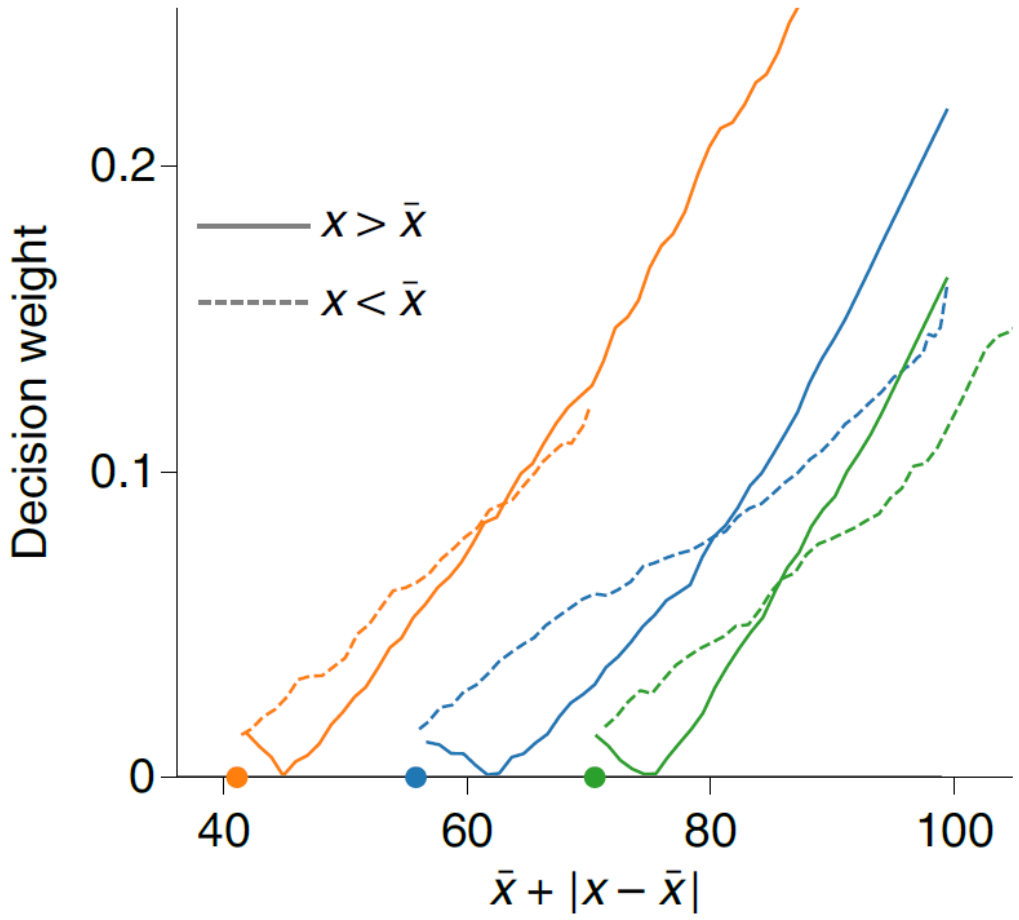
Frydman and Jin (2022)

Under- and over-reaction in probabilistic inference



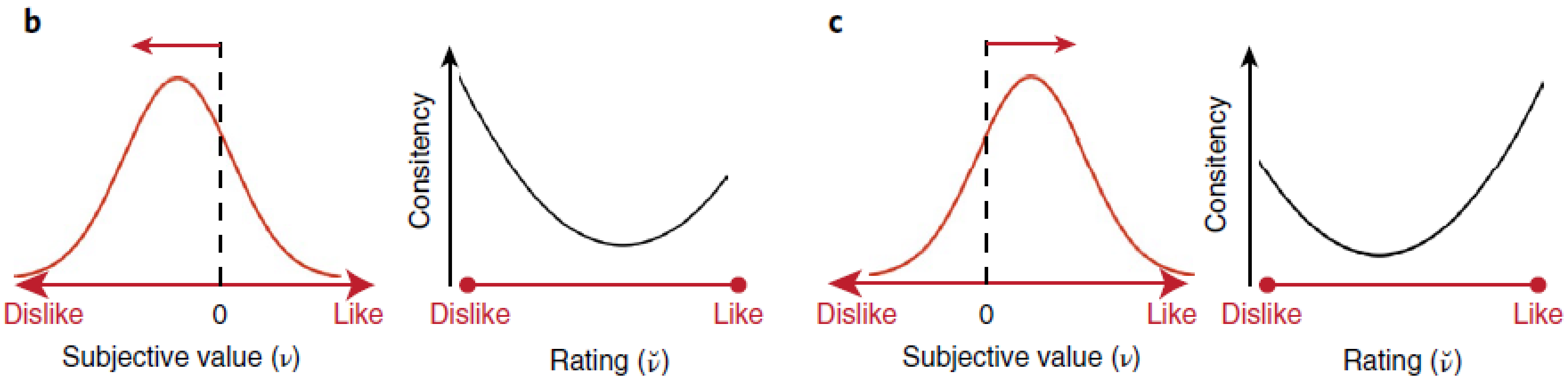
Prat-Carrabin and Woodford (2022, wp)

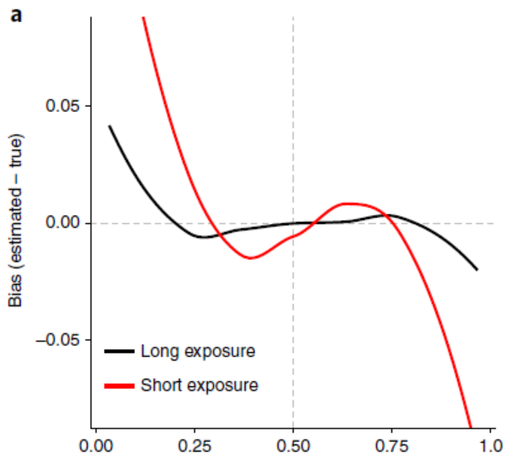
Non-linear weighting of numbers in an average



Prat-Carrabin and Woodford (2022, Nat Hum Behav)

Choice variability in in value-based choice





Polania, Woodford, and Ruff (2019)