Eric Schulz

eric.schulz@helmholtz-munich.de http://hcai-munich.com

Employment History

- 2024- **Director.** Institute for Human-Centered AI. Helmholtz Center for Computational Health, Munich, Germany.
- 2020-2024 **Max Planck Research Group Leader.** Computational Principles of Intelligence Lab, MPI for Biological Cybernetics, Tübingen, Germany.
- 2017-2019 **Data Science Postdoctoral Fellow.** Harvard University, Cambridge, USA.
- 2013 **Volunteer.** Uganda Virus Research Institute, Entebbe, Uganda.
- 2012 – 2013 **Machine Learning Analyst.** Zalando, Berlin, Germany.
- 2008 – 2010 **Student Research Assistant.** MPI for Human Development, Berlin, Germany.
- 2006 – 2007 **Military Service.** United Nations Training Center, Hammelburg, Germany.

Education

- 2014 – 2017 **PhD Experimental Psychology.** University College London, UK.
- 2013 – 2014 **MRes Computer Science.** University College London, UK.
- 2011 – 2012 **MSc Applied Statistics.** University of Oxford, UK.
- 2010 – 2011 **MSc Cognitive and Decision Sciences.** University College London, UK.
- 2007 – 2010 **Vordiplom Psychology.** Humboldt University, Berlin, Germany.

Funding

- 2025-2030 **Wellcome Discovery Award** on cognitive rewards.
- 2024-2029 **ERC Starting Grant** on using cognitive psychology to understand LLMs.
- 2021-2025 **Volkswagen Artificial Intelligence and the Society of the Future Grant** to study curiosity in children and robots. EUR 1,400,000.
- 2022-2024 **BMBF Tübingen AI Center Grant** on teaching machines how to create objects and sketches via human-inspired meta-learning.
- **BMBF Tübingen AI Center Grant** to study multi-task representation learning.
- 2021-2024 **Univerisity of Tübingen Machine Learning Mini Graduate School** to study compositionality in minds and machines.
- 2020-2024 **Max Planck Research Group** on Computational Principles of Intelligence.
- 2020-2023 **Jacobs Early Career Research Fellowship** for Highly talented young scholars working on child development.
- 2017-2019 **Harvard Data Science Postdoctoral Fellowship**.
- 2018 **Robert J. Glushko Award** for Best Doctoral Dissertation in Cognitive Science.
- 2013 **ESPRC scholarship** funding both MRes and PhD at UCL by the Centre for Doctoral Training in Financial Computing and Analytics.
- 2011 **Haniel scholarship** funding MSc at the University of Oxford.
- 2010 **DAAD scholarship** funding MSc at University College London.

Invited Talks

- 2025 **Max Planck Institute for Mathematics in the Sciences.** Topology Seminar.

Invited Talks (continued)

2024
- **Max Planck Institute for Psychiatry.** Departmental Seminar Series.
- **Birkbeck, University of London.** Cognition of Generative AI.
- **Harvard University.** Kempner Institute Seminar Series.
- **University College London.** Gatsby Seminar Series.

2023
- **University of Zurich.** Behavioral Economics Seminar Series.
- **ELLIS Natural Intelligence.** Unit Meeting.
- **Elkana Forum Berlin.** Short Statement.
- **Helmholtz Munich.** Colloquium.
- **Transcontinental Computational Psychiatry Workgroup.**

2022
- **Max Planck Center for Computational Psychiatry.** Colloquium.
- **Max Planck School of Cognition.** Colloquium.
- **Leibniz IWM Tübingen.** Departmental Talk.
- **MPI Leipzig.** Origins of Intelligence Lecture.

2021
- **MPI Berlin.** Department for Humans and Machines.
- **University of Cologne.** Peters Lab.
- **University of Ghent.** Center for Cognitive Neuroscience.
- **TU Darmstadt.** Center for Cognitive Science.

2020
- **University of Tübingen.** Cognitive Science Colloquium.
- **University of New South Wales.** Departmental Colloquium.
- **University of Oxford.** Summerfield Lab Meeting.
- **University of Warwick.** Cognitive Science Group.
- **The University of Edinburgh.** Computational Cognitive Science Group.

2019
- **Stanford University.** FriSem.
- **Max Planck Institute for Human Cognitive and Brain Sciences.** Guest Lecture.
- **Max Planck Institute for Biological Cybernetics.** MPRG Symposium.
- **Cognitive Lunch.** MIT.

2018
- **Ohio State Univeristy.** Brown bag seminar series. Invited by Jay Myung.
- **Early Childhood Cognition Lab.** Lab Meeting at MIT.
- **ONR Science of Autonomy.** Grant Review.
- **Ecole Normale Supérieure.** Workshop organized by Stefano Palminteri.
- **Cognitive Science Conference.** Symposium for Glushko award winners.

2017
- **ConCats seminar series.** New York University.
- **CBB Lunch.** Harvard University.
- **Cognitive Psychology Colloquium.** University of Göttingen.
- **Cognitive Science Colloquium.** University of Onsabrück.

2016
- **London Judgement and Decision Making Seminar.** University College London.
- **Gershman Lab Meeting.** Harvard University.
- **Coffee and Tea Talk.** Max Planck Institute for Human Development.

2015
- **Psychology Seminar Series .** City University.
- **Krause Lab Meeting .** ETH Zürich.
- **Oberauer Lab Meeting .** University of Zürich.
- **Economic Psychology Colloquium .** University of Basel.

Press Coverage

Spiegel	Wenn die KI Angst bekommt, wird sie rassistisch.
Zeit	Captchas in 2025.
Spiegel	Spachmodelle sind egoistisch und nachtragend.
MIT Tech Review	Kann KI Angst haben?
ARD Radio	Machinensturm: Auf der Suche nach Intelligenz.
Deutschlandfunk	Angst vor der KI.
Deutsche Welle	What ChatGPT cannot do.
heise.de	Emotionale KI.
Spektrum	Die kognitiven Kompetenzen des Sprachmodells GPT-3.
Bundesdruckerei	Brücken zwischen Menschen und Maschinen bauen.
Max Magazin	KI ist nicht neugierig.
Radio Eins	Über die Stärken und Schwächen von Chatbots.
Tagesspiegel	Wie denkt Chat-GPT?
Tagblatt	3 Fragen zur KI.
SWR	Wann Menschen Spass empfinden.
Schwäbische	Wie der Südwesten Spitzenforscher ködert.

Supervision

Postdoctoral Researchers

2024-	**Alireza Modirshanechi**. Exploration and surprise in reinforcement learning.
2024-	**Konstantinos Voudouris**. Cognitive benchmarks for AI systems.
2024-	**Milena Rmus**. Code-generation from scientific data.
2021-	**Mirko Thalmann**. Memory-efficient generalization.
2021-	**Marcel Binz**. Foundation models of cognition.

Doctoral Students

2024-	**Marvin Mathony**. Theories of exploration in RNNs.
2024-	**Elif Akata**. Interacting LLMs.
2023-	**Can Demircan**. Internal representations of LLMs.
2023-	**Luca Schulze-Buschoff**. Cognition in multi-modal language models.
2022-	**Kristin Witte**. Exploration, anxiety, and LLMs.
2022-	**Julian Coda-Forno**. Using cognitive psychology to understand LLMs.
2021-	**Tobias Ludwig**. Multi-task reinforcement learning.
2021-	**Susanne Haridi**. Scaling laws of human inference.
2021-	**Akshay Jagadish**. Meta-learning on LLM-generated data.
2021-	**Tankred Saanum**. Efficient deep reinforcement learning.
2020-	**Alexander Kipnis**. Meta-benchmarks for LLMs.
2020-2024	**Franziska Brändle**. Now Postdoc at Oxford.
2020-2024	**Shuchen Wu**. Now Postdoc at Allen Institute.
2020-2023	**Lion Schulz** (secondary supervisor). Now Data Scientist at Bertelsmann.
2020-2023	**Lara Bertram** (secondary supervisor). Now Postdoc at Cambridge.

Selected Publications

Full list: `https://scholar.google.com/citations?user=74Cj5GYAAAAJ`
I have published more than 100 articles. My H-index is 38.

Selected Journal Articles

1. Akata, E., Schulz, L., Coda-Forno, J., Oh, S. J., Bethge, M., & Schulz, E. (2025). Playing repeated games with large language models. *Nature Human Behaviour*.
2. Binz, M., Akata, E., Bethge, M., Brändle, F., Callaway, F., Coda-Forno, J., … Schulz, E. (2025). Centaur: a foundation model of human cognition. *Nature*.
3. Binz, M., Alaniz, S., Roskies, A., Aczel, B., Bergstrom, C. T., Allen, C., … Schulz, E. (2025). How should the advent of large language models affect the practice of science? *Proceedings of the National Academy of Sciences*.
4. Buschoff, L. M. S., Akata, E., Bethge, M., & Schulz, E. (2025). Visual cognition in multimodal large language models. *Nature Machine Intelligence*.
5. Wu, C. M., Meder, B., & Schulz, E. (2025). Unifying principles of generalization: past, present, and future. *Annual Review of Psychology*, *76*.
6. Allen, K., Brändle, F., ..., & Schulz, E. (2024). Using games to understand the mind. *Nature Human Behaviour*.
7. Binz, M., Dasgupta, I., Jagadish, A. K., Botvinick, M., Wang, J. X., & Schulz, E. (2024). Meta-learned models of cognition. *Behavioral and Brain Sciences*.
8. Binz, M. & Schulz, E. (2023). Using cognitive psychology to understand GPT-3. *Proceedings of the National Academy of Sciences*.
9. Brändle, F., Stocks, L. J., Tenenbaum, J. B., Gershman, S. J., & Schulz, E. (2023). Intrinsic exploration as empowerment in a richly structured online game. *Nature Human Behaviour*.
10. Garvert, M. M., Saanum, T., Schulz, E., Schuck, N. W., & Doeller, C. F. (2023). Hippocampal spatio-predictive cognitive maps adaptively guide reward generalization. *Nature Neuroscience*.
11. Giron, A. P., Ciranka, S., Schulz, E., van den Bos, W., Ruggeri, A., Meder, B., & Wu, C. M. (2023). Developmental changes in exploration resemble stochastic optimization. *Nature Human Behaviour*.
12. Tomov, M., Schulz, E., & Gershman, S. J. (2021). Multi-task reinforcement learning in humans. *Nature Human Behaviour*.
13. Schulz, E., Bhui, R., Love, B. C., Brier, B., Todd, M. T., & Gershman, S. J. (2019). Structured, uncertainty-driven exploration in real-world consumer choice. *Proceedings of the National Academy of Sciences*.
14. Schulz, E., Speekenbrink, M., & Krause, A. (2018). A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of mathematical psychology*.
15. Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018). Exploration and generalization in vast spaces. *Nature Human Behaviour*.

Selected Conference Proceedings (Major AI Conferences)

1. Buschoff, L. M., Voudouris, K., Akata, E., Bethge, M., Tenenbaum, J. B., & Schulz, E. (2025). Testing the limits of fine-tuning to improve reasoning in vision language models. In *International Conference on Machine Learning*.

2. Demircan, C., Saanum, T., Jagadish, A. K., Binz, M., & Schulz, E. (2025). Sparse autoencoders reveal temporal difference learning in large language models. In *International Conference on Learning Representations*.
3. Kipnis, A., Voudouris, K., Buschoff, L. M. S., & Schulz, E. (2025). Metabench–a sparse benchmark to measure general ability in large language models. In *International Conference on Learning Representations*.
4. Wu, S., Thalmann, M., Dayan, P., Akata, Z., & Schulz, E. (2025). Building, reusing, and generalizing abstract representations from concrete sequences. In *International Conference on Learning Representations*.
5. Binz, M. & Schulz, E. (2024). Turning large language models into cognitive models. In *International conference on learning representations*.
6. Coda-Forno, J., Binz, M., Wang, J. X., & Schulz, E. (2024). Cogbench: a large language model walks into a psychology lab. In *International Conference on Machine Learning*.
7. Demircan, C., Saanum, T., Pettini, L., Binz, M., Baczkowski, B. M., Doeller, C. F., … Schulz, E. (2024). Evaluating alignment between humans and neural network representations in image-based learning tasks. In *The thirty-eighth annual conference on neural information processing systems*.
8. Jagadish, A. K., Coda-Forno, J., Thalmann, M., Schulz, E., & Binz, M. (2024). Ecologically rational meta-learned inference explains human category learning. In *International Conference on Machine Learning*.
9. Saanum, T., Dayan, P., & Schulz, E. (2024). Simplifying latent dynamics with softly state-invariant world models. In *The thirty-eighth annual conference on neural information processing systems*.
10. Schubert, J. A., Jagadish, A. K., Binz, M., & Schulz, E. (2024). In-context learning agents are asymmetric belief updaters. In *International Conference on Machine Learning*.
11. Coda-Forno, J., Binz, M., Akata, Z., Botvinick, M., Wang, J., & Schulz, E. (2023). Meta-in-context learning in large language models. In *Advances in Neural Information Processing Systems*.
12. Saanum, T., Elteto, N., Dayan, P., Binz, M., & Schulz, E. (2023). Reinforcement learning with simple sequence priors. In *Advances in Neural Information Processing Systems*.
13. Salewski, L., Alaniz, S., Rio-Torto, I., Schulz, E., & Akata, Z. (2023). In-context impersonation reveals large language models' strengths and biases. In *Advances in Neural Information Processing Systems*.
14. Binz, M. & Schulz, E. (2022). Exploration with a finite brain. In *Advances in Neural Information Processing Systems*.
15. Wu, S., Élteto, N., Dasgupta, I., & Schulz, E. (2022). Learning structure from the ground up: hierarchical representation learning by chunking. In *Advances in Neural Information Processing Systems*.