

# Konstantinos Voudouris

[k.voudouris@helmholtz-munich.de](mailto:k.voudouris@helmholtz-munich.de) | [LinkedIn](#) | [GitHub](#) | [Scholar](#) | [Website](#)

## EXPERIENCE

<b>Research Scientist, Cognitive Science</b> <i>Alignment Team—UK AI Security Institute</i>	London, UK Dec 2025–Present
<b>Post-doctoral Research Scientist, Helmholtz Munich</b> <i>Institute for Human-Centered AI (PI: Eric Schulz)</i>	Munich, Germany Jul 2024–Nov 2024
<b>Post-doctoral Research Assistant/Associate, University of Cambridge</b> <i>Major Transitions in the Evolution of Cognition (PIs: M. Halina, A. Barron, C. Klein)</i>	Cambridge, UK Nov 2023–Jul 2024

## EDUCATION

<b>PhD (Psychology, Philosophy of Science, Artificial Intelligence)</b> <i>Passed with no corrections (Thesis: <a href="#">From Animals to Algorithms</a>).</i>	University of Cambridge Oct 2020–Mar 2024
<b>Visiting Scholar</b> <i>Psychology Department (Sponsor: Alison Gopnik)</i>	University of California, Berkeley Jul–Sep 2023
<b>MPhil (Biological Sciences: Psychology)</b> <i>Passed with minor corrections</i>	University of Cambridge Oct 2019–Sep 2020
<b>MA Cantab. (Theoretical and Applied Linguistics)</b> <i>Starred Double First, 1st in cohort</i>	University of Cambridge Oct 2016–Jun 2019

## SELECTED PEER-REVIEWED PUBLICATIONS (\*FIRST AUTHOR, †SENIOR AUTHOR)

Slater, B., Mecattaf, M. G., Tešić, M., Prunty, J., **Voudouris, K.**† & Cheke, L. G.† (2025). A little less conversation, a little more action, please: Investigating the physical common-sense of LLMs in a 3D embodied environment. *PRICAI 2025*. ([Pre-Print](#))

**Voudouris, K.**\* (forthcoming). Cognitive Simplicity As An Idealization. *Animal Behavior & Cognition*. ([Pre-Print](#))

**Voudouris, K.**\* (2025). Analogies And The Associative-Cognitive Distinction In Comparative Psychology. *Biology & Philosophy*, 40, 19. ([Pre-Print](#))

Binz, M., ..., **Voudouris, K.**, ..., **Schulz, E.** (2025) A foundationmodel to predict and capture human cognition. *Nature* 644, 1002–1009.

**Voudouris, K.**\*, Schulze-Buschoff, L. M.\* , Akata, E., Bethge, M., Tenenbaum, J., & Schulz, E. (2025) Testing the limits of fine-tuning for improving visual cognition in vision language models. *42nd Intl Conference on Machine Learning*. ([Pre-Print](#))

Pacchiardi, L., **Voudouris, K.**, Slater, B., Martinez-Plumed, F., Zhou, L., Hernández-Orallo, J., & Schellaert, W. (2025) PredictaBoard: Benchmarking LLM Score Predictability. *Findings of the ACL*. ([Pre-Print](#))

**Voudouris, K.**\*, Cheke, L. G., & Schulz, E. (2025) Bringing comparative cognition approaches to AI systems. *Nature Reviews Psychology*.

**Voudouris, K.**\*, ..., & Cheke, L. G. (2025) The Animal-AI Environment: A Virtual Laboratory For Comparative Cognition and Artificial Intelligence Research. *Behavior Research Methods*, 57:107.

Kipnis, A., **Voudouris, K.**, Buschoff Schulze, L., & Schulz, E. (2025) metabench—A sparse benchmark for reasoning and knowledge in large language models. *13th Intl Conference on Learning Representations*. (*ICLR*) 2025. ([Pre-Print](#))

**Voudouris, K.**\*, Farrar, B. G., Cheke, L. G., & Halina, M. (2025) Morgan’s Canon and the Associative-Cognitive Distinction Today: A Survey of Practitioners. *Journal of Comparative Psychology* ([Pre-Registration](#); [Pre-Print](#)) (JCP Best Paper Award Finalist).

**Voudouris, K.**\*, Liu, J. D., Siwinska, N., Schellaert, W., & Cheke, L. G. (2024). Investigating Object Permanence in Deep Reinforcement Learning Agents. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 46).

**Voudouris, K.**\*, Cheke, L. G., & Halina, M. (2024). The Future is Computational Comparative Cognition. *Comparative Cognition & Behavior Reviews*, 19.

**Voudouris, K.**\*, Donnelly, N., Rutar, D., Burnell, R., Burden, J., Hernández-Orallo, J., & Cheke, L. G. (2022). Evaluating Object Permanence in Embodied Agents Using the Animal-AI Environment. *EBeM’22: Workshop on AI Evaluation Beyond Metrics*, July 25, 2022, Vienna, Austria.

**Voudouris, K.**\*, Crosby, M., Beyret, B., Hernández-Orallo, J., Shanahan, M., Halina, M., & Cheke, L. (2022). Direct Human-AI Comparison in the Animal-AI Environment. *Frontiers in Psychology* 13:711821.

## RESEARCH PROJECT SUPERVISION

---

<b>PhD Student Assistant Supervisor</b> <i>Helmholtz Munich - 2 students</i>	Jul 2024—Present <i>Munich, Germany</i>
<b>Psychology &amp; Behavioural Sciences Dissertation Supervisor</b> <i>Department of Psychology (Statistics, AI Evaluation, Cognitive Modelling) - 4 Students</i>	Oct 2021—Jun 2023 <i>Cambridge, UK</i>

## SELECTED INVITED TALKS, WORKSHOPS, & SYMPOSIA (‡Co-PRESENTED)

---

- What Are AI Capabilities & How Can We Measure Them?** *Science of Intelligence Seminar, March 20 2025, TÜ Berlin.*
- Generalization in Fine-Tuned Vision Language Models** *Kinds of Kinds of Intelligence, January 31 2025, University of Cambridge.*
- Intuitive Physics in Animals and Machines** *Intuitive Physics Across Species and Development Workshop, December 17-18 2024, MPI for Evolutionary Anthropology, Leipzig, Germany.*
- Measurement Layouts for Capability-oriented AI Evaluation**‡ *AAAI-24, February 20 2024, Vancouver, Canada.*
- Studying Artificial and Natural Intelligence with the Animal-AI Environment** *ARCANI Workshop, March 28 2024, Macquarie University, Sydney, Australia.*
- Oops, I Did It Again: A typology of phonological iterativity**‡ *Penn Linguistics Conference March 20 2021.*

## WORKSHOPS, TUTORIALS, & SYMPOSIA

---

- Co-organiser and co-presenter of the *Measurement Layouts for Capability-oriented AI Evaluation* Lab at AAAI-24 (Vancouver, Feb 2024).
- Co-organiser of the *Predictable AI* Conference in Valencia, Spain (8th March 2023).

## RESEARCH GRANTS

---

**Voudouris, K.**.. Graham, C. R. Scimeca, L, Roll, N. Methods for evaluating short-cut learning in transformer-based ASR systems and its implications. (**£3,120**). *Cambridge Language Sciences Incubator Fund 2021.*

**Voudouris, K.**.. Graham, C. R., and Shinohara, Y. AccentBench: A Framework for evaluating ASR generalisation from native to non-native speech. (**£2,960**). *Cambridge Language Sciences Incubator Fund 2022.*

## TEACHING

---

*Cambridge University Undergraduate Supervision: Cognitive Neuroscience & Experimental Psychology* (Part IB, second-year; 2022-2023), **Philosophy of Cognitive Science** (Part II, third-year; 2022-2023), **Comparative Cognition** (Part II, third-year; 2022-2023), **Phonological Theory** (Part II, second- and third-year; 2020-2022). **Introduction to Linguistics** (2020, 2021; [Syllabus](#)).

## MEDIA & PROFESSIONAL SERVICES

---

Interview with Greg Goth for *Communications of the ACM* (Sept 2022) [Is AI Becoming Sentient? Podcast with Kyle Polich](#) for *DataSkeptic* (May 2024) [Evaluating AI Abilities](#). [CFI Blog](#) (June 2024) [Cognitive Psychology for AI Evaluation](#). Reviewer for *Synthese*, *Linguistics Vanguard*, *ICLR 2025*, *British Journal of Philosophy of Science*, *CogSci 2025*, *CCN 2025*, *AI & Society*.

## TECHNICAL SKILLS

---

**Programming Languages:** Python, Julia, R, PowerShell, SQL, LATEX.

**Libraries and Frameworks:** git, Docker, Lux.jl, MySQL, SLURM, tidyverse, torch, transformers, trl, unsloth.

**Natural Languages:** English (native), Greek (C1), French (B2), German (B1), Turkish (A2), Mandarin Chinese (A2).

## REFERENCES

---

Eric Schulz Postdoc Advisor ([eric.schulz@helmholtz-munich.de](mailto:eric.schulz@helmholtz-munich.de))

Lucy Cheke PhD Supervisor ([lgc23@cam.ac.uk](mailto:lgc23@cam.ac.uk))

Marta Halina PhD Supervisor ([mh801@cam.ac.uk](mailto:mh801@cam.ac.uk))

José Hernández-Orallo Academic Collaborator ([jorallo@upv.es](mailto:jorallo@upv.es))

Nicola Clayton PhD Examiner ([nsc22@cam.ac.uk](mailto:nsc22@cam.ac.uk))

Cameron Buckner PhD Examiner ([cjbuckner@uh.edu](mailto:cjbuckner@uh.edu))