Bayesian Statistics and Hierarchical Bayesian Modeling for Psychological Science

Lecture 09

Lei Zhang

Social, Cognitive and Affective Neuroscience Unit (SCAN-Unit)

Department of Cognition, Emotion, and Methods in Psychology

lei.zhang@univie.ac.at

lei-zhang.net

@lei_zhang_lz

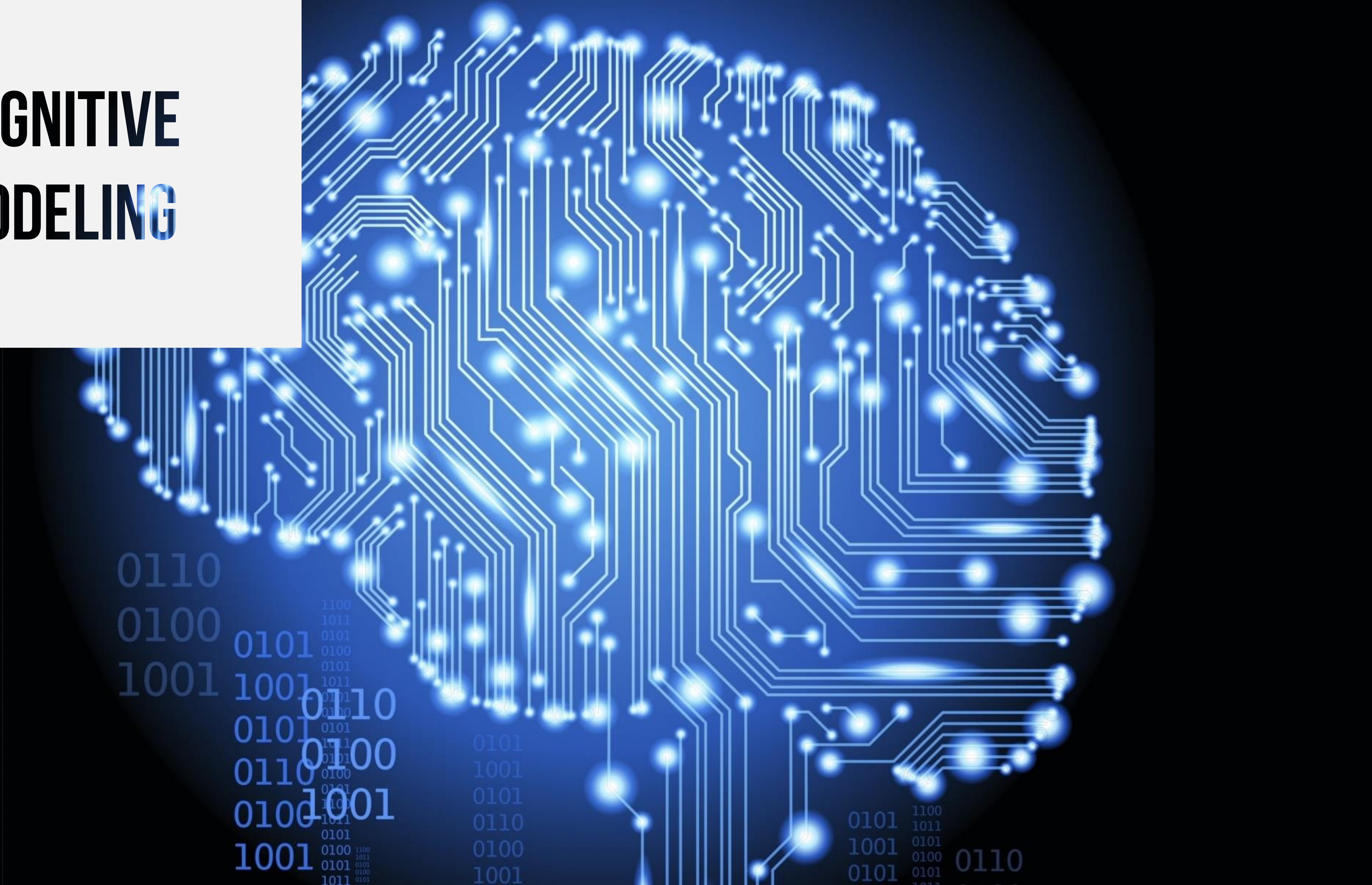
https://github.com/lei-zhang/BayesCog_Wien

universität wien

Fakultät für Psychologie

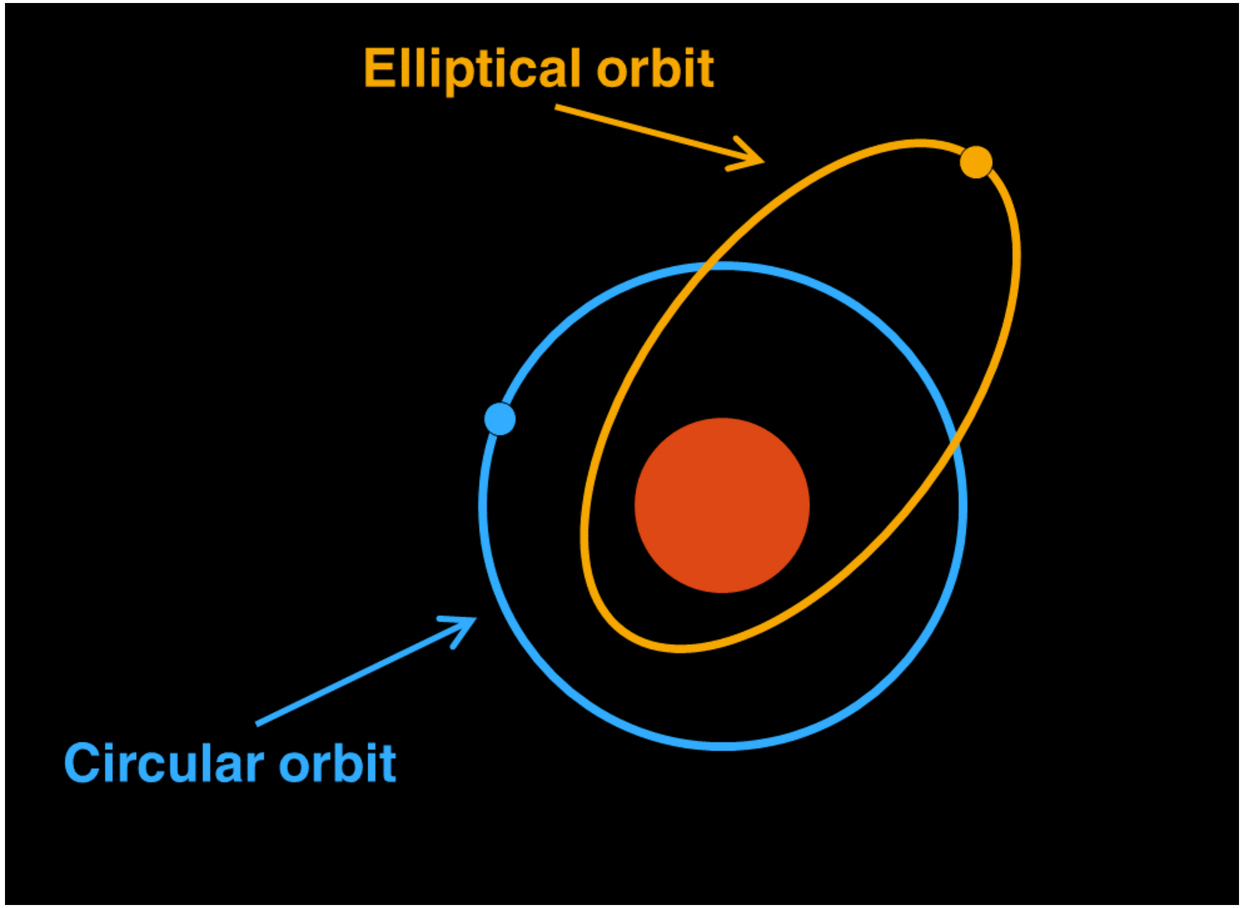
Bayesian warm-up?

COGNITIVE MODELING

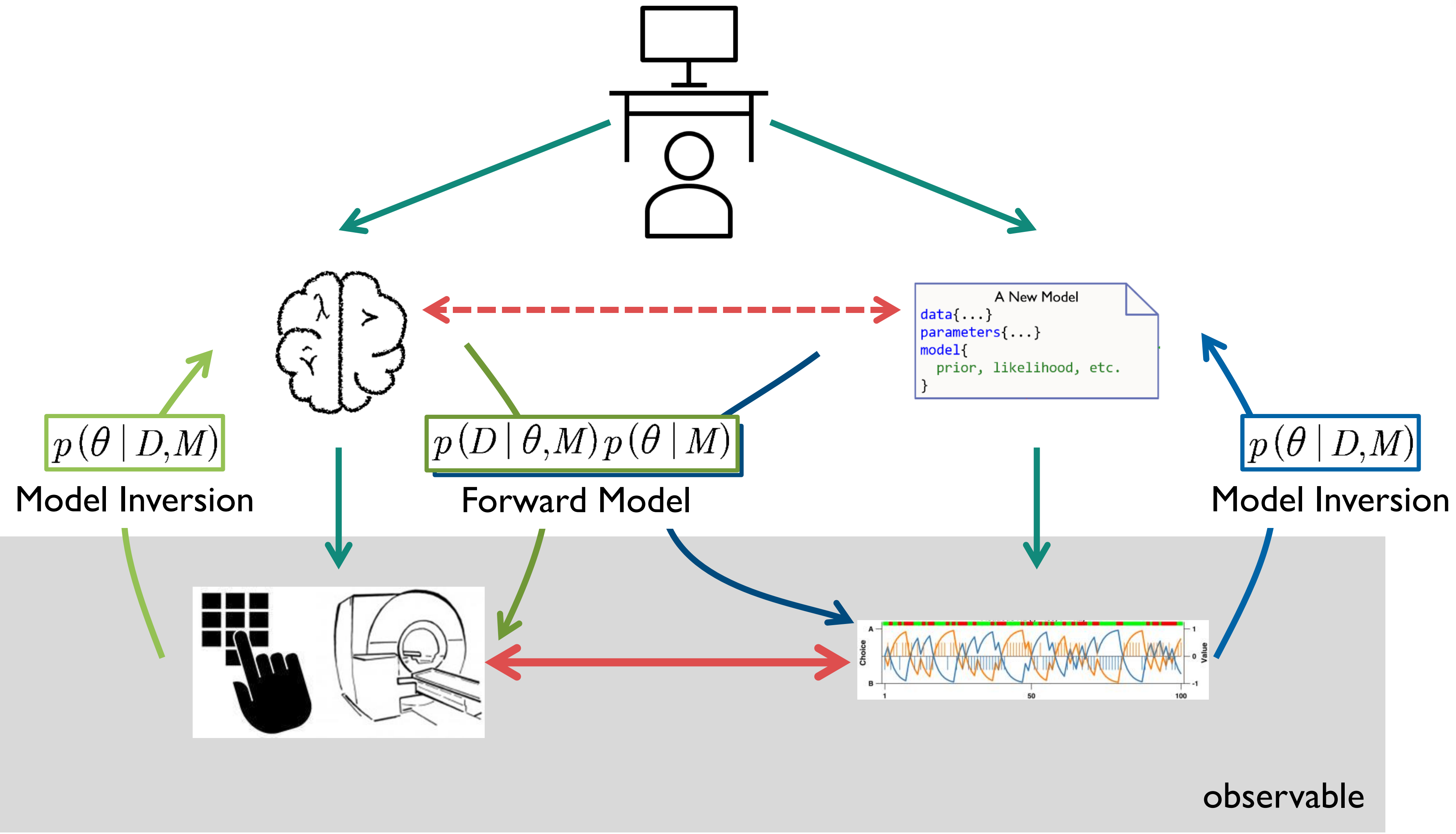


The idea of computational modeling is never new

Scientists use mathematical models to approximate certain processes (physical or mental), in order to explain and to predict.



What is Cognitive Modeling?

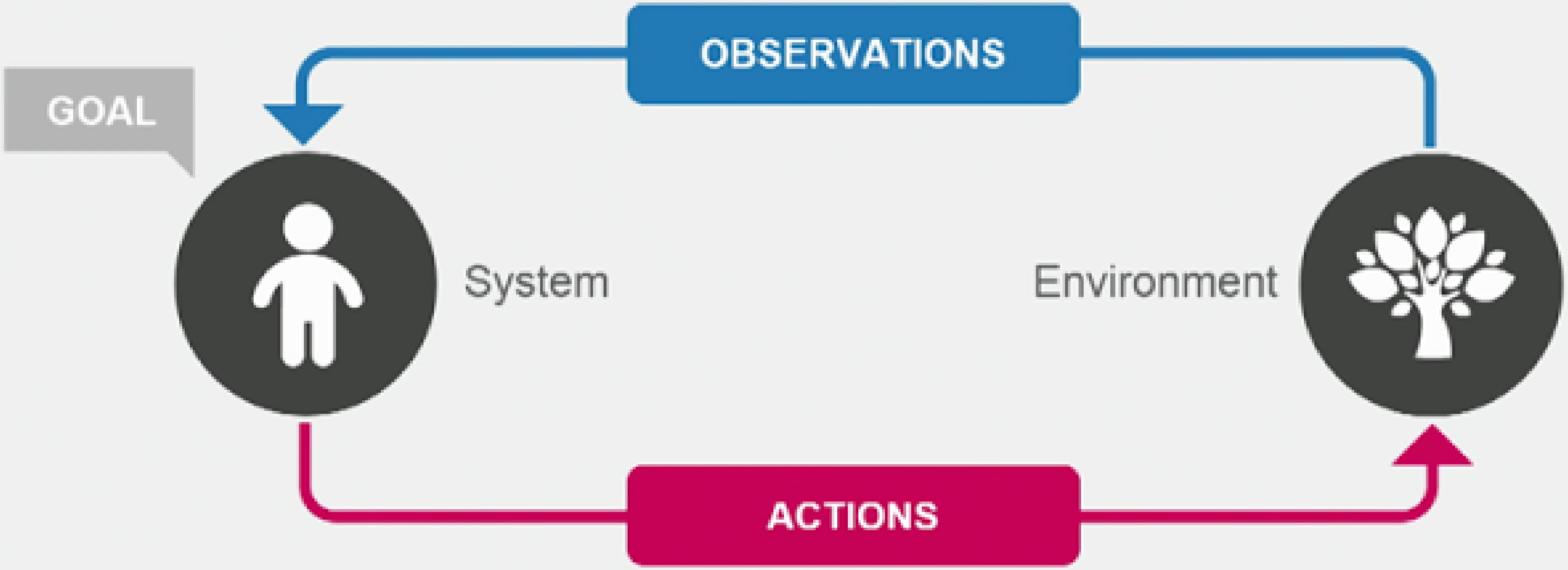


Lewandowsky & Farrell (2011); Stephan et al. (2016); Ahn et al., (2017) Essentially, all the models are wrong, but some are useful.

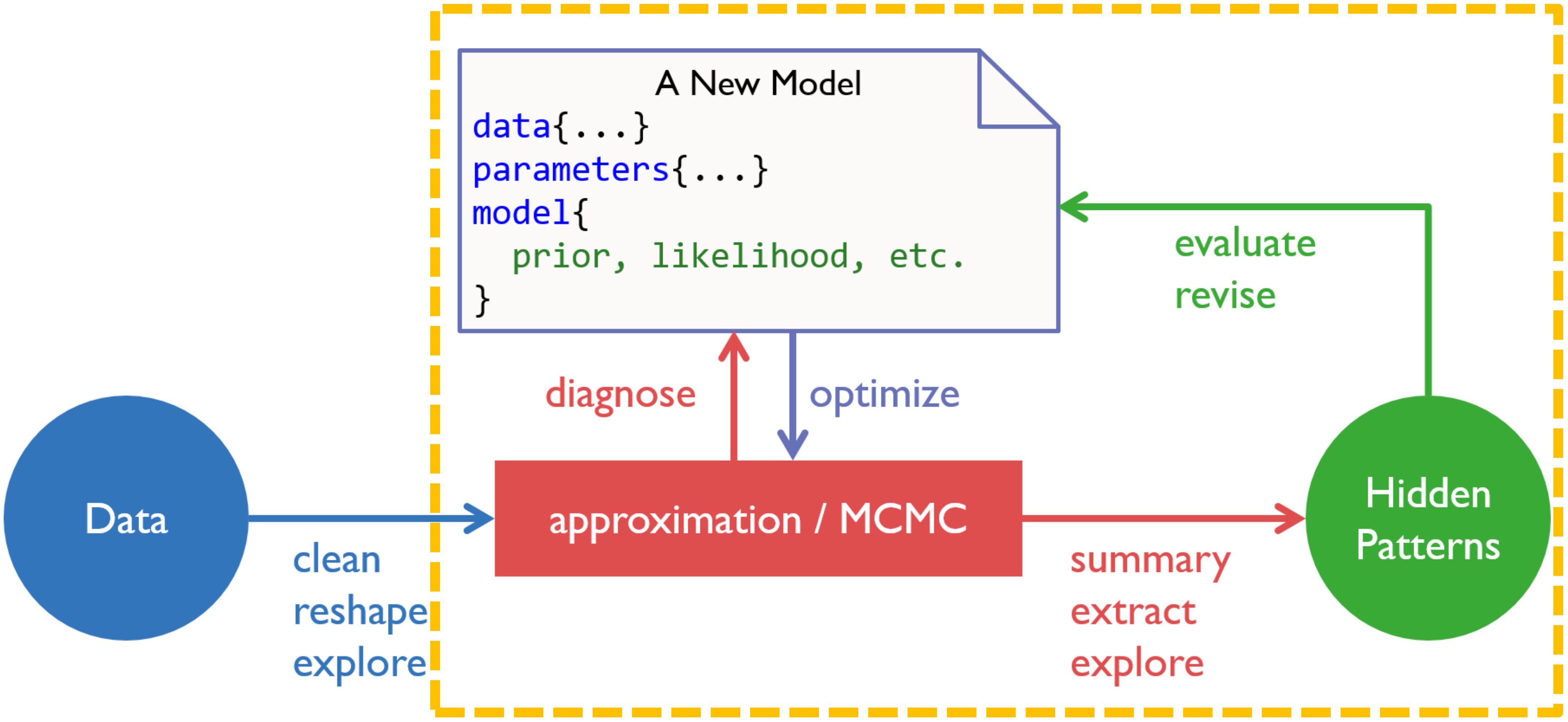
– George E. P. Box

Essentially, all the models are ~~wrong~~ imperfect, but some are useful.

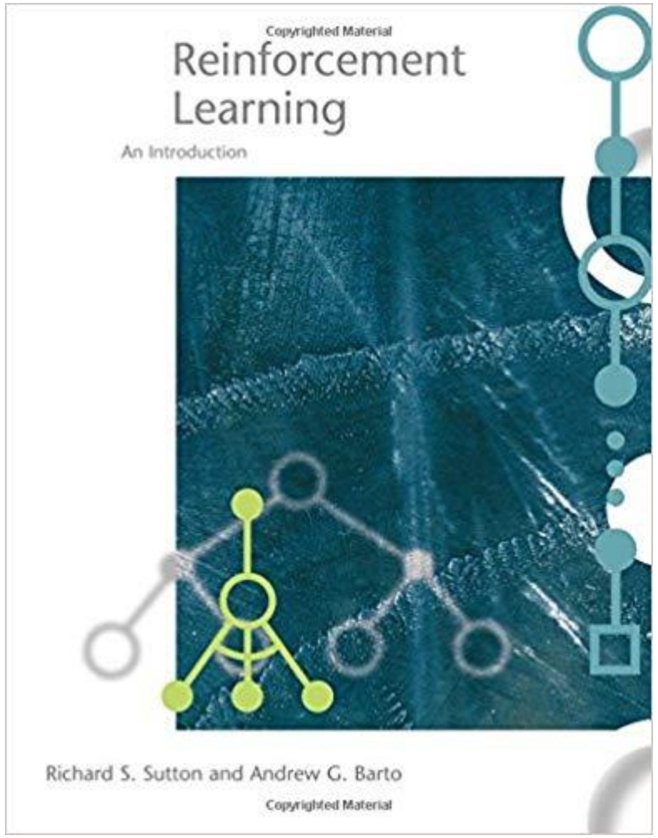
REINFORCEMENT LEARNING FRAMEWORK



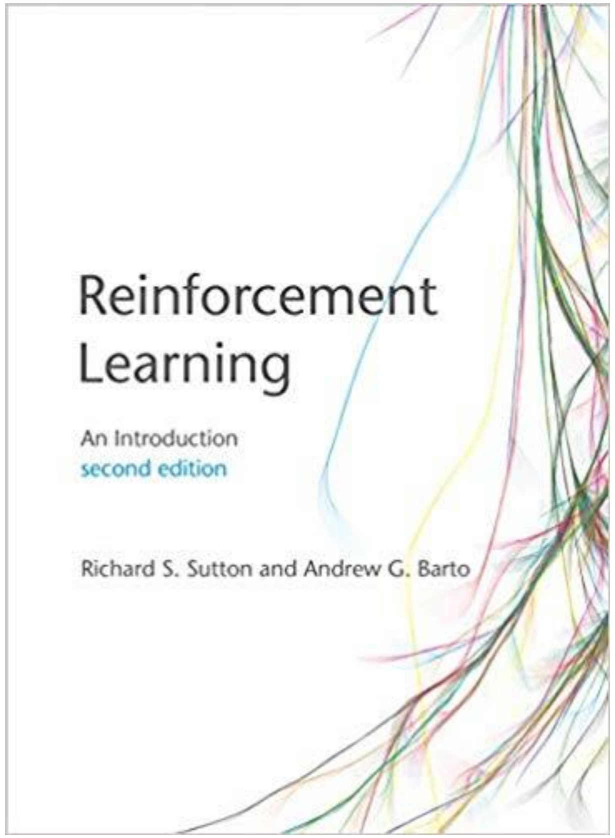
cognitive model
statistics
computing



The very short history

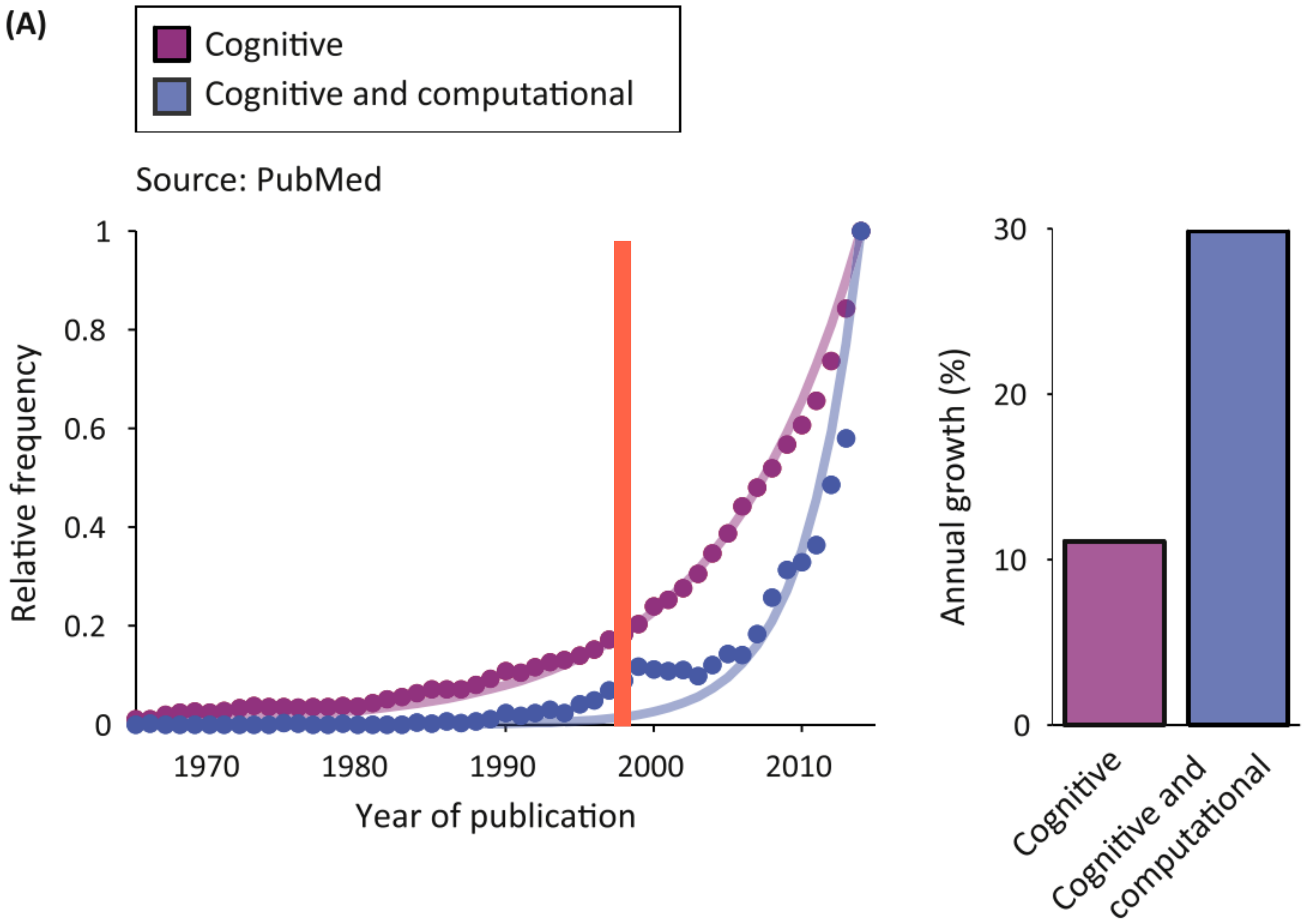


1998



2018

Boom in Cognitive Modeling



Palminteri et al. (2017)

Very recent examples

REPORT

Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work

A. Westbrook[1,2,3,*], R. van den Bosch[2,3], J. I. Määttä[2,3], L. Hofmans[2,3], D. Papadopetraki[2,3], R. Cools[2,3,†], M. J. Frank[1,4,†]

+ See all authors and affiliations

Science 20 Mar 2020:
Vol. 367, Issue 6484, pp. 1362-1366
DOI: 10.1126/science.aaz5891

Neuron

Available online 17 March 2020

In Press, Corrected Proof

CellPress

Article

A Neuro-computational Account of Arbitration between Choice Imitation and Goal Emulation during Human Observational Learning

Caroline J. Charpentier [1, 2], Kiyohito Iigaya [1], John P. O'Doherty [1]

3 out of 4 focused on Reinforcement Learning models!

nature reviews neuroscience

Review Article | Published: 12 March 2020

The neural and computational systems of social learning

Andreas Olsson, Ewelina Knapska & Björn Lindström

Nature Reviews Neuroscience **21**, 197–212(2020) | Cite this article

Translational Psychiatry

Article | Open Access | Published: 17 March 2020

Neurocomputational mechanisms underpinning aberrant social learning in young adults with low self-esteem

Geert-Jan Will, Michael Moutoussis, Palee M. Womack, Edward T. Bullmore, Ian M. Goodyer, Peter Fonagy, Peter B. Jones, NSPN Consortium, Robb B. Rutledge & Raymond J. Dolan

2-armed bandit task

a simple task often used in the laboratory:

- repeated choice between N options (N-armed bandit)
- …whose properties (reward amounts, probabilities) are learned through trial-and-error
- …with a goal in mind: maximize the overall reward

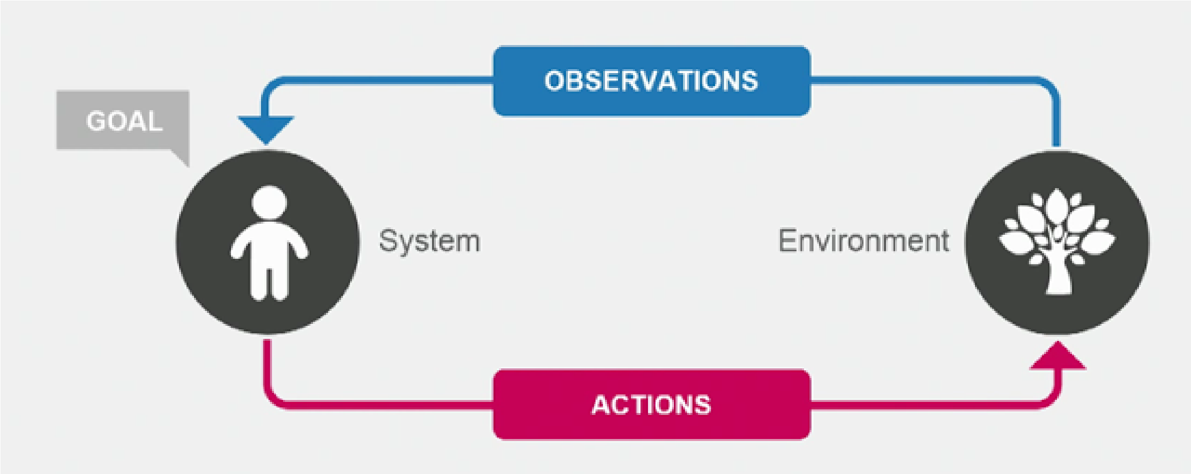
2-armed bandit task

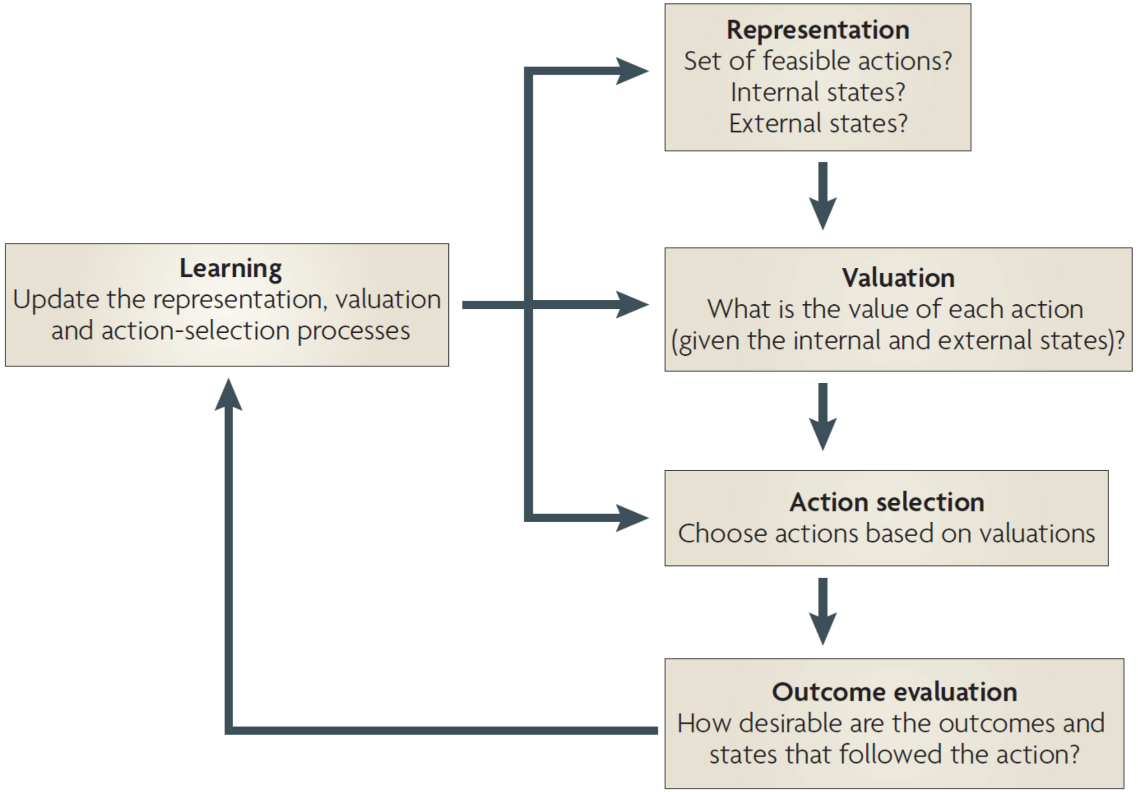
What can be your strategies:

1. predict the value of each deck
2. choose the best
3. learn from outcome to update predictions

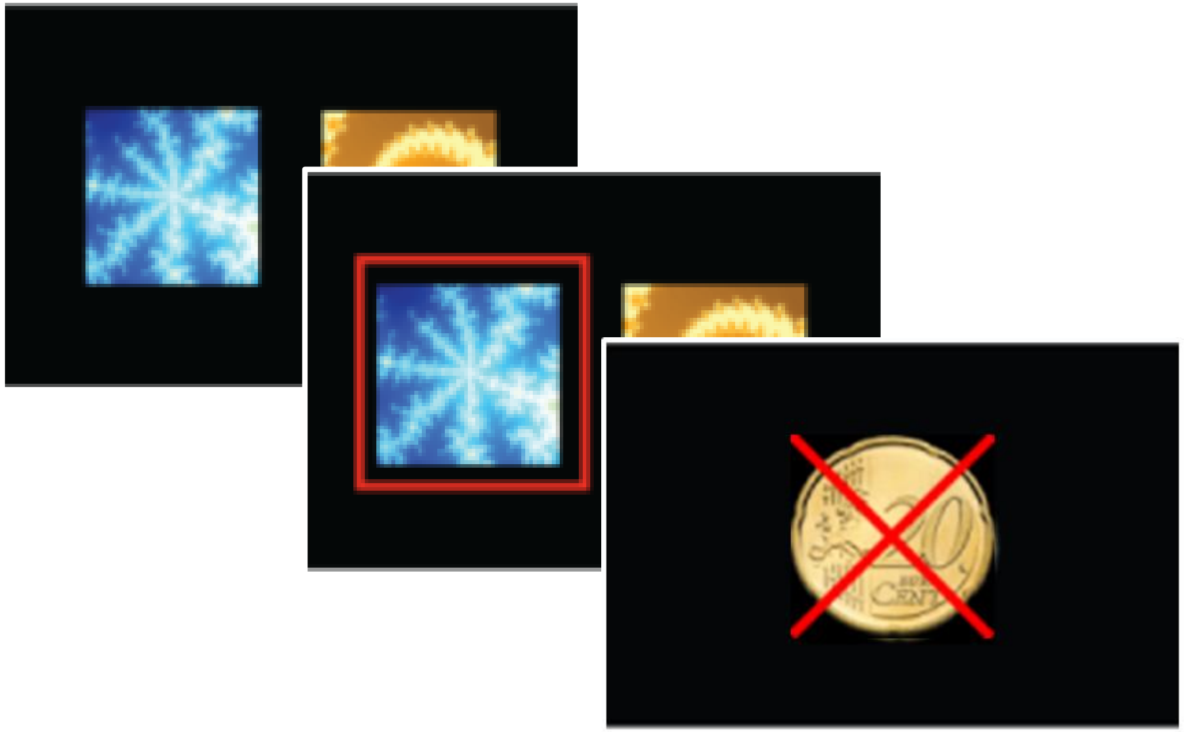
(repeat)

How prediction is shaped by learning?



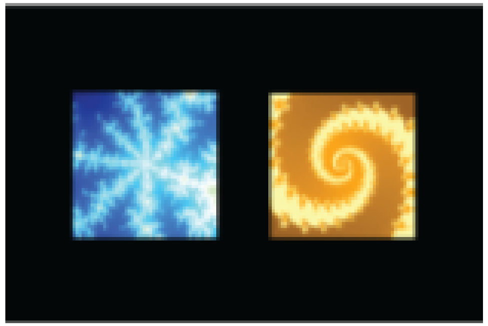


Sutton & Barto (1998); Rangel et al (2008); [Daw & Niv workshop 2015]

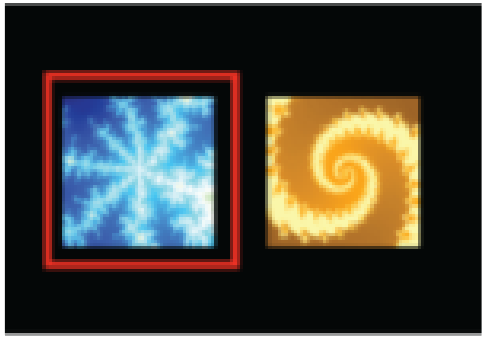




Winnings
$ 10

One simple experiment: two choice task



choice presentation



action selection

outcome

what do we know?

what can we measure?

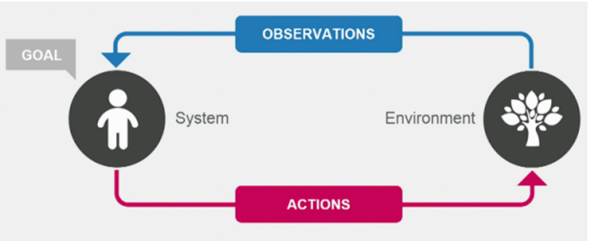
what do we not know?

Data: choice & outcome

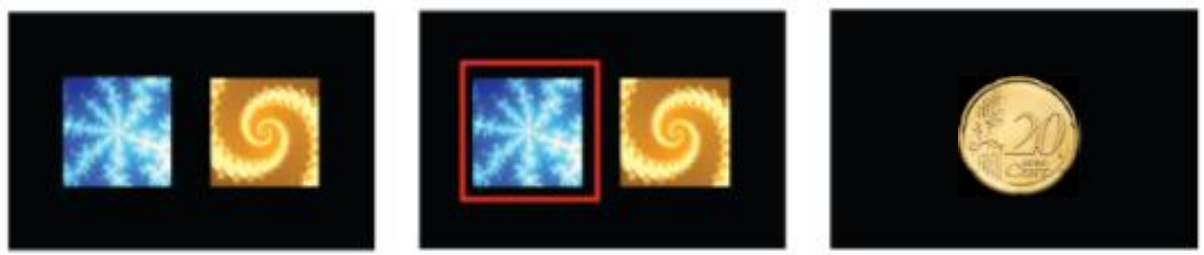
Summary stats: choice accuracy

Learning algorithm: RL update

p(choosing the better option)



Rescorla-Wagner Value Update



Cognitive Model
- cognitive process
- using internal variables and free parameters

Observation Model (Data Model)
- relate model to observed data
- has to account for noise

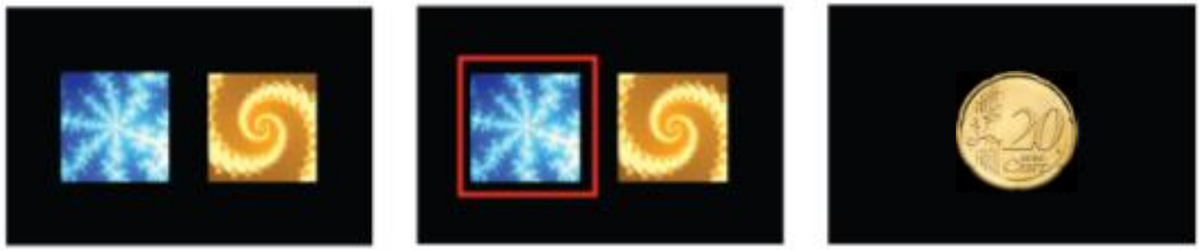
Rescorla & Wagner (1972)

Rescorla-Wagner (1972)

- The idea: error-driven learning
- Change in value is proportional to the difference between actual and predicted outcome

Robert A. Rescorla

Allan R. Wagner



Value update: $V_t = V_{t-1} + \alpha * \mathrm{PE}_{t-1}$

Prediction error: $\mathrm{PE}_{t-1} = R_{t-1} - V_{t-1}$

α – learning rate
PE – reward prediction error
V – value
R – reward

*Expectations on the next trial = the expectation on the current trial + learning rate * prediction error (reward – current expectation)*

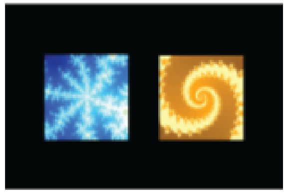
Rescorla & Wagner (1972)

Understand the learning rate

cognitive model
statistics
computing

Value update: $V_t = V_{t-1} + \alpha * \mathrm{PE}_{t-1}$

Prediction error: $\mathrm{PE}_{t-1} = R_{t-1} - V_{t-1}$



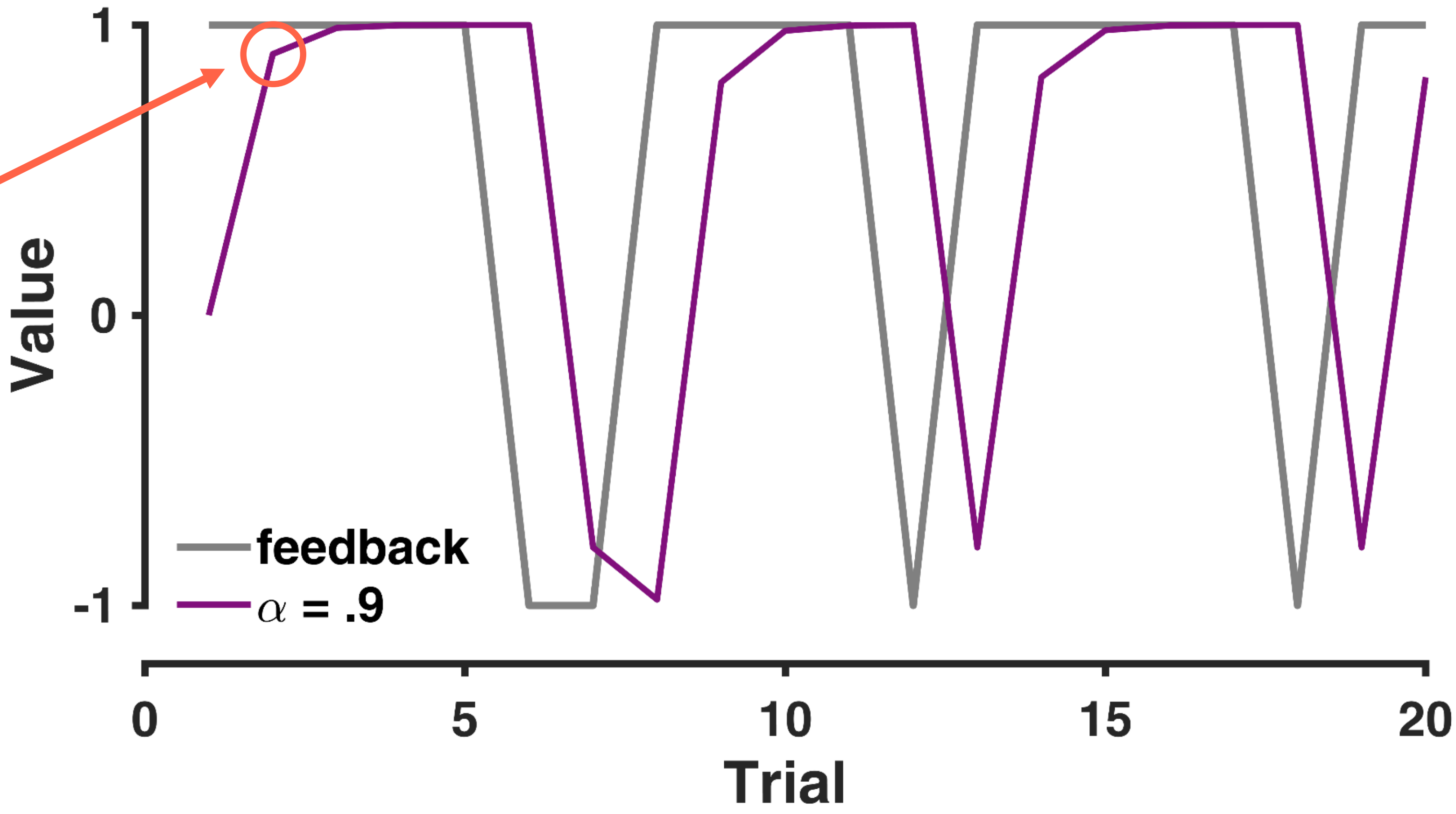
if $\alpha = 0.9$

$V_1 = 0$

$V_2 = V_1 + 0.9 * (1 - V_1)$

$= 0 + 0.9 * (1 - 0)$

$= 0.9$

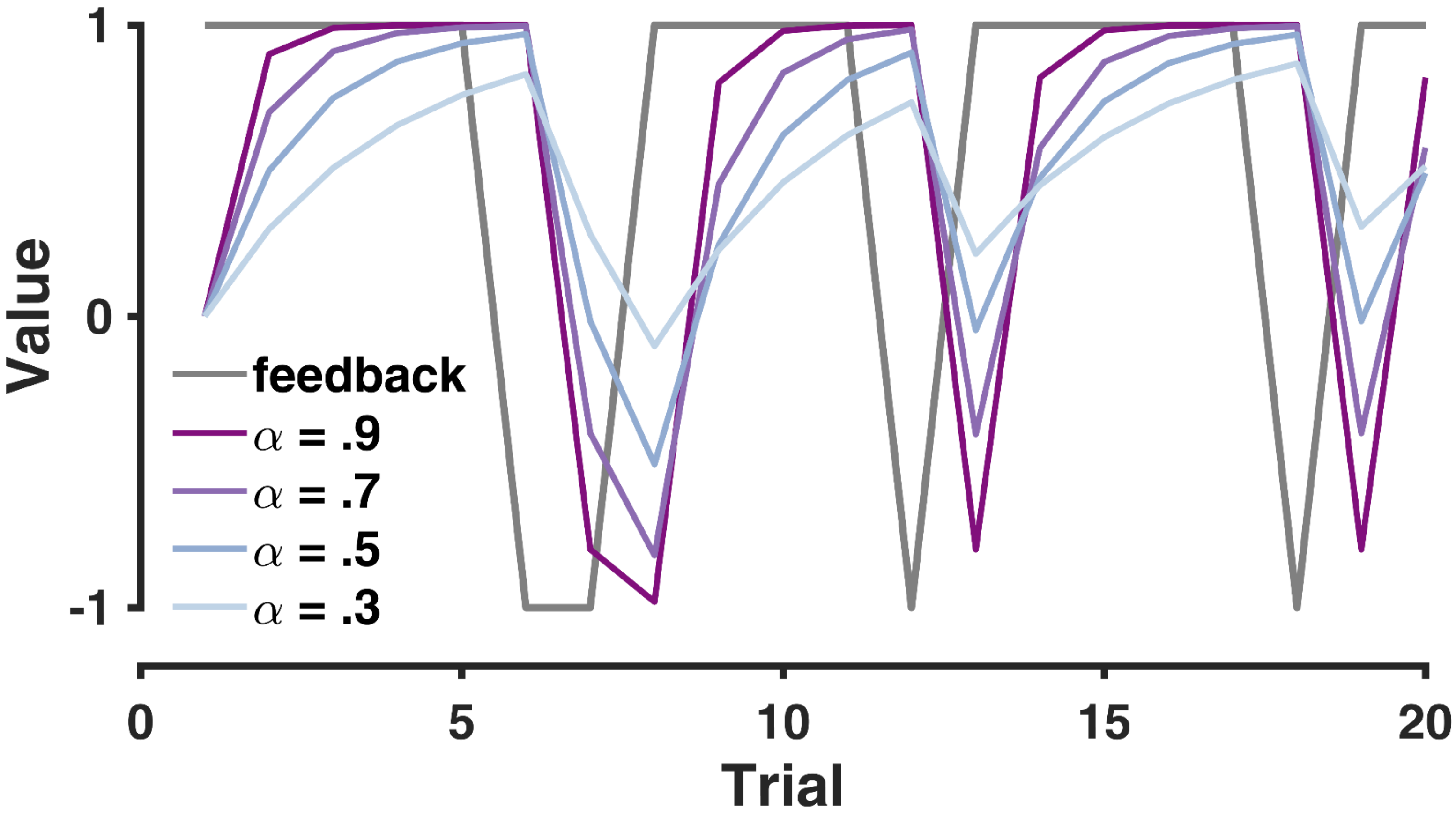


reward contingency – 80:20

Understand the learning rate

Value update: $V_t = V_{t-1} + \alpha * \mathrm{PE}_{t-1}$

Prediction error: $\mathrm{PE}_{t-1} = R_{t-1} - V_{t-1}$



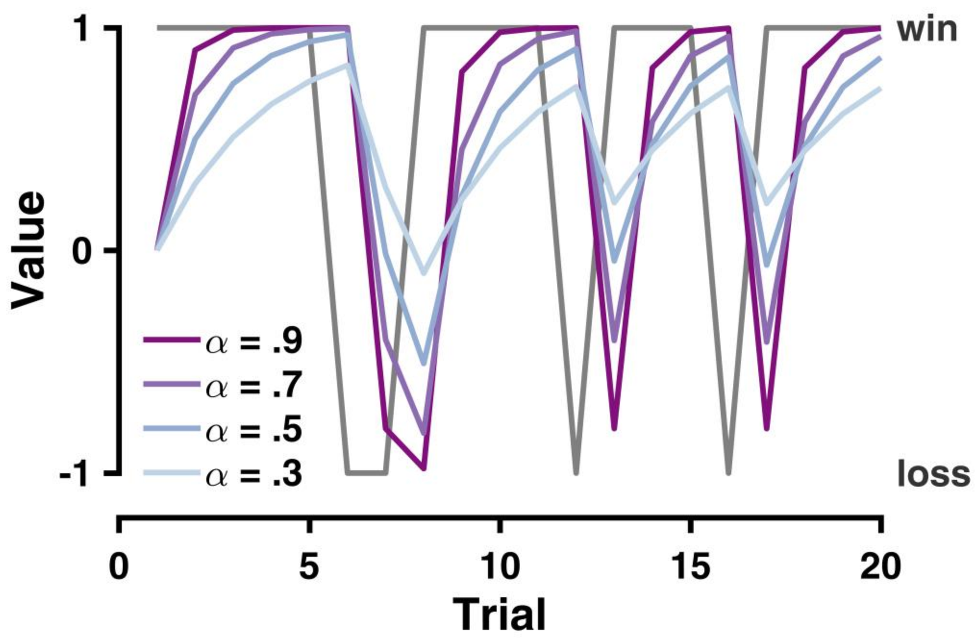
reward contingency – 80:20

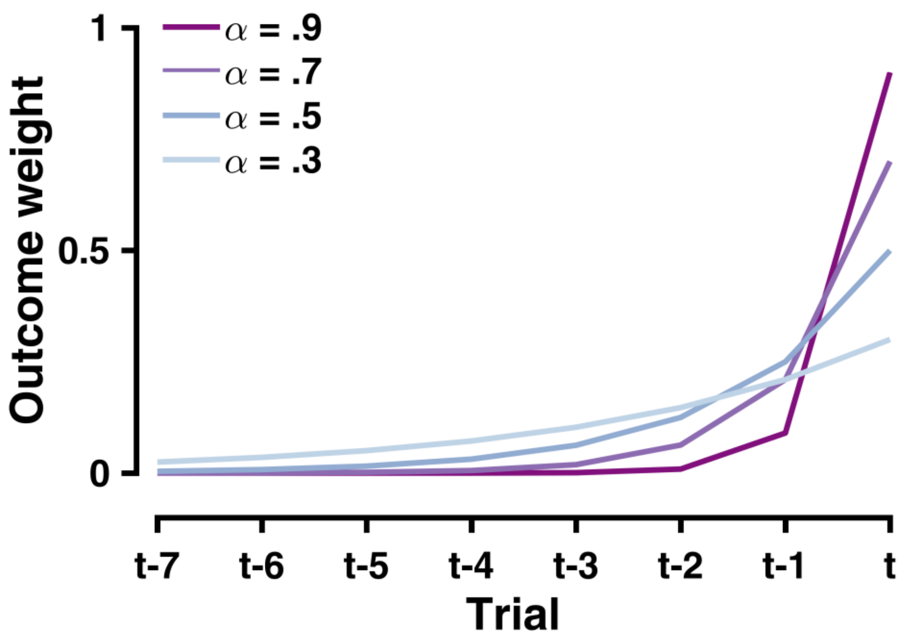
Understand the learning rate

Value update: $V_t = V_{t-1} + \alpha * \mathrm{PE}_{t-1}$

Prediction error: $\mathrm{PE}_{t-1} = R_{t-1} - V_{t-1}$

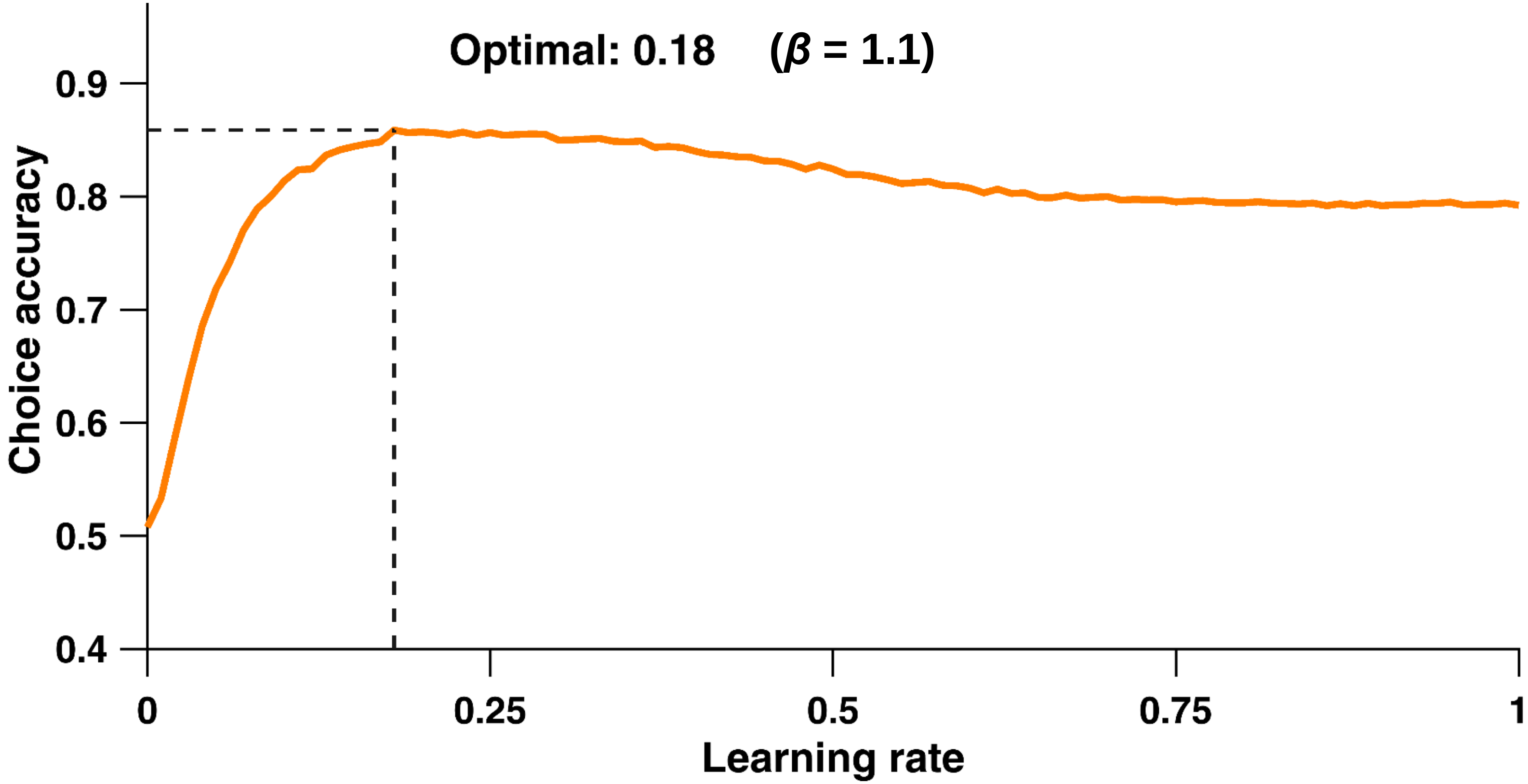
$$\begin{aligned} V_t &= (1-\alpha)V_{t-1} + \alpha R_{t-1} \\ &= (1-\alpha)(V_{t-2} + \alpha(R_{t-2} - V_{t-2})) + \alpha R_{t-1} \\ &= (1-\alpha)^t V_0 + \boxed{\sum_{i=1}^{t} (1-\alpha)^{t-i} \alpha R_i} \end{aligned}$$

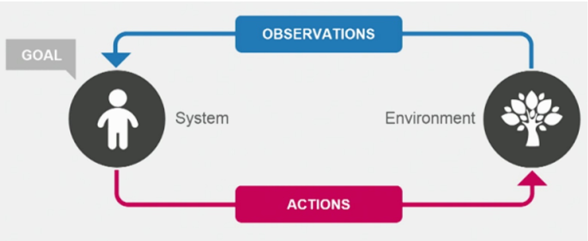




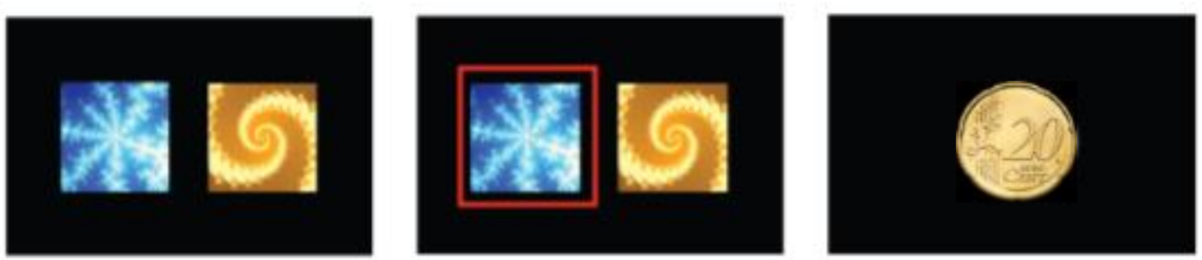
Optimal learning rate?

cognitive model
statistics
computing





Rescorla-Wagner Value Update



Value update:

$$V_{t+1} = V_t + \alpha * PE_t$$

Prediction error:

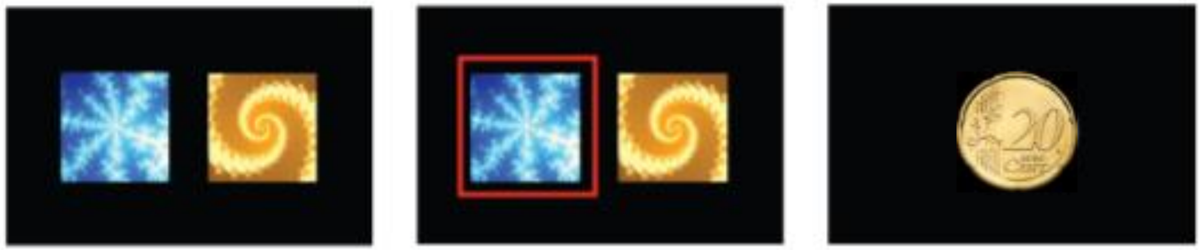
$$PE_t = R_t - V_t$$

choice rule:

greedy / ε-greedy / softmax

Rescorla & Wagner (1972)

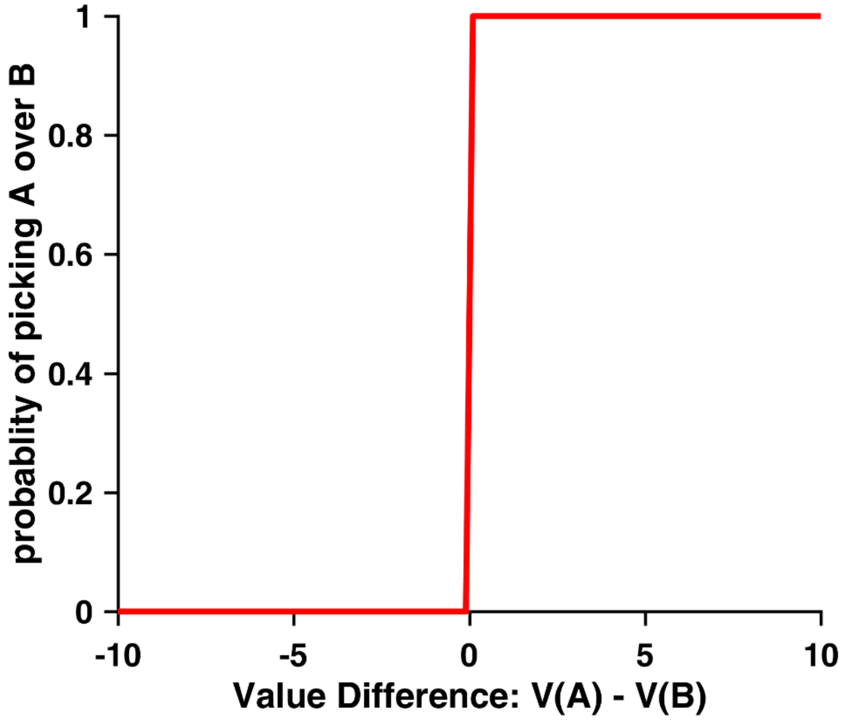
Choice rule: greedy



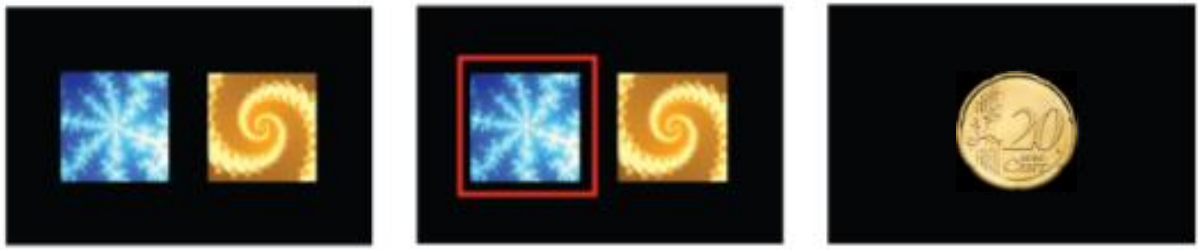
V()

V()

$$p(C=a)=\begin{cases}1, V(a)>V(b)\\0, V(a)<V(b)\end{cases}$$

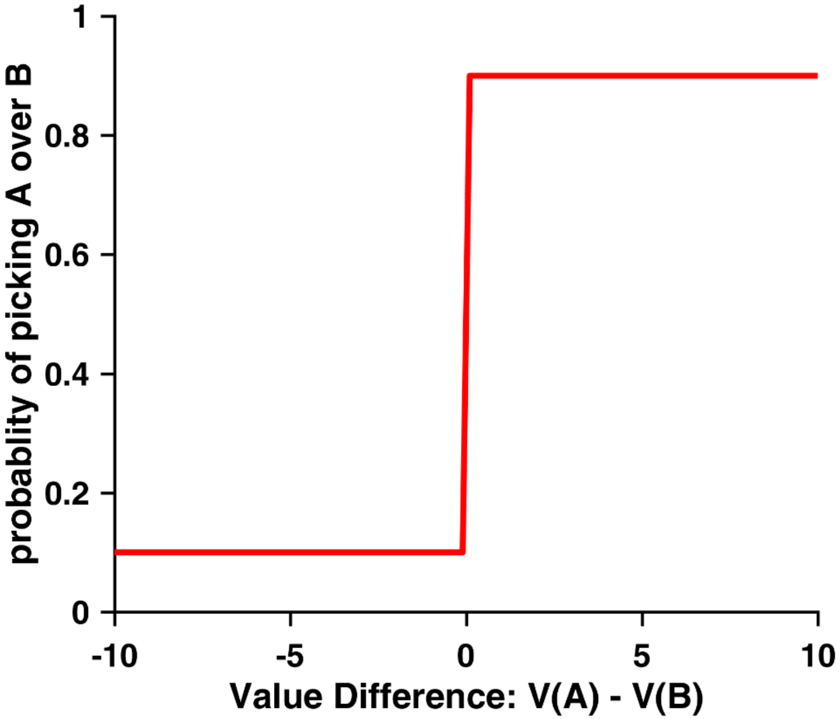


Choice rule: ε-greedy

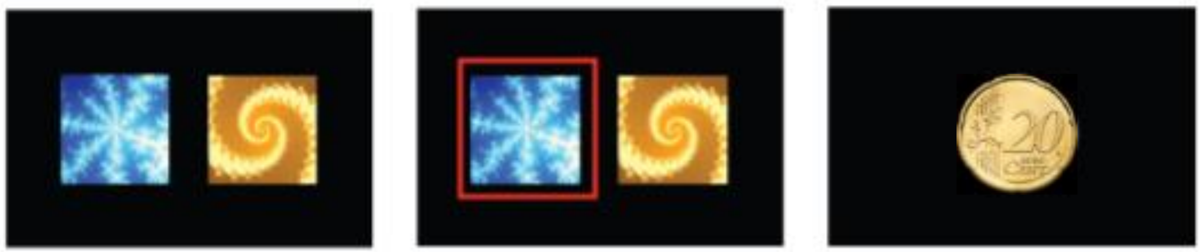


V()
V()

$$p(C=a)=\begin{cases}1-\varepsilon, & V(a)>V(b)\\ \varepsilon, & V(a)<V(b)\end{cases}$$



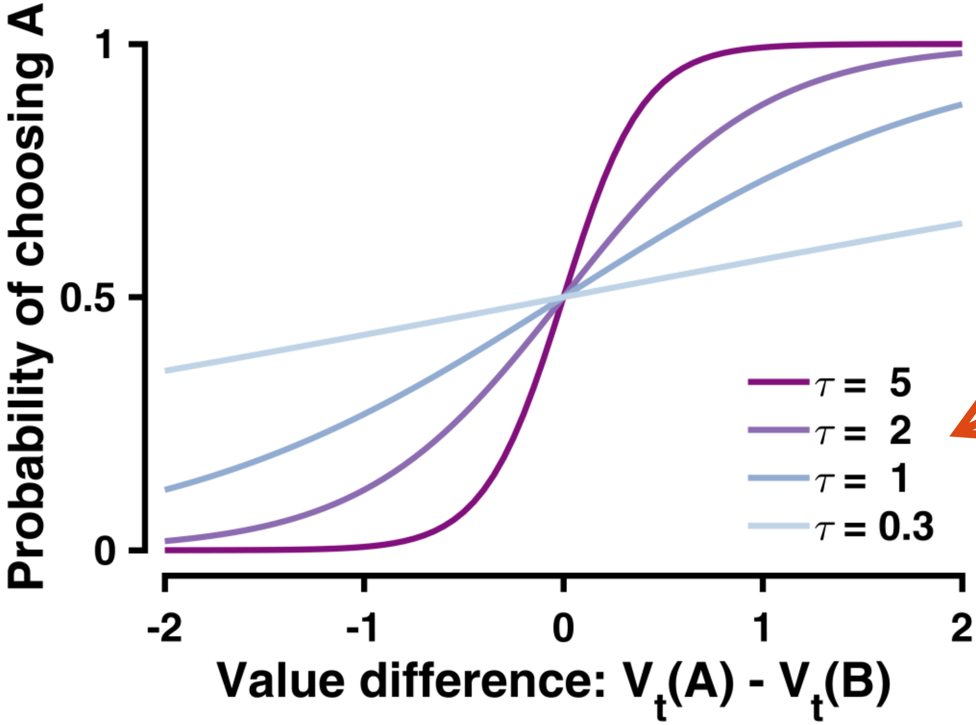
Choice rule: softmax



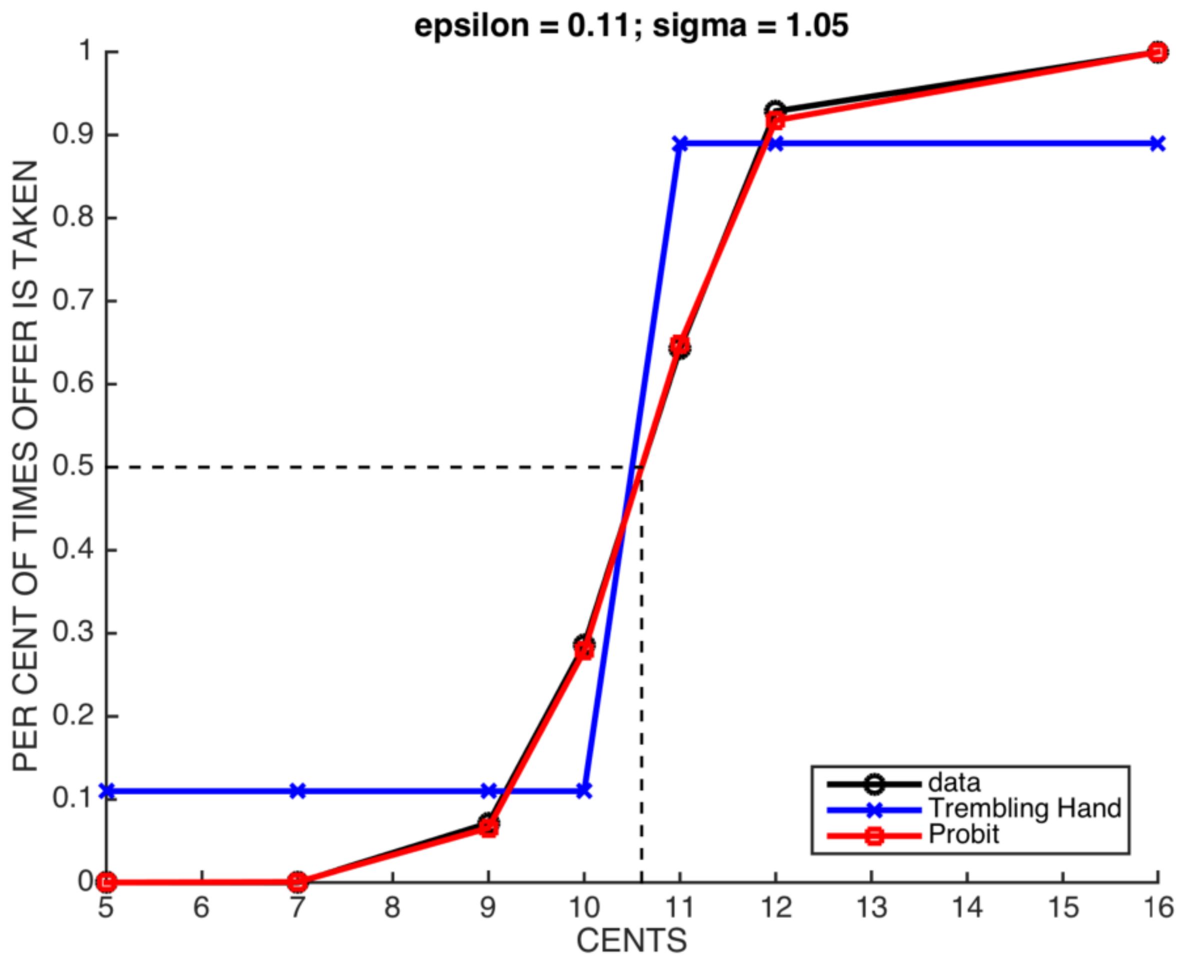
V()
V()

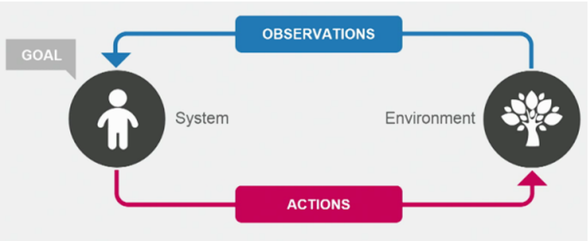
$$\boxed{p_t(A) = \frac{e^{\tau * V_t(A)}}{e^{\tau * V_t(A)} + e^{\tau * V_t(B)}}}$$

$$= \frac{1}{1 + e^{-\tau * (V_t(A) - V_t(B))}}$$

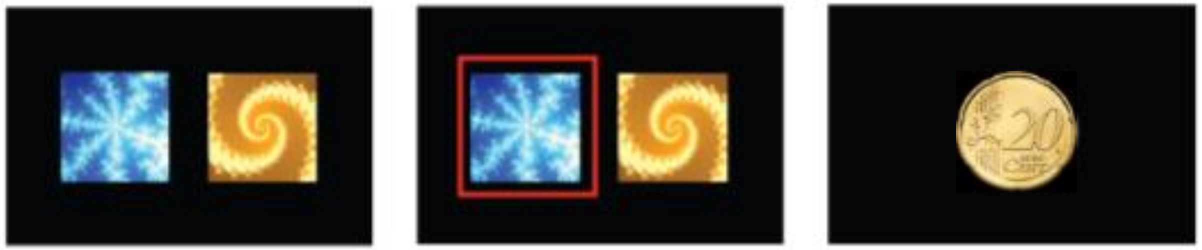


Choice rule: direct comparison





Rescorla-Wagner Value Update



Value update:

$$V_{t+1} = V_t + \alpha * PE_t$$

Prediction error:

$$PE_t = R_t - V_t$$

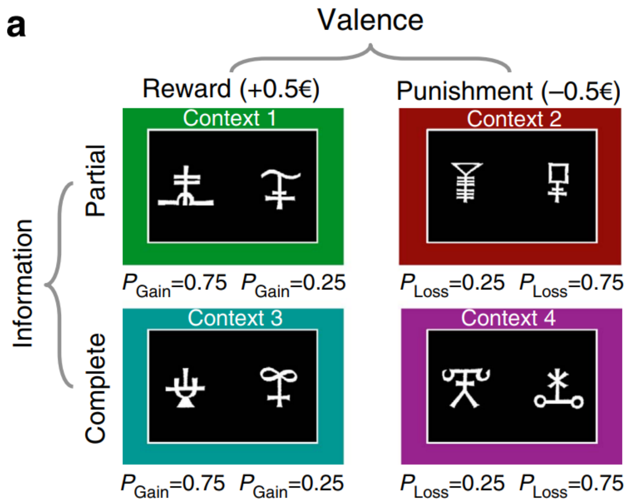
choice rule (sigmoid /softmax):

$$p(C{=}a) = \frac{1}{1+e^{\tau*(v(b)-v(a))}}$$

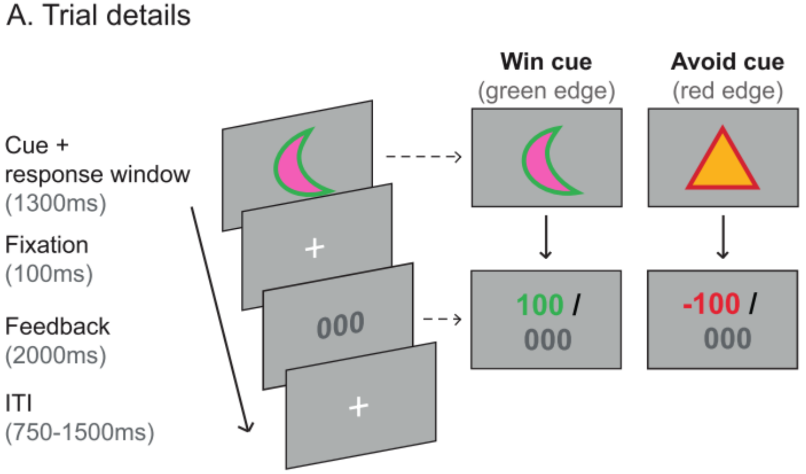
```
α  - learning rate
PE - reward prediction error
V  - value
R  - reward
τ  - softmax temperature
```

Rescorla & Wagner (1972)

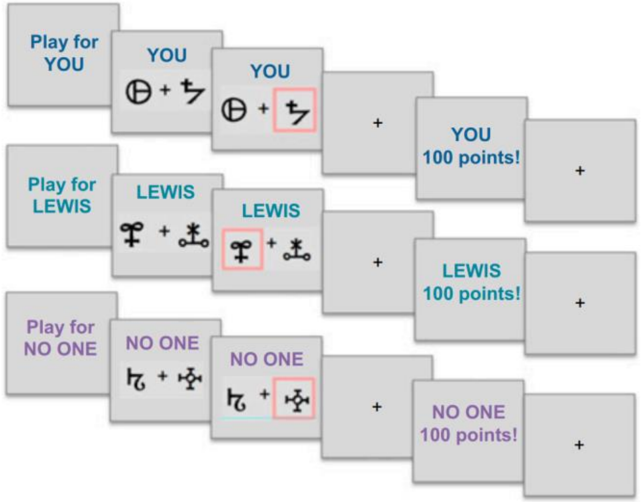
Generalizing RL framework



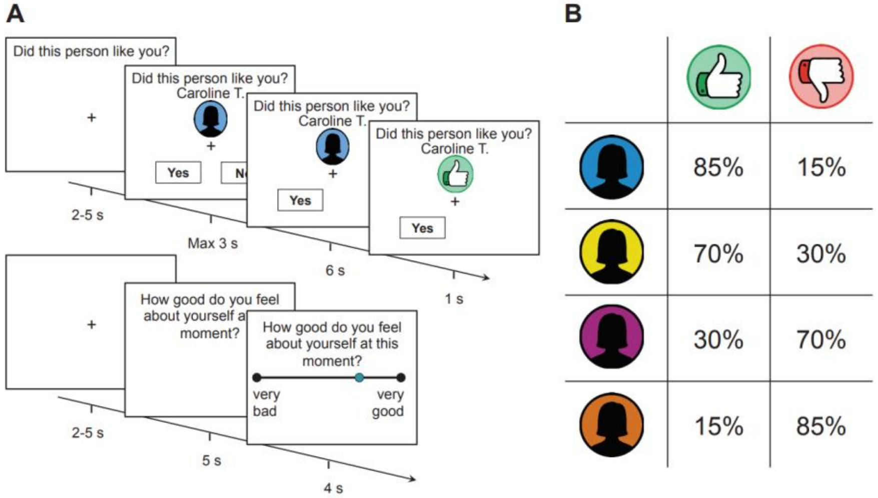
Palminteri et al. (2015)



Swart et al. (2017)



Lockwood et al. (2016)



Will et al. (2017)


ANY QUESTIONS ?

Happy Computing!