

Pranav Mahajan

+44 (0) 7826 325 428 | pranav.mahajan@ndcn.ox.ac.uk | Oxford, U.K.
[Personal Website](#) | [LinkedIn](#) | [Google Scholar](#) | [GitHub](#)

Summary

My research bridges AI and neuroscience, combining methods from AI (e.g., reinforcement learning, neural networks, and LLMs) with cognitive and computational neuroscience (e.g. rigorous human-subject experiments). I aim to develop an algorithmic understanding of learning and decision-making in humans and use it to develop human-centric AI applications, with a particular focus on chronic pain.

I have a strong record of publishing in leading journals (eLife, Brain, Neural Computation, PLOS Computational Biology, PAIN and more) and conferences (CCN, COSYNE, RLDM), giving invited talks, establishing new collaborations, mentoring students, and contributing to diversity, equity, and inclusion (DEI) within academic departments.

Education

University of Oxford

10/2021 – 02/2026

DPhil. (Ph.D.) in Computational Neuroscience and Machine Learning

Advisors: Prof. Ben Seymour (neuroscience, biomed eng.) and Prof. Ioannis Havoutis (robotics, ML)

Thesis: Safe learning in humans and machines

BITS Pilani University, India

08/2017 – 07/2021

B.E. in Electronics and Communications Engineering

CGPA: 9.22/10 (Distinction)

Merit Scholarship, BITS Pilani University (2018)

Experience

Postdoctoral Researcher, University of Oxford

02/2026 – now

- Improving patient-clinician interaction by developing a clinician-copilot using hybrid, knowledge graph-retrieval augmented generation (KG-RAG) systems, with applications to precision medicine.
- Applied for Arthritis UK Career Development Fellowship (£550k)

Guest Scientist, Max Planck Institute for Biological Cybernetics, Tübingen

07/2025 – now

- Applied concepts from reinforcement learning (POMDPs) to computational neuroscience of homeostasis, with applications to personal health and recovery agents (Host: Prof. Peter Dayan)

Doctoral Researcher, University of Oxford

10/2021 – 02/2026

- Led research on reinforcement learning (RL) for computational neuroscience, neuro-inspired robotics and applied machine learning for digital health.
- Developed algorithmic models of adaptive/maladaptive learning and decision-making using RL (safe exploration, multi-objective RL, Linear MDPs, POMDPs) and biologically plausible deep learning models (PyTorch).
- Developed virtual reality experiments to test these predictions (Unity) and performed hierarchical Bayesian model fitting to human behaviour (PyStan).
- Trained a Franka Panda arm in PyBullet simulations for pick and place tasks using Learning from Demonstrations (LfD), capable of fault detection and reactive fault-tolerant control.

- Developed an easy-to-use video-to-3D pose estimation-based movement tracking tool for longitudinal tracking of patient recovery.

SPAR AI Fellow

09/2025 – 01/2026

- Designed value-dilemma stress tests to quantify the gap between stated and revealed preferences in LLMs, and contributed to the science of evaluations by separating weak and strong preferences.
- Rigorously evaluated how this gap varied with model capabilities and whether we can align LLM's revealed preferences to their stated preferences by prompt engineering.

Deep Learning Intern, NVIDIA, Bangalore

04/2020 - 07/2020

- Addressed performance scalability issues with synchronous stochastic gradient descent (SGD) in multi-GPU deep learning by implementing asynchronous SGD for data parallel training in PyTorch.

Skills

Project management, experiment design and data collection (Online: Prolific, Pavlovia; Lab-based: VR)
 Machine learning (Python, SciKit, PyTorch, Tensorflow), Model-fitting (SciPy, PyStan),
 LLMs (OpenRouter), Robotics (PyBullet), Scientific Writing (Overleaf).
 Cloud computing (GCP, Linux shell, Docker, Nvidia Apex, Nsight)
 FSL and fMRI analysis (MRI Graduate Programme at Oxford Centre for Integrative Neuroimaging)

Open-Source Software Development

HOI Toolbox (Python, Google Summer of Code), LatentDiffEq.jl (Julia), SyncBox (MATLAB) 2020-2021

Talks

Invited talks

EPIONE Retreat (Oxford, Cambridge, UCL collaboration)	Jun 2026
Prof. Flavia Mancini and Prof. Mate Lengyel labs, CBL, Cambridge	Nov 2025
Prof. Chris Summerfield's lab, University of Oxford	Sept 2025
Neuromod+ Ethics in Neurotechnology Workshop at Imperial College	April 2025
1st Oxford Health BRC Pain Conference	March 2024
Prof. Boris Gutkin lab, ENS Paris	Sept 2023
Prof. Guangyu Robert Yang lab, MIT	Oct 2022

Contributed talks

`Quantitative Movement Testing (QMT)', Podium Institute Conference	Sept 2024
`Quantitative Aversive Cognitive Testing (QACT)', Neuromatch Conference 4.0	Dec 2022

Grants and Awards

Oxford BRC Health Pain Pump Priming (P3) Award. (£3000 grant)	2024
---	------

Teaching Assistant Experience

Co-organiser and tutor for the Oxford Virtual Reality for Pain workshop	2022
Computational Neuroscience course and Deep Learning courses at the Neuromatch Academy	2021
Neuroscience and AI, Computer Architecture, and Nonlinear Dynamics and Chaos at BITS Pilani	2020

Research Mentorship

Vardhan Palod (BITS Pilani, now MS at Arizona State University), Tianjin Ed Li (Oxford, now PhD at Yale).

Peer Reviewing

Supervised co-reviewing (with Prof. Ben Seymour): Nature Neuroscience, Neuron.

Ad-hoc reviewing: Nature Scientific Reports, Cognitive Computational Neuroscience (CCN) conference.

Diversity, Equity and Inclusion

Developed a network analysis tool for the Athena SWAN award (measuring inclusivity) and served as Departmental Welfare Representative (2023-2025).

List of Publications

Including full-length journal and conference articles & non-archival extended abstracts (denoted by †).

Computational Cognitive Neuroscience

Safe reinforcement learning and pain neuroscience

- Entropy regularised reinforcement learning reconciles aversive and action prediction errors in the tail of the striatum
Pranav Mahajan, Ben Seymour
In preparation
- Composing the value signal for dopamine-mediated learning
Pranav Mahajan, Ben Seymour
Reinforcement Learning and Decision Making Conference (RLDM 2025†) [[poster](#)]
bioRxiv [[doi](#)]
- Balancing safety and efficiency in human decision making
Pranav Mahajan, Shuangyi Tong, Sang Wan Lee, Ben Seymour
eLife (2025) (reviewed preprint: "important", "convincing") [[doi](#)] [[eLife digest](#)]
& *Computational and Systems Neuroscience Conference (COSYNE 2022†)* [[poster](#)]
- Homeostasis after injury: How intertwined inference and control underpin post-injury pain and behaviour
Pranav Mahajan, Peter Dayan, Ben Seymour
PLOS Computational Biology (2025) [[doi](#)]
& *Cognitive Computational Neuroscience conference (CCN 2025†)* [[poster](#)]
- Neural Associative Skill Memories for safer robotics and modelling human sensorimotor repertoires
Pranav Mahajan, Mufeng Tang, T. Ed Li, Ioannis Havoutis, Ben Seymour
Neural Computation (2025) [[doi](#)]
& *International Workshop on Active Inference (IWAI 2024, Spotlight†)* [[poster](#)]
- Natural and Engineered Safe Reinforcement Learning
Pranav Mahajan, Ben Seymour
Review paper, under review at Communications Psychology
- Forward and reverse engineering the pain system: from computational neuroscience to neuro-engineering

Pranav Mahajan, Ben Seymour
PAIN (2025) [[doi](#)] [*review paper*]

- Enhanced behavioural and neural sensitivity to punishments in chronic pain and fatigue
Flavia Mancini, Pranav Mahajan, Anna á V Guttesen, ..., Ben Seymour
Brain (2024) [[doi](#)]

Earlier work on computational models of addiction and biological neural networks

- Discounting and drug seeking in biological hierarchical reinforcement learning
Vardhan Palod, Pranav Mahajan, Veeky Baths, Boris Gutkin
Cognitive Computational Neuroscience Conference (CCN 2025, full paper) [[openreview](#)][[poster](#)]
- Doing what's not wanted: Conflict in incentives and misallocation of behavioural control can lead to drug-seeking despite adverse outcomes
Pranav Mahajan, Veeky Baths, Boris Gutkin
Addiction Neuroscience (2023) [[doi](#)]
- Quantifying synchronisation in a biologically inspired neural network
Pranav Mahajan, Advait Rane, Swapna Sasi, Basabdatta Sen Bhattacharya
IEEE-International Joint Conference on Neural Networks (IJCNN 2021, full paper) [[doi](#)][[pdf](#)]

AI for Healthcare

- Quantitative Movement Testing (QMT): Measuring Chronic Pain Patient Movements from a Single Smartphone Video
Pranav Mahajan, Amanda Wall, Eleonora Maria Camerone, Julie Stebbins, Eoin Kelleher, Shuangyi Tong, Annina Schmid, Katja Wieche, Anushka Soni, Ben Seymour
Annual Podium Conference on Sports Medicine & Technology (2024†) [[poster](#)]
& *arXiv* [[doi](#)] [*submitted*]
- Acoustic and language-based deep learning approaches for Alzheimer's dementia detection from spontaneous speech
Pranav Mahajan, Veeky Baths
Frontiers in Aging Neuroscience (2021) [[doi](#)]

AI Safety

- Mind the Gap: Measuring and mitigating the stated-revealed preference gap in LLMs
Alexander Andonov, Pranav Mahajan, Ihor Kendiukhov, Hussain Syed, Lydia Nottingham
ACL 2026 Eval Eval Workshop and 3rd Technical AI Safety Conference (TAIS 2026†) [[arXiv](#)]