William Dorrell

06/2018 BA, Physics, 1st Emmanuel College, Cambridge

RESEARCH

PhD, Gatsby Unit, w/ T. Behrens, P. Latham, J. Whittington

09/2020 – 12/2025

Developed normative theories of neural representations, specifically:
1. Theory of Grid cells as optimal structural encoding of 2D space
2. Prefrontal cortex representations as optimal encoding of structured sequences
3. Theory of task features that make optimal neurons mixed-selective or modular.
Each theory predicts unexplained aspects of single neuron responses, makes predictions, and is in the process of being experimentally tested.

Inductive Bias Developed meta-learning tools, and used them to understand connectomic patterns in cerebellar-like networks via their inductive bias.

Experimental Collaborations:
Designed, built, and helped run experiments (including spike sorting and analysing grid data) to test each theory and further to investigate grid cells in curvy 2D space. Modified a point process method to find procedural replay in Dorsolateral Striatum.

Bogue Fellowship @ Stanford University

01/2023 – 03/2023

Machine Learning rotation w/ Chelsea Finn & Kyle Hsu:
Developed disentangling methods from biological perspective.

RA @ Okinawa Institute of Science and Technology

02/2020 – 08/2020

w/ Erik de Schutter & Sergio Veduzco-Flores:
Designed a biologically plausible hierarchical reinforcement learning agent.

Herchel Smith Scholar @ Harvard University

04/2019 – 12/2019

w/ Cengiz Pehlevan: Modelling & analysis demonstrated structured connections in olfactory cortex using Julien Grimaud & Venkatesh Murthy's data.

08/2018 – 03/2019

w/ J. Hoffman & H. Pirie: replicated van der Waals forces in acoustic metamaterials

AWARDS

2024/5 UCL-Wits Grants - £4,700 for a collaborative visit to University of Witswatersrand
2022/23 Bogue Fellowship - £5,350 for a research stay at Stanford
2018/19 Herchel Smith Scholarship - $80,000 to attend Harvard for a year
2016 + 2017 Davies Senior Scholarship & Mainhood Prize – for university exam performance
2015 British Chemistry Olympiad Roentgenium Award – highest performance

Teaching

Jan '25, Jan '26	Teacher on Imbizo Computational Neuroscience Summer school
Summer '23, '24, '25	Developed and taught 5-week linear algebra course to biologists.
09/2021 – 09/2022	Mentored undergraduate through Simons program, Maria Yuffa, for a year, teaching Neuro and ML, and pursuing research that was published at ICML.
09/2021 – 04/2022	Teaching Fellow: ML and Theoretical & Systems Neuroscience for PhD students Crafted new tutorials to give prerequisites to many backgrounds (cell bio to maths)
09/2019 – 12/2019	Teaching Fellow, Applied Maths: Neural Computation for 20 graduate students
10/2016 – 05/2017	Volunteer teacher in local Cambridge School for GCSE Science

Misc.

01/2020	Attendant, Imbizo Computational Neuroscience Summer School.
Contributor	Minor contribution to openly available Point Process Model of Neural Sequences
Languages	English (native), French (B2)
Reviewer	Neurips ('22, '23, '24), Neurocomputing ('22), Current Biology ('24), eLife ('24)

PUBLICATIONS

Normative Theories of Neural Representations

"A Theory of Prefrontal Sequential Working Memory Representations"
W Dorrell, P Latham, T Behrens, J Whittington, (in preparation)
"Convex Efficient Coding"
W Dorrell, P Latham, J Whittington (in review) 2025.
"Range, not Independence, Drives Modularity in Biologically Inspired Representations"
W Dorrell, K Hsu, L Hollingsworth, J Lee, J Wu, C Finn, P Latham, T Behrens, J Whittington, Arxiv, ICLR, '25
"Actionable Neural Representations: Grid Cells from Minimal Constraints",
W Dorrell, P Latham, T Behrens, J. Whittington, Arxiv, ICLR (2023).
"Disentangling with Biological Constraints: A Theory of Functional Cell Types",
J Whittington, W Dorrell, S Ganguli, T Behrens, Arxiv, ICLR (2023). *Honourable Mention for Best Paper!*
"On prefrontal working memory and hippocampal episodic memory: Unifying memories stored in weights and activity slots." J Whittington, W Dorrell, T Behrens, S Ganguli, M El-Gaby, BioRxiv, Neuron, (2024).

Other Theoretical Neuroscience

"Meta-Learning the Inductive Biases of Simple Neural Circuits",
W Dorrell, M Yuffa, P Latham, Arxiv, ICML, (2023).
"A Differential Hebbian Framework for Biologically-Plausible Motor Control".
S Verduzco-Flores, W Dorrell, E De Schutter, Arxiv, Neural Networks, (2022).

Experimental Neuroscience/Machine Learning Collaborations

"Replay of procedural experience is independent of the hippocampus"
E. Thompson et al., BioRxiv, in review, (2024)
"A Cellular Basis for Mapping Behavioural Structure"
M. El-Gaby et al., BioRxiv, Nature, (2023)
"Bilateral Alignment of receptive fields in the olfactory cortex points to non-random connectivity",
J Grimaud, W Dorrell, C Pehlevan, V Murthy, BioRxiv, eNeuro, (2020).
"Disentangling via Latent Quantization",
K. Hsu, W. Dorrell, J. Whittington, J. Wu, C. Finn, Arxiv, Neurips (2023).