



Bayesian Statistics and Bayesian Cognitive Modeling

Lei Zhang

Institute of Systems Neuroscience, University Medical Center Hamburg-Eppendorf

Recap

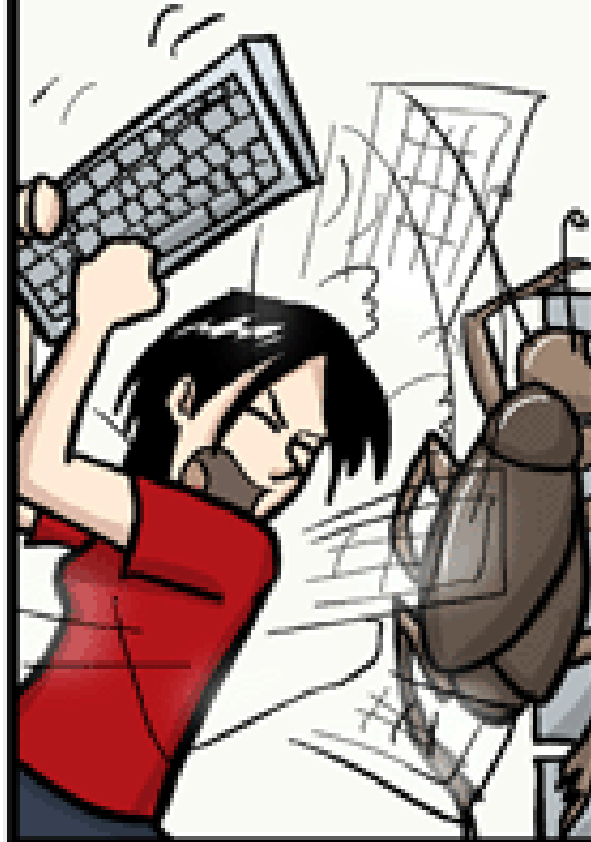
What we've learned...

- Linear Regression Model
- Predictive Check
- Cognitive Modeling
- Reinforcement Learning Model
- Hierarchical Modeling
- Optimizing Stan Codes
- Model Comparison

DAY3

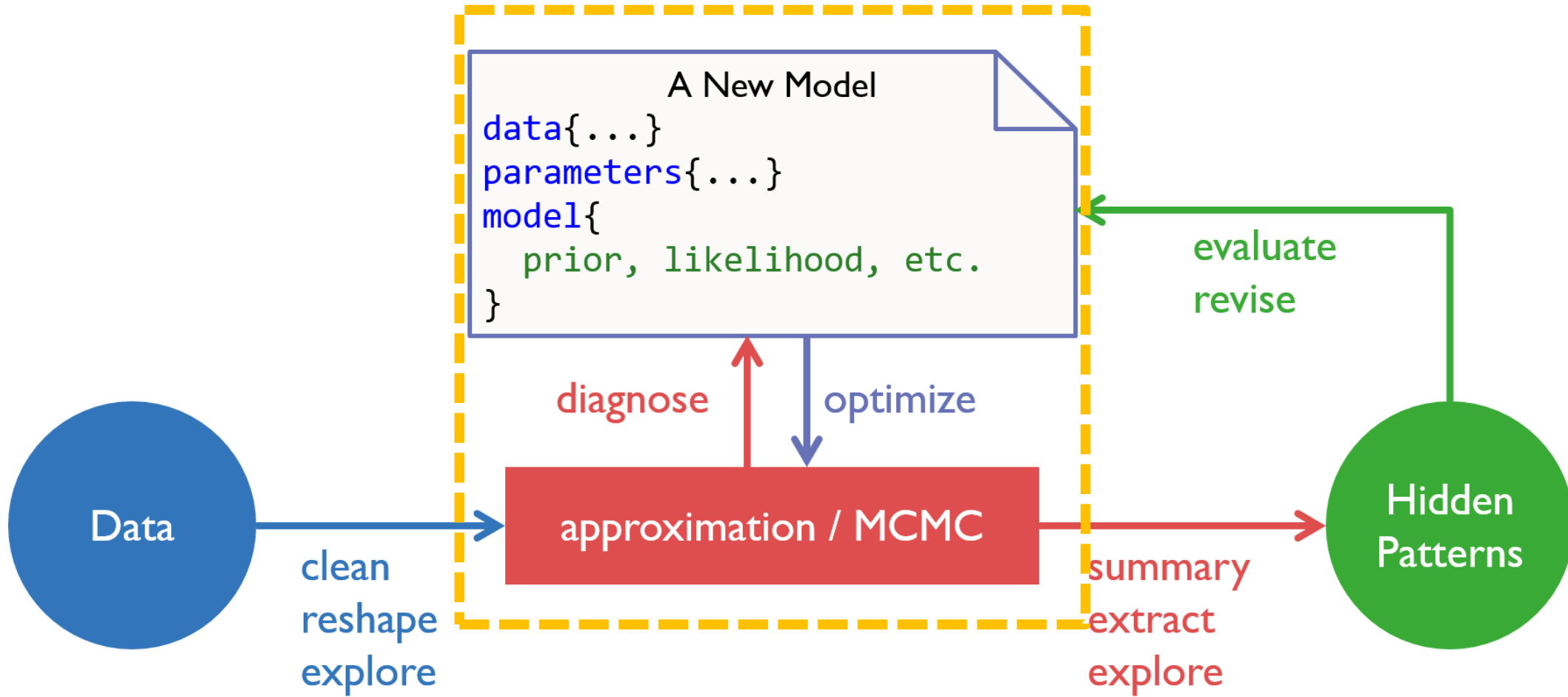
09:00 – 09:30	Implementing Model Comparison
09:30 – 10:15	Stan Style Tips and Debugging
10:15 – 11:00	Introduction to Model-based fMRI
11:00 – 11:15	Coffee break
11:15 – 12:15	Capstone Project: Delay Discounting
12:15 – 13:00	Summary, Misc., Q&A

STAN DEBUGGING



JORGE CHAM © 2005





Stan Style Tips

cognitive model

statistics

computing

Make it Reproducible

- Scripts are good documentation!
- Save your seed (not cross platform*)

Make it Readable

- Choose a consistent style
- Give meaningful variable names

Start with Simulated Data

Design Top-Down, Code Bottom-Up

Write Comments

- Code never lies!



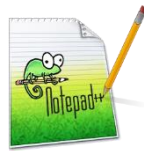
* Stan seed depends on hardware etc.

The Editor of your Choice

cognitive model

statistics

computing



```
data {  
  int<lower=0> w;  
  int<lower=0> N;  
}
```

```
parameters {  
  real<lower=0,upper=1> p;  
}
```

```
model {  
  p ~ uniform(0,1);  
  w ~ binomial(N, p);  
}
```

```
data {  
  int<lower=0> w;  
  int<lower=0> N;  
}
```

```
parameters {  
  real<lower=0,upper=1> p;  
}
```

```
model {  
  p ~ uniform(0,1);  
  w ~ binomial(N, p);  
}
```

```
data {  
  int<lower=0> w;  
  int<lower=0> N;  
}  
  
parameters {  
  real<lower=0,upper=1> p;  
}  
  
model {  
  p ~ uniform(0,1);  
  w ~ binomial(N, p);  
}
```

```
data {  
  int<lower=0> w;  
  int<lower=0> N;  
}  
  
parameters {  
  real<lower=0,upper=1> p;  
}  
  
model {  
  p ~ uniform(0,1);  
  w ~ binomial(N, p);  
}
```

* Click each logo to visit their homepage.

** [Comparison](#)

Common Error / Warning Types

cognitive model

statistics

computing

ERRORS

- forget “ ; ”
- mis-indexing: mismatch or constant support mismatch
- improper constrain
- improper dimension declaration
- vectorizing when not supported
- wrong data type
- wrong distribution names
- forget priors
- miss spelling

WARNINGS


- forget last blank line
- use earlier version of Stan
- numerical problems
- divergent transitions
- hit max_treedepth
- BFMI too low
- improper prior

Debugging in Stan

cognitive model

statistics

computing

- always use a *.stan file
- press  in RStudio
- use `lookup()`
- start with simulated data
- be careful with copy/paste
- run 1 chain, 1 sample
- debugging by printing

```
for (s in 1:1) {  
  vector[2] v;  
  real pe;  
  v <- initV;  
  
  for (t in 1:nTrials) {  
    choice[s,t] ~ categorical_logit( tau[s] * v );  
  
    print("s = ", s, ", t = ", t, ", v = ", v);  
  
    pe <- reward[s,t] - v[choice[s,t]];  
    v[choice[s,t]] <- v[choice[s,t]] + lr[s] * pe;  
  }  
}
```

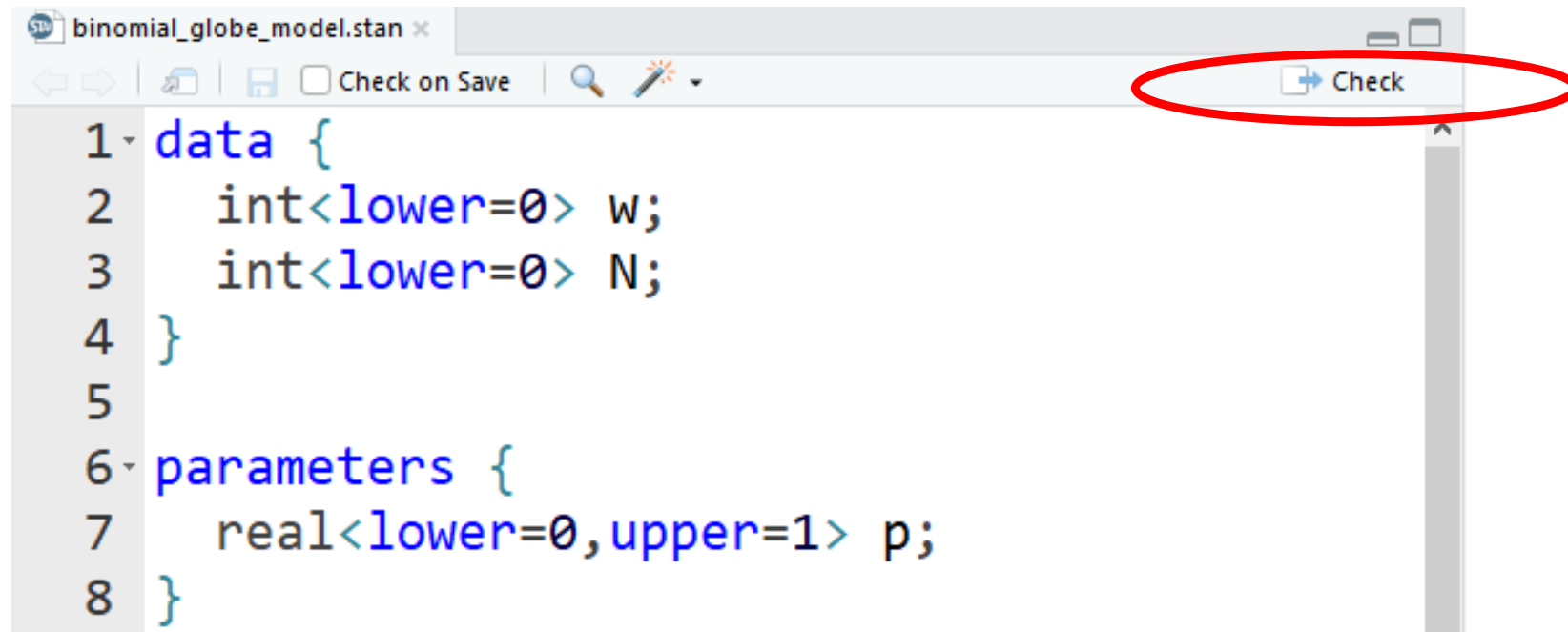
```
> lookup(dnorm)  
StanFunction Arguments ReturnType Page SamplingStatement  
344      normal      (reals mu, reals sigma)      real   369          TRUE  
348  normal_log (reals y, reals mu, reals sigma)      real   369          FALSE
```

Debugging Stan in RStudio

cognitive model

statistics

computing

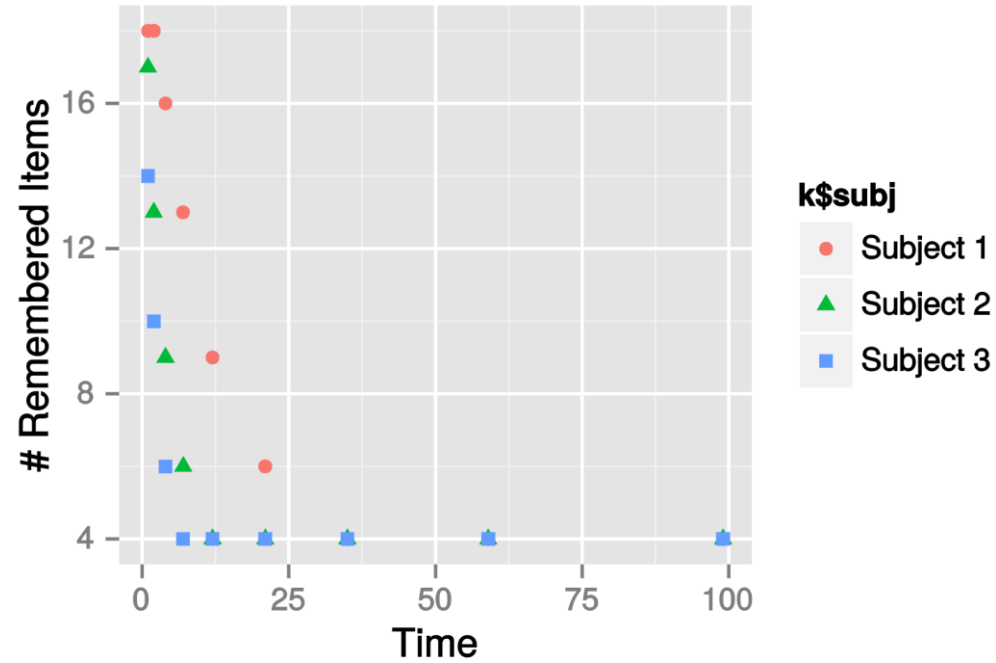


```
binomial_globe_model.stan x
1 data {
2   int<lower=0> w;
3   int<lower=0> N;
4 }
5
6 parameters {
7   real<lower=0,upper=1> p;
8 }
```

```
rstan::rstudio_stanc("_scripts/binomial_globe_model.stan")
```



Example: Memory Retention



Subject	Time Interval								
	1	2	4	7	12	21	35	59	99
1	18	18	16	13	9	6	4	4	4
2	17	13	9	6	4	4	4	4	4
3	14	10	6	4	4	4	4	4	4

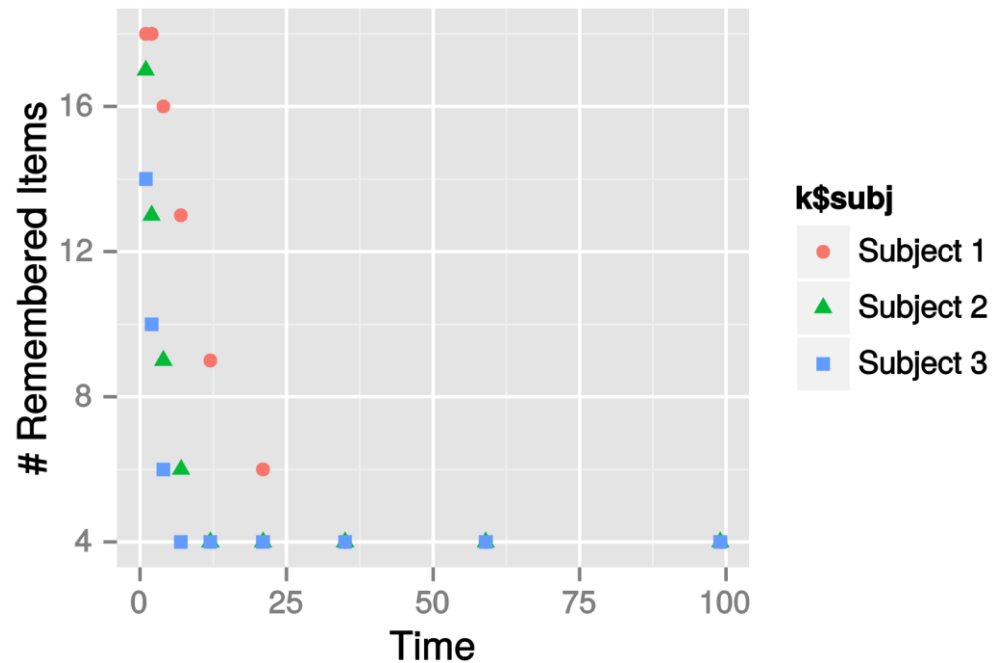


Simple Exponential Decay Model

cognitive model

statistics

computing



$$\theta_t = \min(1.0, \exp(-\alpha t) + \beta)$$

$p(\text{remember})$

decay rate

baseline

Exercise I

cognitive model

statistics

computing

.../BayesCog/09.debugging/_scripts/exp_decay_main.R

TASK: Debugging the Memory retention model

≥ 9 errors!

Viel Spaß!

```
> dataList
$`k`
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9]
[1,]   18   18   16   13    9    6    4    4    4
[2,]   17   13    9    6    4    4    4    4    4
[3,]   14   10    6    4    4    4    4    4    4

$nItem
[1] 18

$intervals
[1] 1 2 4 7 12 21 35 59 99

$ns
[1] 3

$nt
[1] 9
```

Satisfied with the results?

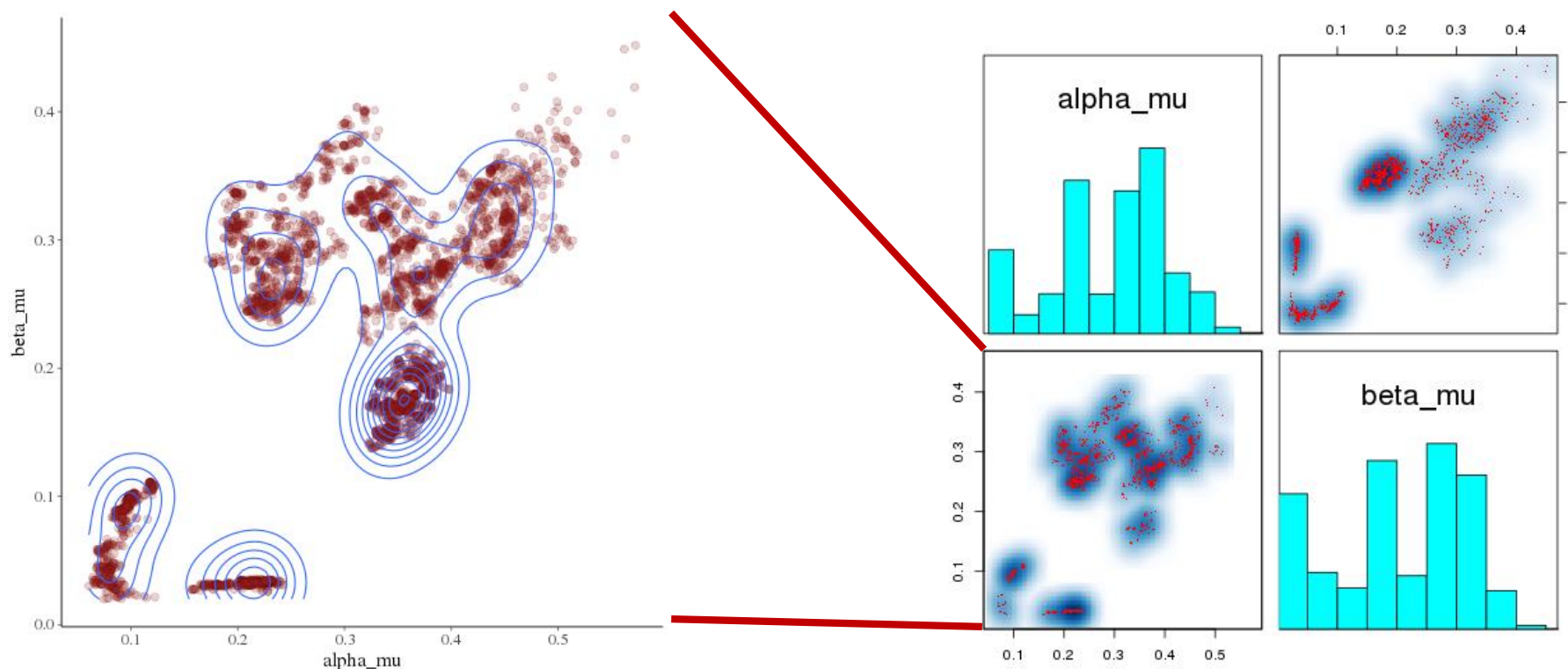
cognitive model

statistics

computing

Warning messages:

1: There were 3998 divergent transitions after warmup. Increasing adapt_delta above 0.8 may help. See <http://mc-stan.org/misc/warnings.html#divergent-transitions-after-warmup>
2: Examine the pairs() plot to diagnose sampling problems



Why Stan Fails?

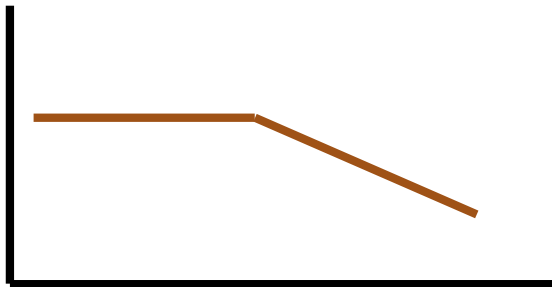
cognitive model

statistics

computing

```
for (s in 1:ns) {  
  for (t in 1:nt) {  
    theta[s,t] = fmin(1.0, exp(-alpha[s] * intervals[t]) + beta[s]);  
    k[s,t] ~ binomial(nItem, theta[s,t]);  
  }  
}
```

Non-differentiable link (likelihood) functions are bad news, particularly in Stan, which relies on derivatives.



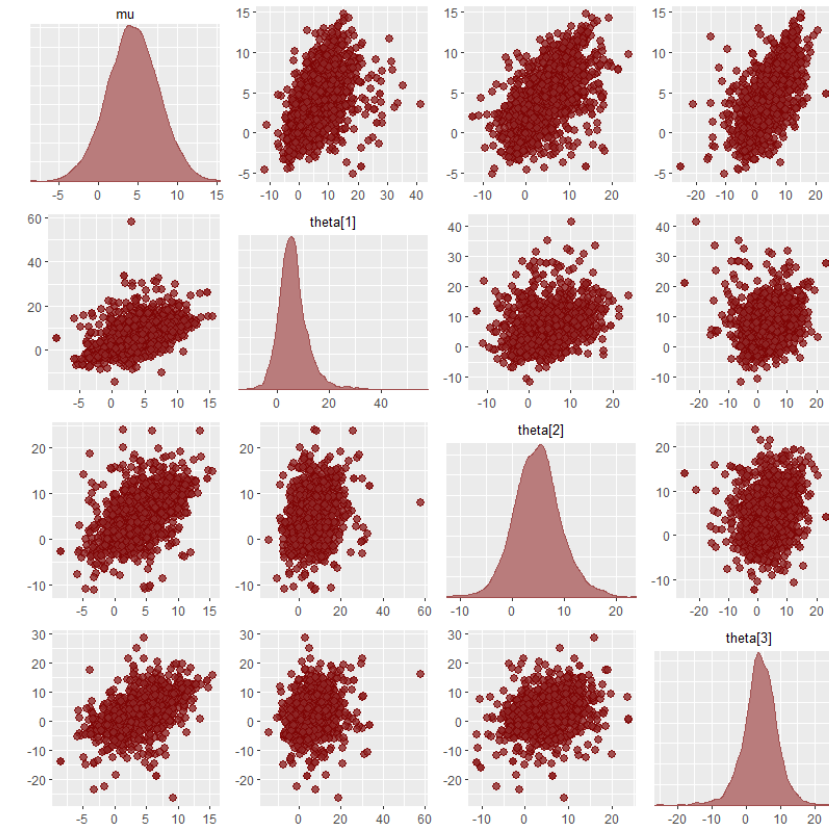
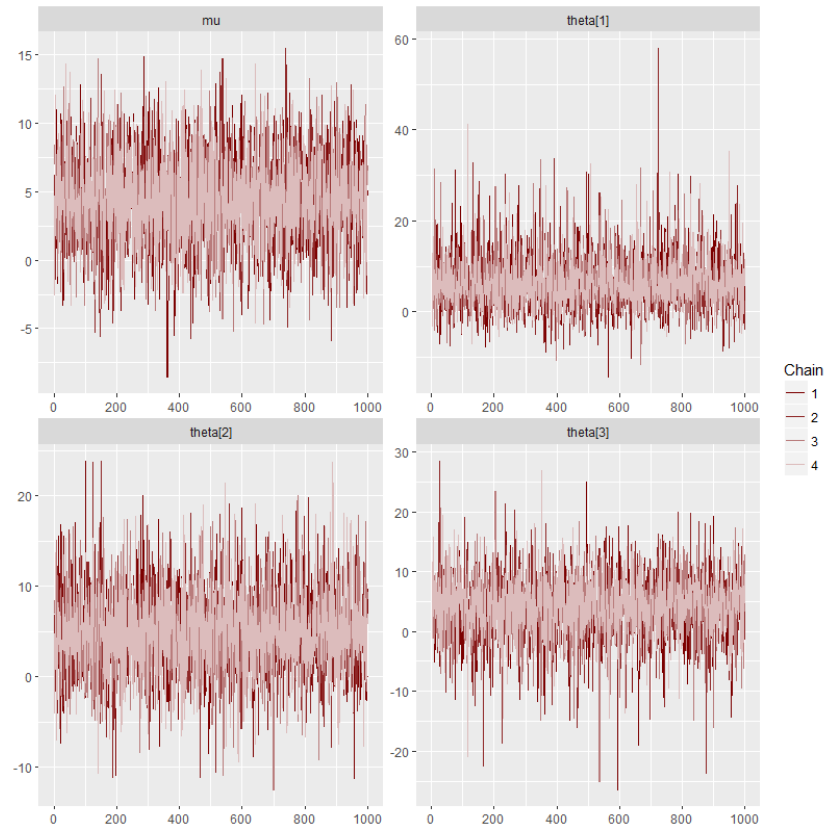
What to look for?

cognitive model

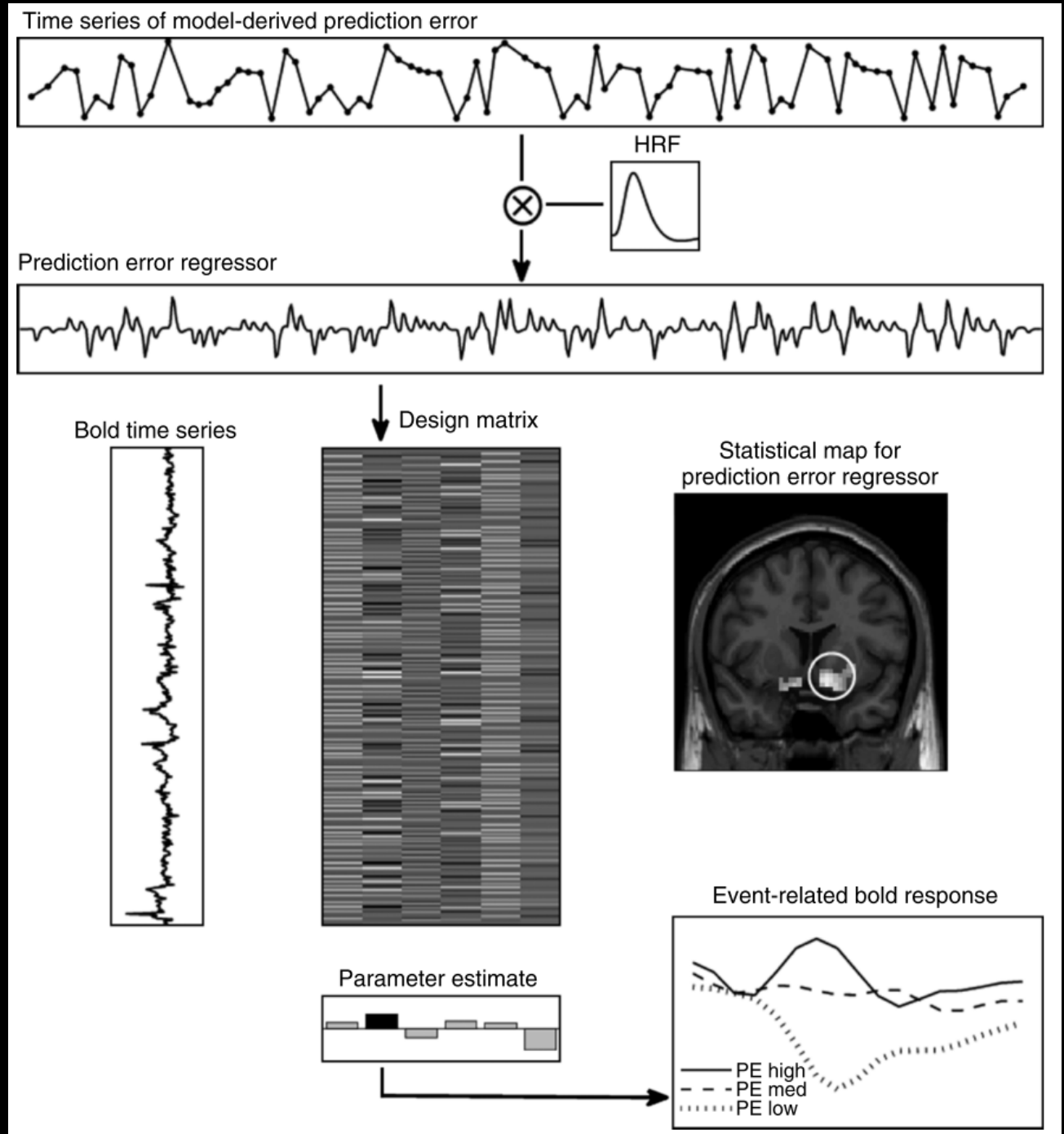
statistics

computing

```
> source('stan_utility.R')
> check_all_diagnostics(fit)
[1] "n_eff / iter looks reasonable for all parameters"
[1] "Rhat looks reasonable for all parameters"
[1] "0 of 4000 iterations ended with a divergence (0%)"
[1] "0 of 4000 iterations saturated the maximum tree depth of 10 (0%)"
[1] "E-BFMI indicated no pathological behavior"
```



INTRODUCTION TO MODEL-BASED FMRI

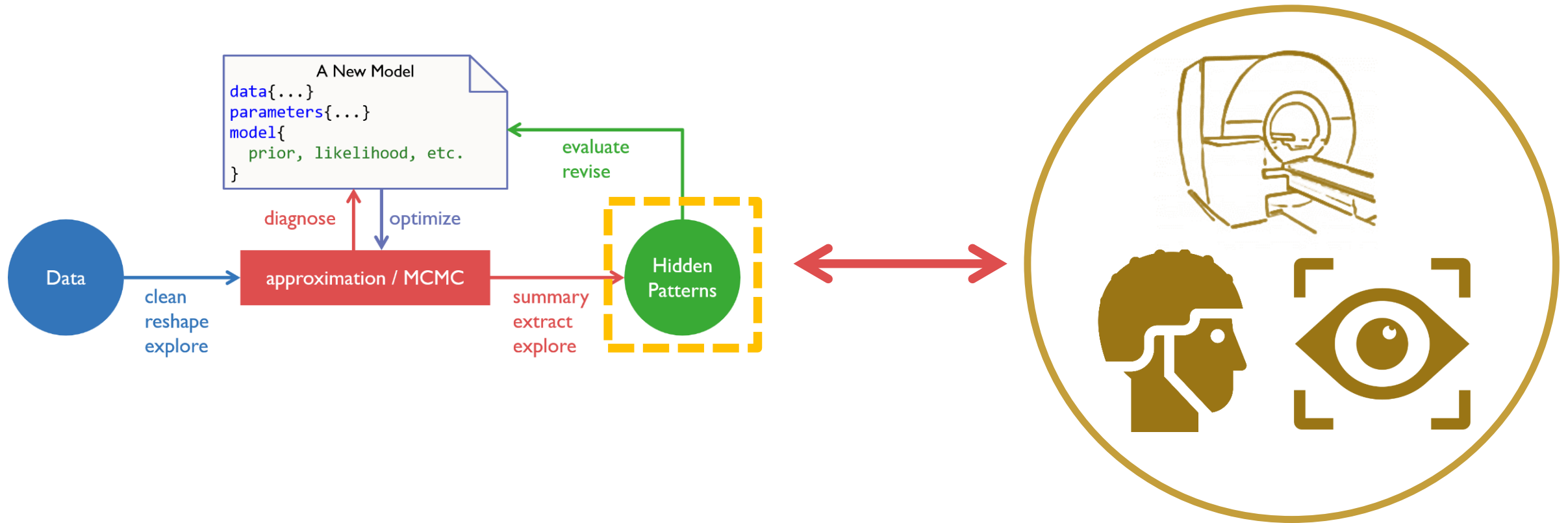


Model-based Analysis

cognitive model

statistics

computing



Explore Hidden Patterns in Stan

cognitive model

statistics

computing

```
generated quantities {  
  ...  
  
  vector[2] v[nSubjects, nTrials+1];  
  real vc[nSubjects, nTrials]; //chosen value  
  real pe[nSubjects, nTrials];  
  
  {  
    for (s in 1:nSubjects) {  
      log_lik[s] <- 0;  
      v[s,1] <- initV;  
  
      for (t in 1:nTrials) {  
        log_lik[s] = log_lik[s] + categorical_logit_log(choice[s,t], tau[s] * v[s,t] );  
  
        vc[s,t] = v[s,t,choice[s,t]];  
  
        pe[s,t] = reward[s,t] - v[s,t,choice[s,t]];  
  
        v[s,t+1] = v[s,t];  
        v[s,t+1,choice[s,t]] = v[s,t,choice[s,t]] + lr[s] * pe[s,t];  
      }  
    }  
  }  
}
```

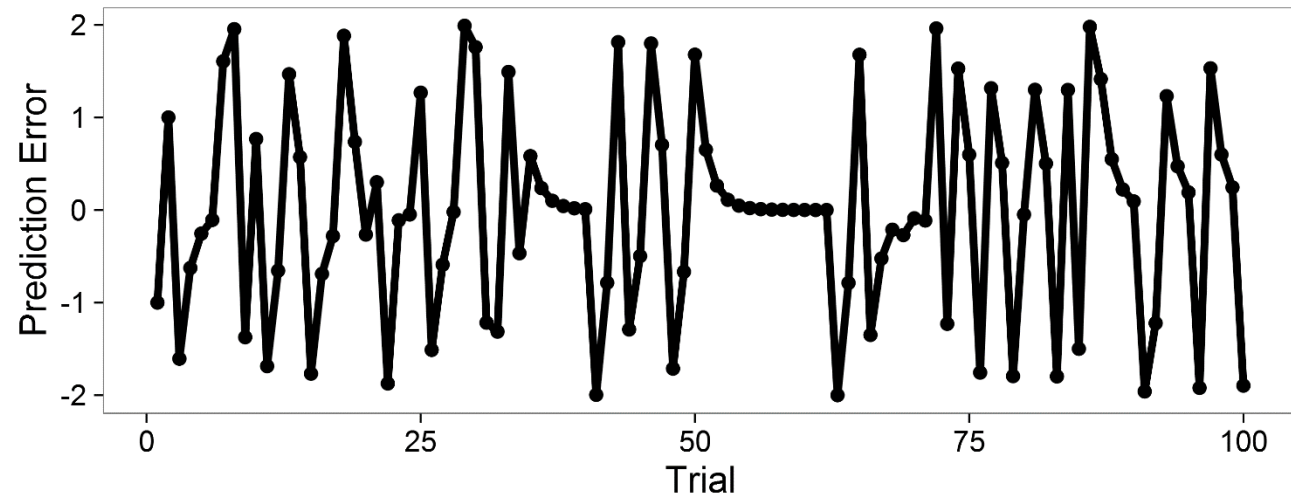
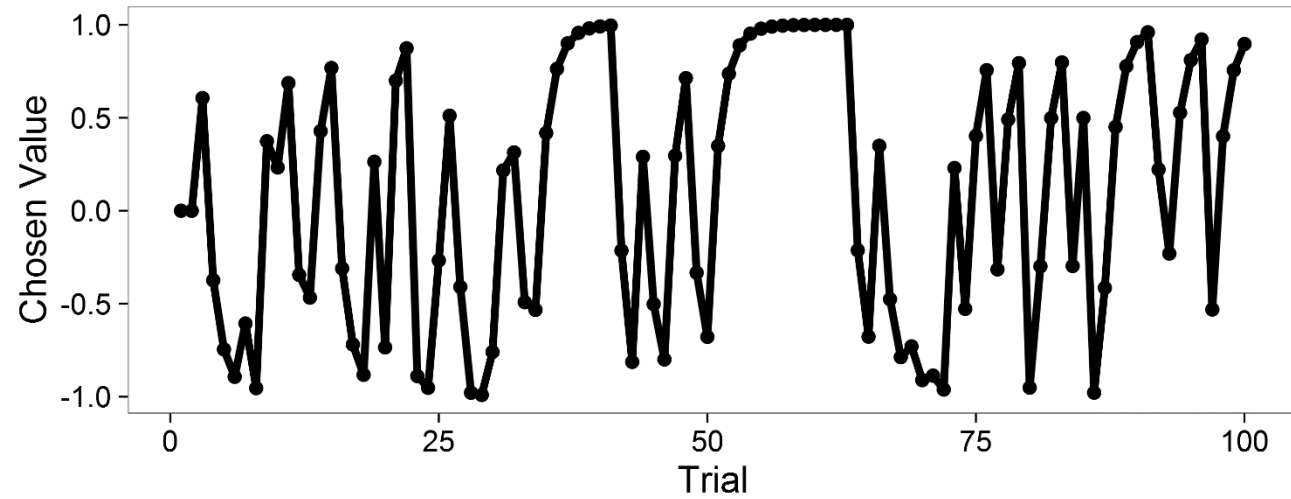
Obtain Decision Variables

subject01

cognitive model

statistics

computing

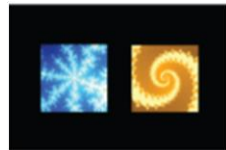


Perform Model-based fMRI

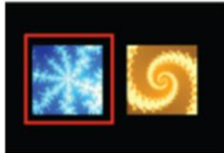
cognitive model

statistics

computing



choice
presentation



action
selection

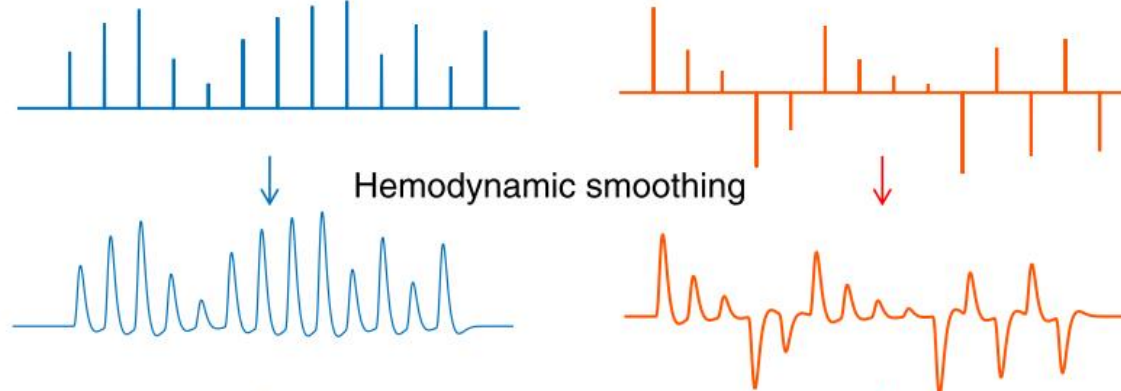


outcome

Computational model

$$V_{t+1} = V_t + \alpha \delta_t$$

Time series of variables



Hemodynamic smoothing

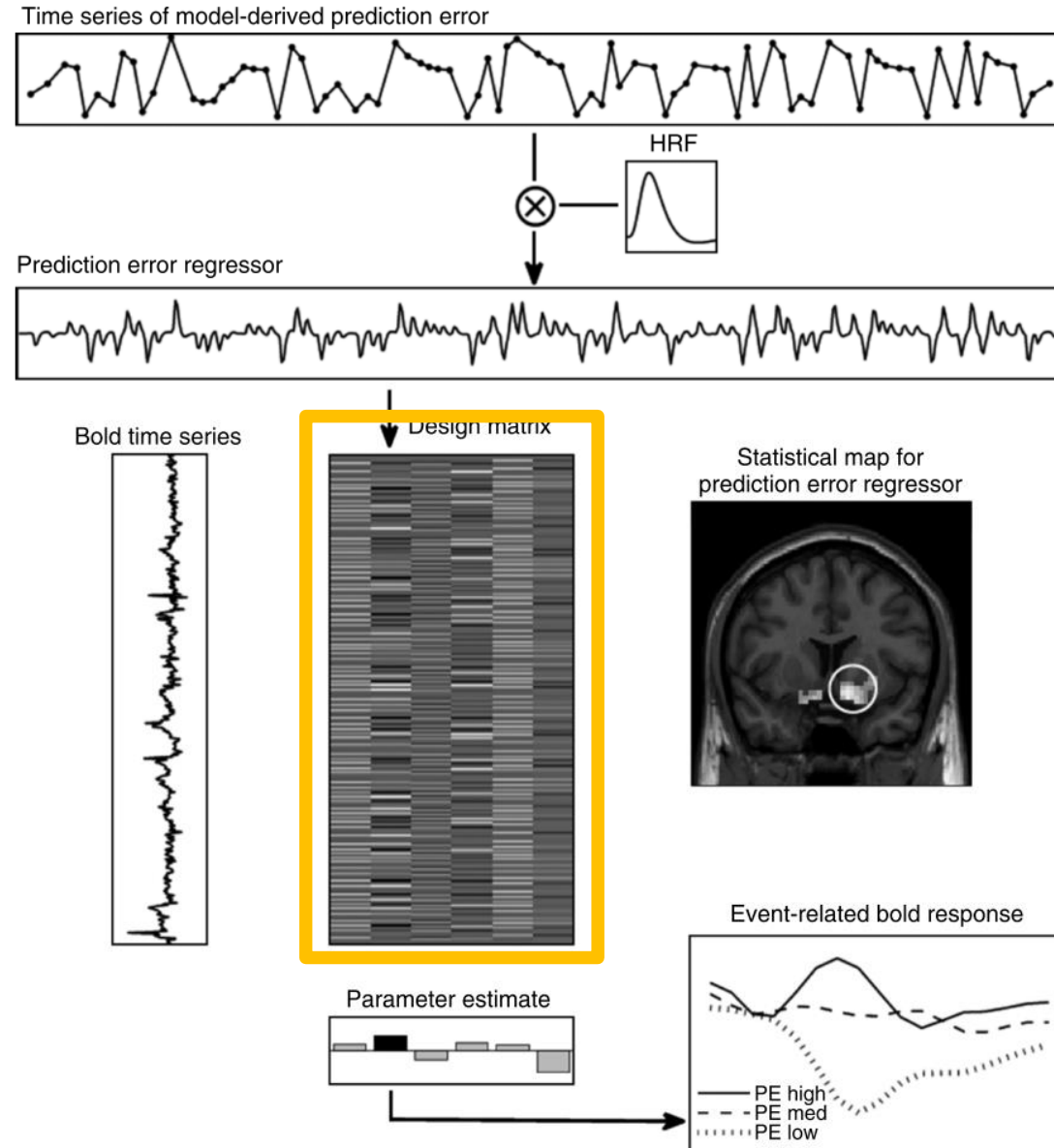
Correlated regions

Perform Model-based fMRI (cont.)

cognitive model

statistics

computing

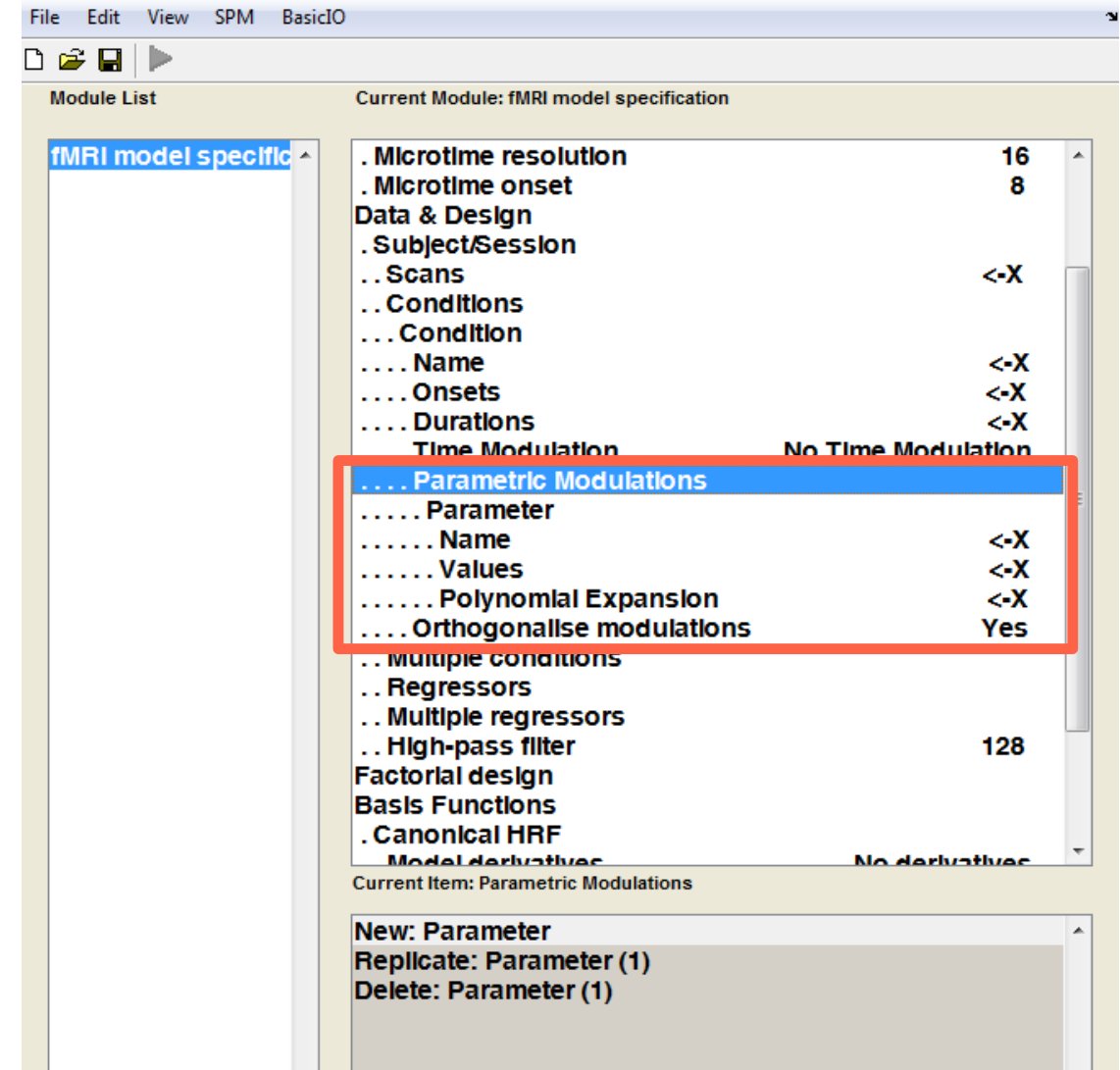
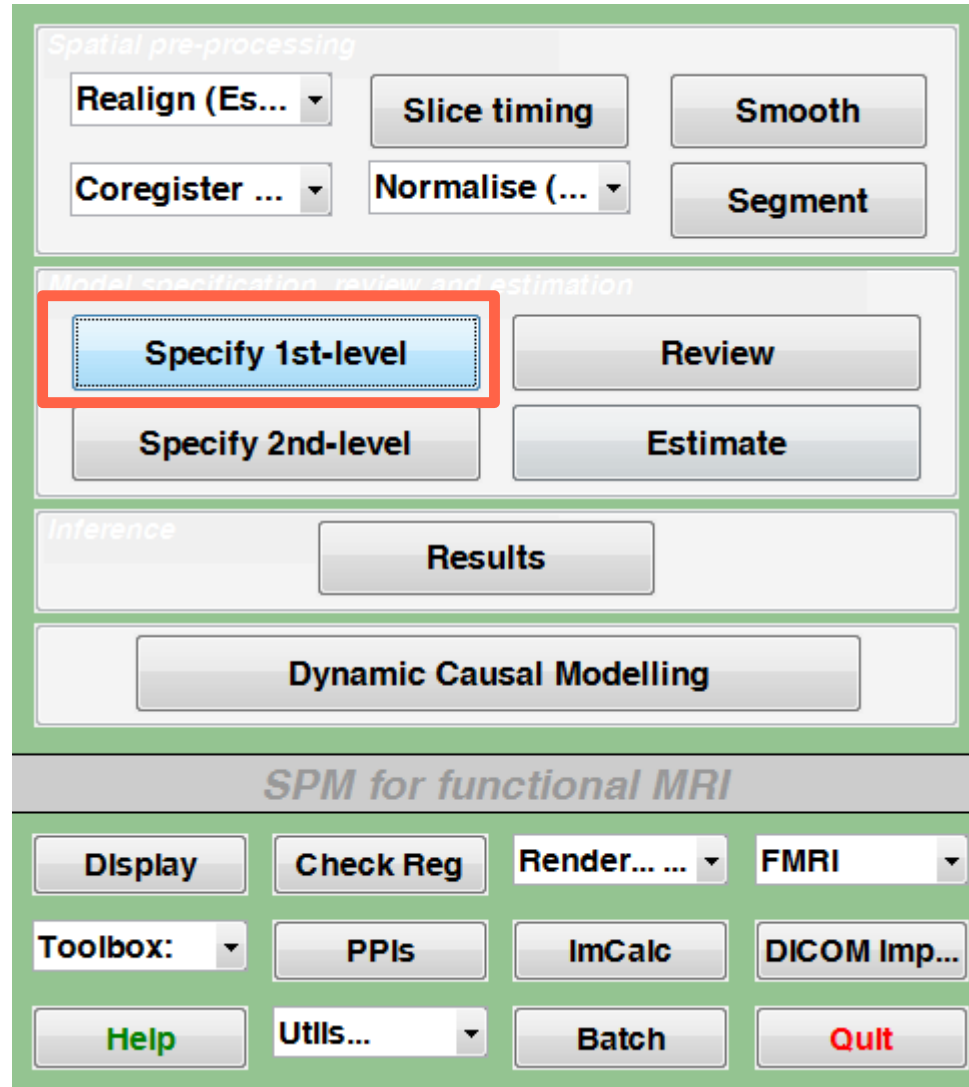


Implementing in SPM12

cognitive model

statistics

computing



SPM12 – batch scripting

cognitive model

statistics

computing

```
matlabbatch{1}.spm.stats.fmri_spec.ssess.cond(cnt).name = 'onsetPE';  
matlabbatch{1}.spm.stats.fmri_spec.ssess.cond(cnt).onset = onscat.sub(i_sub).cueoutcome;  
matlabbatch{1}.spm.stats.fmri_spec.ssess.cond(cnt).duration = 0;  
matlabbatch{1}.spm.stats.fmri_spec.ssess.cond(cnt).tmod = 0;  
matlabbatch{1}.spm.stats.fmri_spec.ssess.cond(cnt).pmod.name = 'PE';  
matlabbatch{1}.spm.stats.fmri_spec.ssess.cond(cnt).pmod.param = pe(i_sub);  
matlabbatch{1}.spm.stats.fmri_spec.ssess.cond(cnt).pmod.poly = 1;  
matlabbatch{1}.spm.stats.fmri_spec.ssess.cond(cnt).orth = 0;
```

make sure: `length(onset) == length(PE)`

A closer look at PE

cognitive model

statistics

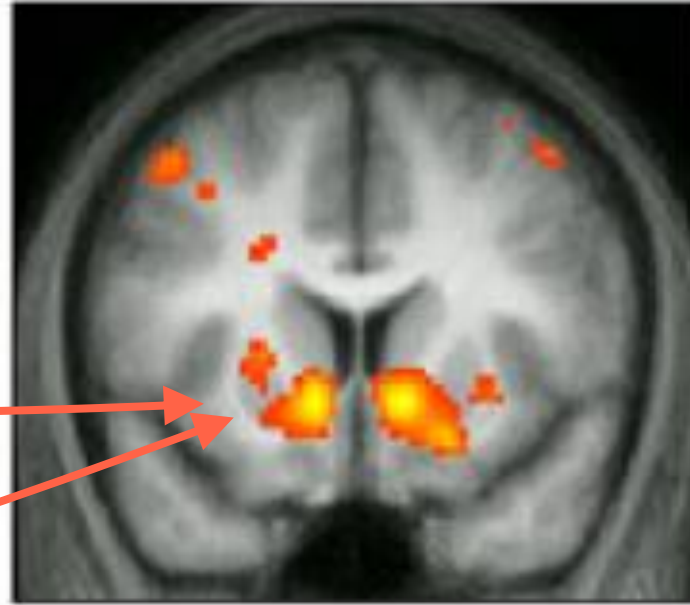
computing

Prediction error:

$$PE = R_t - V_t$$



outcome

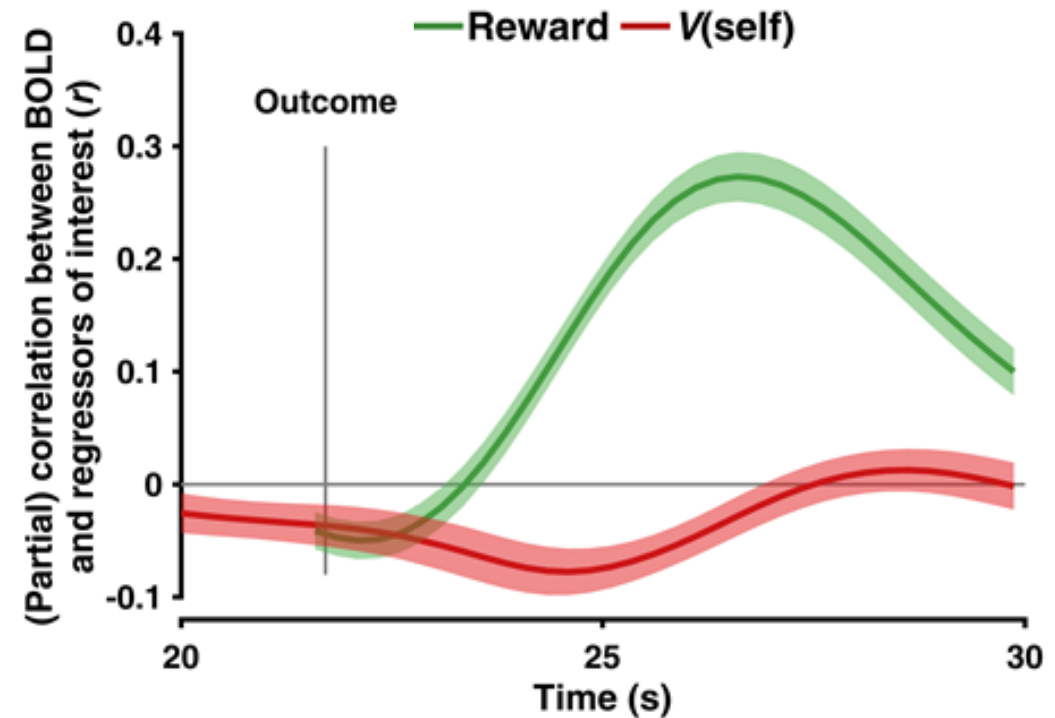
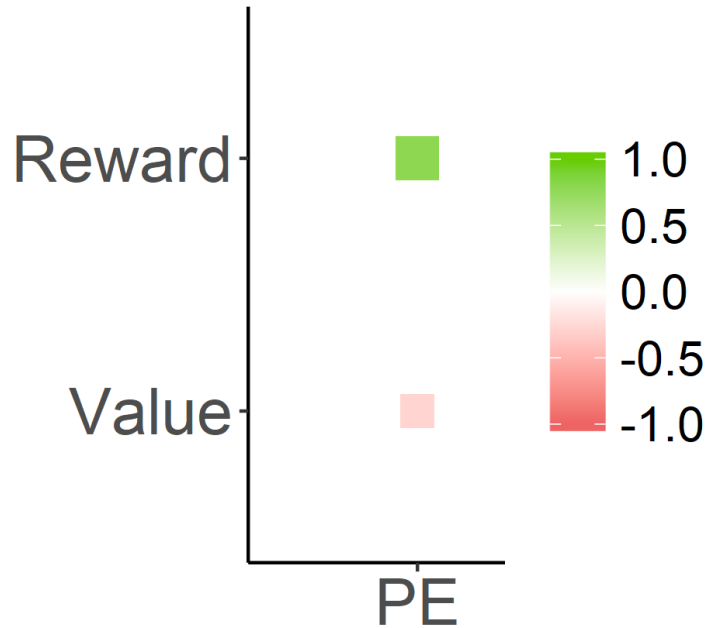
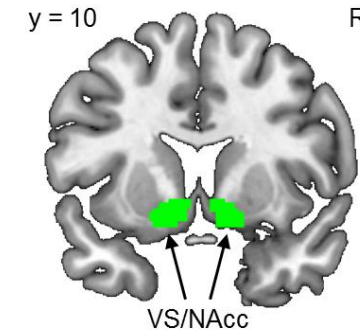


Q: how to justify the striatal activity is indeed associated with PE, rather than reward?

A closer look at PE

Prediction error:

$$PE = R_t - V_t$$



Exercise II

cognitive model

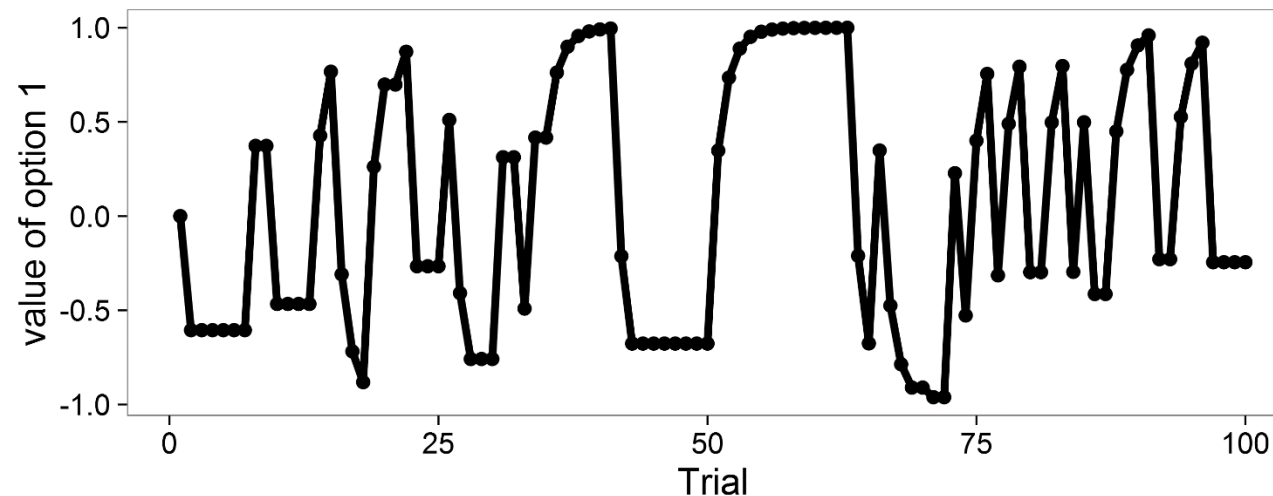
statistics

computing

```
.../BayesCog/10.model_based/_scripts/reinforcement_learning_model_based_main.R
```

TASK: extract and plot $V(\text{option}=1)$, for subject01 (from L65/L105)

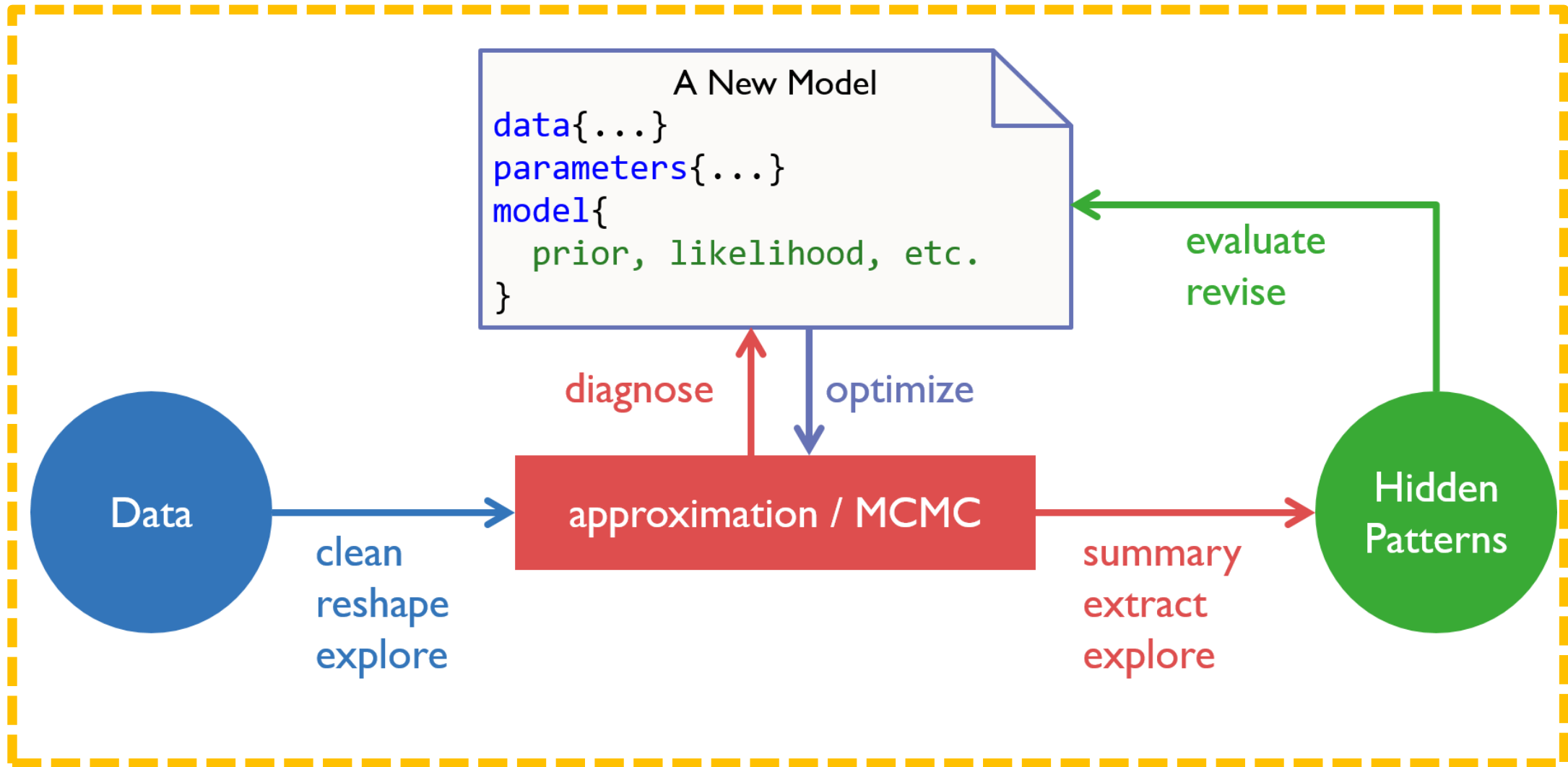
TIP: `fit_r1 <- readRDS('_outputs/fit_r1.RData')`



*Value of the first option, instead of the chosen value

DELAY DISCOUNTING



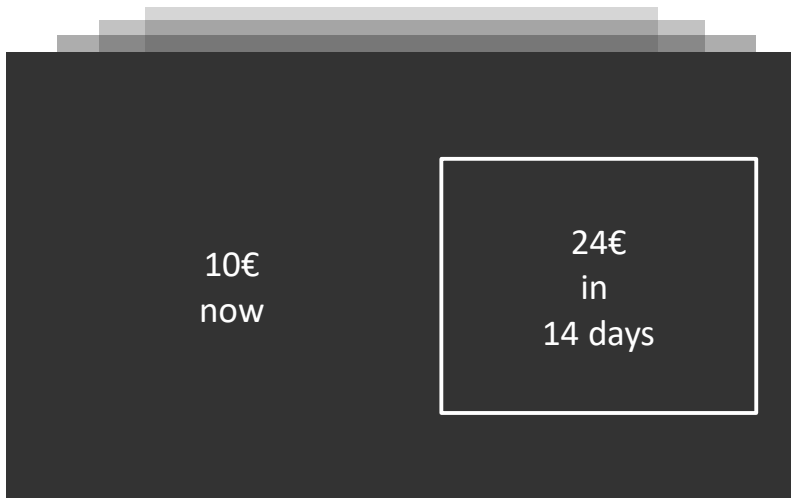


Delay Discounting Task and Models

cognitive model

statistics

computing



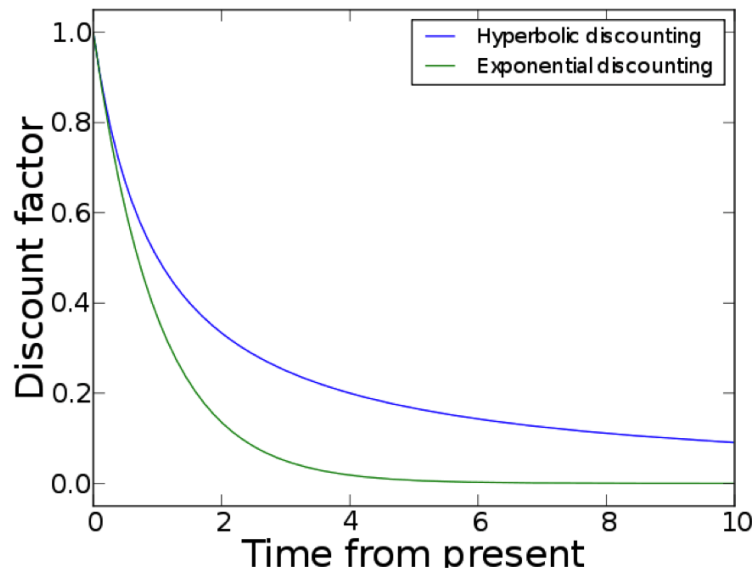
Hyperbolic Discounting Model

$$SV = \frac{A}{1 + k * delay}$$

Exponential Discounting Model

$$SV = A * \exp(-r * delay)$$

$$p(LL) = \frac{1}{1 + \exp^{temp(v(SS) - v(LL))}}$$



LL - late large option

SS - soon small option

Exercise III

cognitive model

statistics

computing

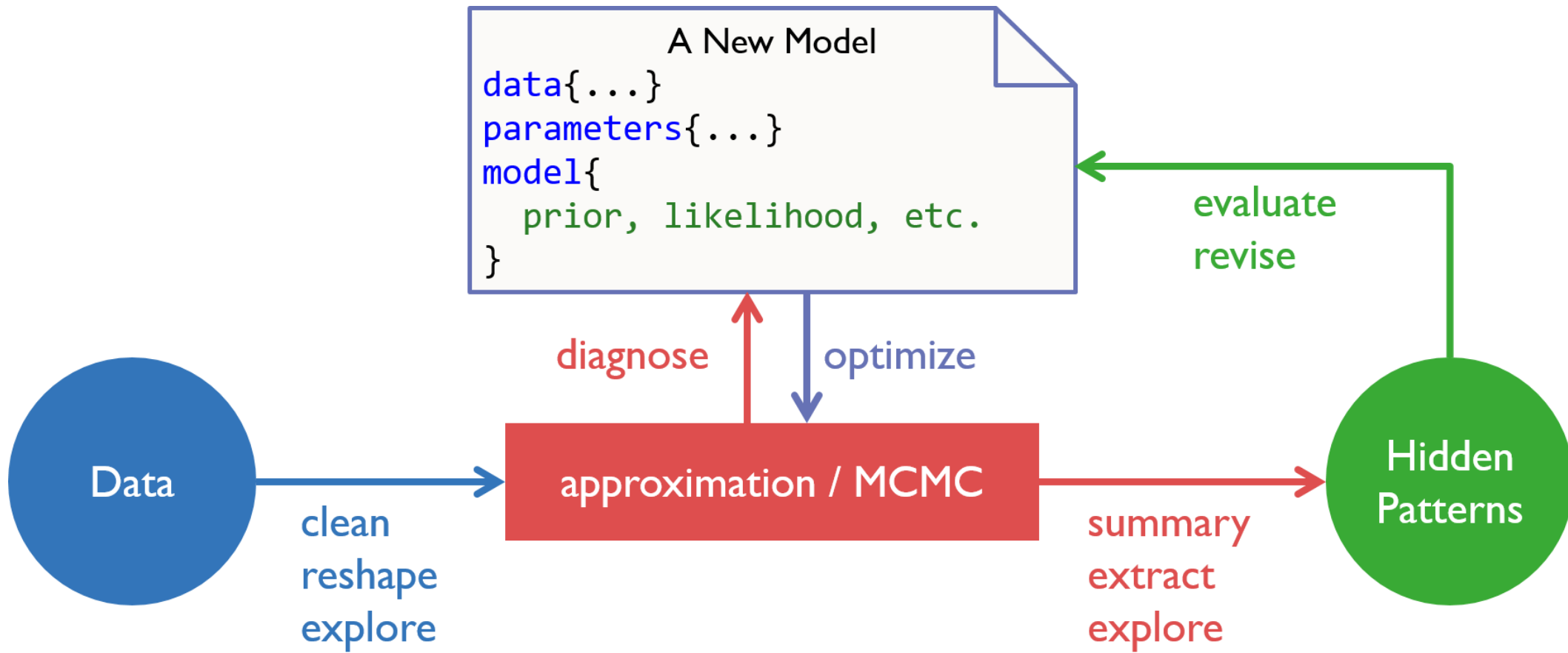
```
.../BayesCog/11.delay_discounting/_scripts/delay_discounting_main.R
```

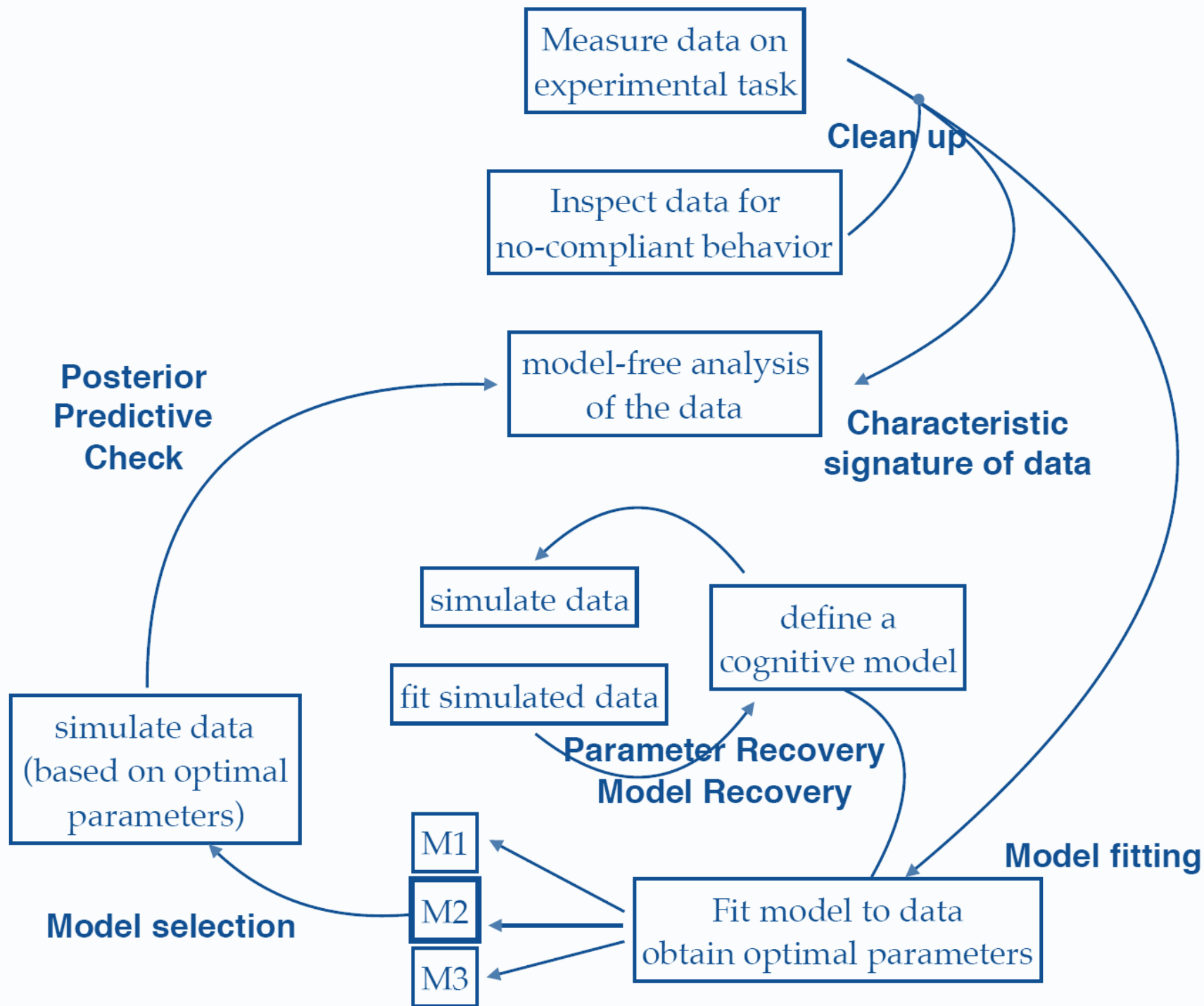
TASK:

- (1) understand how to deal with missing trials
- (2) complete and fit both models
- (3) complete the main script for comparing the two models

Summary

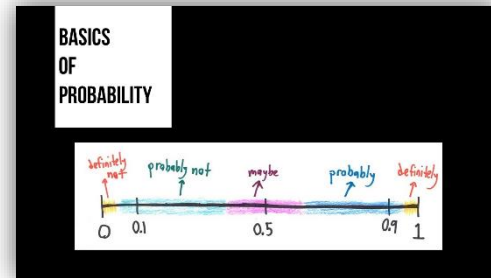






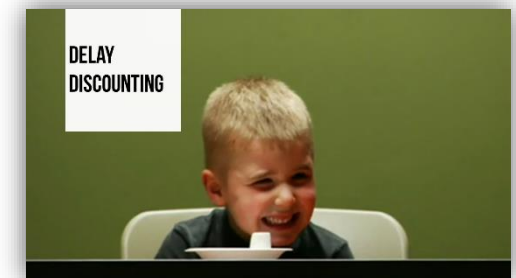
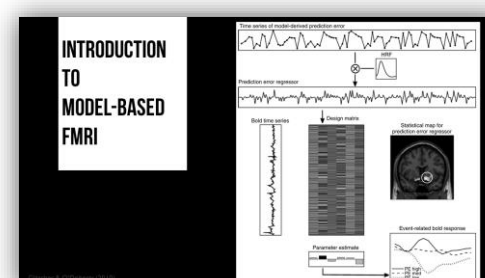
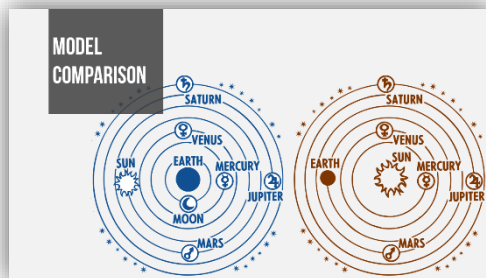
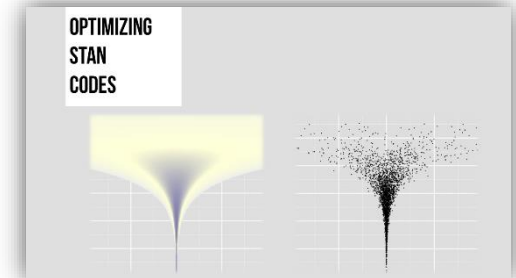
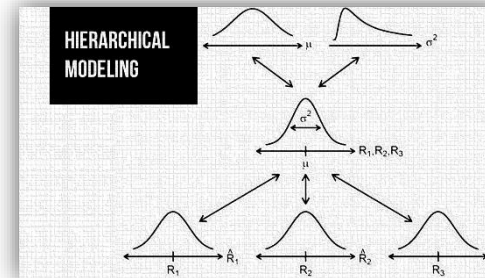
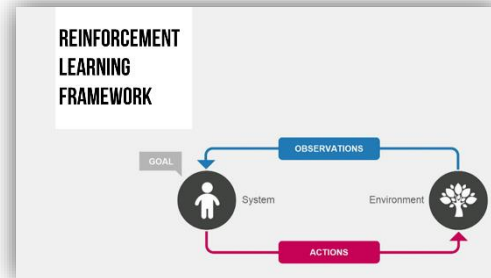
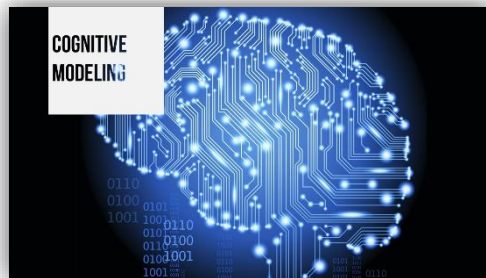
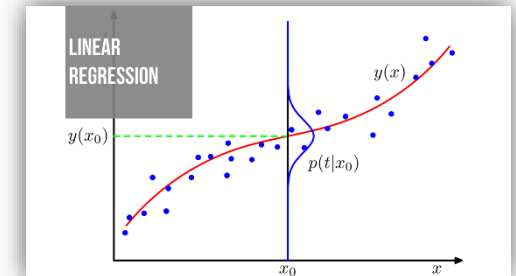
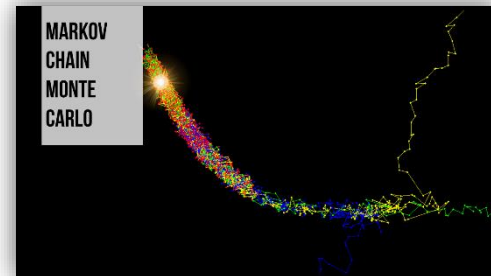
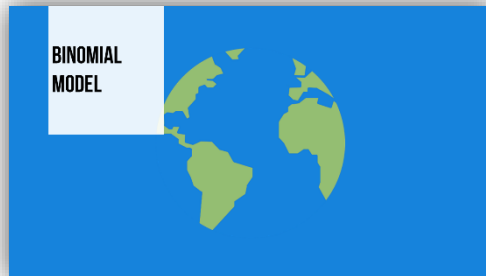
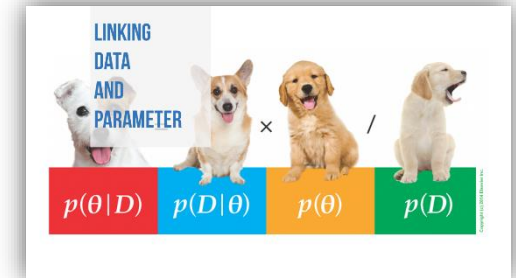
Adapted from Jan Gläscher's workshop

Summary of Topics



BAYES' THEOREM

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



Summary of Examples/Exercises

FOLDER	TASK	MODEL
01.R_basics	NA	NA
02.binomial_globe	Globe toss	Binomial Model
03.bernoulli_coin	Coin flip	Bernoulli Model
04.regression_height	Observed weight and height	Linear regression model
05.regression_height_poly		
06.reinforcement_learning	2-armed bandit task	Simple reinforcement learning (RL) model
07.optm_rl		
08.compare_models	Probabilistic reversal learning task	Simple and fictitious RL models
09.debugging	Memory Retention	Exponential decay model
10.model_based	2-armed bandit task	Simple RL model
11.delay_discounting	Delay discounting task	Hyperbolic and exponential discounting model

After the Workshop, you...

cognitive model

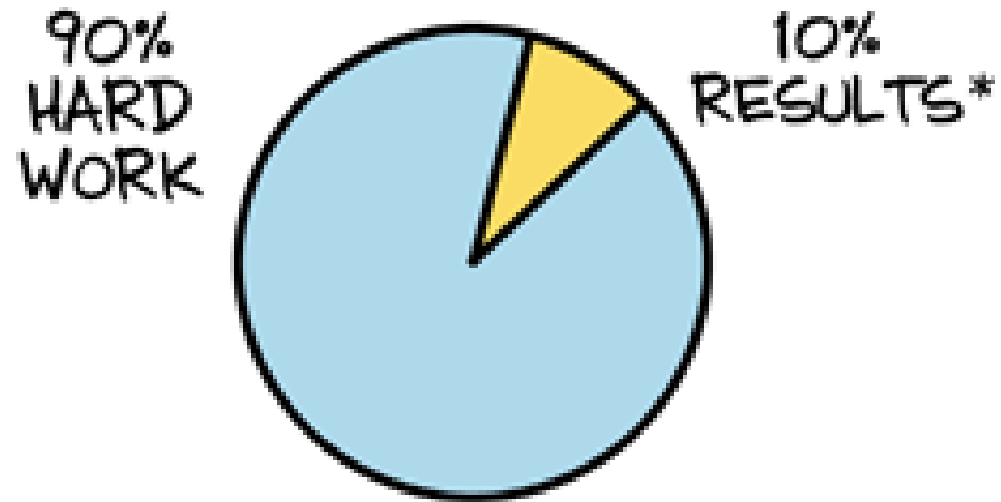
statistics

computing

- ...are able to implement your own model
- ...feel comfortable with reading mathematical equations
- ...consider the implementation of the “computational modeling” section
- ...gain insightful understanding of Bayesian stats and modeling
- ...take it as a good start and work on it later

Remember: practice makes perfect!

DOING RESEARCH:



* BEST CASE SCENARIO

WRITING ABOUT RESEARCH:



Write Your Own Tutorial Paper!

cognitive model

statistics

computing



RESEARCH

Revealing Neurocomputational Mechanisms of Reinforcement Learning and Decision-Making With the hBayesDM Package

Woo-Young Ahn¹, Nathaniel Haines¹, and Lei Zhang²

¹Department of Psychology, The Ohio State University, Columbus, OH

²Institute for Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Keywords: Reinforcement learning; Decision-making, Hierarchical Bayesian modeling, Model-based fMRI

Revealing Neurocomputational Mechanisms of Reinforcement Learning and Decision-Making With the hBayesDM Package

Woo-Young Ahn¹, Nathaniel Haines¹, and Lei Zhang²

¹Department of Psychology, The Ohio State University, Columbus, OH

²Institute for Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Keywords: Reinforcement learning; Decision-making, Hierarchical Bayesian modeling, Model-based fMRI

Task (alphabetical order)	Model name	hBayesDM function	References (see below for full citations)
Balloon Analogue Risk Task	4 parameter model	bart_4par	Wallsten et al. (2005)
Choice reaction time Task	Drift diffusion model	choiceRT_ddm	Ratcliff (1978)
	Linear Ballistic Accumulator model	choiceRT_lba	S. Brown & Heathcote (2008) Annis et al. (2017)
Choice under Risk and Ambiguity (CRA) Task	Linear model	cra_linear	Levy et al. (2009)
	Exponential model	cra_exp	
Delay Discounting Task	Constant-Sensitivity (CS) model	dd_cs	Ebert & Prelec (2007) Samuelson (1937) Mazur (1987)
	Exponential model	dd_exp	
	Hyperbolic model	dd_hyp	
Iowa Gambling Task (IGT)	Prospect Valence Learning-DecayRI	igt_pvl_decay	Ahn et al. (2011; 2014) Ahn et al. (2008) Worthy et al. (2013) Haines et al. (in press)
	Prospect Valence Learning-Delta	igt_pvl_delta	
	Value-Plus-Perseverance (VPP)	igt_vpp	
	Outcome-Represent. Learning (ORL)	igt_orl	
Orthogonalized Go/Nogo Task	RW+noise	gng_m1	Guitart-Masip et al. (2012) Guitart-Masip et al. (2012) Guitart-Masip et al. (2012) Cavanagh et al. (2013)
	RW+noise+go bias	gng_m2	
	RW+noise+go bias+Pav. bias	gng_m3	
	M5 (see Table 1 of the reference)	gng_m4	
Peer influence task	Other-conferred utility (OCU)	peer_ocu	Chung et al. (2015)
Probabilistic Reversal Learning (PRL) Task	Experience-Weighted Attraction	prl_ewa	Ouden et al. (2013) Gläscher et al. (2009) Ouden et al. (2013)
	Fictitious update	prl_fictitious	
	Reward-Punishment (Rew.-Pun.)	prl_rp	
	Fictitious + Rew.-Pun.	prl_fictitious_rp	
	Fictitious + Rew.-Pun. w/o alpha	prl_fictitious_rp_woa	
	Fictitious w/o alpha	prl_fictitious_woa	
Probabilistic Selection Task	Q-learning with two learning rates	pst_gainloss_Q	M. J. Frank et al. (2007)
Risk-Aversion Task	Prospect Theory (PT)	ra_prospect	Sokol-Hessner et al. (2009)
	PT without loss aversion (LA)	ra_noLA	
	PT without risk aversion (RA)	ra_noRA	
			Tom et al. (2007)
Risky Decision Task	Happiness model	rdt_happiness	Rutledge et al. (2014)
Two-Armed Bandit (Experience-based) Task	Rescorla-Wagner (delta) model	bandit2arm_delta	Erev et al. (2010) Hertwig et al. (2004)
Two Step (TS) Task	7 parameter model	ts_7par	Daw et al. (2011)
	6 parameter model	ts_6par	
	4 parameter model	ts_4par	
			Wunderlich et al. (2012)
Four-Armed Bandit (Experience-based) Task	Fictive upd.+rew/pun sens.	bandit4arm_4par	Seymour et al. (2012) Seymour et al. (2012)
	Fictive upd.+rew/pun sens.+lapse	bandit4arm_lapse	
Ultimatum Game	Ideal Bayesian observer model	ug_bayes	Xiang et al. (2013) Gu et al. (2015)
	Rescorla-Wagner (delta) model	ug_delta	
Wisconsin Card Sorting Task	Sequential learning model	wcs_sql	A. J. Bishara et al. (2010)

cognitive model

statistics

computing

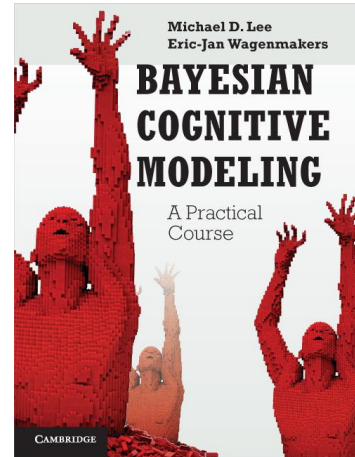
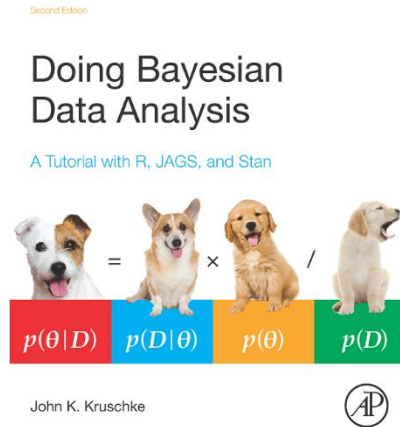
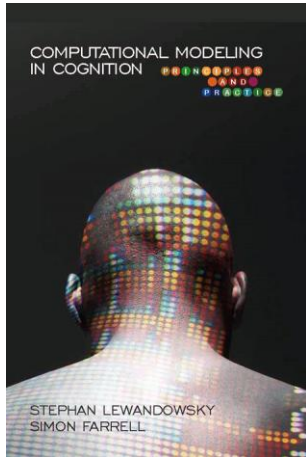


Resources

cognitive model

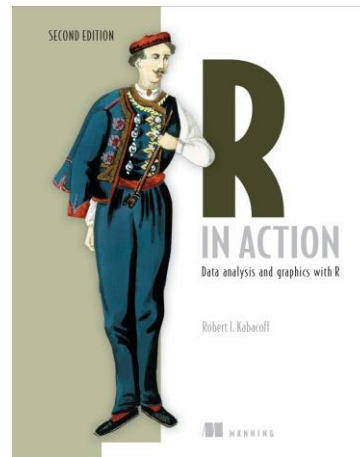
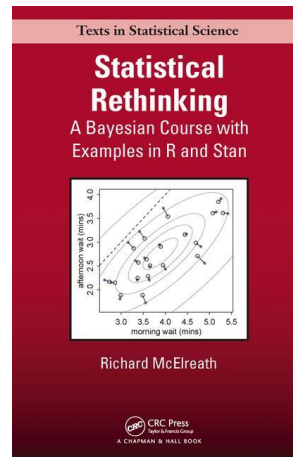
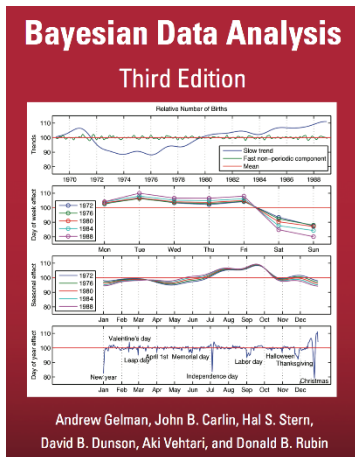
statistics

computing



help & discussion

<https://discourse.mc-stan.org/>



References

- Ahn, W.-Y., Haines, N., & Zhang, L. (2017). Revealing Neurocomputational Mechanisms of Reinforcement Learning and Decision-Making With the hBayesDM Package. *Computational Psychiatry*, 1, 24–57.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, 76(1).
- Cohen, J. D., Daw, N., Engelhardt, B., Hasson, U., Li, K., Niv, Y., ... & Willke, T. L. (2017). Computational approaches to fMRI analysis. *Nature Neuroscience*, 20(3), 304.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis (3rd ed.)*. New York, NY: CRC Press.
- Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7, 457–472.
- Gläscher, J., Hampton, A. N., & O'Doherty, J. P. (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cerebral Cortex*, 19(2), 483-495.
- Gläscher, J. & O'Doherty, J. P. (2010). Model-based approaches to neuroimaging: combining reinforcement learning theory with fMRI data. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(4), 501-510.
- Hampton, A. N., Adolphs, R., Tyszka, J. M., & O'Doherty, J. P. (2007). Contributions of the amygdala to reward expectancy and choice signals in human prefrontal cortex. *Neuron*, 55(4), 545-555.
- Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.
- Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge university press.
- Lewandowsky, S., & Farrell, S. (2010). *Computational modeling in cognition: Principles and practice*. Sage Publications.
- McElreath, R. (2018). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. CRC Press.
- O'Doherty, J. P., Hampton, A., & Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Annals of the New York Academy of Sciences*, 1104(1), 35-53.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, 2, 64-99.
- Sutton, R. S., Barto, A. G., & Bach, F. (1998). *Reinforcement learning: An introduction*. MIT press.
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413-1432.

Acknowledgement

Antonius Wiehler (MMB, Paris)

Jan Gläscher (UKE, Hamburg)

Woo Young Ahn (SNU, Seoul)

Nate Haines (OSU, OH)

ANY
QUESTIONS
?

Happy Computing!