

Paul de Font-Reaulx

CONTACT

✉ paul.defontreaulx@gmail.com
🌐 pauldfr.com

RESEARCH

SPECIALIZATION Moral Psychology, Philosophy of Cognitive Science, Philosophy of AI

COMPETENCE Logic, Decision Theory, PPE

EDUCATION

PhD in Philosophy, University of Michigan, Ann Arbor, 2020-
Committee: Chandra Sripada (Chair), Peter Railton, James Joyce, and Rick Lewis
Certificate in Cognitive Science
Department Visitor, New York University, 2024-2026

BPhil (Master) in Philosophy, Trinity College, University of Oxford, 2018-2020
Grade: Distinction
Clarendon Scholar (Full funding awarded ~1% of graduate students)

BA in Philosophy, Politics, and Economics, Magdalen College, University of Oxford,
2014-2017
Grade: First Class Degree
Anscombe Thesis Prize (Best philosophy thesis of the year)

PUBLICATIONS

“Motivation and Valence” (with Chandra Sripada), *Philosophia*, Invited
Symposium on Carruthers’s *Explaining Our Actions*

“What Makes Discrimination Wrong?”, *Journal of Practical Ethics*, 5(2):105-113. 2017.

IN PROGRESS

R&R “Do Expected Utility Maximizers Have Commitment Issues?”, Minor Revisions at *Philosophy and Phenomenological Research*

Draft “What Stands to Desire as Perceptions Stands to Belief?”
“A Conflation of Valences”

AWARDS

Final Prize Winner, AI Alignment Awards, 2022
\$5,000 for “Generative Theory of Mind and the Value Misgeneralization Problem”

AI Risk Analysis Award Winner, NeurIPS ML Safety Workshop, 2022
~\$2,300 for “Alignment as a Dynamic Process”

GRANTS

Short Term Residency Grant, Institute for Humane Studies, 2024
\$12,000 to visit NYU for an academic year

Early Career Grant, Open Philanthropy, 2023
\$44,770 to buy out teaching for 2 semesters for the project “Alignment and the Structure of Human Values”

TALKS

(Selected)

“Do We Want Technology to Do What We Want?”
IHS Junior Fellowship, Charlotte, 2025

“A Conflation of Valences”
LSE Foundations of Animal Sentience Working Group, London (Online), 2025
NYU Mind, Ethics, and Policy Summit, New York, 2025

“What Stands to Desire as Perception Stands to Belief?”
NYU Mind Seminar Group, New York, 2024
NYU Graduate Work in Progress Group, New York, 2024
Michigan Candidacy Seminar, Michigan (Online), 2024

“The Paradox of Self-Criticism”
IHS Junior Fellowship, Raleigh, 2024

“Machine Theory of Mind and the Structure of Human Values”
NeurIPS MP2 Workshop, New Orleans, 2023 (Declined due to illness)

“Reinforcement Learning as a Model of Human Evaluative Cognition”
Southern Society for Philosophy and Psychology, Louisville, 2023

“Penelope and the Drinks”
Central APA, Denver, 2023
Second Lake Como Summer School on Economic Behaviours, Como, 2022
Filosofidagarna, Lund, 2022
Rocky Mountains Philosophy Conference, Boulder (Online), 2021
Mark L. Shapiro Graduate Philosophy Conference, Brown (Online), 2021

“Alignment as a Dynamic Process”
NeurIPS ML Safety Workshop, Online poster, 2022

“Why Do We Spontaneously Cooperate?”
PPE Society Sixth Annual Meeting, New Orleans, 2022

“Is Reward Enough for Humans?”
The Weinberg Institute Cognitive Science Seminar Series, Michigan, 2022

SCHOLARSHIPS

Junior Fellowship, Institute for Humane Studies, 2025 (\$6,000)
Junior Fellowship, Institute for Humane Studies, 2024 (\$6,000)
Adam Smith Fellowship, Mercatus Center, 2023 (\$10,000)

Global Priorities Fellowship Renewal, Forethought Foundation, 2023 (~\$6,300)
 Global Priorities Fellowship, Forethought Foundation, 2022 (~\$6,300)
 Clarendon Scholarship, University of Oxford, 2018 (~\$70,000)
 Erik and Goran Ennerfelt International Scholar, EGE-Fonden, 2018 (~\$14,000)

TEACHING

Primary Introduction to Formal Logic, University of Michigan, Ann Arbor, Summer 2024
Instructor
Teaching Critical Reasoning, University of Michigan, Ann Arbor, Winter 2023
Assistant Critical Reasoning, University of Michigan, Ann Arbor, Fall 2022
 Introduction to Political Economy, University of Michigan, Ann Arbor, Spring 2022
 Minds and Machines, University of Michigan, Ann Arbor, Fall 2021

SERVICE

Academic Service Admissions Committee Graduate Representative, 2023-2024
 Ethics Bowl Coach, 2023-2024
 Philosophers' Annual Co-Editor, 2023
 Graduate Student Working Group Organizer, 2022-2023
 Picnic Organizer, 2021-2022
 Social Chair, 2021-2022
 COMPASS Organizer, 2020-2021
 Founder and President, Project Access, Sweden, 2016-2017

Refereeing *Philosophy Compass, Synthese, Moral Philosophy and Politics*

**NON-ACADEMIC
 EMPLOYMENT** Research Intern, Governance of AI Project, Future of Humanity Institute, University of Oxford, 2018
 Substitute Teacher, Public schools in Gavle, Sweden, 2013-2016
 Intern, Accenture, Sydney, Australia, 2014