Bayesian Statistics and Hierarchical Bayesian Modeling for Psychological Science

Lecture 07

Lei Zhang

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

Department of Basic Psychological Research and Research Methods

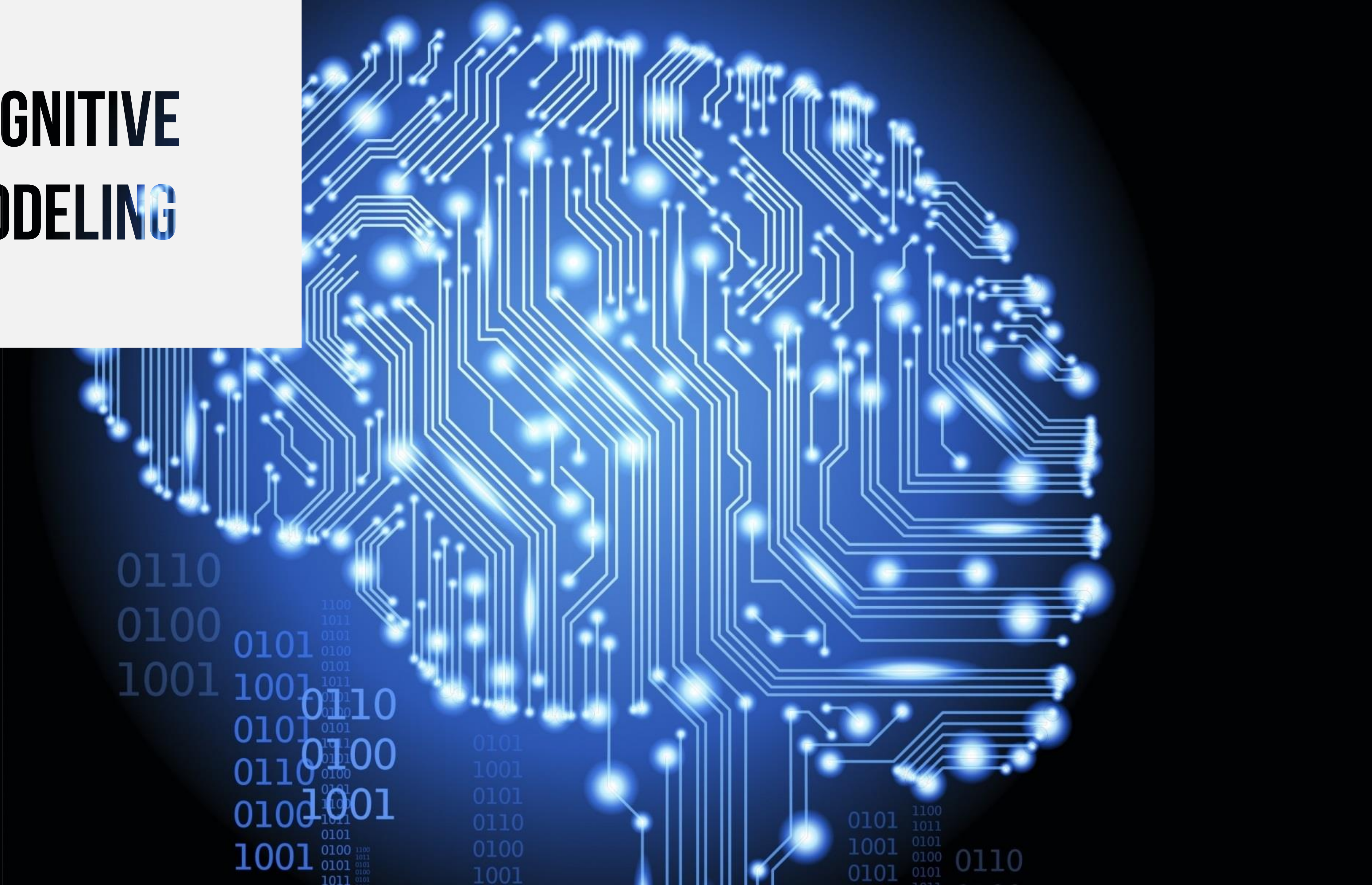
lei.zhang@univie.ac.at
lei-zhang.net
@lei_zhang_lz

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

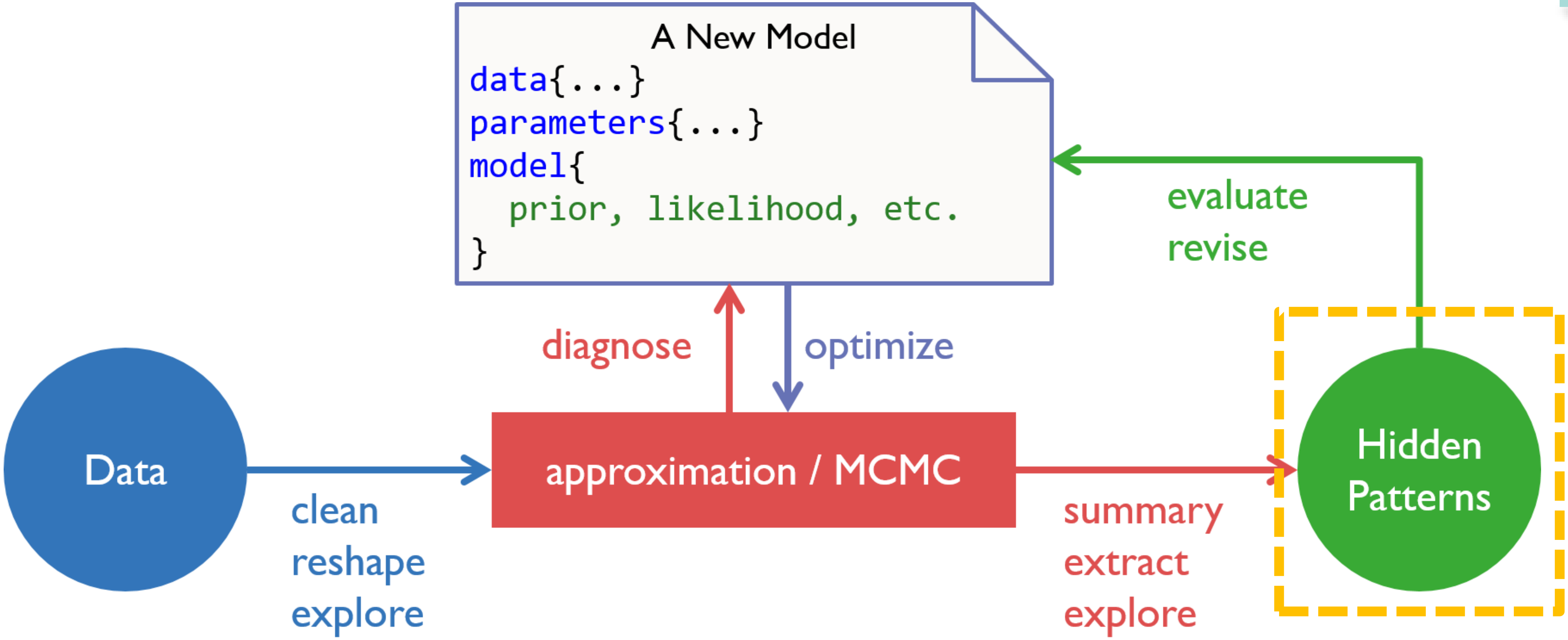
universität wien

Fakultät für Psychologie

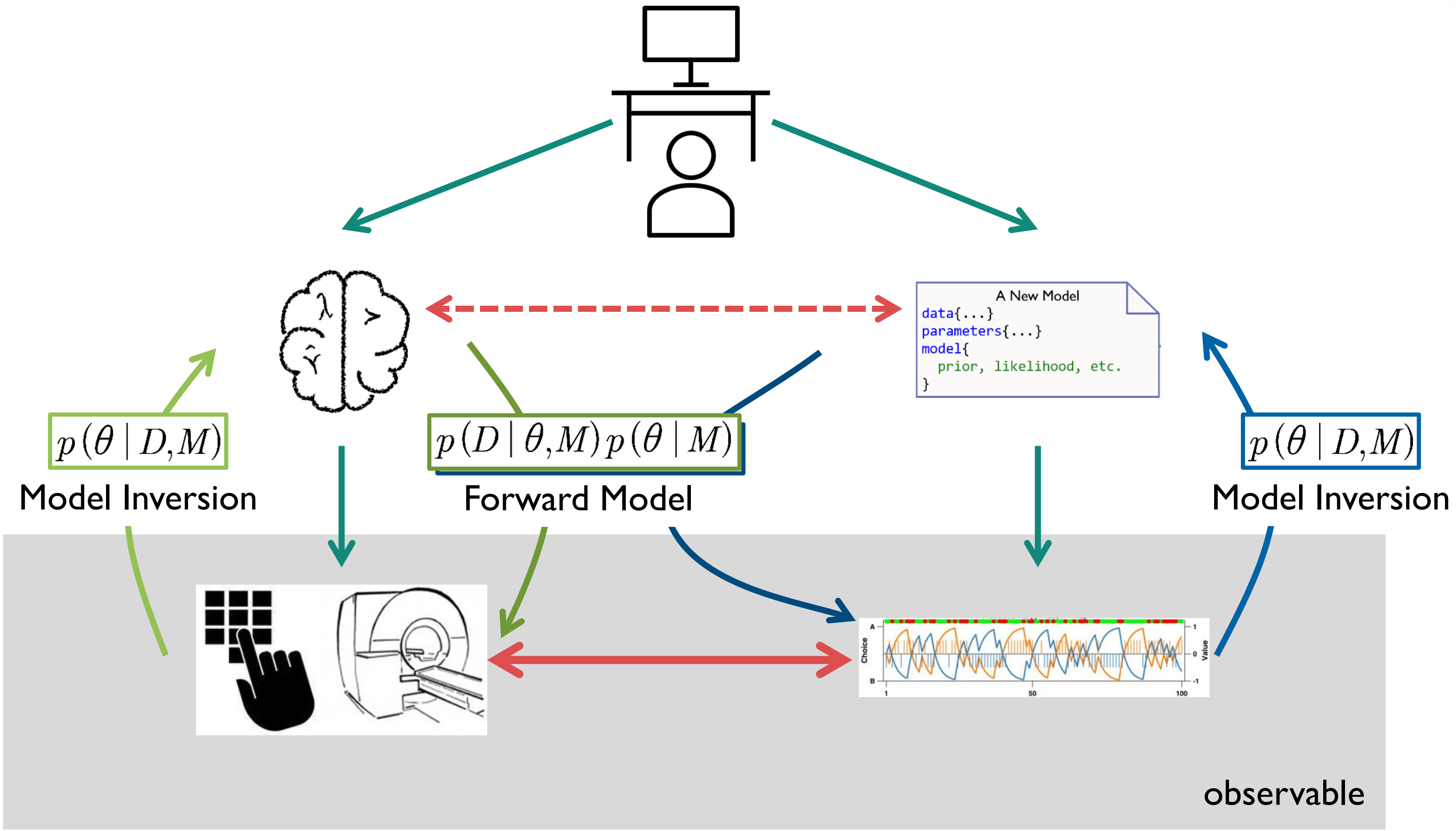
COGNITIVE MODELING



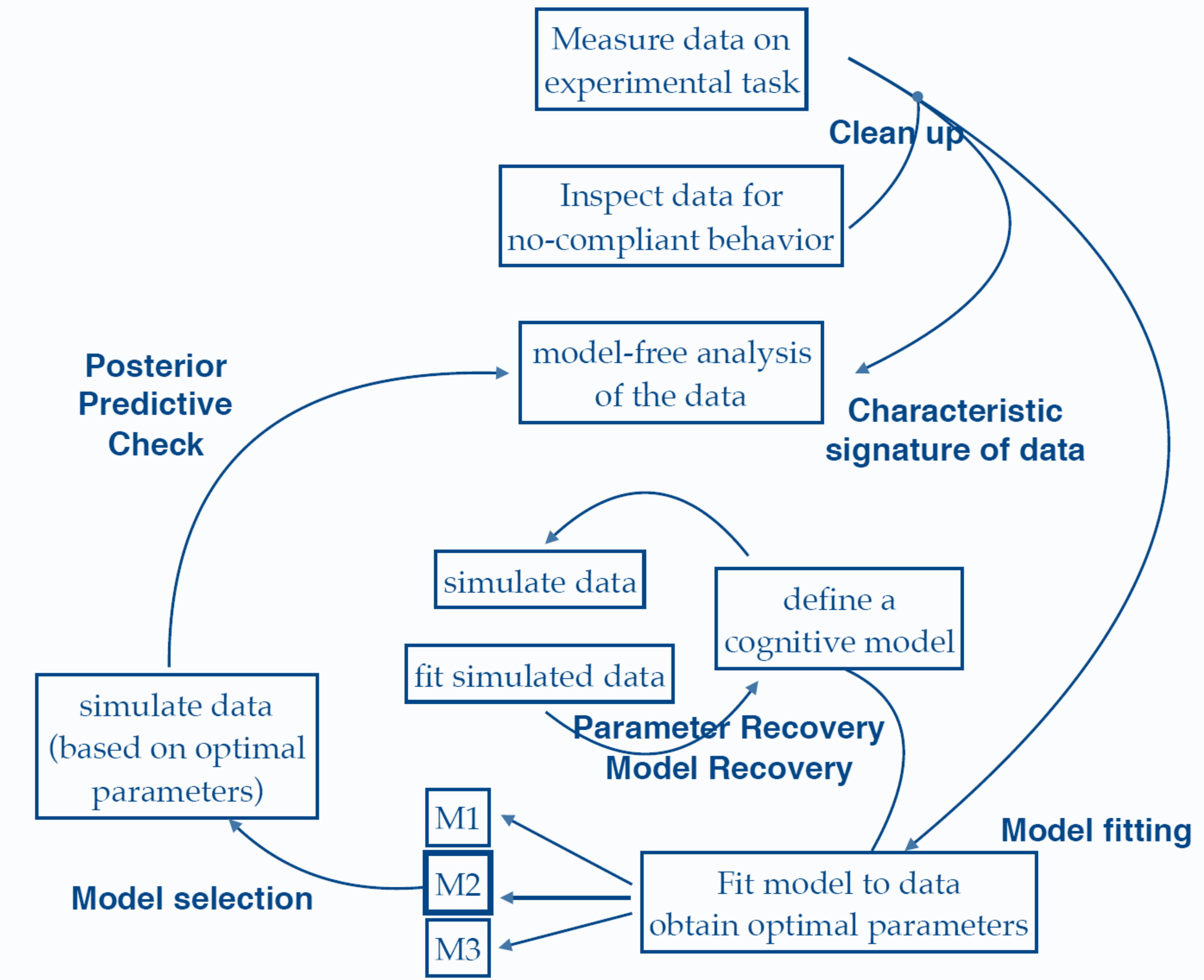
cognitive model
statistics
computing



What is Cognitive Modeling?



Lewandowsky & Farrell (2011); Stephan et al. (2016); Ahn et al., (2017)



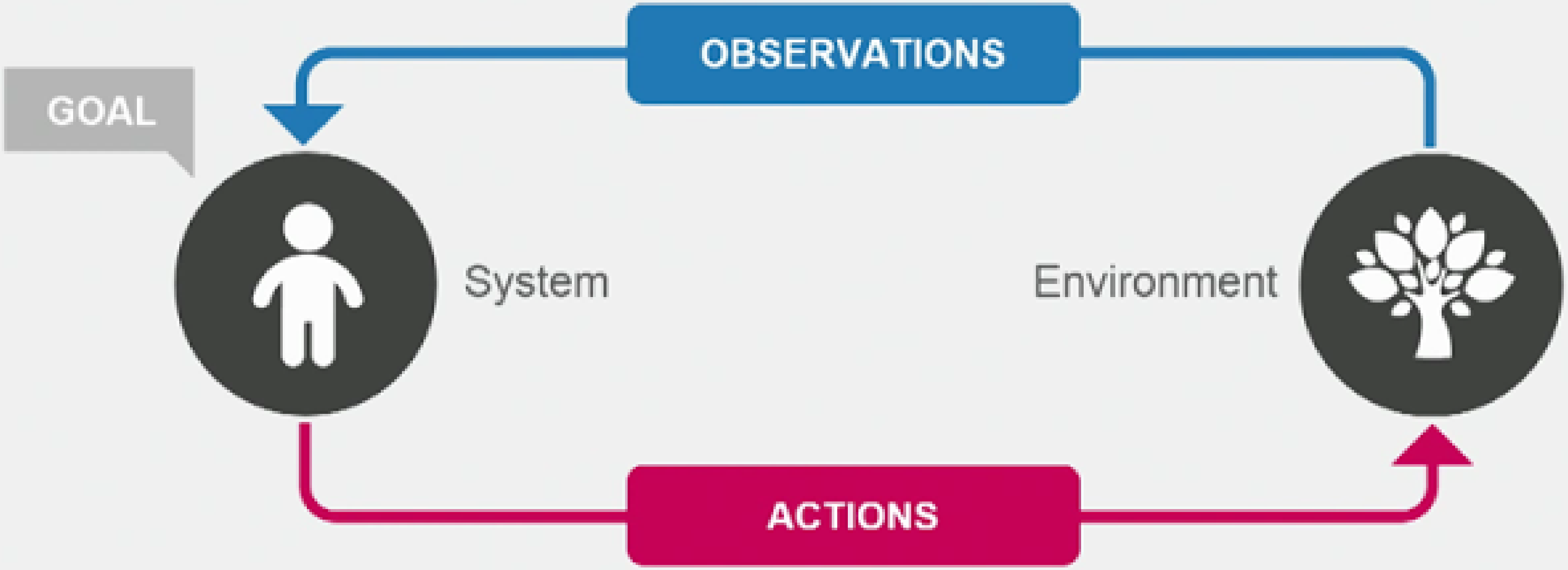
cognitive model
statistics
computing

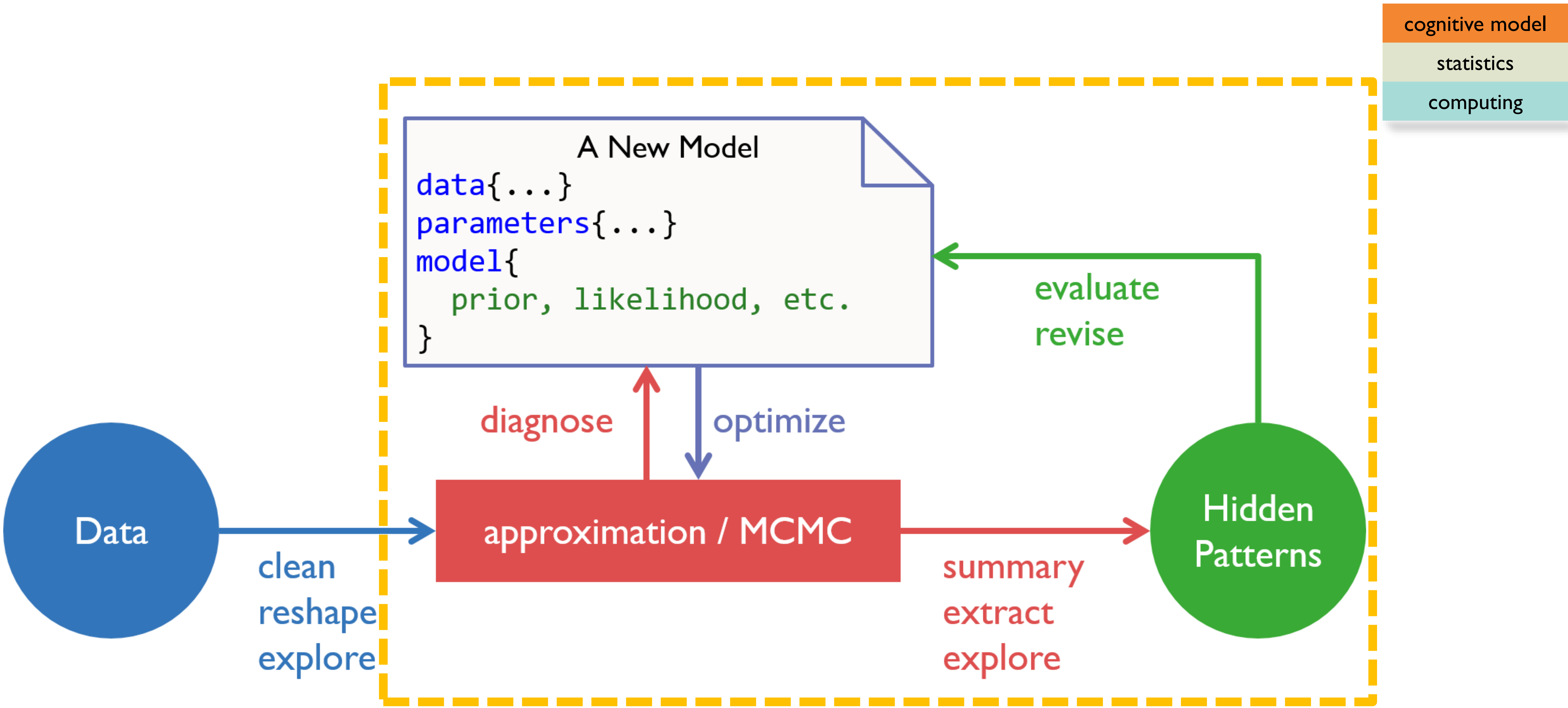
Essentially, all the models are wrong, but some are useful.

– George E. P. Box

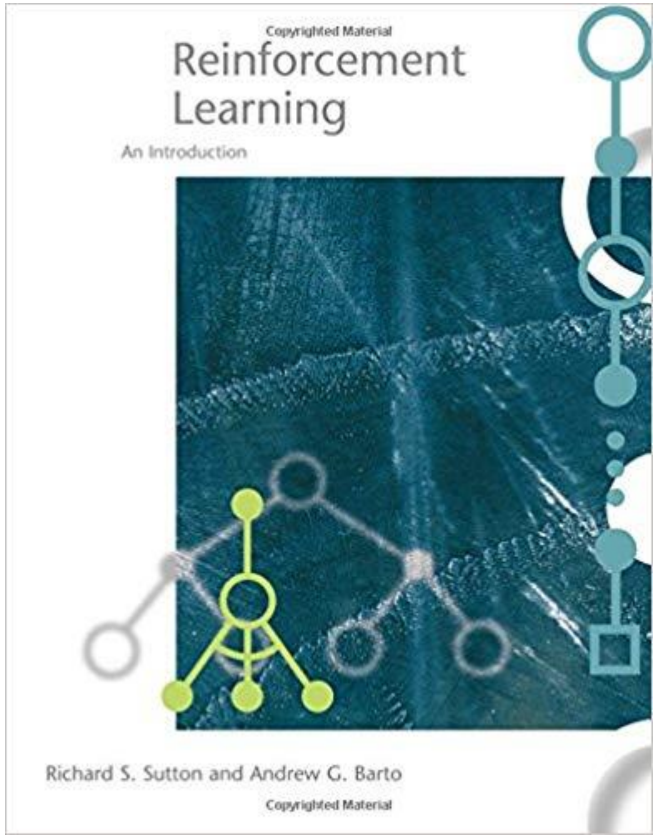
Essentially, all the models are ~~wrong~~ imperfect, but some are useful.

REINFORCEMENT LEARNING FRAMEWORK

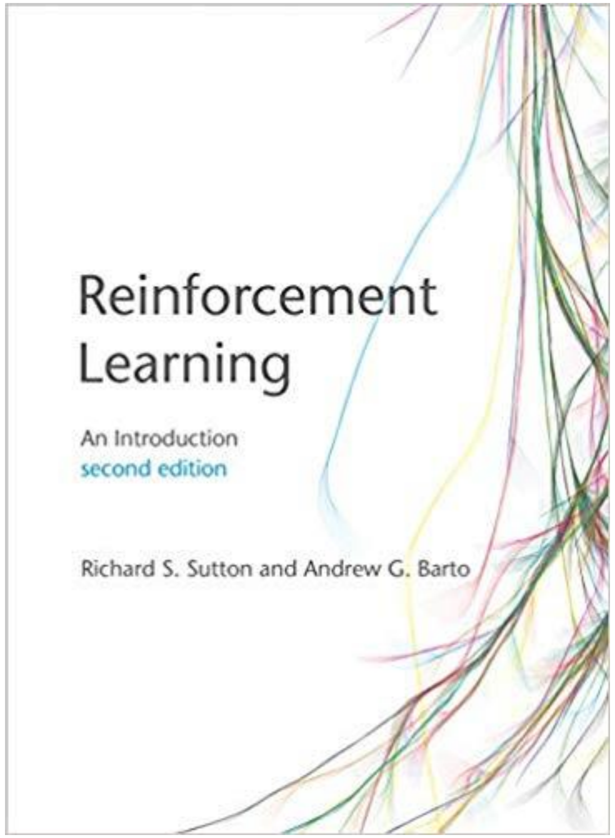



cognitive model
statistics
computing
A New Model
data{...}
parameters{...}
model{
prior, likelihood, etc.
}
evaluate
revise
diagnose
optimize
Data
approximation / MCMC
Hidden Patterns
clean
reshape
explore
summary
extract
explore

The very short history

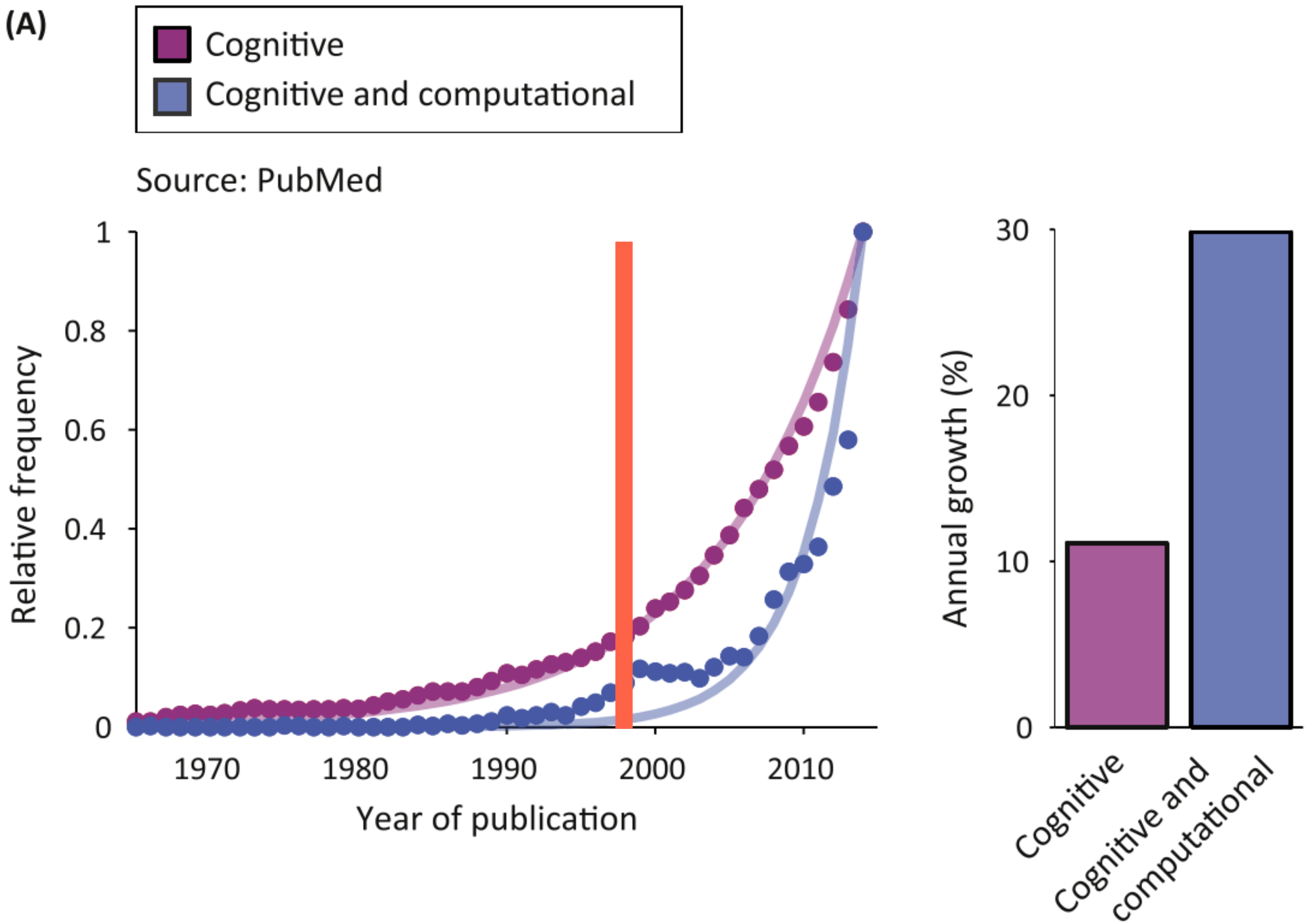


1998



2018

Boom in Cognitive Modeling



Palminteri et al. (2017)

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

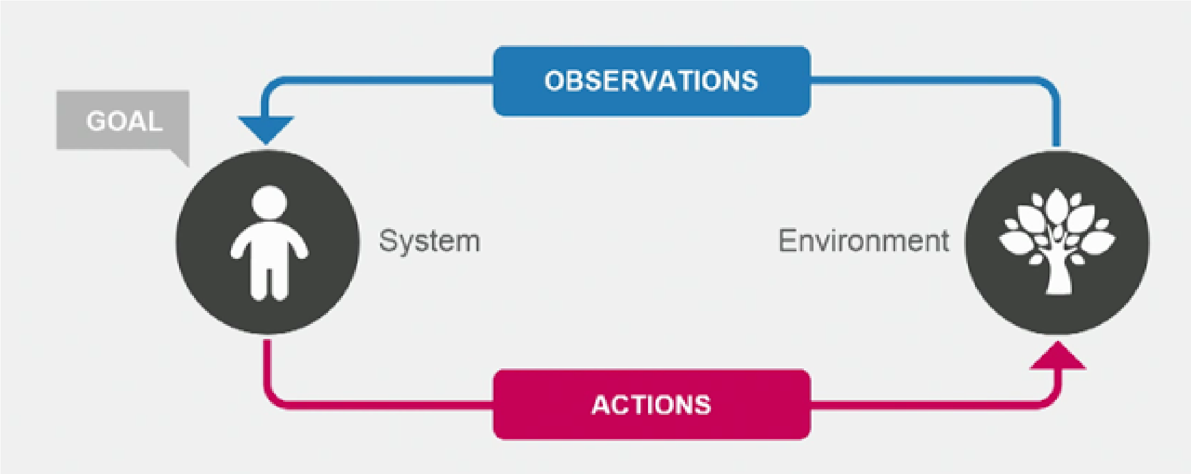
2-armed bandit task

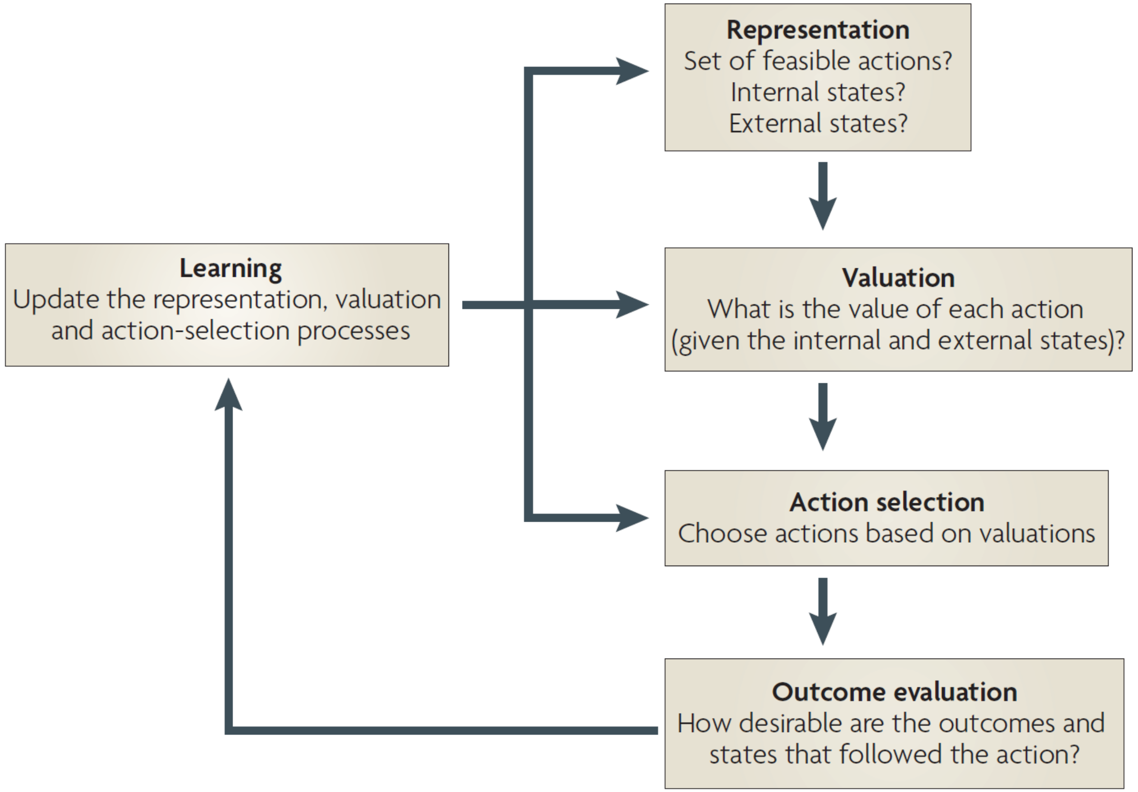
What can be your strategies:

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?





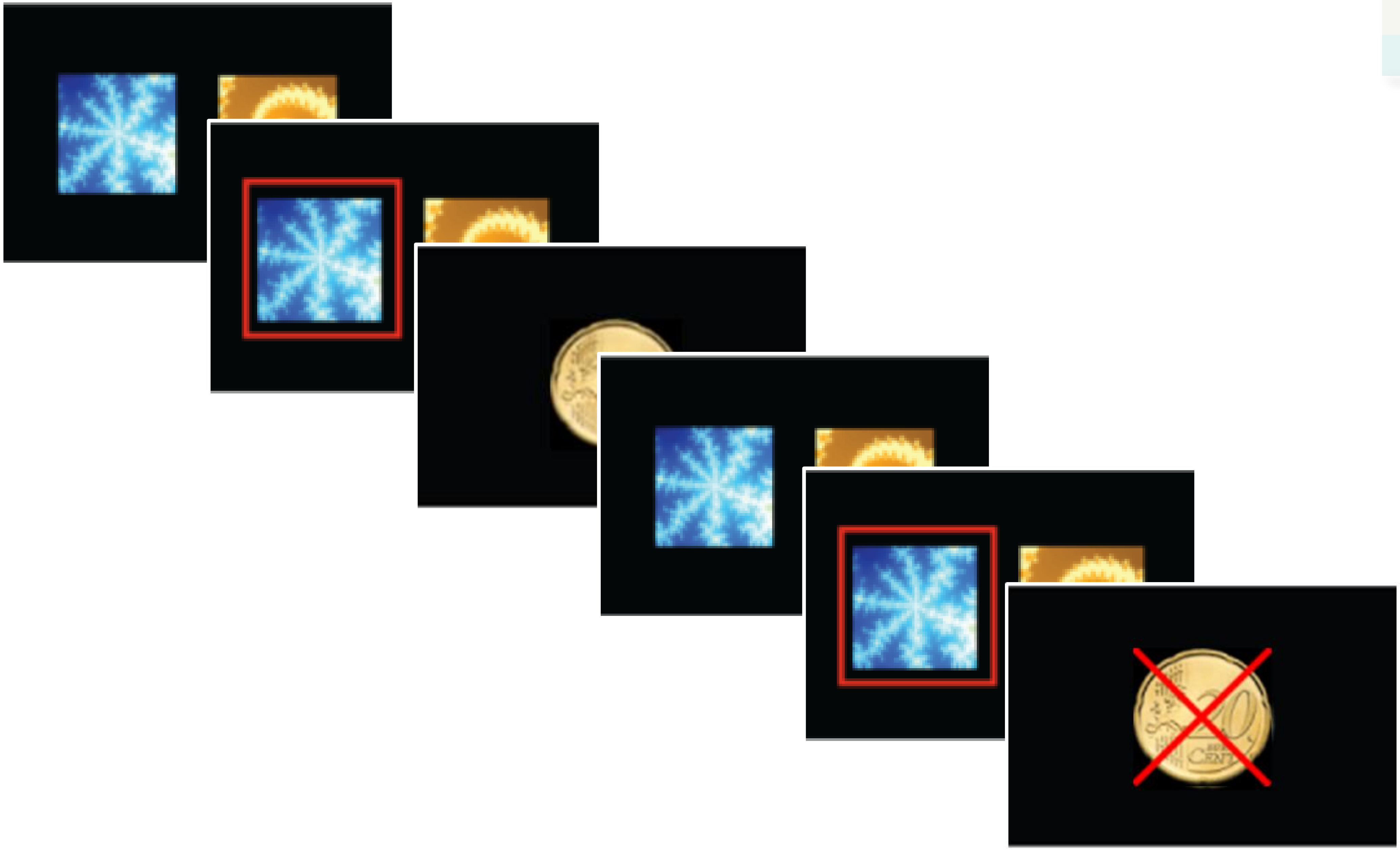
Sutton & Barto (1998); Rangel et al (2008); [Daw & Niv workshop 2015]

Modeling the 2-armed bandit task

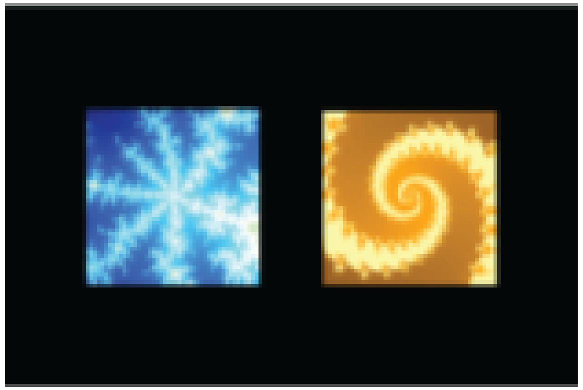
how do you suggest to model this learning process?

suppose we ran this experiment on a person

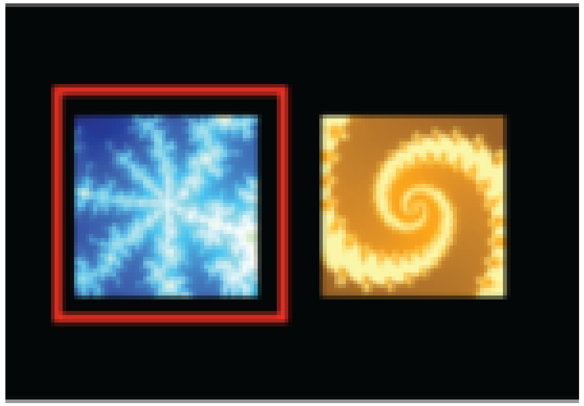
our models are basically detailed hypotheses about behavior and about the brain… we can test these hypotheses!



One simple experiment: two choice task



choice presentation



action selection

outcome

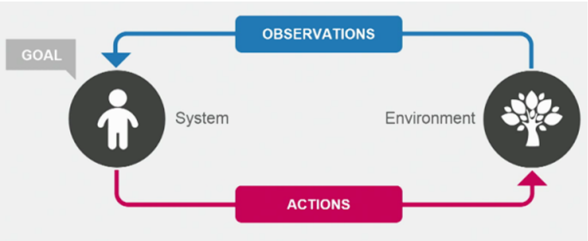
reward contingency – 80:20

Elements

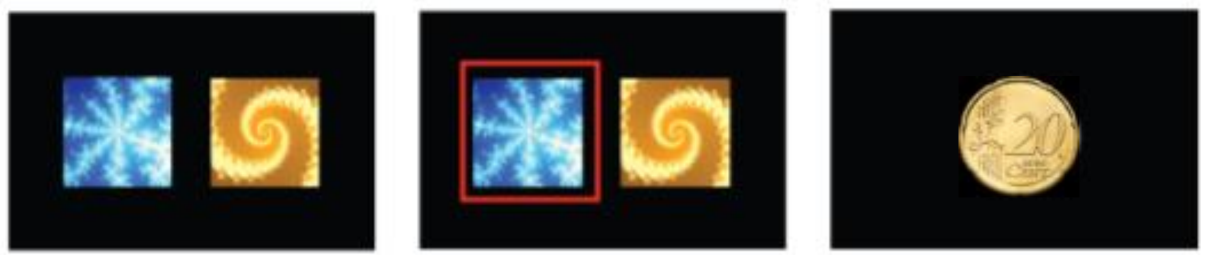
what do we know? — Data: choice & outcome

what can we measure? — Summary stats: choice accuracy

what do we not know? — Learning algorithm: RL update



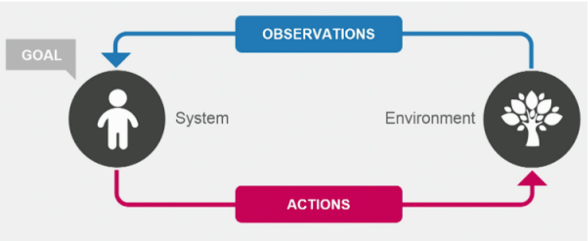
Rescorla-Wagner Value Update



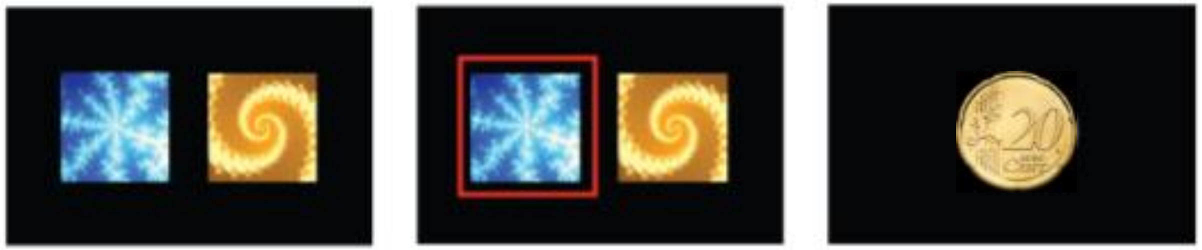
Cognitive Model
- cognitive process
- using internal variables and free parameters

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

Rescorla & Wagner (1972)



Rescorla-Wagner Value Update



Value update:

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

Prediction error:

$$PE = R_t - V_t$$

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

Rescorla & Wagner (1972)

Understand the learning rate

Value update:

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

Prediction error:

$$PE = R_t - V_t$$

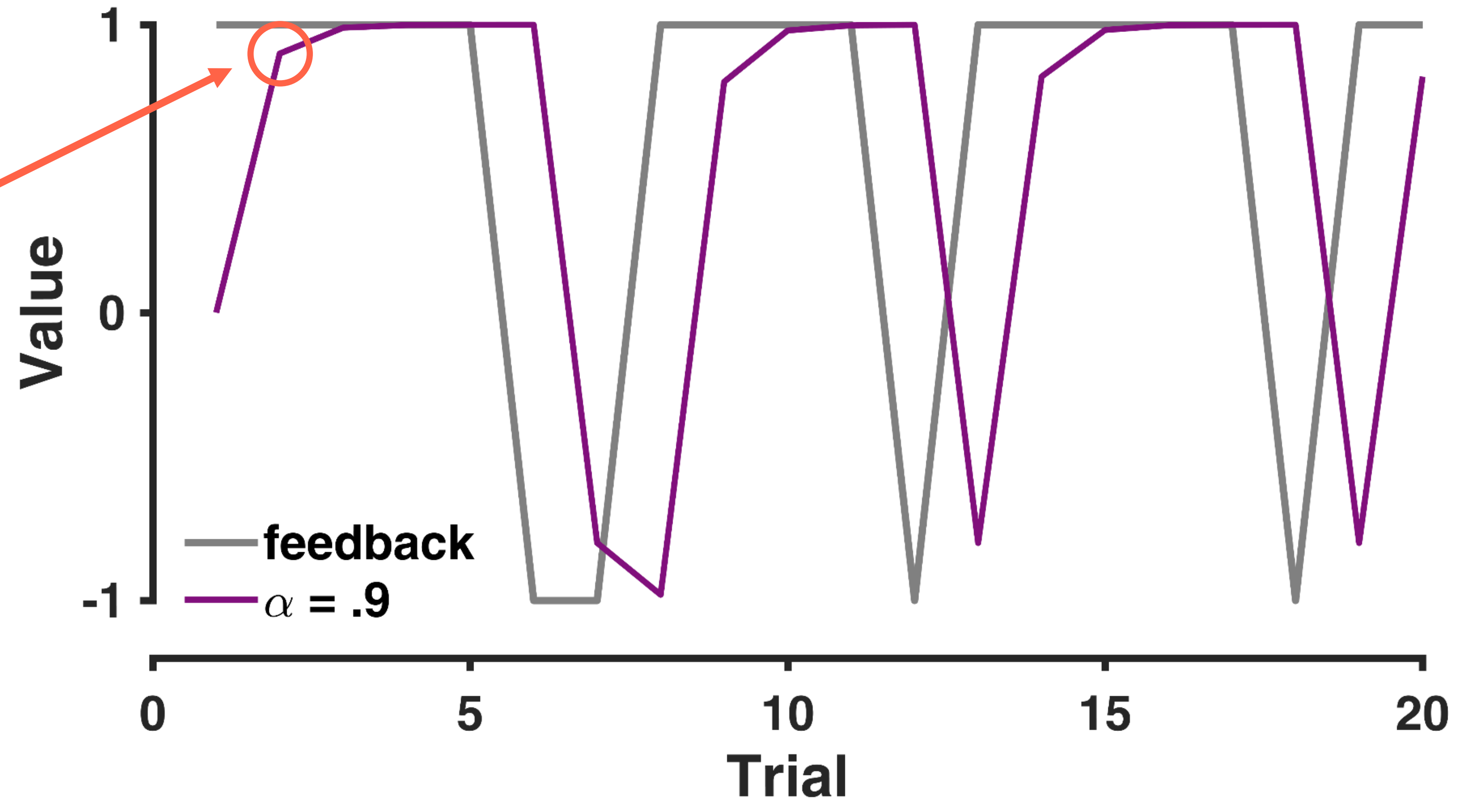
if $\alpha = 0.9$

$$V_1 = 0$$

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

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

$$= 0.9$$



reward contingency – 80:20

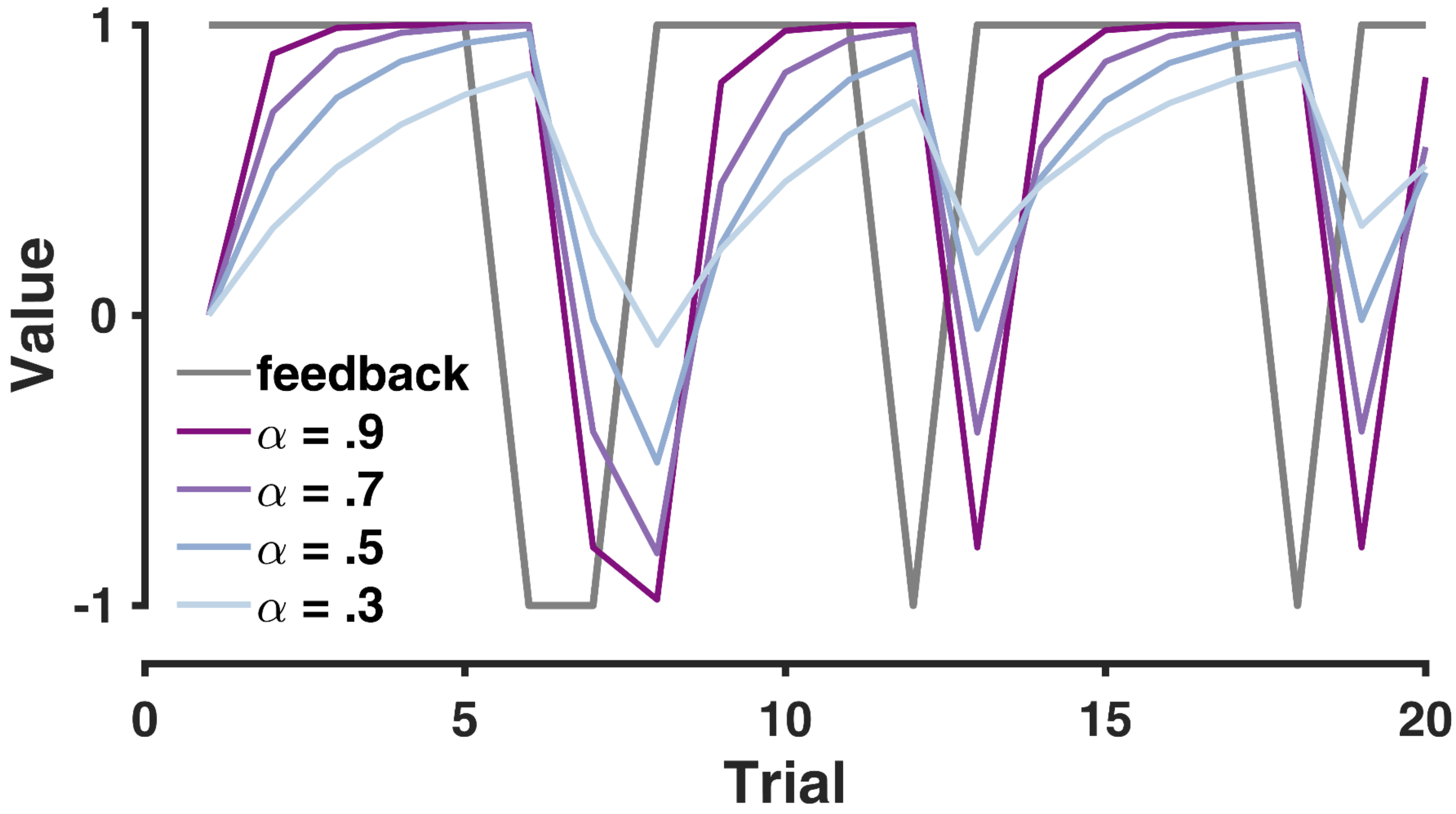
Understand the learning rate

Value update:

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

Prediction error:

$$PE = R_t - V_t$$



reward contingency – 80:20

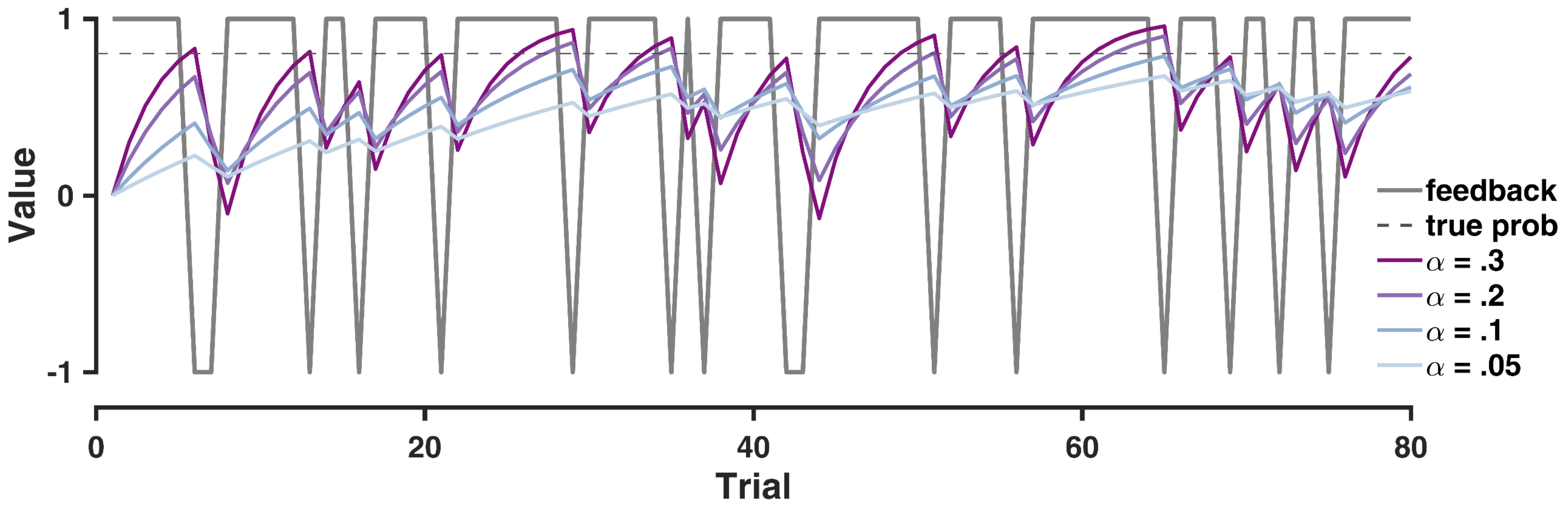
Understand the learning rate

Value update:

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

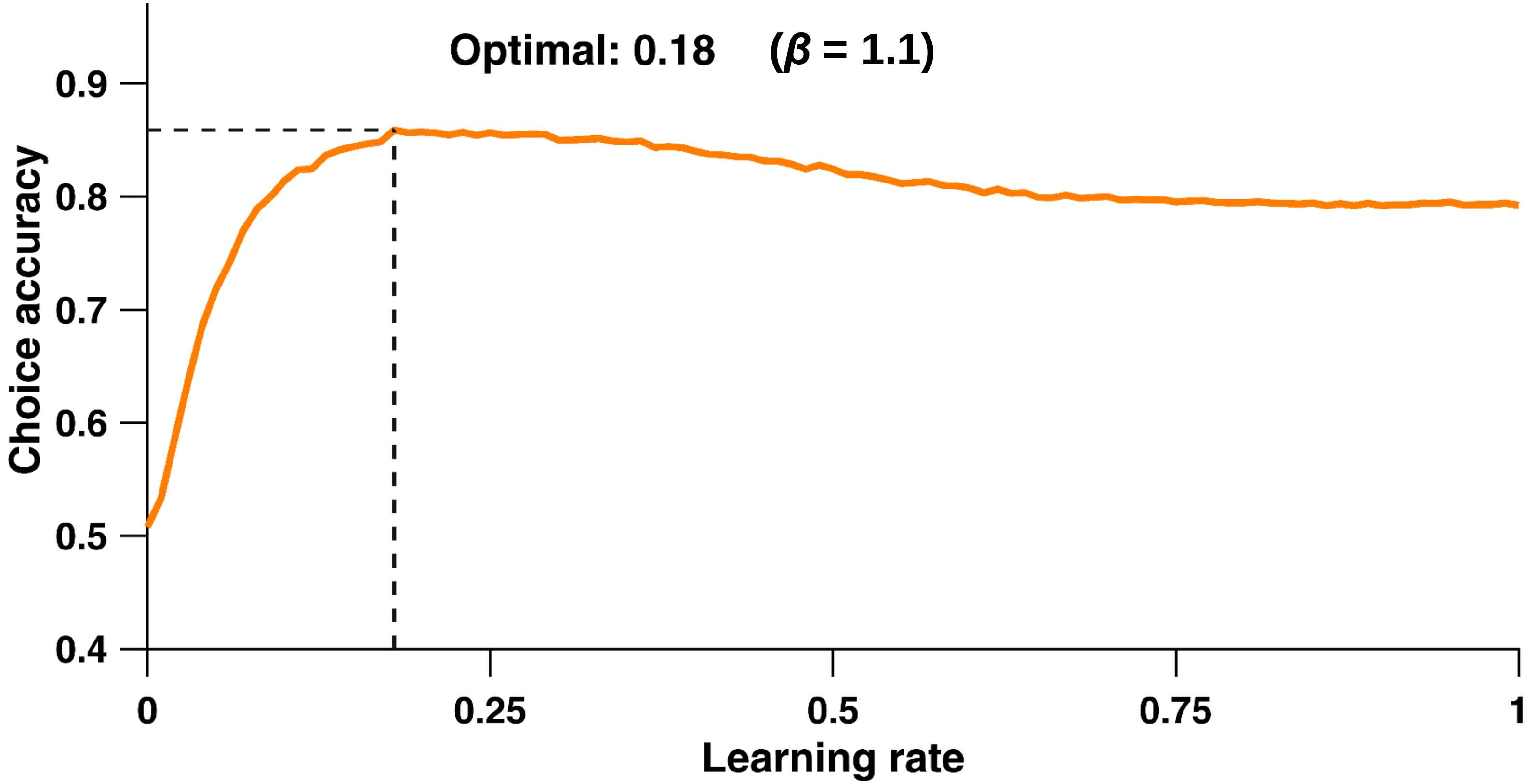
Prediction error:

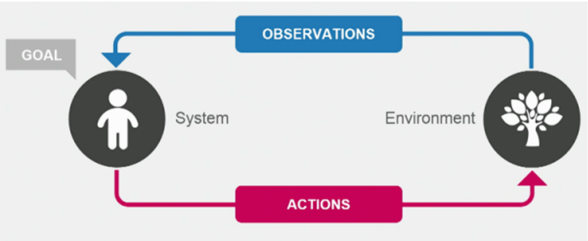
$$PE = R_t - V_t$$



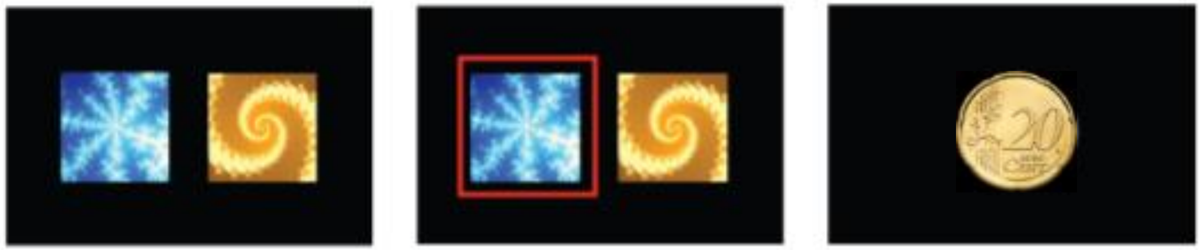
reward contingency – 80:20

Optimal learning rate?





Rescorla-Wagner Value Update



Value update:

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

Prediction error:

$$PE = R_t - V_t$$

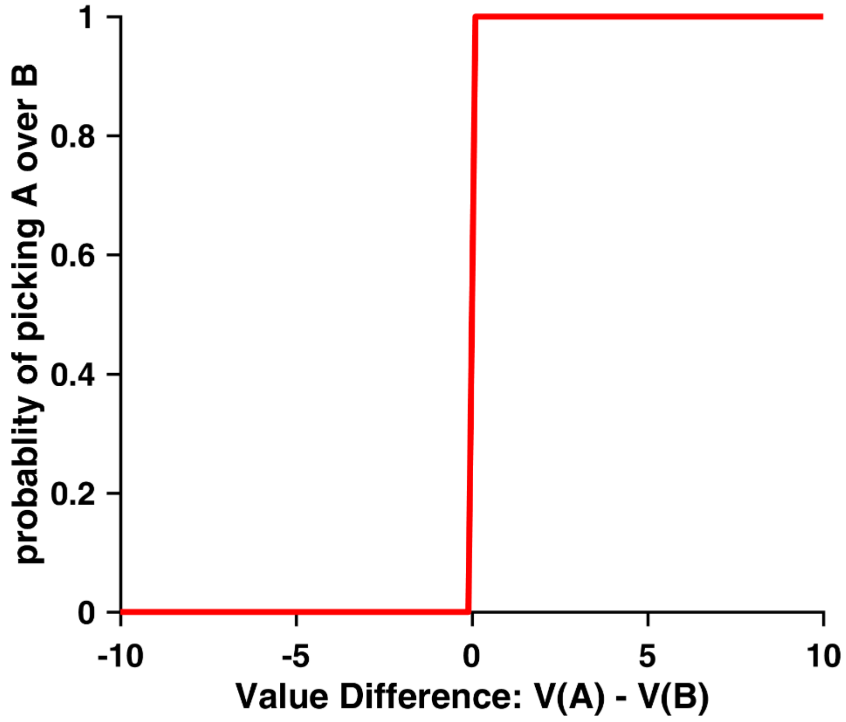
choice rule:

greedy / ε-greedy / softmax

Rescorla & Wagner (1972)

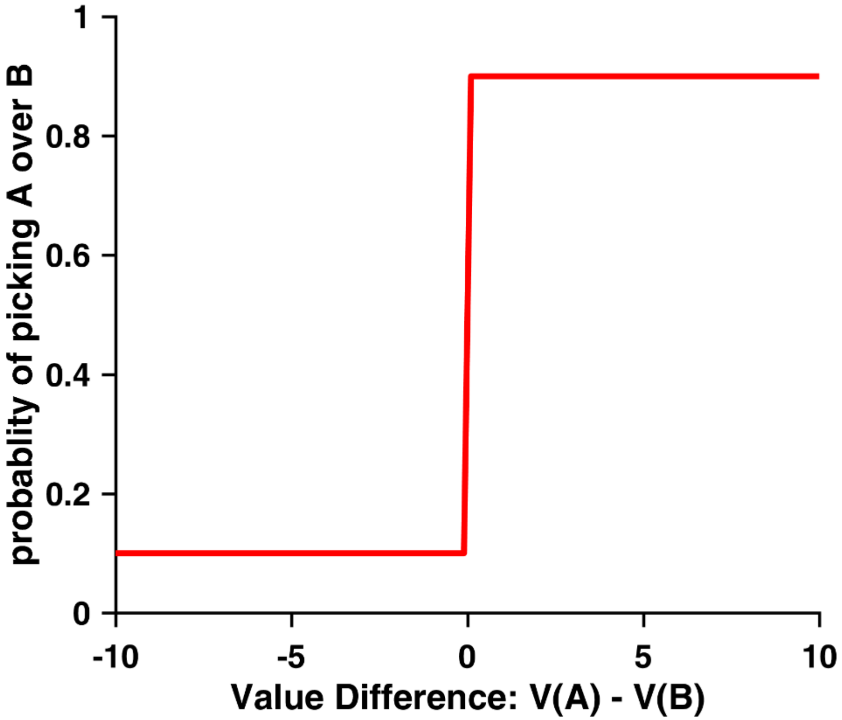
Choice rule: greedy

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



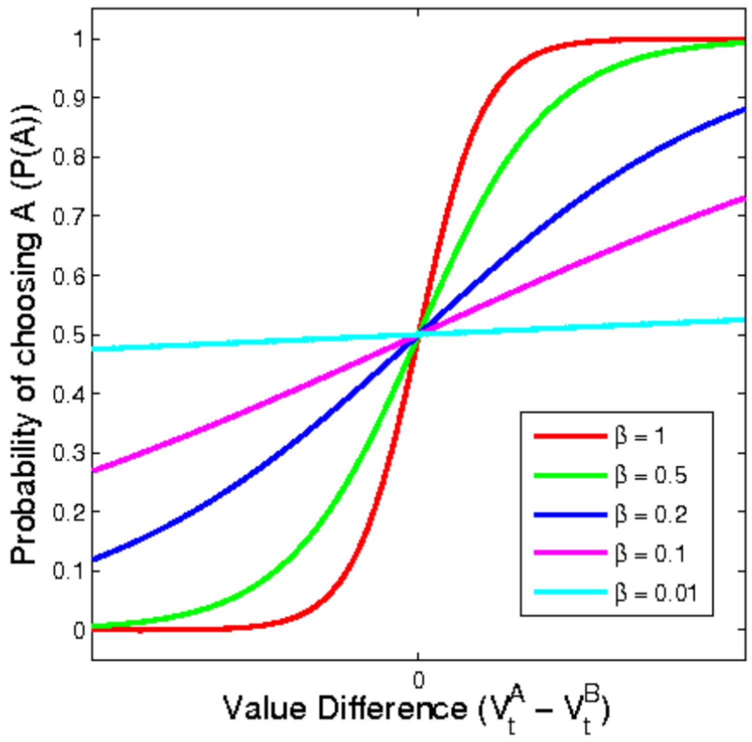
Choice rule: ε-greedy

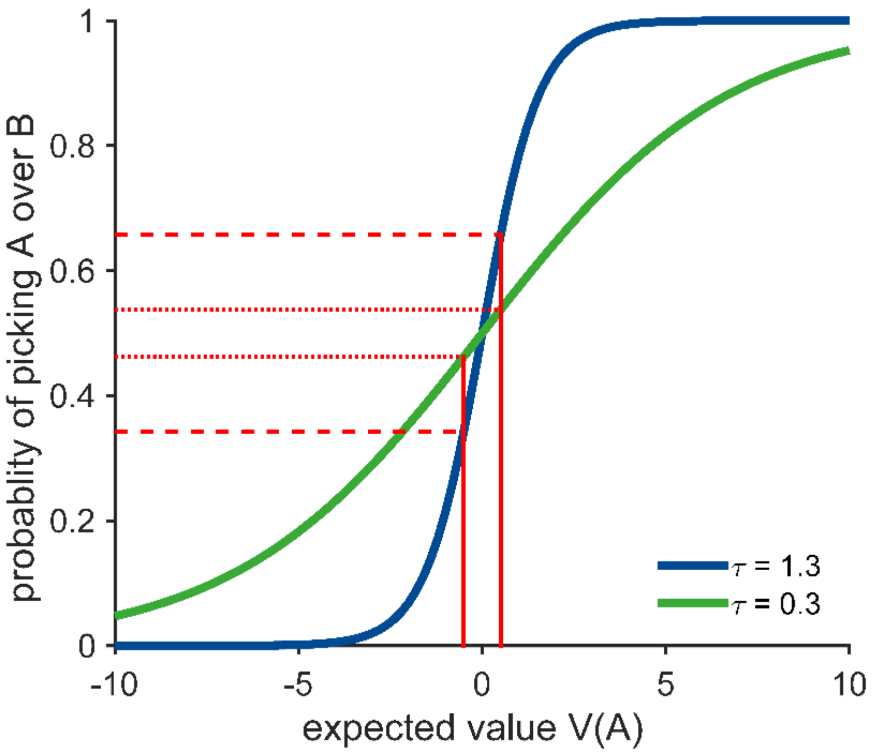
$$p(C=a)=\begin{cases}1-\varepsilon, & V(a)>V(b)\\ \varepsilon, & V(a)<V(b)\end{cases}$$



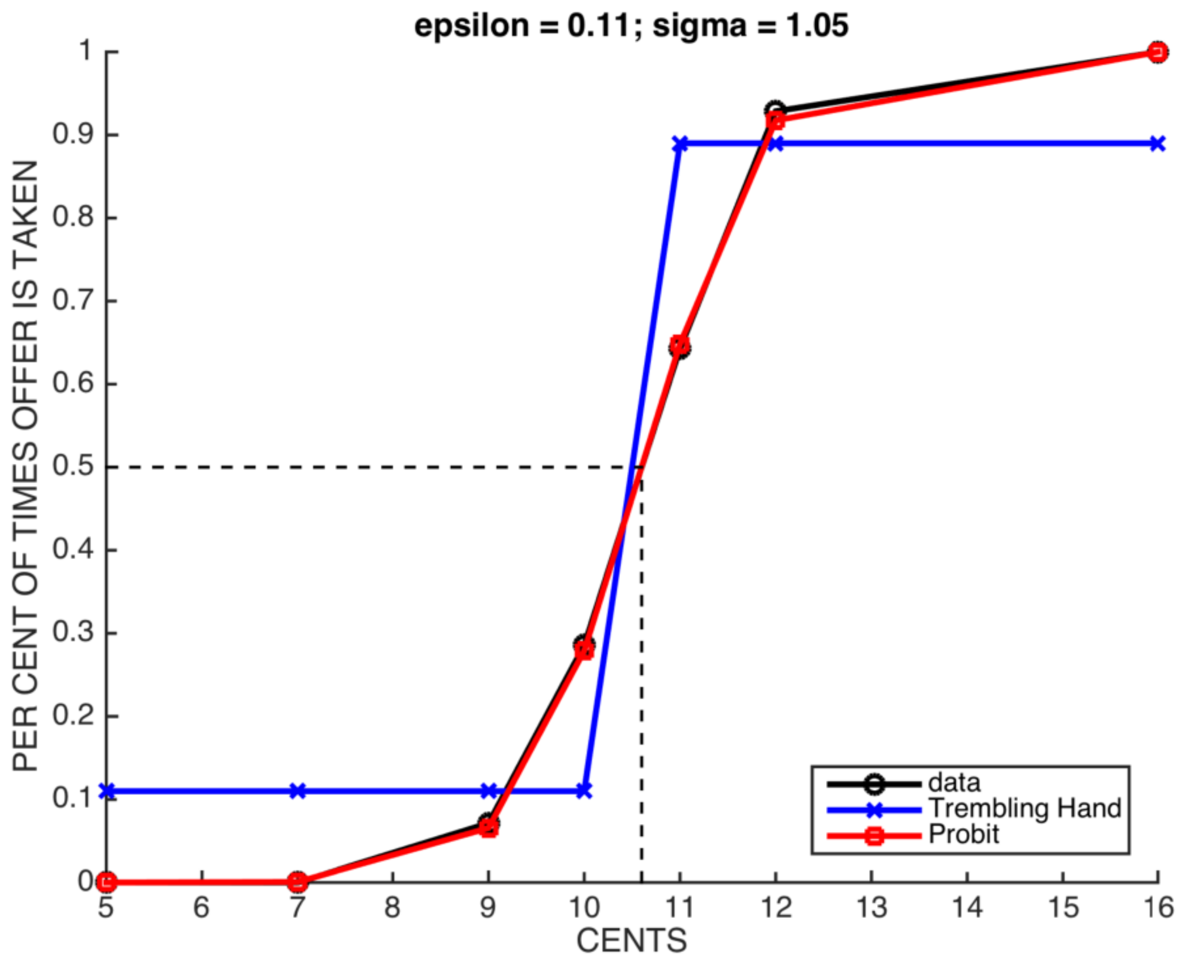
Choice rule: softmax

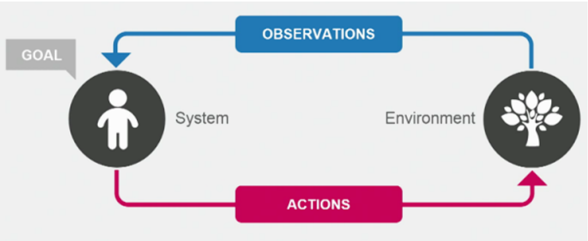
$$p(C=a)=\frac{e^{\tau*V(a)}}{e^{\tau*V(a)}+e^{\tau*V(b)}}=\frac{1}{1+e^{-\tau*(V(a)-V(b))}}$$



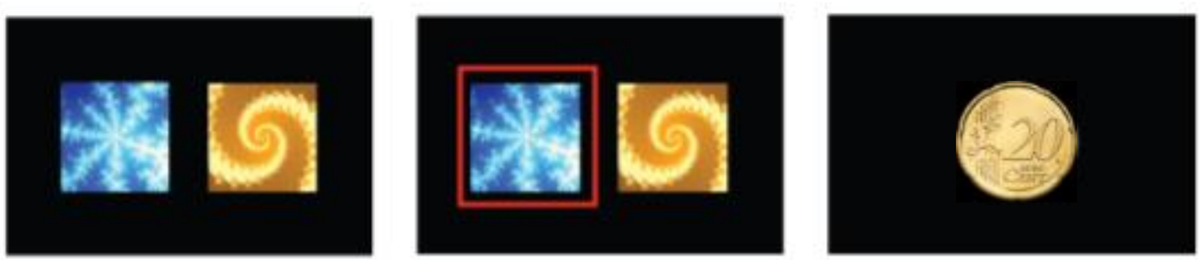


Choice rule: direct comparison





Rescorla-Wagner Value Update



Value update:

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

Prediction error:

$$PE = R_t - V_t$$

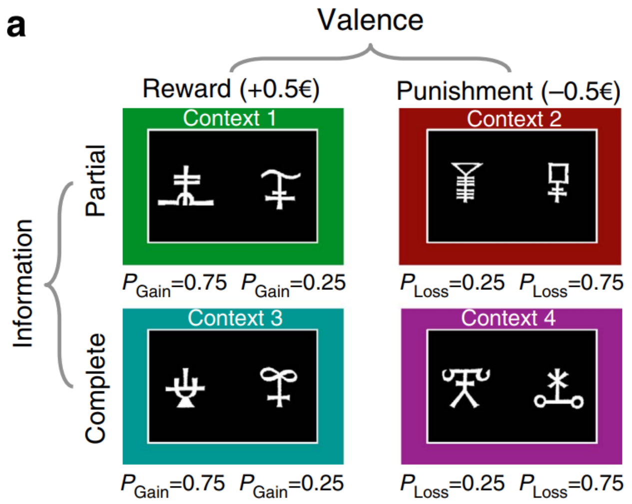
choice rule (sigmoid /softmax):

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

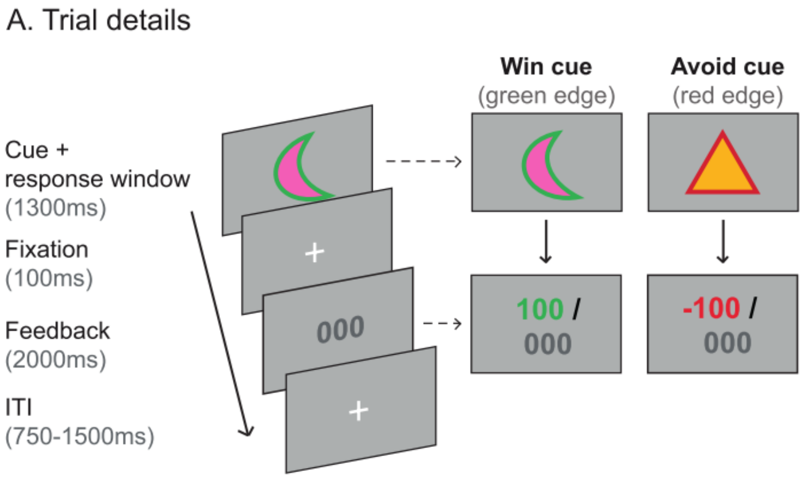
```
α  - learning rate
PE - reward prediction error
V  - value
R  - reward
τ  - softmax temperature
```

Rescorla & Wagner (1972)

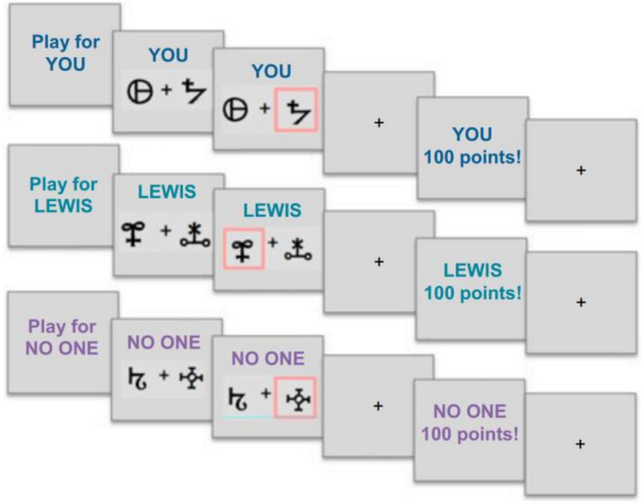
Generalizing RL framework



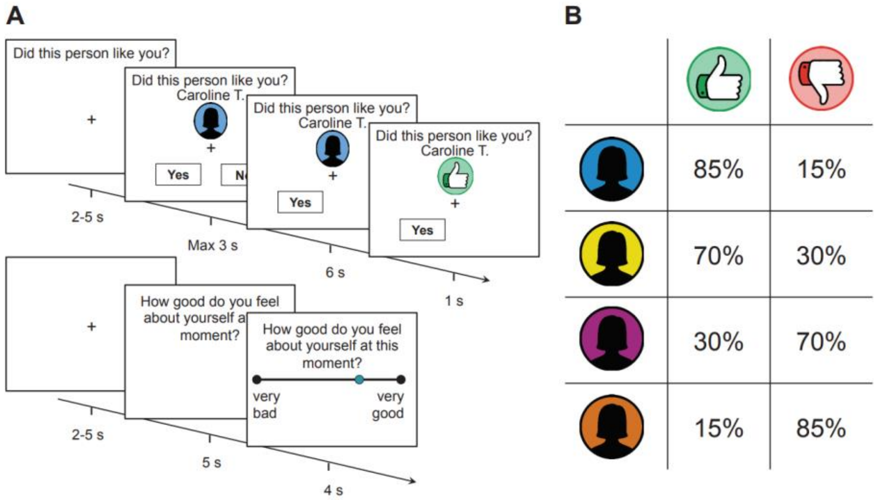
Palminteri et al. (2015)



Swart et al. (2017)



Lockwood et al. (2016)



Will et al. (2017)