

FACULTY OF PHILOSOPHY  
UNIVERSITY OF OXFORD



# PHILOSOPHY LECTURE PROSPECTUS

TRINITY TERM 2023

**NOTES:**

- The normal duration of an event is one hour. Where the class or lecture lasts longer than an hour, the start time and end time will be given.
- By convention, in-person lectures at Oxford begin at 5 minutes past the hour and end at 5 minutes before the hour.
- Unless otherwise specified, the lectures and classes are given for all of weeks 1 to 8.
- Teaching is now taking place in person. You should not expect recordings to be made available on a general basis.
- Every effort is made to ensure that the information contained in this Prospectus is accurate at the start of term, but sometimes errors persist. If you think you have found a mistake, please contact James Knight ([james.knight@philosophy.ox.ac.uk](mailto:james.knight@philosophy.ox.ac.uk)).

# Lectures for the First Public Examination

Students preparing for their First Public Examination (Prelims or Mods) should attend the following lectures this term:

*PPE, Philosophy and Modern Languages, Philosophy and Theology, Psychology and Philosophy: Moral Philosophy, and General Philosophy*

*Mathematics and Philosophy, Physics and Philosophy, Computer Science and Philosophy: Elements of Deductive Logic, and General Philosophy*

*Literae Humaniores*: any listed Prelims/Mods lecture that corresponds to their chosen Philosophy option for Mods

## **Alan Turing on Computability and Intelligence**

Prof Peter Millican – W. 10 – 12 (*weeks 1 to 4*), Hertford College (Ferrar Room)

These lectures, designed for the first year course in Computer Science and Philosophy, start with the background to Alan Turing's 1936 paper "On Computable Numbers", including Hilbert's programme, Gödel's incompleteness theorem, and Cantor's results concerning the countability of infinite sets. They then work in detail through the 1936 paper, using Charles Petzold's book *The Annotated Turing* (which contains the entire paper, together with comprehensive discussion) as a basis. Finally, the last three lectures will turn to Turing's 1950 paper "Computing Machinery and Intelligence", discussing some of the philosophical issues arising from the Turing Test and Searle's Chinese Room thought-experiment.

## **Frege: Foundations of Arithmetic**

Prof James Studd – M. T. 12 (*weeks 1 to 4*), Radcliffe Humanities (Lecture Room)

These are the core lectures for first-year mathematics and philosophy students. We'll consider, among other things, Frege's attack on Mill's empiricism, Frege's views on number ascriptions, the 'Julius Caesar' problem, and Frege's attempt at a logicist reduction of arithmetic to Hume's Principle, and ultimately to his ill-fated theory of extensions.

Set Text: Frege, *Foundations of Arithmetic* (trans. J. L. Austin)

## **The Leibniz-Clarke Correspondence**

Dr Henrique Gomes – Th. 12 (*weeks 1, 2, 4, other weeks TBC*) – Radcliffe Humanities (Lecture Room)

This course will introduce the centuries-old debate about the nature of space and time. One main question will be whether space is absolute or relative; and indeed what are the various meanings of these two words. A key text in this debate is the correspondence between Samuel Clarke---representing the ideas of Isaac Newton---and Gottfried Leibniz. We will start with the background to the debate in the works of Galileo and Descartes. We will then see how both Newton and Leibniz responded to this background; and finally, we will contrast

their arguments, while investigating Leibniz's metaphysical views in more detail. The course is primarily aimed at Physics & Philosophy students, but all are welcome.

# Lectures for the Honour Schools

Lectures listed in this section are **core lectures** for the papers in the Honour Schools: that is, these are lectures intended especially for students taking those papers at Finals. Questions set in Finals papers usually take the content of core lectures into account to some extent. It is therefore very much in your interest if you are a finalist to attend as many relevant core lectures as your schedule permits.

Students should also refer to the section *Other Lectures*, following. Lectures listed there are not official core lectures, but sometimes cover topics of relevance to the Finals papers.

## **104 Philosophy of Mind**

Prof Michael Martin – F. 12, Radcliffe Humanities (Lecture Room)

These seven lectures are concerned with the relation between minds and objects in the world; and more specifically with the notion of acquaintance.

We'll first look at debates about acquaintance, originally at the beginning of the twentieth century; in the revival of the debate in the 1970s; and more recent contributions.

We'll then look at the question whether acquaintance with the self plays a central role in first person thought.

We'll then look at imagination and memory and ask whether acquaintance plays a central role in explaining their commonalities and differences.

NOTE there is no lecture in Week 5

### Reading List

#### Week 1

#### Essential Reading

Russell, Bertrand. "Knowledge by Acquaintance and Knowledge by Description." *Proceedings of the Aristotelian Society* 11 (1910): 108–28. <http://www.jstor.org/stable/4543805>.

#### Further Reading

Gareth Evans, 1982 *The Varieties of Reference*, Oxford: Clarendon Press, Chs. 1 & 2

Russell, Bertrand. "On Denoting." *Mind* 14, no. 56 (1905): 479–93. <http://www.jstor.org/stable/2248381>.

John Hawthorne and David Manley. 2012 *The Reference Book*, Oxford: Clarendon Press, Ch. 1

## Week 2

### Essential Reading

Gareth Evans, 1985. "Understanding Demonstratives". In *Collected Papers*, Oxford: Clarendon Press, 291 – 321

### Further Reading

Gareth Evans, 1982, *The Varieties of Reference*, Ch. 6

## Week 3

### Essential Reading

Mark Sainsbury, 2020 "Varieties of Singularity". In *Singular Thought & Mental Files*, Goodman, Genone, Kroll, edd. Oxford: Clarendon Press

### Further Reading

Robin Jeshion, 2010, 'Singular Thought: Acquaintance, Semantic Instrumentalism & Cognitivism', in *New Essays in Singular Thought*, ed. Jeshion, Oxford: Clarendon Press

Francois Recanati, 2010, "Singular Thought: In Defence of Acquaintance", in *New Essays in Singular Thought*, ed. Jeshion, Oxford: Clarendon Press

## Week 4

### First-Person without Acquaintance

#### Essential Reading

GEM Anscombe, 1975, "The First Person". In *Mind & Language*, Sam Guttenplan, ed., Oxford: Clarendon Press, pp. 45 – 66.

#### Further Reading

Gareth Evans, 1982 *The Varieties of Reference*, Oxford: Clarendon Press, Ch. 7

Rumfitt, Ian. "Frege's Theory of Predication: An Elaboration and Defense, with Some New Applications." *The Philosophical Review* 103, no. 4 (1994): 599–637. <https://doi.org/10.2307/2186099>.

## Week 5

No meeting

## Week 6

### First-Person with Acquaintance

#### Essential Reading

Saul Kripke, 2011, "The First Person", In *Philosophical Troubles: Collected Papers vol 1*, Oxford: Clarendon Press, pp. 292 – 321

#### Further Reading

Gottlob Frege, 1984, "Thoughts", in *Collected Papers on Mathematics, Logic and Philosophy*, McGuinness, ed., Oxford: Blackwell, pp. 351 – 372

Bertrand Russell, 1986, *The Theory of Knowledge Manuscript*, London: Routledge, Ch. III

## Week 7

### Essential Reading

Christopher Peacocke, 1985, 'Imagination, Experience, and Possibility: A Berkeleian View Defended' in *Essays on Berkeley: A Tercentennial Celebration*, Foster and Robinson, eds., Oxford: Clarendon Press, pp. 19 – 36

### Further Reading

Bernard Williams, 1973, "Imagination and the Self", in *Problems of the Self*, Cambridge: Cambridge University Press, pp. 26 – 45

## Week 8

Soteriou, Matthew. "The Epistemological Role of Episodic Recollection." *Philosophy and Phenomenological Research* 77, no. 2 (2008): 472–92.  
<http://www.istor.org/stable/40041244>.

### Further Reading

John Campbell, 1997 *European Journal of Philosophy* 5:2 ISSN 0966–8373 pp. 105–118

### **108 Philosophy of Logic and Language**

Prof Natalia Waight Hickman – Th.10, Radcliffe Humanities (Lecture Room)

These lectures will focus on the philosophy of language. They will cover some of the most vexed problems for semantic analysis, including vagueness, non-substitutivity, and reference failure; and also, foundational questions about the general nature of linguistic meaning, its relation to speakers and to the world.

### **109 Aesthetics and the Philosophy of Criticism**

Prof Louise Hanson – T. 2 – 4 (*weeks 1 to 4*), online (link to be posted on Canvas)

### **113 Post-Kantian Philosophy: Sartre**

Prof Joseph Schear – W. 12 (*except week 5*), Christ Church (Lecture Room 2)

This course of lectures is primarily devoted to Jean-Paul Sartre's 1943 book, *Being and Nothingness: An Essay on Phenomenological Ontology*. (Sartre's shorter 1937 work 'The Transcendence of the Ego: A Sketch for a Phenomenological Description' is to be discussed as well, for this essay outlines some of the key arguments subsequently developed in the 1943 book.) The intended audience are undergraduates taking the Post-Kantian philosophy paper but anyone is welcome to attend. There will be no lecture in week 5. Please use the recent translation by Sarah Richmond. Provisional schedule:

1. Introduction: Phenomenology, Ontology, 'Existentialism'
2. The Transcendence of the Ego
3. The Cogito and the Problem of Nothingness
4. Bad faith
5. No lecture
6. Being-for-Others I
7. Being-for-Others II
8. Freedom



## Other Lectures (suitable for all audiences)

### **The 2023 John Locke Lectures: *Recognizing Knowledge: Intuitive and Reflective Epistemology***

**Prof Jennifer Nagel** – M. W. 5 – 7 (*weeks 6 to 8*), Keble College (H B Allen Lecture Theatre)

*The Faculty is delighted to welcome the esteemed 2023 John Locke lecturer, Prof Jennifer Nagel. Prof Nagel is a Professor of Philosophy at the University of Toronto.*

Humans have a remarkable capacity to track what others do and do not know. This capacity guides us in everyday social navigation, for example as we switch between the roles of telling and asking in conversation. It also provides raw data to epistemology, in the form of intuitive judgments about possible cases of knowledge. Over the years, philosophers, psychologists, and sociologists have discovered a variety of cross-culturally robust patterns of epistemic intuition, patterns that are attractively systematic, but often disturbingly paradoxical. This series of talks examines the natural origins and functions of our capacity to detect knowledge, in search of a better analysis of the data guiding epistemology, and ultimately a clearer view of knowledge itself.

#### **Lectures**

Lecture 1: The strange divergence between intuitive and reflective knowledge attribution

Lecture 2: Knowledge and surprise

Lecture 3: Mental state recognition among animals

Lecture 4: Distinctively human mindreading

Lecture 5: Knowledge possession and knowledge transmission

Lecture 6: Knowledge detection and the nature of knowledge

### **The 2023 Gareth Evans Memorial Lecture: *TBC***

**Prof John Campbell** (Berkeley) – T. 5 (*week 5*), TBC

### **Gödel's Incompleteness Theorems and their philosophical implications**

Fabian Pregel – Th. 11 (*weeks 1 to 4*), Radcliffe Humanities (Lecture Room)

No more than an introductory logic course will be required to follow this lecture course. The goal is to acquaint students with both the philosophical as well as technical aspects of the theorems. Planned lecture content:

**Lecture 1** Technical background—covering Peano Arithmetic, consistency, completeness, Gödel numbering, Provability Predicate, Hilbert-Bernay conditions

**Lecture 2** Incompleteness Theorems—Diagonal Lemma, First Incompleteness Theorem, Second Incompleteness Theorem

**Lecture 3** Philosophical Implications—Hilbert’s program, Formalism, Deductivism

**Lecture 4** Philosophical implications for Logicism—ambitions of the Logicist program, potential responses to incompleteness and challenges with those responses

*Gödelian incompleteness* is the phenomenon that, for any consistent, recursively enumerable axiom system sufficiently strong to derive certain parts of arithmetic (Robinson Arithmetic), a sentence in the language of arithmetic exists such that neither the sentence nor its negation is formally provable from the axiom system. This phenomenon is of broad philosophical interest across logic and the philosophy of mathematics.

Gödel’s incompleteness theorems are traditionally seen to have ended Hilbert’s program, underpin a major argument against Deductivism and impose limits on what Logicism can hope to achieve. More broadly, they are a result that often strikes people as highly surprising.

# Graduate Classes

Graduate classes are, except where otherwise indicated, intended for the Faculty's graduate students. (The BPhil Pro-Seminar is restricted to first-year BPhil students.) Other students may attend Faculty graduate classes, and are welcome, provided they first seek and obtain the permission of the class-giver(s).

## **BPhil Pro-Seminar: History of Philosophy**

Various class-givers and locations – F. 11 – 1

The Pro-seminar introduces students to study, practice, and standards in graduate-level philosophy. Every starting BPhil student will attend four sessions with one class-giver, then change group midway through term for four sessions with another class-giver. Seminars in Trinity Term will cover history of philosophy. Class-givers will contact their groups, specifying readings and confirming the class time, in advance of term.

## **Memory in Hellenistic Philosophy and Late Antiquity**

Prof Luca Castagnoli and Prof Tobias Reinhardt – M. 10 – 12, Corpus Christi College (Seminar Room)

In this class we will examine some of the complex ways in which the Platonic and Aristotelian insights into the functioning and role of memory in human cognition and life were received, criticised, transformed and added to in the Hellenistic age and in late antiquity, from the early Epicureans and Stoics (late 4<sup>th</sup> century BCE) to Augustine (3<sup>rd</sup>/4<sup>th</sup> century CE). The texts which we will examine, and the study of memory in this period, raise a number of exegetical and philosophical questions which span a variety of different but interrelated areas, including psychology, epistemology, ethics, and scientific and philosophical method.

There will be opportunities for students to give short presentations in weeks 2 to 8. A set of bibliographies and selected texts will be circulated at the first session in week 1 and made available on Canvas.

Week 1: Introduction: Pre-Hellenistic Philosophers on Memory (mainly Plato and Aristotle)

Week 2: Epicureans on memory

Week 3: Stoics on memory (and their critics)

Week 4: Rationalist, Empiricist and Methodist doctors on the use of memory in the art of medicine

Week 5: Galen on the physiology and pathologies of memory

Week 6: Plotinus on memory

Week 7: Plotinus on whether the stars have memory

Week 8: Augustine on memory, recollection and forgetting

## **Philosophy of Language (MS class)**

Prof Ian Rumfitt – Th. 2 – 4, All Souls College (Hovenden Room)

I have a draft book—provisionally entitled *If Truth Be Told*—which seeks to situate semantic theory within a broader account of speech acts. In this seminar, I shall present the first ten chapters. While aimed chiefly at graduate students, the seminar addresses foundational issues, so it may also be of interest to undergraduates who have studied some philosophy of language. I envisage it as a ‘reading class’: while I will begin each session by outlining the main line of argument, participants will benefit if they have read the prescribed part of the book, which will be available on Orlo.

At least until the end of Trinity Term, I retain copyright of the material, and I ask participants not to quote or circulate any of it without my permission.

### **Plan**

Week One	Thursday 27 April	Foundations
Reading:	Chapter 1 (‘Introduction’) and Chapter 2 (‘An Approach to the Theory of Meaning’)	
Week Two	Thursday 4 May	Relational Speech Acts
Reading:	Chapter 3 (‘The Structure of Speech Acts’) and Chapter 4 (‘Directed Speech’)	
Week Three	Thursday 11 May	Acts of Telling
Reading:	Chapter 5 (‘Telling someone that something is the case’)	
Week Four	Thursday 18 May	Unfolding the Concept of Truth
Reading:	Chapter 6 (‘Truth as Told’)	
Week Five	Thursday 25 May	Semantics for a Simple Language
Reading:	Chapter 7 (‘Internal and External Truth’)	
Week Six	Thursday 1 June	Restricted Quantifiers
Reading:	Chapter 8 (‘Restricted Quantifiers’)	
Week Seven	Thursday 8 June	The Proper Goals of Empirical Semantics
Reading:	Chapter 9 (‘Semantic Theories for Natural Languages’)	
Week Eight	Thursday 15 June	Questions
Reading:	Chapter 10 (‘Questions and the Meaning of Interrogative Sentences’)	

## **Understanding**

Prof Alison Hills and Prof Alex Prescott-Couch – T. 9 – 11, Radcliffe Humanities (Ryle Room)

Please refer to the Canvas site for updates.

## **Epistemology**

Prof Timothy Williamson – W. 2 – 4, Radcliffe Humanities (Lecture Room)

Week 1 (26 April)      Aesthetic and epistemic feelings in theory evaluation  
(to be presented by Mariona Miyata-Sturm)

Reading for week 1:

Bird, Alexander (2020) "How Can Loveliness Be a Guide to Truth? Inference to the Best Explanation and Exemplars" in *The Aesthetics of Science: Beauty, Imagination and Understanding* (eds. Milena Ivanova and Steven French) Abingdon: Routledge.

[http://www.alexanderbird.org/Research/How\\_can\\_loveliness\\_be\\_a\\_guide\\_to\\_truth.pdf](http://www.alexanderbird.org/Research/How_can_loveliness_be_a_guide_to_truth.pdf)

Reading materials for weeks 2-7 will be available on Professor Williamson's Faculty webpage (<https://www.philosophy.ox.ac.uk/people/timothy-williamson>) in advance.

Week 2 (3 May)      Heuristics in philosophy

Week 3      (10 May)      Overfitting in philosophy

Week 4 (17 May)      Epistemological consequences of Frege puzzles

Week 5 (24 May)      Kinds of rationality

Week 6 (31 May)      Internalist and externalist justification of belief and action

Week 7 (7 June)      Acting on knowledge-how

Week 8 (14 June)      Unexceptional moral epistemology

## **Feminist Theory**

Dr Sebastian Bishop – W. 9 – 11, Radcliffe Humanities (Ryle Room)

### Course overview

When philosophers discuss whether a person is living 'autonomously', they are generally interested in whether that person is living a self-governing life. In particular, philosophers often seek to determine whether that person is living a life that involves successfully pursuing values and goals that are authentically their own.

'Relational autonomy' is the name that has been given to a particular kind of feminist reconceptualisation of autonomy. Relational theorists argue that the classic liberal approaches to autonomy (associated with the likes of Harry Frankfurt, John Christman, Isaiah Berlin, John Rawls, as well as Faden and Beaucham) are guilty of being too individualistic. Various it is suggested that liberal approaches to autonomy err in associating autonomy too narrowly with e.g., self-sufficiency, making consensual decisions, freedom from outside interference, value-neutrality, rationality.

This postgraduate course reflects on how, in contradistinction to the liberal approach, relational theorists argue that our understanding of autonomy must incorporate the fundamental insight that we are social beings that find ourselves embedded within important social relations. Relational theorists invite us to reflect on the importance of e.g., our personal relationships (friendships, parent-child relationships, romantic relationships), social institutions (such as education and professional work), our social standing, how important aspects of our identity (such as our race, gender, religion, sexual orientation) are understood by those around us, the societal norms and expectations surrounding these aspects of our identity.

Overall this course aims to immerse students in contemporary, cutting-edge debates around relational autonomy. Students are invited to reflect on the various ways in which our social relations might both promote and hinder our living autonomously. Finally, the course considers a number of challenges to the relational approach to autonomy. Students are called upon to critically reflect on the extent to which relational approaches to autonomy can survive these challenges.

#### Where does relational autonomy fit into the wider feminist literature?

Feminism is first and foremost a political project that aims to liberate women from patriarchal oppression. Modern feminism, however, doesn't just focus on women, but instead considers how people of all identities may be subject to patriarchal oppression. Naturally, this is a wide-ranging, even daunting project. One helpful way of understanding feminism is, therefore, to break the feminist project down into two related objectives.

First, to identify and analyse oppressive patriarchal forces, and to aim to better understand how different oppressive forces function in the world. They reflect on where these oppressive forces come from. In what way do they set-back our interests? Why do these oppressive forces still exist? How are these forces maintained?

Second, having identified particular oppressive forces in the world, feminists ask how we might respond to these forces? How might we fight against these forces? What political and institutional changes might we try and make? Feminists are also interested in the fundamental question of how we should respond, emotionally speaking, to oppression.

Relational theorists are interested in both of these feminist objectives. As we will see during this course, relational theorists suggest that our autonomy can be threatened by e.g., internalised oppression, taking part in highly asymmetrical relationships, poor social support structures, poor institutional support structures, lack of support for our (valid) emotional responses to oppression. Relational approaches to autonomy help to elucidate the dangers of living in highly unequal, patriarchal societies.

Having shed light on some of the threats to our autonomous flourishing, relational theorists then consider what changes we might make to better promote the autonomy of people. Among other suggestions, this course considers: how we might redesign clinical practice and

research, how we might change what kind of support we receive from our government and surrounding institutions, how we might fight internalised oppression, how we might fight socially conditioned biases, whether we should take a more permissive attitude towards

certain kinds of paternalist interference, the importance of making space for and facilitating certain kinds of emotional responses.

### **Schedule and Weekly Readings:**

#### Week 1 – The Liberal Approach to Autonomy

We begin the course by examining the liberal approach to autonomy. In particular we consider the importance of critical reflection and consent. Should we count a person's decision as autonomous, so long as they have critically reflected upon their options? Should we still respect a person's autonomous decision when we know this same decision will undermine their autonomy in the future?

Primary:

- John Christman (1991) "Autonomy and Personal History", *Canadian Journal of Philosophy* 21 (1), pp.1-24.
- Tom Beauchamp (2001) "Informed Consent: Its History, Meaning, and Present Challenges", *Cambridge Quarterly of Healthcare Ethics* 20 (4), pp.515-523.

Secondary:

- Isaiah Berlin (1969). "Two Concepts of Living", in *Four Essays On Liberty* (Oxford: Oxford University Press) pp.118-172.
- Jonathan Pugh (2020). "Informed Consent, Autonomy, and Beliefs", in *Autonomy Rationality, and Contemporary Bioethics* (Oxford: Oxford University Press).
- Martha Nussbaum (1995). "The Feminist Critique of Liberalism", *The 1997 Lindley Lecture*.
- Harry Frankfurt (1971). "Freedom of the Will and the Concept of a Person", *The Journal of Philosophy* 68 (1), pp. 5-20.
- Gerald Dworkin (1988) *The Theory and Practice of Autonomy* (Cambridge: Cambridge University Press).

#### Week 2 – The Relational Critique of Liberal Approaches to Autonomy

This week we consider how relational theorists have critiqued classic, liberal approaches to autonomy. To what extent do liberal approaches fail to take seriously the social nature of our lives? Why do feminists like Natalie Stoljar worry that informed consent is neither sufficient, nor necessary, for autonomous decision-making?

We also begin to consider how relational theorists have sought to improve upon the liberal approach – especially in the context of informed consent and medical ethics.

Primary:

- Natalie Stoljar (2011). "Informed Consent and Relational Conceptions of Autonomy", *Journal of Medical Philosophy* 36 (4), pp.375-384.
- Catriona Mackenzie and Natalie Stoljar (2000). "Introduction: Autonomy refigured" in *Relational Autonomy: Feminist Perspectives on Autonomy, Agency and the Social Self* (New York: Oxford University Press), eds. Catriona Mackenzie and Natalie Stoljar.



Secondary:

- Natalie Stoljar (2018). "Feminist Perspectives on Autonomy", *The Stanford Encyclopedia of Philosophy*, ed. Edward Zalta, <<https://plato.stanford.edu/archives/win2018/entries/feminism-autonomy/>>.
- Susan E Kelly, Edward S Dove, Federica Lucivero, Mavis Machirori, Sandi Dheensa, and Barbara Prainsack (2017). "Beyond Individualism: Is There a Place for Relational Autonomy in Clinical Practice and Research?", *Clinical Ethics* 12 (3), pp.150-165.
- Carlos Gómez-Vírveda and Rafael Amo Usanos (2021). "Relational Autonomy: Lessons from COVID-19 and Twentieth-Century Philosophy", *Medical Health Care and Philosophy* 24, pp.493-550.
- Jennifer Walter and Lane Friedman Ross (2014). "Relational Autonomy: Moving Beyond the Limits of Isolated Individualism", *Pediatrics* 133 (1), pp.16-23.
- Carolyn Ells, Matthew Hunt, and Jane Chambers-Evans (2011). "Relational Autonomy as an Essential Component of Patient-Centered Care", *International Journal of Feminist Approaches to Bioethics* 4 (2), pp.79-101.

Week 3 – Social Relations that Promote Autonomy

This week we reflect on the ways in which our social relations can help us to live autonomously. We consider the ways in which relational theorists have argued that social relations can be instrumentally valuable in promoting autonomy – for instance, how social relations of the right kind can help us to critically reflect on our values, develop self-respect, develop our emotional capacities, and more successfully pursue our goals. We also reflect on the complicated idea that good social relations might form a constitutive part of an autonomous life.

Primary:

- Catriona Mackenzie (2008). "Relational Autonomy, Normative Authority and Perfectionism", *Journal of Social Philosophy* 39, pp.512-533
- Seana Valentine Shiffrin (2014). "A Thinker-Based Approach to Freedom of Speech", in *Speech Matters: On Lying, Morality, and the Law* (Princeton: Princeton University Press).

Secondary:

- Tineke Abma, Barth Oeseburg, Guy Widdershoven, and Marian Verkerk (2009). "The Quality of Caring Relationships", *Psychology Research and Behavior Management* 2, pp.39-45.
- Carolyn McLeod & Emma Ryman (2020). "Trust, Autonomy, and the Fiduciary Relationship", in *Fiduciaries and Trust: Ethics, Politics, Economics, and Law* (Cambridge: Cambridge University Press), eds. Paul Miller & Matthew Harding.
- Andrea Westlund (2012). "Autonomy in Relation", in *Out From the Shadows: Analytical Feminist Contributions to Traditional Philosophy* (Colorado: Westview Press), eds. Sharon Crasnow & Anita Superson.

- Marilyn Friedman (2013). "Relational autonomy and individuality", *The University of Toronto Law Journal* 63, (2), pp. 327-341.
- Carolyn McLeod (2020). "My Relational Autonomy and My Relationship with Susan Sherwin", *International Journal of Feminist Approaches to Bioethics* 13 (2), pp.9-11.

#### Week 4 – Internalised Oppression

One of the most striking (and controversial) claims that relational theorists make, is that even when we are dealing with mature adults who are making (seemingly) free decisions about what goals and values to pursue, certain goals and values are nonetheless incompatible with being autonomous. In this context, relational theorists are especially worried about the prospect of internalised oppression.

Consider, for instance, the case of a highly deferential, self-abnegating housewife who takes their main goal in life to be making their husband happy even at their own expense. To what extent might the housewife's goals be the product of internalised oppression? How might internalised oppression be the product of bad social relations? To what extent, if any, does internalised oppression pose a threat to autonomy?

Primary:

- Marina Oshana (1998). "Personal Autonomy and Society", *Journal of Social Philosophy* 29 (1), pp.81-102.
- Natalie Stoljar "Autonomy and the Feminist Intuition", in *Relational Autonomy: Feminist Perspectives on Autonomy, Agency and the Social Self* (New York: Oxford University Press), eds. Catriona Mackenzie and Natalie Stoljar.

Secondary (on internalised oppression)

- Marina Oshana (2014). "Is Social-Relational Autonomy a Plausible Ideal?", in *Personal Autonomy and Social Oppression Personal Autonomy and Social Oppression* (New York: Routledge).
- Ji-Young Lee (2022). "Normative competence, autonomy, and oppression", *Feminist Philosophy Quarterly* 8 (1), pp.1-17.
- Rebekah Johnston (2017). "Personal Autonomy, Social Identity, and Oppressive Social Contexts", *Hypatia* 32, pp.312-328.
- Laura Richards talking about coercive control and domestic abuse: <https://www.thelaurarichards.com/resources/coercivecontrol>

Secondary (on the way we inherit our concepts and words from others)

- Sally Haslanger (2017). "Culture and Critique", *Aristotelian Society Supplementary* 91 (1), pp.149-173.
- Filipa Melo Lopes (2019). "Perpetuating the Patriarchy: Misogyny and (Post-)Feminist Backlash", *Philosophical Studies* 176 (9), pp.2517-2538.
- Quill Kukla (2020). "Misogyny and Ideological Logic", *Philosophy and Phenomenological Research* 101 (1), pp.230-235.
- Sandra Lee Bartky (1990). "Foucault, Femininity, and the Modernization of Patriarchal Power", in *Femininity and Domination* (New York: Routledge).

- Sally Haslanger (2007). "But Mom, Crop-Tops Are Cute!" Social Knowledge, Social Structure and Ideology Critique", *Philosophical Issues* 17 (1), pp.70–91.
- Jennifer Nedelsky (1989). "Reconceiving Autonomy: Sources, Thoughts and Possibilities", *Yale Journal of Law and Feminism* 1, pp.7-36.

### Week 5 –Vulnerability

This week we reflect on the extent to which persons are vulnerable. Are certain types of vulnerability unavoidable? Does the vulnerability of citizens, as Mackenzie supposes, impose obligations on the government? What kind of obligations does she have in mind?

We also consider Kittay's work on vulnerability. We reflect on the extent to which even highly asymmetrical relations may be a source of autonomy. If highly asymmetrical relations may be sources of autonomy, this may leave relational theorists with a tricky dilemma: how do we draw a distinction between the (autonomous) self-sacrificial labour of a loving mother for her vulnerable child, and the (autonomy undermining) self-sacrificial labour of a highly deferential housewife towards her husband. How should we understand the differences between these cases?

Primary:

- Catriona Mackenzie (2013). "The Importance of Relational Autonomy and Capabilities for an Ethics of Vulnerability", in *Vulnerability: New Essays in Ethics and Feminist Philosophy* (Oxford: Oxford University Press), eds. Catriona Mackenzie, Wendy Rogers, and Susan Dodds.
- Eva Feder Kittay (2011). "The Ethics of Care, Dependence, and Disability", *Ratio Juris* 24 (1), pp.49-58.

Secondary:

- Pamela Sue Anderson (2003). "Autonomy, Vulnerability and Gender", *Feminist Theory* 4 (2), pp.149-164.
- Marilyn Friedman (2013). "Relational autonomy and individuality", *The University of Toronto Law Journal* 63, (2), pp. 327-341.
- Martha Fineman (2008). "The Vulnerable Subject: Anchoring Equality in the Human Condition" *Yale Journal of Law and Feminism* 20 (1), pp.1-23.
- Martha Nussbaum (2000) *Women and Human Development: The Capabilities Approach* (Cambridge: Cambridge University Press).

### Week 6 – Emotions (especially emotional responses to oppression)

What feelings are appropriate in response to e.g., studying oppression, personally experiencing oppression, seeing others experiencing oppression? One natural response is to feel angry. But what if it turns out that our anger is unproductive (or even harmful to our political goals)? Would we, nonetheless, have good reason to feel angry? This week we consider the value of anger, and how it relates to relational autonomy.

We also consider Betzler's suggestion that empathy between agents is both instrumentally valuable (insofar as it e.g., increases self-confidence and self-trust) and constitutively valuable (insofar as it e.g., involves valuable vindication and intimacy).



Primary:

- Amia Srinivasan (2018). "The Aptness of Anger", *Journal of Political Philosophy* 26 (2), pp.123-144.

Secondary:

- Monika Betzler (2019). "The Relational Value of Empathy", *International Journal of Philosophical Studies*, 27 (2), pp.136-161
- Laura Luz Silva (2021). "The Efficacy of Anger: Recognition and Retribution", in *The Politics of Emotional Shockwaves* (London: Palgrave Macmillan), eds. Ana Falcato & Sara Graça da Silva, pp.27-55.
- Regina Rini (2018). "How to Take Offense: Responding to Microaggression", *Journal of the American Philosophical Association*, 4 (3), pp.332-351.
- Jolanda van Dijke, Inge van Nistelrooij, Pien Bos, Joachim Duyndam (2020). "Towards a Relational Conceptualization of Empathy", *Nursing Philosophy* 21 (3), e12297

### Week 7 – Feminist Critiques of the Relational Approach

Some feminists register serious worries with the relational approach. This week we consider whether it is objectionable to dismiss well thought-out goals as being 'the result of internalised oppression and therefore lacking in autonomy'. Does the relational approach risk disrespecting (the choices of) those labelled as 'oppressed'? Is the relational approach too demanding? How seriously should we take Khader's worry that relational accounts risk 're-marginalising' those they seek to help?

Primary:

- Serene Khader (2020). "The Feminist Case Against Relational Autonomy", *Journal of Moral Philosophy* 17 (5), pp.499-526.
- Serene Khader (2016). "Beyond Autonomy Fetishism: Affiliation with Autonomy in Women's Empowerment", *Journal of Human Development and Capabilities* 17 (1), pp.125-139.

Secondary:

- John Christman (2014). "Relational Autonomy and the Social Dynamics of Paternalism", *Ethical Theory and Moral Practice* 17 (3), pp.369-382.
- Ranjoo Seodu Herr (2018). "Islamist Women's Agency and Relational Autonomy", *Hypatia* 33 (2), pp.195-215.
- Quill Kukla (2021). "A Nonideal Theory of Sexual Consent", *Ethics* 131 (2), pp.270-292.

### Week 8: Paternalism Defended?

Maternalism describes a special type of intervention that is distinct from standard paternalism. Maternalism refers to paternalist interferences that take place within a relationship of mutual trust (the relational condition), where the paternalizer knows the paternalizee well enough to judge that the intervention will support the paternalizee's autonomy (the epistemic condition).

To what extent should we embrace paternalism so long as it will promote autonomy in the long-term? We also reflect more on whether relational theorists are disrespectful towards

those they label as lacking in autonomy. Is it necessarily insulting to suggest that people sometimes make bad decisions, or that complex social forces can pose a threat to our autonomy? To what extent can we take people's choices seriously, while nonetheless framing them as flawed and lacking in autonomy?

Primary:

- Laura Specker Sullivan and Fay Niker (2018). "Relational Autonomy, Paternalism, and Maternalism" *Ethical Theory and Moral Practice*, 21 (3), pp. 649–667.
- Sarah Conly (2013). "Coercive Paternalism in Health Care: Against Freedom of Choice", *Public Health Ethics* 6 (3), pp.241-245.

Secondary:

- Rosa Terlazzo (2016). "Conceptualizing Adaptive Preferences Respectfully: An Indirectly Substantive Account", *Journal of Political Philosophy* 24 (2), pp.206-226.
- Rosa Terlazzo (2020). "(When) Do Victims Have Duties to Resist Oppression?", *Social Theory and Practice* 46 (2), pp.391-416.
- Jennifer Blumenthal-Barby "On the Concept and Measure of Voluntariness: Insights From Behavioral Economics and Cognitive Science", *American Journal of Bioethics* 11 (8), pp.25-26.
- Jennifer Blumenthal-Barby "Between Reason and Coercion: Ethically Permissible Influence in Health Care and Health Policy Contexts", *Kennedy Institute of Ethics Journal* 22 (4), pp.345-366.
- Eun-Jung Katherine Kim (2012). "On the Burka Ban", *Public Affairs Quarterly* 26 (4), pp.293-312.

#### Bonus Topic – Gender Bias in Philosophy:

During week 4 we considered how things like our language, our concepts, and our conceptual frameworks are shaped by our surrounding social relations. Through looking at Haslanger, in particular, we considered how 'ideology' functions to maintain oppression, and how might we go about resisting oppressive 'social schemas'. Building on this, we might also consider the various way in which our surrounding 'ideology' contributes to the gender bias we find in philosophy. The readings also reflect on how might we go about fighting against this bias.

Primary:

- Cheshire Calhoun (2009). "The Undergraduate Pipeline Problem", *Hypatia* 24 (2), pp.216-223.
- Jennifer Saul (2013). "Scepticism and Implicit Bias", *Disputatio* 5 (37), pp.243-263.
- Kirstie Dotson (2012). "How is This Paper Philosophy", *Comparative Philosophy* 3 (1), pp.3-29.
- Molly Paxton, Carrie Figdor, and Valerie Tiberius (2012). "Quantifying the Gender Gap: An Empirical Study of the Underrepresentation of Women in Philosophy", *Hypatia* 27 (4), pp.949-957.

## **Perceptual objectivity**

Prof Michael Martin – W. 11 – 1 (*not on in weeks 1 and 5 – extra session in week 9*),  
Radcliffe Humanities (Ryle Room)

In *Individuals*, and later in *Bounds of Sense*, PF Strawson introduced a debate about 'objective experience' (the term is coined by Bennett reviewing Strawson, not Strawson himself). Strawson contrasts objective experience with solipsistic consciousness and with the view of experience favoured by the 'pure sense-datum theory'.

This term will be looking at some of the historical antecedents of Strawson's puzzle within the empiricist tradition; Strawson's own response; and Evans's criticisms; before looking at Burge's attacks on both Strawson and Evans.

NOTE: this seminar will start in WEEK TWO. There will be no meeting in Week 5; there will be an additional meeting in Week 9.

### **Week 1**

No meeting

### **Week 2**

Essential Reading

George Berkeley, *First Dialogue between Hylas and Philonous*  
, *Treatise on the Principles of Human Knowledge* secs 1 – 24

Further Reading

John Campbell & Quassim Cassam, 2014, *Berkeley's Puzzle*, Oxford: Clarendon Press, Chs. 1 & 5

PF Strawson, *Individuals*, 1957, London: Methuen, Ch. 2

Jonathan Bennett, 'Strawson on Kant', *The Philosophical Review*, vol 77, No 3 (Jul 1968) pp. 340 – 349

### **Week 3**

Essential Reading

David Hume, *A Treatise of Human Nature*, BkI Pt IV, sec ii  
*An Enquiry concerning Human Understanding*, sec. xii

Further Reading

Donald Ainslie, *Hume's True Scepticism*, Ch. 3

Annemarie Butler, 'Hume on Believing the Vulgar Fiction of Continued Existence', *History of Philosophy Quarterly*, Vol. 27, No. 3 (JULY 2010), pp. 237-254

Donald L. M. Baxter, 'Hume on Steadfast Objects and Time', *Hume Studies* Volume XXVII, Number 1 (April, 2001) 129-148

#### **Week 4**

##### Essential Reading

GE Moore, 'Some Judgments of Perception', *Proceedings of the Aristotelian Society* 19:1--29 (1918 – 19)

, 'Visual Sense-data', 1957 *British Philosophy in Mid-Century*. C. Mace, ed. . London, George Allen & Unwin.

##### Further Reading

PF Snowdon, 'G.E. Moore on Sense-Data and Perception', in S. Nuccetelli & G Seay eds, *Themes from GE Moore: New Essays in Epistemology and Ethics*, Oxford: Clarendon Press  
Thomas Baldwin, *G.E. Moore*, 1990 London:Routledge, Ch. VIII

#### **Week 6**

##### Essential Reading

Bertrand Russell, 1986 (1913) *The Theory of Knowledge Manuscript*, London: Routledge Chs. II & III

, 1940 *An Inquiry into Meaning & Truth*, Ch. VII

##### Further Reading

Thomas Nagel, 'What is it Like to be a Bat?', *The Philosophical Review*, Vol. 83, No. 4 (Oct., 1974), pp. 435-450

#### **Week 7**

##### Essential Reading

PF Strawson, *Individuals*, 1957, London: Routledge, Ch. 2

, *The Bounds of Sense*, 1966, London: Methuen, Pt One sec 2, pp. 24 – 33; Pt Two II secs 5 – 9 pp. 89 – 117

##### Further Reading

Richard Rorty, 'Strawson's Objectivity Argument', *The Review of Metaphysics*, Vol 24, No 2 (Dec 1970) pp. 207 – 244

#### **Week 8**

##### Essential Reading

Gareth Evans, 'Things without the Mind', in van Straaten, ed., *Philosophical Subjects*, Oxford: Clarendon Press, 1980; reprinted in *Collected Papers*, Oxford: Clarendon Press, 1985.

##### Further Reading

PF Strawson, 'Reply to Evans', in van Straaten, ed., *Philosophical Subjects*, Oxford: Clarendon Press, 1980, pp.273 – 282

Gareth Evans, 1985, 'Molyneux's Question' in *Collected Papers*, Oxford: Clarendon Press

John McDowell, 1985, 'Values & Secondary Qualities', in In Ted Honderich (ed.), *Morality and Objectivity*. London: Routledge. pp. 110-129.



## **Week 9**

### Essential Reading

Tyler Burge, 2010. *The Origins of Objectivity*, Oxford: Clarendon Press, Ch. 6, pp. 154 – 210

### **Topics in Minds and Machines: Perception, Cognition, and ChatGPT**

Prof Will Davies and Prof Philipp Koralus – T. 11 – 1, Radcliffe Humanities (Ryle Room)

This course will provide a graduate-level introduction to a range of philosophical topics concerning minds and machines. The topics form three main clusters: cognition, perception, and large language models (LLMs):

1. Topics related to cognition will include reasoning and decision-making; neural network models and their relationship to learning.
2. Topics related to perception will include whether there is a border between perception and cognition, and if so, how to mark it; looking beyond language at iconic or imagistic forms of representation, and assessing their role in perception and cognition; examining the nature of bias in perceptual systems, and whether this can constitute forms of prejudice; and the study of perception in non-human animals, with connections to the question of animal consciousness.
3. We will also have an in-depth look at LLMs like ChatGPT, with visitors from computer science; discuss how LLMs relate to foundational issues about cognition, and consider how to integrate this technology with philosophy and cognitive science.

Links to readings will be posted on Canvas Thursday of 0<sup>th</sup> week. Please email [philipp.koralus@stcatz.ox.ac.uk](mailto:philipp.koralus@stcatz.ox.ac.uk) and [will.davies@philosophy.ox.ac.uk](mailto:will.davies@philosophy.ox.ac.uk) to express your interest in joining the seminar. Developments relevant to our discussion of GPT (the area is moving quickly!) can also be followed at [twitter.com/PhilippKoralus](https://twitter.com/PhilippKoralus)

### ***Tentative Schedule***

Week 1. Aspects of Human Reason and Judgment

Week 2. The perception-cognition border

Week 3. Neural network models and learning

Week 4. Large language models and ChatGPT

Week 5. Beyond language: iconic/imagistic representation

Week 6. Perceptual prejudice

Week 7. Using philosophy to power human-centered AI

Week 8. Perception and consciousness in non-human animals

## **Subjectivism about Value**

Dr Christopher Frugé, Th. 9 – 11, Radcliffe Humanities (Ryle Room)

This class will explore subjectivist views of wellbeing in light of the objection that valuing attitudes can be mistaken. We'll explore several incarnations of the objection and use them to assess various subjectivist approaches.

Email me at [christopher.fruge@philosophy.ox.ac.uk](mailto:christopher.fruge@philosophy.ox.ac.uk) if you'd like to attend. This will help me organize the sessions.

Here's a tentative list of required readings:

### **Week 1: Desire Theory and Mistakes**

Crisp, Roger. Well-Being. In *The Stanford Encyclopedia of Philosophy*.  
Parfit, Derek. *Reasons and Persons*. Appendix I.

### **Week 2: Adaptive Preferences**

Arneson, Richard. Desire Formation and the Human Good.  
Sen, Amartya. Gender Inequality and Theories of Justice.

### **Week 3: Actual Attitudes**

Heathwood, Chris. The Problem of Defective Desires.  
Heathwood, Chris. Which Desires are Relevant to Well-Being?

### **Week 4: Deep Values**

Tiberius, Valerie. *Well-Being as Value Fulfilment* ch. 1 sec. 4, chs. 2-3.

### **Week 5: Coherent Attitudes**

Street, Sharon. Coming to Terms with Contingency: secs. 2 & 5.  
Street, Sharon. In Defense of Future Tuesday Indifference.

### **Week 6: Idealized Attitudes**

Brandt, Richard. *A Theory of the Good and Right* ch. 6.  
Railton, Peter. Moral Realism.

### **Week 7: Against Idealization**

Enoch, David. Why Idealize?  
Rosati, Connie. Persons, Perspectives, and Full Information Accounts of the Good.

### **Week 8: Defenses of Idealization**

Dorsey, Dale. Idealization at the Heart of Subjectivism.  
Rosati, Connie. Naturalism, Normativity, and the Open Question Argument.

## **Moral Metaphysics**

Dr Umut Baysan – Th. 11 – 1, Radcliffe Humanities (Ryle Room)

This class is on the metaphysics of *moral properties* (and more generally, normative properties). Our focus will be primarily on the *naturalism/non-naturalism* debate in metaethics, but we will also explore questions about the existence of moral properties in the first place and the connection between metaethics and first-order normative ethics.

See below for a week-by-week breakdown of topics and readings. Starred (\*) items are mandatory readings for attendance. That is, I kindly request all participants to have read (at least) the starred items to attend that week's session.

Readings are available online (with one exception, see below) either as journal articles or as e-books (in which case they are available via SOLO), and in some cases as open-access resources easily accessed after a quick Google search.

### **Week 1: The “open question” argument**

G. E. Moore famously rejected naturalism about moral properties on the basis of “the open question argument”. Although this argument is now widely rejected, it remains influential in shaping the naturalism/non-naturalism debate in metaethics. Thus, we will start the term by examining this argument and focus on some naturalist assessments of Moore's attack on naturalism.

#### **Readings:**

- \*Nicholas Sturgeon, “Moore on Ethical Naturalism”, *Ethics*, 2003.
- If you are not familiar with it, read through the first 15 sections of Chapter 1 of G.E. Moore's *Principia Ethica*. (This book is now available open-access on “Project Gutenberg”; but hard copies of it are available in most libraries in Oxford.)
- Peter Railton, “Moral Realism”, *The Philosophical Review*, 1986.

## **Week 2: Moral properties and moral concepts**

It has been customary to reject the open question argument by drawing on a distinction between *properties* and *concepts*. Accordingly, it is sometimes suggested that the open question argument and other non-naturalist attacks on naturalism (e.g., the claim that normative properties are “just too different” from natural properties) can be countered by focusing on different *ways of thinking* about moral properties. This week, we will explore some recent versions of this strategy.

### **Readings:**

- \*David Copp, “Just too different: normative properties and natural properties”, *Philosophical Studies*, 2020.
- Helen Yetter-Chappell and Richard Yetter Chappell, “Mind-body meets metaphysics: a moral concept strategy”, *Philosophical Studies*, 2013.

## **Week 3: The moral error theory**

The moral error theory is sometimes defended on the basis that there is something very peculiar about moral properties, so we should not accept the existence of moral properties. What are the merits of this argument? We will also look at some recent work on how to formulate the moral (and more generally normative) error theory. In particular, we will try to answer: If there are no moral/normative properties, what parts of our moral/normative discourse must be false?

### **Readings:**

- \*J. L. Mackie, *Ethics: Inventing Right and Wrong* (Penguin, 1977), pp. 36-42.
- \*Bart Streumer, *Unbelievable Errors* (OUP, 2017), Chapter 8.
- David Brink, “Moral realism and the sceptical arguments from disagreement and queerness”, *Australasian Journal of Philosophy*, 1984.
- See also Selim Berker’s “Mackie Was Not an Error Theorist” in *Philosophical Perspectives*, 2019, for an interesting interpretation of Mackie.

## **Week 4: Moral supervenience (I)**

Most philosophers working in metaethics (naturalists and non-naturalists alike) accept that moral properties supervene on natural properties. Some philosophers also think that if such supervenience holds, then moral properties must be reducible to natural properties. This week, we will explore the merits of this argument.

### **Readings:**

- \*Bart Streumer, *Unbelievable Errors* (OUP, 2017), Chapter 2.
- David Enoch, *Taking Morality Seriously* (OUP, 2011), Chapter 3, sections 3.6-3.7 and Chapter 6, section 6.2.
- Debbie Roberts, “Why believe in normative supervenience?”, *Oxford Studies in Metaethics* vol. 13, 2018.

## **Week 5: Moral supervenience (II)**

This week, we will continue exploring the prospects of a supervenience thesis in metaethics. If moral properties supervene on natural properties, what is the best of explanation of such supervenience? Must we explain such supervenience, or can we take it as a “brute fact”?

Can we reject the supervenience claim in the first place?

**Readings:**

- \*Tristram McPherson, “Ethical non-naturalism and the metaphysics of supervenience”, *Oxford Studies in Metaethics*, vol. 7, 2012.
- \*Anandi Hattiangadi, “Moral Supervenience”, *Canadian Journal of Philosophy*, 2018.
- Russ Shafer-Landau, *Moral Realism* (OUP, 2003), Chapter 4, section 3
- Alison Hills, “Supervenience and moral realism”, in *Reduction, Abstraction, Analysis*, (Ontos Verlag, 2009).

**Week 6: Moral causation and moral explanation**

Some philosophers think that if moral properties existed, they would be causally and explanatorily redundant, and they use this claim for rejecting the existence of moral properties. Naturalists in metaethics typically respond to this challenge by showing how moral properties can be causally and explanatorily relevant. This week, we will explore the prospects of such a response.

**Readings:**

- \*Nicholas Sturgeon, “Moral Explanations”, in D. Copp and D. Zimmerman (eds.), *Morality, Reason and Truth* (Rowman and Allanheld, 1985)
  - It is not easy to get hold of an electronic copy of this. However, this paper has been reprinted in several edited collections and anthologies, and you can find various copies of these volumes and anthologies in most libraries in Oxford. See: (i) A. Fisher & S. Kirchin (eds.), *Arguing About Metaethics* (Routledge, 2006); (ii) G. Sayre-McCord (ed.), *Essays on Moral Realism*, (Cornell, 1988); (iii) J. Rachels, (ed.), *Ethical Theory* (OUP, 1998). Alternatively, you can read Sturgeon’s more recent article covering similar grounds: “Moral explanations defended” in *Contemporary Debates in Moral Theory*, ed. J. Drieier (Blackwell, 2006); available online via SOLO.
- Ryan Stringer, “Ethical Emergentism and Moral Causation”, *Journal of Moral Philosophy*, 2021.
- I have an unpublished manuscript on this topic, and I am happy to share it with those who are intending to attend this week’s class.

**Week 7: “Post-modal” metaethics**

Contemporary developments in meta-metaphysics suggest the concept of supervenience (which is a purely *modal* concept) is unhelpful in expressing claims in metaphysics, and the “post-modal” notions of *essence* and *ground* are better suited for this job. In this week’s seminar, we will explore examples of a post-modal approach to metaethics.

**Readings:**

- \*John Bengson, Terence Cuneo, and Russ Shafer-Landau, “The Source of Normativity”, *Mind*, 2023.
- Stephanie Leary, “What is Non-Naturalism?”, *Ergo*, 2022.

**Week 8: From metaethics to normative ethics (and back)**

It is sometimes thought that metaethics and normative ethics are isolated from each other and there are very few, if any, first-order normative implications of claims in metaethics. In

this week's seminar, we will look at the works of some contemporary philosophers who think otherwise.

**Readings:**

- \*Max Khan Hayward, "Immoral Realism", *Philosophical Studies*, 2019.
- Melis Erdur, "A Moral Argument Against Moral Realism", *Ethical Theory and Moral Practice*, 2016.
- David Enoch, "Thanks, We're Good: Why Moral Realism is Not Morally Objectionable", *Philosophical Studies*, 2021.

**AI and AGI: Philosophical and Ethical Issues**

Dr Adam Bales and others – W. 2 – 4, Radcliffe Humanities (Ryle Room)

*If you think you're likely to come along then it would be great if you could let us know, via [adam.bales@philosophy.ox.ac.uk](mailto:adam.bales@philosophy.ox.ac.uk), so that we can plan for numbers (however this isn't a requirement; please do feel free to turn up regardless of whether you've emailed).*

The last decade has seen remarkable progress in AI, with particular fanfare around gameplaying systems (like AlphaGo), large language models (like ChatGPT), and text-to-image models (like Stable Diffusion and Dall-e). Further, it seems likely that we will continue to see AI progress in the coming years and decades. More speculatively, some people think that within this timeframe we will develop artificial general intelligence (AGI), roughly AI systems that are as cognitively general as humans and at least as cognitively capable.

Regardless of whether or not we develop AGI, AI systems will have a substantial impact on the world. These systems and this impact raise a range of philosophical and ethical issues, and in this seminar series, we'll explore some of these issues.

The first half of the course will cover issues raised by AI systems in approximately their current form and the second half will focus on issues raised by AGI. (Though the division is somewhat artificial and some weeks will explore issues that are relevant to a wide range of AI systems, including both current systems and AGI.)

The weekly schedule is below, and more detail (including readings) is available [here](#).

Week 1 (Wed 26th April): Introduction to AI and AGI (Adam Bales)

Week 2 (May 3rd): Moral Machines (Charlotte Unruh)

Week 3 (May 10th): Autonomous Weapons (Linda Eggert)

Week 4 (May 17th): Automation and Employment (Charlotte Unruh)

Week 5 (May 24th): Neutrality and Bias (Milo Phillips-Brown)

Week 6 (May 31st): Human Extinction and Disempowerment (Adam Bales)

Week 7 (June 7th): The Singularity Hypothesis (David Thorstad)

Week 8 (June 14th): Moral Status (Andreas Mogensen)