

Jesse Pepijn Geerts

jessegeerts@gmail.com | +44 7534 845 455 | <https://jessegeerts.github.io> | LinkedIn: jesse-geerts-a4b923bb

About

Computational cognitive neuroscientist and AI researcher with 8+ years of experience in scientific research, specializing in cognitive science, neuroscience, and machine learning. My current research focuses on mechanistic interpretability, specifically the generalization properties of foundation models (LLMs). I have extensive experience designing and running machine learning experiments (PyTorch, JAX), human behavioral experiments, and combining both to adjudicate between different cognitive hypotheses. I have a strong track record of disseminating findings through publications in peer-reviewed journals and conferences ranging from cognitive science (Psychological Review) and neuroscience (Cosyne) to machine learning (ICLR). I am experienced in mentoring PhD and Master's students, teaching undergraduate, Master's, and PhD courses, and lecturing at summer schools.

Experience

Research Fellow | Computational Neuroscience Lab, Imperial | January 2024 – Present

- Conducted research on **AI interpretability**, focusing on circuit mechanisms of generalization ability in transformer models.
- Setting up and leading collaborative project between Imperial, Columbia University and DeepMind on understanding relational reasoning and chain-of-thought prompting in **foundation models**.
- Developed **Deep Reinforcement Learning** model of motor learning explaining striatal and cortical neural manifolds.
- **Mentoring** of MSc thesis on deep self-supervised learning, lecturer in [neural dynamics MSc course](#) at Imperial and [Neuroinformatics course](#) at SWC.
- Invited lecturer on Reinforcement Learning at [NeuroAI summerschool in Amsterdam](#).

Postdoctoral Researcher | Space & Memory Lab, UCL | March 2021 – September 2023

- **Lead and managed** multiple research projects **studying how humans, animals and AI models learn** in decision making and navigation tasks.
- Used **machine learning to study neural time series and behaviour data**, developed custom analysis pipelines and computational models to capture animal behaviour.
- Computational modelling for collaborative project with experimental neuroscience study on dopamine prediction errors, now under 2nd round review at Nature.
- Teaching assistant at UCL's [Neural Computation](#) course.
- Collaborated internationally with different academic and industry partners.
- Author for Dutch Review of Books, writing about AI and neuroscience.
- **Mentored and supervised** 2 PhD students.

PhD Researcher | Space & Memory Lab, UCL | September 2017 – March 2021

- Developed a **reinforcement model explaining contextual decision making** under uncertainty and hippocampal remapping, now published in Psych Review.
- Designed and implemented a novel reinforcement learning framework for **modelling spatial cognition**, now published in PNAS.

- Teaching assistant, developing and presenting course material for new courses in [theoretical neuroscience](#) and [python programming](#) at the SWC and Gatsby Unit.
- **Organised** internationally renowned neuroscience **seminar series** [SWC Annual Symposium](#)
- Volunteer teacher at [BrainCamp Kosovo](#).
- **Published 2 peer-reviewed journal publications**, presented 8+ poster and talks at international conferences such as **ICLR**, **Cosyne** and **CCN**.

Education

Ph.D. in Computational Neuroscience

University College London, 2021

Dissertation: "[Hippocampal predictive maps of an uncertain world](#)"

Advisor: *Neil Burgess* | Examiners: *Athena Akrami* & *Peter Dayan*

M.Sc. in Brain and Mind Sciences

University College London & ENS Paris, 2016

Graduated with Distinction

B.Sc. in Natural Sciences & Neuroscience

University of Amsterdam, 2014

Graduated with Honors, minor in Philosophy & Ethics

Selected Publications

- **Geerts JP, Stachenfeld KLS, Chan SCY & Clopath C.** "*Determinants of emergent acquisition of transitive inference during in-context learning.*" in prep.
- **Greenstreet FMJ*, Geerts JP*, Gallego JA & Clopath C.** "*Learned action embeddings explain striatal and cortical representations during motor learning.*" in prep.
- **Geerts JP, Gershman SJ, Burgess N & Stachenfeld KLS,** "*A probabilistic successor representation for context-dependent prediction.*" *Psychological Review*, 2023. [DOI](#)
- **Greenstreet FMJ, [...] Geerts JP [...] Stephenson-Jones, M** "*Action prediction error: a value-free dopaminergic teaching signal that drives stable learning.*" under 2nd round review at *Nature*. [preprint](#).
- **Geerts JP*, Chersi F*, Stachenfeld KLS & Burgess N.** "*A general model of hippocampal and dorsal striatal learning and decision making.*" *PNAS*, 2020. [DOI](#)
- **Pinotsis, DA, Geerts JP, et al.** "*Linking canonical microcircuits and neuronal activity: Dynamic causal modelling of laminar recordings.*" *NeuroImage*, 2017. [DOI](#)