Introduction to R/Rstudio

With general tips and tricks for working with R and RStudio

Lei Zhang Aamir Sohail

Table of contents

0.1 Welcome

This website[1][2] is an adapted version of teaching materials originally made for the award winning* **BayesCog** seminar at the Faculty of Psychology, University of Vienna, as part of the Advanced Seminar for master's students (Mind and Brain track; recorded during Summer Term 2020/2021). Further content from the compact BayesCog workshop at UKE Hamburg have also been added.

Recording: Recordings from the original version of the course are available on YouTube (also see below). The most recent recording from the 2021 summer semester is also available on Youtube.

Outreach: **Twitter thread** (being **liked 700+ times** on Twitter) summarizing the contents of the course.

Award/Recognition: The original course received a commendation award from the Society for the Improvement of Psychological Science (SIPS) (also see a tweet), as well as an ECR Teaching Award from the Faculty pf Psychology, University of Vienna.

0.2 Course summary

Computational modeling and mathematical modeling provide an insightful quantitative framework that allows researchers to inspect latent processes and to understand hidden mechanisms. Hence, computational modeling has gained increasing attention in many areas of cognitive science through cognitive modeling. One illustration of this trend is the growing popularity of Bayesian approaches to cognitive modeling. To this end, this course teaches the theoretical and practical knowledge necessary to perform, evaluate and interpret Bayesian modeling analyses, with a specific emphasis towards modeling latent cognitive processes.

This course is dedicated to introducing students to the basic knowledge of Bayesian statistics as well as basic techniques of Bayesian cognitive modeling. We will use R/RStudio and a newly developed statistical computing language - Stan - to perform Bayesian analyses, ranging from simple binomial models and linear regression models to more complex hierarchical reinforcement learning (RL) models.

[1] Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information. San Francisco: W. H. Freeman.

[2] Adapted from Jan Gläscher's workshop on cognitive modeling.

0.3 Contributors

 **Lei Zhang**

Dr. Lei Zhang is a cognitive computational neuroscientist, Principal Investigator of the Adaptive Learning Psychology and Neuroscience Lab, ALP(e)N Lab, and Associate Professor at the School of Psychology, University of Birmingham. Lei obtained his BSc in Psychology from Beijing Normal University, China, his MSc in Cognitive Neuroscience at the Basque Center on Cognition, Brain and Language, Spain and his PhD (summa cum laude), along with a one-year transition postdoc, with Jan Gläscher at the Institute of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Germany. He was then a Roche intern for Scientific Exchange (RiSE) at F. Hoffmann-La Roche AG, and worked as a postdoctoral fellow with Claus Lamm at the Social Cognitive and Affective Neuroscience Unit (SCAN-Unit), University of Vienna, Austria. Lei joined the Centre for Human Brain Health, Institute of Mental Health, and School of Psychology at University of Birmingham as an Associate Professor in 2022. His research applies knowledge from cognitive neuroscience, psychology, and computational modeling to gain a comprehensive understanding of how the brain computes values and social information when making decisions, and how they are affected in mental health disorders. He is also motivated towards fostering open and collaborative science, being a founding member of the grassroots China Open Science Network (COSN) and is the recipient of multiple awards including the BNA Individual Researcher Credibility Prize in 2024.

Aamir Sohail

Aamir Sohail is an MRC Advanced Interdisciplinary Methods (AIM) DTP PhD student based at the Centre for Human Brain Health (CHBH), University of Birmingham, where he is supervised by Lei Zhang and Patricia Lockwood. He completed a BSc in Biomedical Science at Imperial College London, followed by an MSc in Brain Imaging at the University of Nottingham. He then worked as a Junior Research Fellow at the Centre for Integrative Neuroscience and Neurodynamics (CINN), University of Reading with Anastasia Christakou. His research interests involve using a combination of behavioural tasks, computational modeling and neuroimaging to understand social decision-making, and using this knowledge to inform the precision-based treatment of mental health disorders. Outside of research, he is also passionate about facilitating inclusivity and diversity in academia, as well as promoting open and reproducible science.

0.4 License

0.5 Citing

If you use materials from this course in your work or research, please cite it as:

Zhang, L., & Sohail, A. (2025). BayesCog: Bayesian Statistics and Hierarchical Bayesian Modeling for Psychological Science [Online course]. Zenodo. https://doi.org/[DOI_HERE]

For BibTeX users:

```
@online{zhang_sohail_2025,
    title = {BayesCog: Bayesian Statistics and Hierarchical Bayesian Modeling for Psychol
    author = {Zhang, Lei and Sohail, Aamir},
    year = {2025},
    publisher = {Zenodo},
    doi = {DOI_HERE},
    url = {ZENODO_URL},
    note = {Online course},
    repository = {https://github.com/REPOSITORY_URL}
}
```

Note: Once this course is published on Zenodo, this citation information will be updated with the corresponding DOI and URL.

0.6 Contact

For bug reports, issues or comments, please contact Lei Zhang, or Aamir Sohail, or open a thread on the GitHub repository.

0.6.1 Aims and philosophy

This course was initially developed for master's students at the University of Vienna, but is intended for anyone interested in learning about Bayesian statistics, and using Bayesian methods to both build and apply cognitive models. As a result, students at all levels, postdocs, as well as senior faculty have completed this course!

The term 'Bayesian' is not limited to modeling, and reflects a more general approach to probability observed in many research areas. Therefore, to avoid confusion beforehand, this course is not about Bayes in the brain (Bayesian brain hypothesis), or why Bayesian statistics is a better alternative to frequentist statistics (even though it is).

Instead, this course aims to develop understanding and experience in using Bayesian statistics to analyse cognitive processes by constructing models.

The course's approach to understanding cognition through modeling is guided by David Marr's influential framework of three levels of analysis[3]:

1. **The Computational Level (Why):** This addresses what problem the cognitive system is trying to solve and why. It focuses on the goals and logic of the computation.
2. **The Algorithmic Level (What):** This specifies the representation and algorithm used to solve the computational problem. It describes the rules and strategies that implement the solution.
3. **The Implementation Level (How):** This details how the algorithm is physically realized in neural circuits and brain structures.

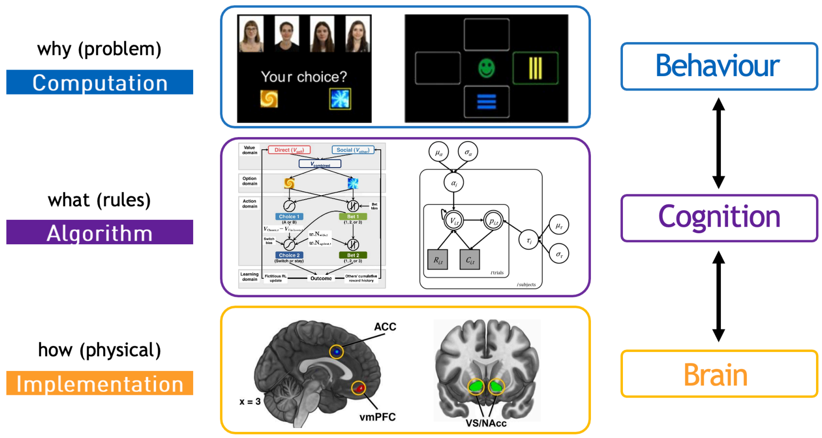
Throughout this course, we'll use cognitive modeling as a bridge between these levels, particularly focusing on how algorithmic-level models can help us understand behaviour:

Marr's framework as applied to cognitive modeling

The goals of this course are to:

- Build a foundational knowledge in cognitive behaviour and model-building

[3] Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information. San Francisco: W. H. Freeman.



- Learn practical `R` programming
- Build cognitive models using RStan
- Engage in open-source methods using `git` and GitHub (optional)

Subsequently, after completing the course, you …

- feel comfortable with reading mathematical equations
- consider the implementation of the "computational modeling" section
- gain insightful understanding of Bayesian stats and modeling
- can apply computational modeling in your own experiments

This course particularly places a strong emphasis on **hands-on experience modeling in Stan**, gaining experience with the modeling workflow depicted below[4]:

The modeling workflow covered in this course

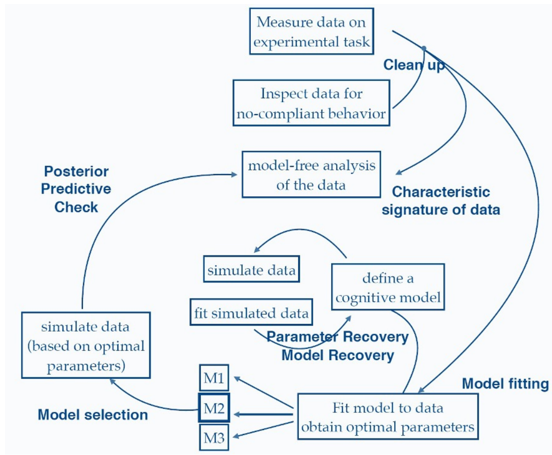
Throughout this course, you will learn to implement parameter and model recovery, perform posterior predictive checks, and assess competing models using model selection. Don't worry if none of this is currently familiar; hopefully it will be after completing this course!

0.7 Pre-requisites

To fully benefit from the materials, you are expected to have:

- Some basic understanding of frequentist statistics
- Some basic knowledge of programming

[4] Adapted from Jan Gläscher's workshop on cognitive modeling.

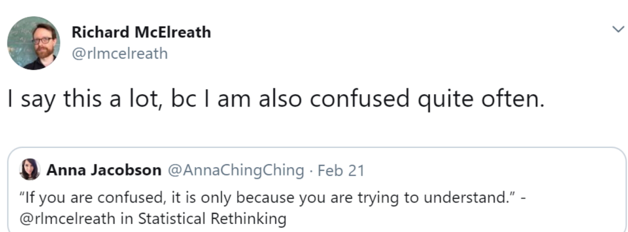


> **ℹ Pre-requisites**
>
> Any advanced knowledge of programming or statistics, whilst useful, is not required!

The course develops in complexity across the workshops, with the later material on complex models and model comparison being more challenging. However, do not feel as if you need to complete all workshops within a specific time-frame! These materials should consist part of your education in cognitive modeling.

Being in part a course in statistics and mathematical modeling, math equations feature throughout. Understanding the mathematical equations underlying computational models is necessary for understanding the relationship between data and parameters. However, some sections of the course feature additional math (e.g., derivations, statistical distributions) which are not strictly necessary to understand. **As a general rule of thumb, most important is that you try to understand the equations describing your data and parameters, and how they are represented mathematically!**

In any case, do not worry if you struggle during the course!



0.7.1 List of folders and contents

The workshops are split into a several distinct sub-folders …

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	Observed weight and height	Linear regression model
06.reinforcement_learning	2-armed bandit task	Simple reinforcement learning (RL)
07.optm_rl	2-armed bandit task	Simple reinforcement learning (RL)
08.compare_models	Probabilistic reversal learning task	Simple and fictitious RL models
09.debugging	Memory Retention	Exponential decay model
10.model_based	WIP	WIP
11.delay_discounting	WIP	WIP

0.8 Set-up

0.8.1 R packages

This project uses `renv` to manage package dependencies. To set up the environment:

1. Install `renv`: `install.packages("renv")`
2. Open the project in `RStudio`
3. Run `renv::restore()` to install all required packages

After this initial setup, the project environment will load automatically whenever you open the project.

> Querying packages
>
> You can always check if you are missing a certain package by clicking on the ‘Packages’ tab (next to Files/Plots tab) or by running `library()`.

0.9 Exercises

In addition to following along with the taught material, there are a number of exercises that you can work through. The code and solutions to the exercises in some cases will be worked through, but will not in others.

Specifically, each workshop (located within the `/workshops` folder) will have two sub-folders: `_data` and `_scripts`. Each `_scripts` folder contains separate files with and without the `master` suffix.

```
R_basics.Rproj
_data
_scripts
    R_basics.R
    R_basics_master.R
```

The base script will contain the necessary code without the solutions, whilst the `master` script also contains the solutions.

> **Exercise is optional**
>
> You do not have to work through the exercises to benefit from and complete this course, but is recommended if you are wanting to gain practical experience with programming in Stan.

1 Workshop 1: R Basics

Welcome to the first workshop of the BayesCog course!

The first workshop is designed to primarily introduce those without any previous experience to `R` and RStudio so that everyone is on the same page for the subsequent workshops. We will also cover tips and tricks with working in `R`, so it may also be useful for those already familiar!

Topics for this workshop include:

- Understanding `R` and RStudio's interface and functionality
- Basic R operations and data types
- Working with variables and functions
- Data structures and manipulation
- Control flow and logical operators
- Using `R` packages
- Data visualization with `ggplot2`

Working directory for this workshop

Model code and `R` scripts for this workshop are located in the (`/workshops/01.R_basics`) directory. Remember to use the `R.proj` file within each folder to avoid manually setting directories!

The copy of this workshop notes can be found on the course GitHub page.