Marko Tešić

Leverhulme Centre for the Future of Intelligence
University of Cambridge
Level 1, 16 Mill Lane, Cambridge, CB2 1SB, UK

✉ mt961@cam.ac.uk
🌐 markotesic.org
in, ☍, ☌

Research Experience

Research Associate at Leverhulme Centre for the Future of Intelligence, University of Cambridge — *July 2023 – Present*
- Working on evaluating AI cognitive capabilities by adapting physiological behavioural tests and the methods from psychometrics.
- Provided scientific advice on the Future of Work and AI progress as part of the OECD expert group.
- Managing a team of researchers working on estimating the impact of AI on the human workforce.

Royal Academy of Engineering UK IC Postdoctoral Research Fellow at Birkbeck, University of London — *Mar. 2021 – Feb. 2023*
- Principal Investigator on a project on understanding the impact of AI explanations on human beliefs and decision-making.
- Worked with and presented research to the members of the UK, the US and Australian intelligences communities.
- Organised a workshop Workshop on the Human Behavioral Aspects of (X)AI, bringing together cognitive scientists and ML researchers from academia, industry and government working on and with (explainable) AI.

Researcher on *The Bayesian Approach to Robust Argumentation Machines* project at MCMP, LMU, Munich & Birkbeck, University of London — *Sep. 2021 – Feb. 2023*
- Automated argument generation and evaluation from Bayesian network models.

Data Study Group (DSG) Principal Investigator at the Alan Turing Institute — *Oct. 2022 – April 2023*
- Led the scoping and execution of a data science challenge with the UK Department for Transport, supervising participants and ensuring high-quality solutions and reporting.

Research Intern at BlackRock, Factor Based Strategies Group — *Oct. 2019 – Mar. 2020*
- Developed causal Bayesian models of investment factors and ESG criteria.

Member of the Translation Team UK on the project *Bayesian Argumentation via Delphi (BARD)* within IARPA at Birkbeck, University of London & UCL — *Oct. 2017 – Nov. 2018*
- Created intelligence gathering-inspired situations
- Built Bayesian network models of these situations
- Fully designed, ran, and analyzed experiments testing people's evidential, causal, and probabilistic reasoning

Education

Ph.D. in Psychology — *2020*
Department of Psychological Sciences, Birkbeck, University of London, UK
  Thesis: *Explanation and Argument*
  Focus: causal-probabilistic reasoning, Bayesian networks, psychology of explanations
  Supervisors: Ulrike Hahn and David Lagnado

M.A. in Logic and Philosophy of Science — *2016*
Munich Center for Mathematical Philosophy, Ludwig Maximilian University, Munich, Germany

B.A. in Philosophy — *2014*
University of Belgrade, Serbia

Selected Work

- **Marko Tešić***, John Burden*, Lorenzo Pacchiardi*, José Hernández-Orallo (2025). Paradigms of AI Evaluation: Mapping Goals, Methodologies and Culture. *Preprint*

- **Marko Tešić***, Lorenzo Pacchiardi*, Lucy Cheke, José Hernández-Orallo (2024). Leaving the barn door open for Clever Hans: Simple features predict LLM benchmark answers. *Preprint*

- Matteo Gabriel Mecattaf*, Ben Slater*, **Marko Tešić**, Jonathan Prunty, Konstantinos Voudouris, Lucy Cheke (2024). A little less conversation, a little more action, please: Investigating the physical common-sense of LLMs in a 3D embodied environment. *Preprint*

- Rakshit S. Trivedi et al. (2024). Melting Pot Contest: Charting the Future of Generalized Cooperative Intelligence. *NeurIPS 2024 Track on Datasets and Benchmarks.*

- **Marko Tešić**, & Janet Slesinski (2024). Benchmark Design Criteria for Mathematical Reasoning in LLMs. *Center for Curriculum Redesign.*

- **Marko Tešić** & Ulrike Hahn (2023). The impact of explanations as communicative acts on belief in a claim: The role of source reliability. *Cognition, 240*(105586).

- Ulrike Hahn & **Marko Tešić** (2023). Argument and Explanation. *Philosophical Transactions of the Royal Society A, 381*(2251). Theme issue on *Cognitive Artificial Intelligence.*

- **Marko Tešić** & Ulrike Hahn (2022). Can counterfactual explanations of AI systems' predictions skew lay users' causal intuitions about the world? If so, can we correct for that? *Patterns, 3*(12).

- **Marko Tešić** (2021). On the transferability of insights from the psychology of explanation to explainable AI. Human Centered AI workshop at NeurIPS 2021.

- **Marko Tešić** & Ulrike Hahn (2021). Explanation in AI systems. In S. Muggleton & N. Chater (Eds.), *Human-Like Machine Intelligence* (pp. 114–136). Oxford University Press.

- **Marko Tešić***, Alice Liefgreen*, & David Lagnado (2020). The propensity interpretation of probability and diagnostic split in explaining away. *Cognitive Psychology, 121.*

- Nicole Cruz, Saoirse Desai, Stephen Dewitt, Ulrike Hahn, David Lagnado, Alice Liefgreen, Kirsty Phillips, Toby Pilditch & **Marko Tešić** (2020). Widening access to Bayesian problem solving. *Frontiers in Psychology, 11*, 660.

- Alice Liefgreen*, **Marko Tešić***, & David Lagnado (2018). Explaining away: Significance of priors, diagnostic reasoning, and structural complexity. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *Proceedings of the 40th Annual Meeting of the Cognitive Science Society* (pp. 2047–2052). Austin, TX: Cognitive Science Society.

* *equal contribution*; Full list of publications available via Google Scholar.

Selected Presentations

Robust evaluation of Generative AI. I presented a tutorial on evaluating large language models at the European Association for Data Science summer school on generative AI.   *June 20, 2024*

AAAI-24 tutorial: Measurement Layouts for Capability-Oriented AI Evaluation. I co-organized a tutorial on Bayesian hierarchical models for learning the AI capability and presented my work on inferring LLM capabilities.   *Feb. 20, 2024*

Supervision and Teaching

Managing a Postdoc and Research Assistant on an Accenture-backed project assessing AI's core cognitive capabilities for workplace tasks. CFI, University of Cambridge   *Jan. 2024 – Present*

Visiting Lecturer for the M.A. courses *Computational Approaches to Mind* and *Fundamental Debates in Cognitive Science*. Department of Psychological Sciences, Birkbeck, University of London   *Jan. 2023 – Apr. 2023*

Visiting Lecturer for the M.A. course *Cognitive and Economic Science of Rational Choice*. Department of Psychology and Economics, City, University of London   *Oct. 2020 – Dec. 2020*

Honors and Awards

The Alan Turing Institute Post-Doctoral Enrichment Award	*July 2022 – Jan. 2023*
The Royal Academy of Engineering UK IC Postdoctoral Research Fellowship (£200,000)	*Mar. 2021 – Feb. 2023*
Ph.D. studentship (IARPA & Birkbeck, University of London)	*2017 – 2020*
Dositeja scholarship for graduate studies	*2017/18; 2015/16; 2014/15*
BAYHOST scholarship for graduate studies	*2015/16; 2014/15*