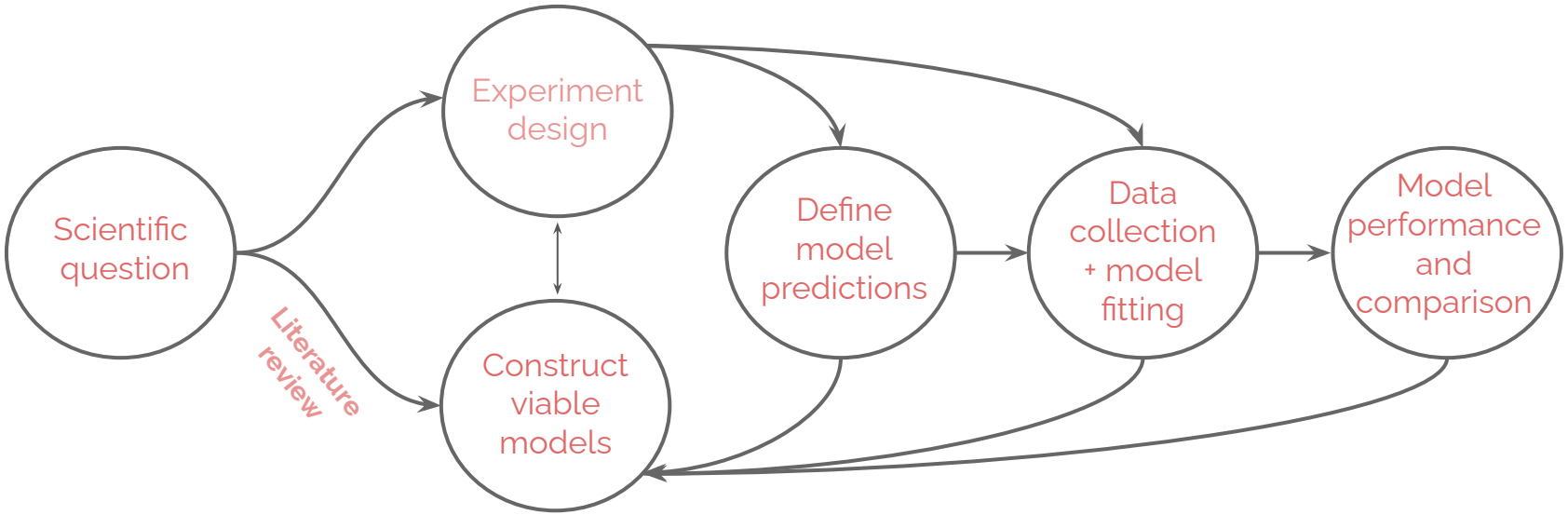
HELMHOLTZ MUNICH

# Generating Computational Cognitive Models using Large Language Models

Milena Rmus*, **Akshay Jagadish*,** Marvin Mathony,
Tobias Ludwig, & Eric Schulz

NeurIPS-2025 (main track)


Scientific question
Literature review
Experiment design
Construct viable models
Define model predictions
Data collection + model fitting
Model performance and comparison

# Precise but expensive

Handcrafting


Time

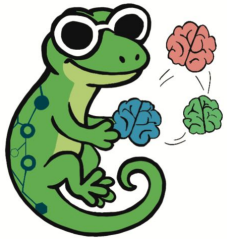
Personal bias


Large Language Models can help address this!

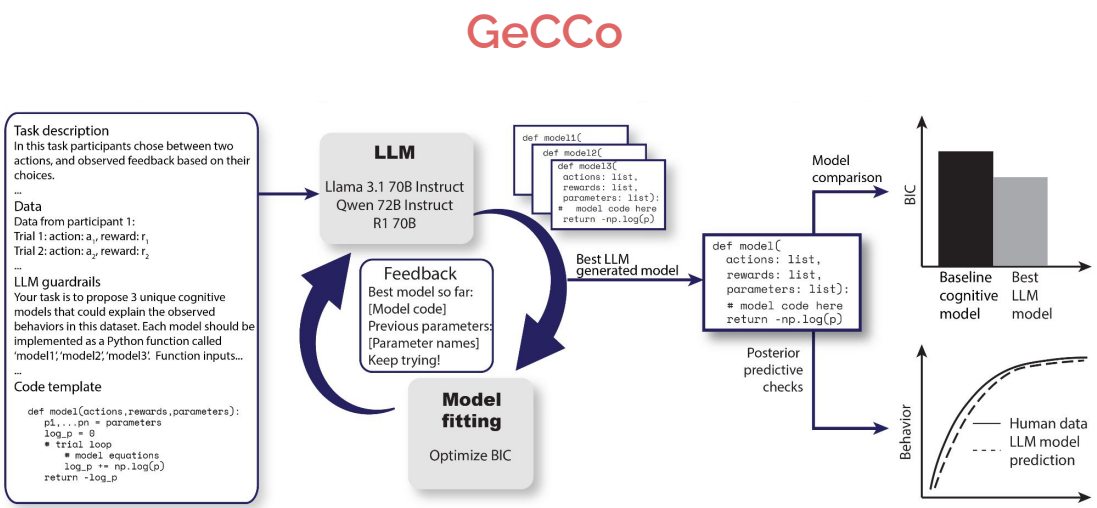
# Introducing the pipeline for **Ge**nerating **C**omputational **Co**gnitive Models (GeCCo)

**Task description**
In this task participants chose between two actions, and observed feedback based on their choices.
...
**Data**
Data from participant 1:
Trial 1: action: $a_1$, reward: $r_1$
Trial 2: action: $a_2$, reward: $r_2$
...
**LLM guardrails**
Your task is to propose 3 unique cognitive models that could explain the observed behaviors in this dataset. Each model should be implemented as a Python function called 'model1', 'model2', 'model3'. Function inputs...
...
**Code template**

```
def model(actions,rewards,parameters):
    p1,...pn = parameters
    log_p = 0
    # trial loop
        # model equations
        log_p += np.log(p)
    return -log_p
```



# Results



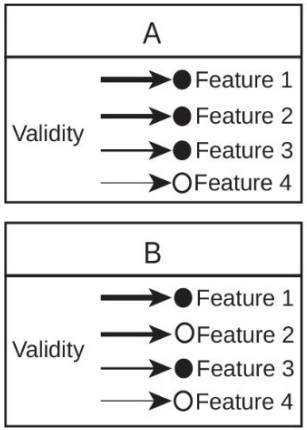

## Base LLMs

1. Llama-3 70B
2. Qwen-2.5 72B
3. R1 70B

## Cognitive domains

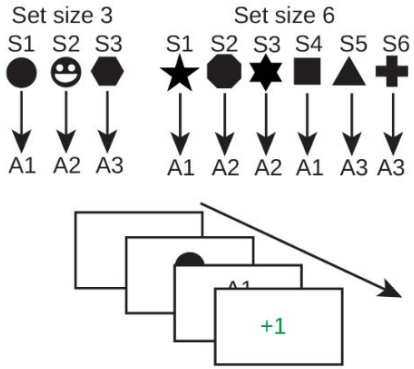
1. Decision making
2. Learning
3. Planning
4. Working memory

# Results: Decision making




Hilbig & Moshagen (2014)

# Results: Working memory




Rmus et al (2023)

# Control experiments

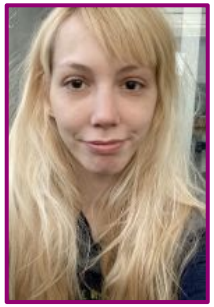
# Conclusion

- A scalable pipeline using LLMs for generating computational cognitive models, called GeCCo.

- GeCCo successfully generates interpretable models across multiple domains:
  1. Decision-making
  2. Reinforcement Learning
  3. (Model-based) Planning
  4. Working Memory

- LLMs have the potential to democratize complex scientific discovery and accelerate pace of scientific research in cognitive science.

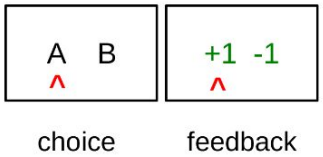
# Acknowledgments



HCAI

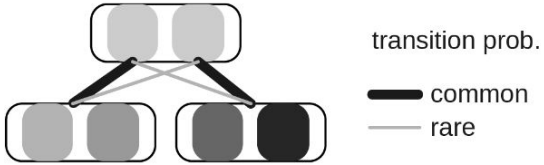
# Results: Learning

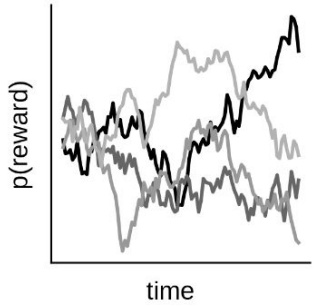
Chambon et al. 2020
full feedback (exp. 2)




Chambon et al. (2020)

# Results: Planning




drifting reward probabilities for 2nd-stage states to encourage exploration