

Marcel Binz

Ingolstädter Landstraße 1, 85764 Oberschleißheim, Germany

✉ marcel.binz@helmholtz-munich.de 🌐 marcelbinz.github.io 📄 marcelbinz ☎ (+49)15203605574

RESEARCH INTERESTS

Cognitive Science; Machine Learning; Meta-Learning; Resource Rationality; Large Language Models; Deep Learning; Bayesian Inference; Information Theory; Decision-Making; Reinforcement Learning

POSITION

Helmholtz Munich , Institute of Human-Centered AI Research scientist and deputy head	12/2023 - present
--	--------------------------

EXPERIENCE

Max Planck Institute for Biological Cybernetics , PI: Dr. Eric Schulz Postdoctoral researcher	02/2021 - 11/2023
Harvard University , PI: Prof. Samuel Gershman Research visit	09/2019 - 12/2019
Facebook Inc. Research internship	06/2016 - 12/2016
Eberhard Karls Universität Tübingen , PI: Prof. Martin Butz Research assistant	04/2015 - 08/2015

EDUCATION

Philipps-Universität Marburg , PI: Prof. Dominik Endres Dr. rer. nat. (Psychology)	2018 - 2021
KTH Royal Institute of Technology, Stockholm M.Sc. (Machine Learning)	2015 - 2018
Eberhard Karls Universität Tübingen B.Sc. (Cognitive Science)	2012 - 2015

AWARDS

German Cognitive Science Society Best Publication Award Best publication in cognitive science by a young investigator	2018 - 2022
EuroCogSci 2019 Best Poster Award Best poster presentation	2019
DMV-Abiturpreis Excellent performance in high school mathematics	2010

PUBLICATIONS

Binz, M. and Schulz, E., 2024. Turning large language models into cognitive models. *International Conference on Learning Representations (ICLR)*.

- Binz, M.**, Dasgupta, I., Jagadish, A., Botvinick, M., Wang, J. X. and Schulz, E., 2023. Meta-learned models of cognition. *Behavioral and Brain Sciences*.
- Coda-Forno, J., **Binz, M.**, Akata, Z., Botvinick, M., Wang, J. X. and Schulz, E., 2023. Meta-in-context learning in large language models. *37th Conference on Neural Information Processing Systems (NeurIPS 2023)*.
- Saanum, T., Éltető, N., Dayan, P., **Binz, M.** and Schulz, E., 2023. Reinforcement Learning with Simple Sequence Priors. *37th Conference on Neural Information Processing Systems (NeurIPS 2023)*.
- Schulze Buschoff, L. M., Schulz, E. and **Binz, M.**, 2023. The Acquisition of Physical Knowledge in Generative Neural Networks. *Fortieth International Conference on Machine Learning (ICML 2023)*.
- Binz, M.** and Schulz, E., 2023. Using cognitive psychology to understand GPT-3. *Proceedings of the National Academy of Sciences*.
- Binz, M.** and Schulz, E., 2022. Reconstructing the Einstellung Effect. *Computational Brain & Behavior*.
- Binz, M.** and Schulz, E., 2022. Modeling Human Exploration Through Resource-Rational Reinforcement Learning. *36th Conference on Neural Information Processing Systems (NeurIPS 2022)*. **Selected as Oral**.
- Binz, M.**, Gershman, S.J., Schulz, E. and Endres, D., 2022. Heuristics From Bounded Meta-Learned Inference. *Psychological Review*.
- Brändle, F., **Binz, M.** and Schulz, E., 2022. Exploration Beyond Bandits. *The Drive for Knowledge: The Science of Human Information Seeking*. Cambridge University Press.

PREPRINTS

- Jagadish, A. K., Coda-Forno, J., Thalmann, M., Schulz, E. and **Binz, M.**, 2024. Ecologically rational meta-learned inference explains human category learning.
- Schubert, J. A., Jagadish, A. K., **Binz, M.** and Schulz, E., 2024. In-context learning agents are asymmetric belief updaters.
- Binz, M.**, Alaniz, S., Roskies, A., Aczel, B., Bergstrom, C., Allen, C., Schad, D., Wulff, D. U., West, J., Zhang, Q., Shiffrin, R., Gershman, S. J., Popov, V., Bender, E. M., Marelli, M., Botvinick, M. M., Akata, Z. and Schulz, E., 2023. How should the advent of large language models affect the practice of science? *In review at PNAS*.
- Hussain, Z., **Binz, M.**, Mata, R. and Wulff, D. U., 2023. A tutorial on open-source large language models for behavioral science. *In review at Behavior Research Methods*.
- Jagadish, A. K., **Binz, M.**, Saanum, T., Wang, J. X. and Schulz, E., 2023. Zero-shot compositional reinforcement learning in humans. *In review at Nature Human Behaviour*.
- Demircan, C., Saanum, T., Pettini, L., **Binz, M.**, Baczkowski, B. M., Kaanders, P., . . . and Schulz, E., 2023. Language Aligned Visual Representations Predict Human Behavior in Naturalistic Learning Tasks. *In review at PNAS*.
- Coda-Forno, J., Witte, K., Jagadish, A. K., **Binz, M.**, Akata, Z. and Schulz, E., 2023. Inducing anxiety in large language models increases exploration and bias. *In revision at Nature*.

NON-ARCHIVAL PUBLICATIONS

- Schubert, J. A., Jagadish, A. K., **Binz, M.** and Schulz, E., 2023. A Rational Analysis of the Optimism Bias using Meta-Reinforcement Learning. *Conference on Cognitive Computational Neuroscience (CCN)*

2023).

Jagadish, A. K., Saanum, T., Wang, J. X., **Binz, M.** and Schulz, E., 2022. Probing Compositional Inference in Natural and Artificial Agents. *5th Multidisciplinary Conference on Reinforcement Learning and Decision Making (RLDM 2022)*.

Demircan, C., Pettini, L., Saanum, T., **Binz, M.**, Baczkowski, B. M., Doeller, C., . . . and Schulz, E., 2022. Decision-making with naturalistic options. *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Binz, M. and Endres, D., 2019. Emulating human developmental stages with bayesian neural networks. *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Binz, M. and Endres, D., 2019. Where do heuristics come from?. *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Butz, M. V., Simonic, M., **Binz, M.**, Einig, J., Ehrenfeld, S. and Schrodtt, F., 2016. Is it Living? Insights from Modeling Event-Oriented, Self-Motivated, Acting, Learning and Conversing Game Agents. *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Binz, M., Otte, S. and Zell, A., 2015. On the applicability of recurrent neural networks for pattern recognition in electroencephalography signals. *Workshop New Challenges in Neural Computation*.

TEACHING

Computational Cognitive Science , Eberhard Karls University of Tübingen Lecturer	2022, 2023
International Interdisciplinary Computational Cognitive Science Summer School Lecturer	2022, 2023
Bayesian Statistics and Machine Learning , Philipps-Universität Marburg Lecturer	2020
Theoretical Neuroscience , Philipps-Universität Marburg Lecturer	2019, 2020
Deep Learning in Data Science , KTH Royal Institute of Technology Teaching assistant	2017

REVIEWING

Open Mind	2024 - present
Nature	2023 - present
Nature Human Behaviour	2023 - present
International Conference on Learning Representations (ICLR)	2023 - present
Conference on Neural Information Processing Systems (NeurIPS)	2023 - present
Behavior Research Methods	2023 - present
Trends in Cognitive Sciences	2023 - present
Conference on Cognitive Computational Neuroscience	2023 - present
Proceedings of the National Academy of Sciences (PNAS)	2022 - present
Psychological Review	2022 - present
Computational Brain & Behavior	2022 - present
Annual Meeting of the Cognitive Science Society	2021 - present

SUPERVISION (PHD STUDENTS)

Julian Coda-Forno (co-supervised with Eric Schulz and Jane Wang) Meta-Learning in Large Language Models	2022 - present
Akshay Kumar Jagadish (co-supervised with Eric Schulz) Reverse-Engineering Adaptive Principles of Cognition	2021 - present

SUPERVISION (MASTER AND BACHELOR STUDENTS)

Johannes Schubert Investigating the Optimism Bias Using Meta-Reinforcement Learning	2023
Luca Schulze Buschoff Development as Decompression	2022
Akshay Kumar Jagadish Compositional Generalization in Meta-Reinforcement Learning	2021
Gwen Hirsch Comparing Meta-Learners with Human Performance in a Continual Learning Framework	2020
Hauke Niehaus Simulating Decision-Making Deficits in a Deep Meta-Reinforcement-Learning Agent	2019

INVITED TALKS

Cognitive Sciences Colloquium , Irvine	2024
Cognition, Brain, & Behavior Research Seminar , Harvard	2023
nEuro-economics seminar series , Paris	2023
Digital Change Symposium , Kloster Seeon	2023
International Interdisciplinary Computational Cognitive Science Summer School	2023
Language Models in Judgment and Decision Making Research Symposium , Vienna	2023
Large Language Models Meet Cognitive Science Workshop , Sydney	2023
Neuro-Cognitive Modeling Group Lab Meeting , Tübingen	2023
International Titisee Conference on NeuroAI , Titisee	2023
Colloquium of the Institute of Cognitive Science , Osnabrück	2023
Reinforcement Learning and Decision-Making Seminar , Tübingen	2023
Conference on Neural Information Processing Systems , New Orleans	2022
Memory, Judgement and Decision-Making Seminar , Mannheim	2022
International Interdisciplinary Computational Cognitive Science Summer School	2022
Conference of the German Cognitive Science Society , Freiburg	2022
Human and Machine Cognition Lab Meeting , Tübingen	2022
Reinforcement Learning and Decision-Making Seminar , Tübingen	2021
Joint Lab Retreat: Summerfield , Schuck, Schulz	2021
Colloquium of the Institute for Neuroinformatics , Bochum	2019

GRANTS

Scientific Inference and Statistical Inference Conference	2023
--	------

Funding for travel and accommodation	
International Titisee Conference on NeuroAI	2023
Funding for travel and accommodation	
German Academic Exchange Service (DAAD) Scholarship	2019
Funding for a three month research visit at Harvard University	
Summer Institute on Bounded Rationality	2019
Funding for travel and accommodation	