



“The point is
to change it”

Essays on
philosophy in
public life



The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds and undertakes rigorous research, encourages innovation and supports the use of sound evidence to inform social and economic policy, and improve people's lives. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics, the Ada Lovelace Institute and the Nuffield Family Justice Observatory.

This volume has been funded by the Nuffield Foundation, but the views expressed are those of the authors and not necessarily the Foundation.

Find out more at: www.nuffieldfoundation.org

© 2025 — all rights reserved - Nuffield Foundation

“The point is
to change it”

Essays on
philosophy in
public life

How to read this volume

The volume begins with a preface by Tim Gardam, former Chief Executive of the Nuffield Foundation. Professor Dave Archard, volume co-editor and former chair of the Nuffield Council on Bioethics, follows with an overview of academic philosophy's engagement with public life in the UK and beyond, with particular attention paid to how this relationship has changed over the last century.

The preface and introduction lay the groundwork for the contributions which follow. We suggest you start with these pieces, before reading the essays, which are grouped into three sections:

- “Prospects” contains varying perspectives on what — if anything — academic philosophy has to offer public life.
- “Reflections” includes accounts from academic philosophers of their own engagement with public life.
- “Interventions” comprises a set of essays engaging directly with pressing contemporary debates in public life, from climate change policy to the role of artificial intelligence in warfare.

The essays themselves are not intended to be read sequentially, but to be enjoyed in any order you see fit — like articles in a literary magazine or academic journal. We encourage you to begin with the contributions that most immediately grab your interest, and continue to explore from there.

Contents

7 Foreword

Tim Gardam, Nuffield Foundation

11 Introduction

Professor David Archard, Queen's University Belfast

Section 1: Prospects

24 No dea(r)th of philosophy

Dr. Shuk Ying Chan, University College London

30 Public philosophy for gremlins

Dr. Lorna Finlayson, University of Essex

37 Identity dialogues

Professor Tariq Modood, University of Bristol

45 Philosophy matters

Professor Debra Satz, Stanford University

51 Philosophy as democratic underlabour

Professor Adam Swift, University College London

59 The marketplace of ideas: who's buying?

Professor James Wilson, University College London

67 The point is to change it

Professor Jonathan Wolff, University of Oxford

Section 2: Reflections

75 Did that answer your question?

Professor Clare Chambers,
University of Cambridge

83 Breaking bread with the enemy

Dr. Anca Gheaus, Central European University

89 Of weasels and women, or, what is public philosophy anyway?

Dr. Katharine Jenkins, Glasgow University

96 On the inequity of ethics

Dr. Winnie Ma, UC Berkeley

103 Philosophy in the flow of political life: realism, moralism and community wealth building

Professor Martin O'Neill, University of York

111 To the shoemakers and the ship-builders

Professor John Tasioulas, University of Oxford

Section 3: Interventions

121 Philosophy protects the climate

Professor John Broome, University of Oxford

130 The pathology of the prison

Dr. Sarah Bufkin, University of Birmingham

139 Call the midwife

Professor Roger Crisp, Dr Katrien Devolder
and Dr Jonathan Pugh, University of Oxford

146 Ours to question why

Professor Cécile Fabre, University of Oxford

153 On the new demise of ethics

Dr. Elke Schwarz, Queen Mary University
of London

160 The world and his wife

Dr. Musab Younis, University of Oxford

166 Author biographies

172 Acknowledgements

Foreword

Foreword

Tim Gardam, Nuffield Foundation

These essays on the role of philosophers in public life are an unusual initiative for the Nuffield Foundation and this foreword is by way of an explanation. In recent years, Nuffield has been primarily a funder of data-driven social science; however, the Foundation's original purpose was to "advance social wellbeing" in the broadest sense and it embraced thinking across the sciences and humanities. Two of Nuffield's Centres, the long-established Nuffield Council on Bioethics and, more recently, the Ada Lovelace Institute, were built on the premise that good policy outcomes derive from the interplay between the clarification of principles and empirical analysis. Both bodies owe their creation in large measure to Onora O'Neill, a previous Chair of the Nuffield Foundation, and one of the philosophers most influential in framing questions of public policy in recent decades.

Nuffield has held a long-standing belief in "the power of evidence to change lives", but there are times when the concentration on data analysis to provide that essential empirical evidence can appear somewhat detached from any underpinning consideration of the questions the data are meant to illuminate. For all the rigour of its methodologies, there can be a narrowness in the way that social science construes the questions it addresses, and a lack of interest in different intellectual traditions dealing with similar questions. This project emerged from a growing sense that the disciplines of social science, robust and valuable in their own terms, require some wider normative framing, especially at this uncertain moment in our history and culture. This objective goes beyond academic aspirations for "interdisciplinarity"; it has a direct bearing on the challenge of translating clear and reasoned thinking about what constitutes social wellbeing into effective policy and finding better practical answers to common concerns.

The project is also a response to a growing sense of crisis. The norms of government and policy making, and the assumptions on which these were constructed mid 20th century, now confront a corrosion in public discourse, an accelerating decline in public trust in institutions and an AI revolution which has overturned the way we process information about the world and understand our place within it. The financial crash of 2008 and the economic stagnation and widening inequalities that have followed have been accompanied by a breakdown in

the relationship between the policy maker’s use of data as evidence and the experiences of those whose lives make up that data. The well-tried tools of social policy — quantitative data analysis, complex trade-offs, evaluation and prioritisation — have been associated with ways of thinking that have at times appeared indifferent to those whose wellbeing they are meant to serve.

One response to this erosion of trust in the authority and fairness of decision-making has been to look to practical ethics as a framework to provide some reassurance of public institutions’ accountability and good practice. Ethical frameworks for policy making are (rightly) prominent in the climate debate, in data and AI governance and in addressing questions of intergenerational equity. However, although reference to ethics in public debate is now commonplace, and to be welcomed, the more difficult thinking, testing the prior questions and principles that should underlie the ethics of policy formulation, has not been so prominent. Public ethics are at risk of being elided into sets of rules to apply to difficult questions in various arenas — the ethics of corporate governance, ethics in public life, ethical investing, data ethics. How we should determine ethical import is much less discussed than whether or not a particular way of acting is permissible. In more debased ways, ethics has become a useful adjunct to corporate branding. “Values” (usually no more than a list of generic adjectives) are displayed on private and public sector websites as self-validated testimony to their own virtue.

The challenge of reaching difficult ethical conclusions is with increasing frequency made synonymous with new models of public consultation — citizen juries, “deep listening” and the like — as if these processes in themselves can resolve questions of moral ambiguity and fundamental value conflicts as they touch on public policy decisions. Though such deliberative models have value, especially at a time when institutional decisions-making has lost authority, without the critical distance and muscular thinking of the philosopher to interrogate them, their claims to be sufficient as ethical arbiters of policy making must be at best limited.

Simultaneously, the rapid acceleration of advances in the life sciences, data science and AI has intensified concern about the ethics of their application. (The urgency of the need for an ethical framing of these technologies is intensified by the fact that they are frequently in the hands of mega-rich individuals who are disdainful of public institutional norms). The AI challenge is often set out in existential terms — the need to govern scientific advances that challenge the future of human agency. But there are also more immediate ethical concerns

— asymmetries of information and power, the end of privacy, the protection of, and limits to free expression; how to mitigate the risk inherent in AI models of exclusion, discrimination, unfair categorisation, pre-judgements based on analysis of proclivity, and so on. The danger here however is that ethics could be construed solely as being a brake on scientific discovery, rather than being the means of governing its beneficial deployment.

There is, without question, much that is admirable in an ethical purposiveness that seeks to calibrate public policy and decision making but, in practice, this approach can too often be reduced to regulatory questions of compliance. We run the risk of using ethics as a tool to close down issues of complexity rather than to open them up.

Hence the need for this project. We asked Professor Dave Archard, Chair of the Nuffield Council on Bioethics from 2017 to 2022, to lead it. During the Covid pandemic, he personified the direct engagement of bioethics, thinking across intellectual disciplines, and then translating that deliberation to guide those making policy under stress. In choosing his title for the project, *“The point is to change it”*, Professor Archard has deliberately echoed Marx to reflect the necessary edge, urgency and responsibility of Philosophy in the current context. This project is not simply a challenge to social scientists to think more broadly about how philosophers address similar agendas; it is also an invitation to philosophers to translate their work and language into terms that will help frame and clarify the thinking of policy makers and navigate paths to good government.

As an academic funder, the Nuffield Foundation does not take a position on any of the arguments set out in these essays. Our interest is in giving space to consider the ways in which philosophers’ modes of thought contribute constructively, (or not), to public policy discourse (and indeed whether they can or should). This volume is therefore focused on *how* philosophers might help us to think about big public themes not what any conclusion on any one theme should be.

Our other aim has been to reconnect the way philosophers think about ethical questions with the wicked questions and trade-offs of social and economic policy. This is more than just an intellectual exercise: better thinking is essential to the defence of humane government at a time of vertiginous uncertainty when both normative reasoning and empirical analysis are at risk in our public culture.

Introduction

Introduction

Professor Dave Archard, Queen's University Belfast

We have taken as our title for this collection the second sentence of one of Karl Marx's most celebrated aphorisms, Thesis Eleven of his *Theses on Feuerbach*: 'The philosophers have only interpreted the world; the point however is to change it.' Written in 1845 it was only published after its author's death, but the quoted words are on his Highgate cemetery tombstone. The Eleventh Thesis has also been much discussed, even while its proper meaning has often been misunderstood.

We take it not to be a rejection of philosophy as such. Marx was criticising those of his contemporary philosophers whose work he believed had no theoretical relevance to the important task of understanding and — crucially — thereby changing the world. It is thus not a plea to act without the benefit of theory; it is a demand that transformative political action be informed by theory. Of course, this demand can be acknowledged and endorsed without favouring Marx's historicist view of both theory and the preconditions for change.

Taken then only as the demand that those who do philosophise about the world take seriously what it means to change that world it is an apt maxim. For, normative philosophy — encompassing practical ethics, jurisprudence, and political theory — speaks about what ought to be the case. It surely matters then whether what should be the case can be the case, and, if so, how it could be the case. Why does that matter now? How could such philosophy change the world, and what should philosophers say about this?

In answer to the first question, four different background considerations come into play: the economic context; the political context; how philosophy understands itself; and how philosophy is understood by others. To take each in turn. First, capitalism, despite Marx's prediction of its demise, has survived and flourished. We all live and work in nation-states which are subject to the constraints and pressures of a global economy in which the market is dominant. Politics is, Marx was right at least about this, shaped by economics. That is no better illustrated than by the actions of the major tech and media corporations pursuing market dominance, ever greater profits, and significant political influence. In such a context,

ethical considerations of policy can easily be squeezed out, and whilst ethics of course matters, the voices of those defending its importance struggle to be heard.

Second, the present political context is one of a widespread distrust of politicians, a crisis of liberal democratic governance including the apparent indifference of populations to elections and the disturbing preference of many voters for authoritarian leaders, the rise of populism, the pervasiveness of misinformation and a general disregard for the truth in political rhetoric. These features have been widely noted and variously explained. Here we do not enter these debates but content ourselves with remarking on the challenge faced, in such a context, by a philosophical enterprise that esteems the free exchange of ideas, rigorous argument, conceptual clarity and the rational scrutiny of all claims.

We will say more about the question of how philosophy understands itself. The history of normative practical philosophy in the English-speaking world in the last hundred years is essentially one of a rebirth or rediscovery. In the 1960s ethics was not generally viewed as a proper part of academic philosophy, to stand on an equal footing with metaphysics, epistemology, and logic. At worst the making of moral judgements was seen as something any thoughtful and intelligent person could engage in. Bertrand Russell, the most widely recognised and feted philosopher of his age, was clear that philosophy proper did not comprise moral commentary on any matter. He relegated this to his popular writings for an interested, intelligent lay audience.

At best when a philosophical appraisal of moral judgements was made, they were either simply insights into the good and the right which, without further argument, were had or were not; or they were the expression of emotions, and, as such, not open to appraisal as justified (or not). A. J. Ayer, who would have been delighted to be viewed as his age's next Russell authored the expressivist theory of ethics, sometimes simply summarised as the 'boo-hurrah' account of what moral language amounts to. Moral judgements — for instance, 'Murder is wrong' - on this account are no more than expressions of how one feels about something — 'Murder? Ugh!' — and are not to be thought of as making claims of what is the case, that murder really is a bad thing. Intriguingly, Ayer was the first President of the Society for Applied Philosophy, but principally because of his well-known public profile.

The proper function of philosophy was seen in the 1950s and 1960s as the careful analysis of language, the exposition of what is said and meant, such that confusion is

thereby avoided. It is not to argue for normative principles. As for political philosophy the editor of a 1956 collection of relevant essays commented — a remark endlessly re-quoted — that for now political philosophy was 'dead'. He meant political philosophy as a grand theoretical enterprise in the tradition of Locke, Hobbes, Rousseau and Marx that addressed the big questions of how we should best shape our political lives together. T. D. Weldon's *The Vocabulary of Politics* (1953), now largely forgotten, defended the view that all the problems of political philosophy up to then could be set to one side as deriving from linguistic confusions and errors, thereby relieving us of unhelpful 'metaphysical lumber'.

Everything changed in the 1970s. There was a 'practical turn' to a moral consideration of real issues. Nowhere was this more evident than in political philosophy. In 1971 John Rawls published *A Theory of Justice*. Now, of course, the death of political philosophy prior to this date can be exaggerated. There were important contributions to political theory — in the English-speaking world from the conservative thinker Michael Oakeshott and from Isaiah Berlin who made seminal contributions to the history of ideas. In Europe, work rather conveniently summarised as 'Western Marxism' — writing by, amongst others, Adorno, Lukács, Gramsci, and Sartre — sought to construct a version of that doctrine fit for the political problems of a world ravaged by war, fascism, colonialism, and the survival of capitalism.

Nevertheless, Rawls is clearly of immense significance. He revived the grand tradition of normative political philosophy, not least by putting the idea of a social contract, previously used by Locke, Hobbes and Rousseau to service in an understanding of what a society could accept as fair terms of cooperation. He displayed the virtues of English-speaking philosophy — conceptual clarity and rigour of argument — in defending a comprehensive normative theory. He made 'justice' a key concept, and he offered an alternative moral grounding to the defence of his views from that — utilitarianism — which was the dominant ethical orthodoxy. His writing, both the early *A Theory* and his later reconsideration of issues, *Political Liberalism*, has provoked voluminous critical review and regenerated interest in what normative theory can and should speak about.

In the 1970s and 1980s there was also a 'practical turn' in ethical theory. In the United States the leading journal *Philosophy and Public Affairs* was founded in 1972. Its title says it all. In the United Kingdom, the Society for Applied Philosophy was founded in 1982, and the first issue of the *Journal of Applied Philosophy* was published in 1984. These journals explicitly,

assiduously, and self-consciously promoted philosophical work on issues of contemporary practical significance: those of evident political importance such as just war, civil disobedience, discrimination, those of juridical importance, such as the justification of punishment, and especially capital punishment; those pertaining to biomedicine, such as the proper scope of research, the duties of medical practitioner; and those of personal moral choice, such as abortion and permissible sexual behaviour.

In due course, I want to say more about this 'practical turn,' but for now I want to turn to the third background consideration in play when we talk about what practical normative philosophy ought to do now. This is the question of how philosophy is viewed by others. For all the grand claims that academic philosophers might make about its value and proper status, philosophy is one of the smaller subject areas in modern Universities. It is usually dwarfed by law, medicine, and the physical sciences. It has also struggled to retain its putatively essential role in any reputable academic institution.

The problem philosophy faces is not unique within the humanities, and, in turn, this more general crisis can largely be explained by the increasing dominance within the academic sector of an economic model. Universities need to secure sustainable funding, and this puts pressure on individual subjects to show that they can justify their existence in economic terms — by, for instance, showing how their graduates might make a distinctive contribution to the economy, or by justifying their research in terms of its economic benefits. Unfortunately, appeals to the value of philosophy in terms of the acquisition of invaluable critical skills of analysis and argument are not in the context entirely persuasive, and most probably not of Government.

However, philosophy has also been subject to intellectual criticisms from within academia from those subjects that would also offer an alternative theoretical approach to issues. The criticisms are those of practical irrelevance, of an idealism that misunderstands the real world, and — most damagingly — of a radical contingency. To explain this last: English-speaking philosophy is criticised for being the product of and shaped by a particular time and place. It is Western, or Eurocentric; it is modern in an era that has moved beyond the certainties of the Enlightenment. Post-modernism in this sense has been hugely influential within social and cultural theory. Philosophy seems sometimes to have little reply to these charges other than to re-assert Enlightenment values, especially that of Truth (with a capital T) and continue with its way of proceeding.

It would be much better to do two things. The first is to make clear the real dangers of moral relativism, chief of which is the ultimate failure to be able to say anything of real normative power. It should matter that we can say, here and now, that some things are just wrong full stop, and others would be better. The second is to point out that those who decry the making of moral judgements, presuming that they can be defended as correct or incorrect, do themselves make such judgements even if they are not always honest about doing so.

It is appropriate now to say something more about the 'practical turn.' The phrase 'applied philosophy' is unfortunate and unhelpful. It conjures up an unappealing view of what is being done as the application of high-level principles or norms to everyday practical problems. The approach is often characterised as 'top down' and proceeds — at its simplest — by means of what logicians would term syllogistic reasoning. The argument in question uses a major premise which affirms a general truth, such as 'The killing of innocent human beings is always wrong,' which when conjoined with a minor premise that states a factual truth, such as 'A foetus is an innocent human being' yields a concluding moral judgment 'Abortion as the killing of a foetus is always wrong'.

This approach is also pejoratively characterised as 'armchair philosophising'. The thought thereby signalled is that the practical normative philosopher can without leaving their study argue for general moral principles and then acquiring from the relevant expert — in, say, medicine or the empirical social science — those statements of fact that are needed argue to a conclusion of moral judgment.

This view of the 'applied moral philosopher is not entirely unwarranted or prejudicial in the early days of the practical turn. However, things have moved on. Many who do practical ethics both prefer that self-description and engage in what, by contrast with the 'top-down' approach, can be termed a 'bottom-up' one. This starts from the often-complex reality of an issue, understands it on its own terms, and seeks, using both a proper understanding of the relevant facts and a knowledge of ethics, to work up to a moral judgement.

This approach is also much more congenial to interdisciplinary work in which the philosopher analyses and evaluates a practical issue *with* those from other disciplines and professions, rather than simply taking from them what is needed for the minor premise. Nowhere is this more obvious than in bioethics of which I will say more shortly.

However, something should now be said about the distinction in the practical turn between political philosophy and practical ethics. It would be a mistake to characterise the former as one instance of the latter — the use of moral deliberation to understand and evaluate political problems. This is not just a matter of scope — business ethics deals with the ethics of business, political philosophy deals with the ethics of politics. It is rather that this approach fails to acknowledge the specificity of politics — what is the particular and unique character of what happens when we form political societies wherein some of us exercise power over the rest. What justifies this outcome and what are the resultant problems? This is what Aristotle addressed in defining humans as 'political animals'; it is what Hobbes and Locke considered in contrasting a 'state of nature' lacking rule with a society that was governed by law; even Marx had suggestive things to say about how the final stage of communism would entail a withering away of the state, and the replacement of the government of persons with 'the administration of things.'

The stronger charge — sometimes described as 'political realism' - of those who draw attention to the special and distinctive nature of politics is that a practical ethical view of political philosophy is simple-minded, naïve, and inadequate to the task of changing our political world. It cannot be that moral philosophers work out what is a better way to arrange our society such that there are, for instance, fairer terms of cooperation, and that the path to making the requisite change is thereby suggested. We evidently need to understand what politics is, how political institutions function, what is and what is not good governance, how laws are made and unmade, and what it means to exercise power.

This is not to deny that we can draw from ethics a better appreciation of relevant concepts like 'obligation,' 'legitimacy,' 'power; and 'democracy.' However, the 'political turn' of some writers in some areas of practical ethics is noteworthy and important. A good example is that of political theorists who wish to give full and proper recognition to the status of nonhuman animals. They have done so by shifting this question of status from one owned by moral philosophers to one that implicates politics, raising questions about who should belong to a political society, how those who are not human might be represented in its systems of governance, and how we as humans should act politically to protect nonhuman animals.

Earlier it was said that bottom-up practical ethics operates best by working with those from non-philosophical disciplines. An excellent example to consider what this means is 'bioethics.' As a subject bio- or medical ethics originates in work done by moral philosophers

in the United States at the end of the 1970s. The classic first text is *Principles of Medical Ethics* by James Childress and Tom Beauchamp which has gone through several editions since its original publication in 1979. This famously outlines, defends and applies four moral principles to biomedical practice. Its authors were clear that what they were doing was practical ethics construed as using moral concepts and norms to think about problems, practices and policies. They distinguished their normative ethics from descriptive ethics done chiefly by social scientists concerned with how people do morally think and behave in respect of these problems. The former asks, what ought to be the case, the latter asks what the facts of the case are.

Since the 1970s bioethics has changed dramatically and in four regards. First, it is now not narrowly understood as a sub-domain of ethics, but much more broadly as an interdisciplinary subject comprising law, medicine, anthropology, politics, sociology, psychology, economics, and history. Second, its original focus on clinical practice and research — considering issues such as the obtaining of consent from patient or research subject — has expanded to include, most significantly, questions of public health, addressing such as what kinds of intervention are warranted to ensure the health of a population and what does it mean to think of health outcomes as fair or unfair across a society. Third, bioethics is global in scope. It addresses problems worldwide. The COVID pandemic made clear that some health issues must be addressed across nations and questions of global injustice (in access to medicines, distribution of resources, differential impact of diseases) are necessarily broached. Fourth, bioethics is not now simply an academic subject but a profession, a consultancy option, embedded in committees, commissions, and institutions of health care and biomedical research.

For some these changes have meant a loss of quality, a lessening of rigour, as individuals rush to secure funding to research or to publish on the latest development. Yet, if this is true, the fault can be laid at the doors of those who approve the funding or the publications, not as one of bioethics as broadly construed. Moreover, the following should be noted. It is true that clinical ethics — of both research and medical practice — can be a relatively straightforward matter of using ethical tools to appraise a dilemma that might arise in these areas. Indeed, there are ethical protocols ready to use by a relevant body, such as a clinical ethics or research committee. The same is not true of the adoption and implementation of laws, regulations and policies, in respect of new developments in medical and biological technology.

Here there are questions not simply or only of what might be ethically permissible, but also of what might be feasible laws, what might be broadly acceptable to the public, what would facilitate scientific research of sufficient value, and what could command the support of Government. The work of Mary Warnock in chairing the committee whose final Report on the matter of fertility treatment and embryo research led to the creation of the Human Fertilization and Embryology Authority is exemplary in this regard. She laboured conscientiously and diligently to make proposals that were both philosophically respectable (even if she did not persuade every philosopher of this) and had a good chance of making it into appropriate legislation.

Here the bioethicist cannot work alone (and least of all seated in an armchair sequestered in their study) to produce good outcomes. They must work with those who appreciate what generates disagreement on a proposal and why, what will satisfy the requirements of good law, and how to engage with the public both to consult and to inform. Above all the function of bioethics is to advise and recommend, not to dictate. Here in the United Kingdom the Nuffield Council on Bioethics offers an excellent instance of what is possible. Over the years it has produced consistently high-quality reports on all the major developments in bioscience, such as gene editing, as well as the social and political context of health policy, such as on the issue of an ageing population or the design of public health interventions. It has commanded the respect of the public, and always got the ear of successive Governments. Notably it is a Council *on* bioethics and not simply *of* bioethicists.

We are brought then to the question, which follows from this discussion of bioethical law and policy and which opened this Introduction, namely what can and should be said by those who do normative practical ethics about what might and could be the case.

We can start at the beginning — in terms of philosophy literally so — with Plato, whose *Republic* argued for rule by the wise, that is philosophers. For Plato, the question of what a good political society is can be answered by a simple statement: the political good of all is best served by having as rulers those who know what is best. These are philosophers who, moreover, will rule in the interests of all and not of themselves because as philosophers they will act from duty and not desire for personal advancement or for selfish reasons.

No-one now would endorse this proposal even if some autocrats might judge themselves to be philosophers. However, Plato challenges the democratic principle whereby all should

participate in some form in the exercise of political power by suggesting that the rule of many is evidently inferior to rule by the incorruptible and sagacious few. The alternative to democracy is *epistocracy*, rule by those in the know.

What exactly is wrong with epistocracy? Some plausible answer is needed if only because if it is not wrong there is a ready defence of what normative philosophers might do to change the world, namely simply rule the non-philosophical majority. It helps to distinguish between the possible replies, and of course some combination of them is possible. First, democracy will be defended not in terms of its outcomes, which we can agree may be stupid or vicious laws, but by its foundation in a principle of equality. Yet, that answer invites other questions, for the scope of citizenship does not extend to everyone subject to the laws of some state. Children, prisoners and the insane are conventional exceptions. And with children at least we seem to have an epistocratic reason for their exclusion, namely that they are not able to understand what it means to be a citizen and know what is involved in voting. However, if children are excluded for this reason, why not many adults who we could reasonably, and regularly, judge do not understand what is at stake in any election?

The case of children is also a way of answering a second response to Plato's anti-democratic argument, namely that there is no such thing as the expertise he thought warranted philosophers' rule. For we can surely say that we know who does not have the required expertise and that presumes we have at least some sense of what that is.

We might respond to Plato's challenge to democratic rule by saying that even if there is expertise of the relevant kind, we cannot reliably and certainly identify who has it. That has some plausibility, but if philosophers do not have it, who else could be a plausible candidate? Of course, we might even concede that not only philosophers have it, but that still allows that at least they do.

A final response to Plato is that even if there is expertise of the relevant kind, and even if those who have it — such as philosophers — can be reliably identified, they should not exercise it. In the personal domain — consider the agony aunt's reply to the troubled friend of the unfaithful wife who worries whether they should tell the husband the truth — the conventional response is that moral expertise should not be deployed because it vitiates the choice of the person advised and it is the making of that choice which gives it its moral worth.

However, are matters as clear in the case of political rule? Here it might be said that the exercise of expertise subverts democracy by taking away from citizens the empowering and educating choices they have otherwise to make. Yet, it would surely make a difference how much is left to the ethical experts, and it seems far-fetched to think that the delegation of some law and policy choices to such experts would undercut democratic governance, especially if the delegation was democratically made ('We the people give you the philosophical few the power to make these decisions'), and if the scope of such delegated powers was carefully delimited.

If normative philosophers are not Plato's rulers, then might they at least offer advice? In the history of the subject there have been philosopher-advisers. Aristotle taught Alexander the Great; John Locke advised the Earl of Shaftesbury and suffered exile in Holland for his pains. In the modern era philosophers have been official advisers in their role on Government committees and commissions. Bernard Williams chaired a Committee on obscenity and film censorship; Mary Warnock, as noted earlier, chaired committees on special education and, most famously, on fertility treatment and embryo research.

Yet we need not appeal to these kinds of example. Indeed, we could represent publications in normative practical ethics as advisory in character. After all, an article arguing that, for instance, assisted dying is morally permissible, amounts to advice to law makers that it should be legally permitted or decriminalised if currently something that is prosecuted. We could say this whilst recognising the important differences between law and morality. If such work can be interpreted as advisory, then there are evidently obligations on their authors. For advice can be bad in at least two ways: first, *what* is advised may be bad and harmful; or, second, *how* the advice is offered may be poor — for instance, by being open to dangerous misinterpretation or being unclear in what exactly it commends — such that what is done as a result is bad and harmful.

Normative practical philosophers may well protest at this point — and some have — that all of this goes far too far. There need be no false modesty on the part of philosophers who say something along the following lines, 'It is no part of what we do as practical philosophers to recommend that the world should be a certain way, and it is certainly not incumbent on us to show how the world might be changed. Least of all should we be charged with the responsibility of making those changes.'

In fact, normative practical philosophers may well say some of the above. They might, for instance, argue that although they do make recommendations for improving the world, it is for others to show how that could be managed in the real world. This suggests a possible division of labour — between the architect or designer of the beautiful edifice and the builders who will erect it. However, whilst there is a distinction between being the builder and the architect, there is surely an obligation upon the latter to show that the building could be built. Those who argue for what ought to be the case must be sensitive to criticism that, since 'ought implies can' (something should be the case only if it could be), they are able to show that what they recommend is possible.

This criticism often comes in the form of a characterisation of contemporary practical normative philosophy as idealistic, other worldly, abstracted from the brutal realities of the world, and as such unlikely to have any influence in and on that world. The criticism deserves to be answered.

To Rawls is owed a first response, namely that a distinction can be drawn between an ideal theory of justice in which the preconditions for an ideally just society, such as the dispositions of its citizens to support and live by the principles of justice, obtain, and non-ideal theory when they do not. However, non-ideal theory is not developed in his writing, albeit much subsequent work, influenced by Rawls' distinction, has been devoted to spelling out exactly what is meant by non-ideal.

Equally, and in the same vein, a lot of work has been published on what might be meant by saying that a normative theory is feasible or realistic. However, none of this offers an elucidation of a feasible or non-ideal state of affairs. Thus, it might make more sense simply to borrow from economists the idea of sub-optimal outcomes and then address the question — using examples — of what the best sub-optimal state of affairs might be in any particular case. For instance, take an ethically contentious issue, such as assisted dying, climate change, gender identity, the categorization of sexual offences, the sale and use of 'soft' drugs, or the age of criminal responsibility. In the absence of a consensus or overwhelming majority in favour of the optimal proposal or faced with the evident unwillingness of a government to legislate as ideally desired philosophers might suggest what could be done that is better than exists if not the best they would wish for in the best of all possible worlds.

Some — the work of Amartya Sen is exemplary in this regard — will argue that we do not even need ideal theory. We do not have to know what is the best in order to identify what is better or worse. We can address injustice — and take steps to ameliorate or eliminate instances of it — without spelling out and defending what would be perfectly just. Such an argument will not persuade those who think that we do need principles or standards by which we can make comparative judgements of better or worse; and that these presume or can be appreciated only by an appeal to ideals.

The debates about non-ideal theory, what it means to say that a normative theory is 'feasible' (what is presumed about the world and about human nature), make a difference to the kinds of normative theory that are and can be done. Nevertheless, it must remain incumbent on practical normative philosophers — those who argue and write about what ought to be the case — to recognise what is owed to their audience in the way of an exposition of what could be the case, and what that would involve. Otherwise, Marx's ghost will remind them that they have missed the point of doing such theory, that of changing for the better the world they otherwise only interpret.

Section 1: Prospects

No dea(r)th of philosophy

Dr Shuk Ying Chan, University College London

Philosophy has had a long line of sceptics going as far back as Socrates' jailers. As the central question of this volume asks, what has philosophy done for and in public life? Although it claims to think deeply about morality, ethics, justice and injustice, what change has philosophy actually made?

Yet the idea that philosophy has had little to offer public life and done little to make change in the world relies on a specific — and I would argue, much too narrow — view of what counts as philosophy. Once we appreciate how ubiquitous philosophical thinking and argumentation are in our social and political life, it becomes impossible to maintain that it has done little for public life. And it also becomes possible to better appreciate the importance of the academic or professional practice of philosophy.

Let's start by making a distinction between academic or professional philosophy and public philosophy. Academic philosophers work in universities and research institutions, publish their work in peer-reviewed books and journals aimed mostly at other colleagues, and teach philosophy in an institutional setting. Ideally, they are afforded the resources and time to engage in close examination of concepts and ideas through writing and debate, sharpening understandings of competing conceptions and their moral and political implications. What I'm loosely calling public philosophers, on the other hand, consist of a broad category of political actors whose primary role is to push for and/or enact change in the world, whether through organising social movements to shed light on neglected problems and to demand change, or through leading a government, proposing and passing new laws and policies, reforming existing institutions, and so on.

The narrow view of philosophy would hold that only what academic philosophers do counts as philosophy. While it would be hard to find someone who explicitly defends this claim, the

claim is implicit in the oft-repeated idea that political philosophy was largely dead in the 1960s until the publication of John Rawls's *Theory of Justice*. As I argue below, such a claim ignores the many works published during that time, works that do not follow the conventional mould of academic philosophy and which are not written by academic philosophers, but nonetheless can and ought to be read as philosophy. The narrow view is also implicit in the longstanding neglect of these types of work in the curricula or 'canon' of academic philosophy and 'analytic' political theory, a trend that is only now starting to change due to pressure from students as well as recent political developments outside of the academy, most notably the Movement for Black Lives.

I think the narrow view of philosophy can obscure the indispensable role that moral and ethical theorising has always played in real world change.

To demonstrate this, let's look at some examples of public philosophy. At the same time in the 1960s that political philosophy was pronounced 'dead', a book titled *The Wretched of the Earth* was published. The author was Frantz Fanon, who is famous today as a figurehead for activists and social movements, especially those engaging in racial justice struggles. Although *The Wretched of the Earth* is based on Fanon's experiences as an activist for the Front de Libération Nationale in the Algerian War of Independence, it would be a mistake to read it as a historical narrative or political manifesto rather than a work of critical and normative philosophy.

From diagnosing the structures and practices of oppression at work in the colony, to proposing a theory of revolution that identifies progressive and reactionary agents, and finally, proposing a vision for a decolonised world in which the humanity of the colonial subject and the coloniser is restored, the book contains important arguments. The arguments cover the nature of justice and injustice, the moral justifiability of violent means of resistance, the role and nature of good political leadership, and the sociopolitical conditions for human liberation.

While these arguments are not couched in the language of (academic) moral and political philosophy — such as rights, duties, principles of justice, etc — they are nonetheless philosophical arguments. These include discussion *about* what rights people have, and which are unjustly denied in the colonial situation; what is owed between coloniser and colonised (for example, reparations) and between the formerly colonised (for example,

democratic participation in state-building); and principles that ought to guide the transition to a just, decolonised world (for example, 'the last shall be first'). To be sure, reconstructing these moral claims is a complex task that involves textual and contextual interpretation, and there is often more than one way to read the text. Yet the same is true of any work of philosophy.

Another book published in the mid-1960s was Kwame Nkrumah's *Neo-colonialism: The Last Stage of Imperialism*. As the first prime minister of Ghana who had just been deposed in a CIA-assisted military coup, Nkrumah wrote *Neo-colonialism* to investigate and expose the relations of global political domination and economic exploitation that persisted beyond formal colonialism, or colonialism in which the coloniser takes control of the colonised territory. Although much of the book consists of detailed empirical analysis of the specific players — such as investment banks, mining companies, etc — who were at the frontline of exploiting the postcolony in Nkrumah's time, there is also a theory of neo-colonialism that runs through the book. This theory understands ideas such as self-determination, popular sovereignty, exploitation etc, in particular ways. These understandings were hugely influential among pan-African leaders and thinkers debating what genuine freedom after formal empire meant, and how it might be attained.

We do not need to look only at historically influential figures. We could equally look at contemporary social movements — take climate activism as an example. Embedded in the political action and programmes of groups like Just Stop Oil and Extinction Rebellion are arguments about the nature of climate injustice, who is responsible for the climate crisis, the moral justifiability of resistance strategies that impose costs on others, and so on. These arguments are not always explicitly articulated in words but expressed through the kinds of actions taken and the targets chosen for those actions. And yet they remain moral-philosophical arguments with specific premises, reasoning, and conclusions about justice and injustice.

Once we take on a broader view of what (moral and political) philosophy is, it becomes apparent that it has always played a significant role in public life. As I noted above, *The Wretched of the Earth* became widely influential among racial justice activists, especially in the era of the Black Power and civil rights movements in the US (whose leaders and participants, it should be noted, were also producing works in the 1960s that can be read as philosophy, such as Kwame Ture and Charles Hamilton's 1967 *Black Power: the*

Politics of Liberation in America). *The Wretched of the Earth* remains widely read amongst activists who see themselves as engaging in anticolonial struggles today.

Neo-colonialism also became a major influence, especially among African anti-imperialist writers and Latin American economists thinking about the problem of dependency. Indeed, the book was considered such a danger to the US that the State Department famously withdrew aid for Ghana following its publication. Finally, the arguments that climate activists are making today — through public dialogue, social media posts and political action — are shifting society’s focus from the previous emphasis on individual behavioural change to tackle climate change, towards centring the role of fossil fuel corporations and other business interests in perpetuating the crisis.

Yet not all these political actors will recognise that what they are doing is a form of philosophy; indeed, many may even recoil at the description and reject it. Activists, for one, may think that describing their manifestos and appeals to the public as a kind of political and moral philosophy is abstracting away from the real-world concerns that motivate their actions. In the narrow understanding of philosophy I have argued against, the enterprise appears to be far removed from the on-the-street actions of obstructing roads and blocking access to government buildings. Some activists may disdain words and emphasise the importance of immediate action, following the oft-cited claim from Marx that the point is not to interpret the world but to change it.

What’s more, political activism requires practices and dispositions that may not always align with the open-ended truth-seeking attitude that many of us associate with philosophy. To motivate and mobilise participants, activists need to carefully design their communications to highlight the issues that foster and direct public anger to what they see as the causes of injustice. While this is not necessarily incompatible with the spirit of open inquiry, when the issue at hand is pressing and the need for action is urgent, activists may have no choice but to suspend further debate and opt for action.

Finally, activists do not always lay out all the relevant moral and ethical considerations for participants to engage with and may instead see their jobs as guiding the latter towards the conclusions they endorse.

Politicians and policymakers, on the other hand, may want to deny that underlying their policies and programmes are specific claims about right and wrong, fairness and unfairness, justice and injustice. Instead, it is common for these political actors to disguise (whether intentionally or not) what are in fact value judgements in seemingly value neutral language such as efficiency, growth, law and legality, public safety, social harmony, and so on. Even when values such as democracy, freedom and equality are invoked, as they inevitably are during election season, politicians often do so without defining these concepts and instead assume a shared understanding. Shying away from explicit moral argumentation appears to avoid controversy or at least glosses over divisions. And in an age where distrust of politicians is already at an all-time high, politicians are aware that attempting to advise voters on morality and ethics is unlikely to be well received.

In short, if philosophy can be found everywhere in public life, it isn't because real-world policymakers clamour for the title of public philosopher — far from it. Rather, it is because through their political rhetoric and political actions, these figures will inevitably make normative claims and promote particular ideas of justice and injustice, and their influence in the world makes it particularly important to recognise and evaluate them.

This is where academic philosophers can and ought to come in. Academic philosophers are trained to detect arguments about morality and value more broadly, even when they are implicit — indeed, even when those making the claims are unaware of their own moral or ethical reasoning. Academic philosophers are experienced in reconstructing the logic behind justifications for a particular action or policy, and in subjecting that reasoning to closer scrutiny. Even though, as I have argued, philosophy is not the exclusive purview of academic philosophers, they have the time and resources to hone their critical and analytic thinking skills. And, some of us also benefit from secure, paid positions that allow us to critique the flawed arguments of the powerful and lend support to the justified claims of the less powerful.

It seems to me, then, that this is one of the key contributions of academic philosophy in relation to public life: drawing out, reconstructing and scrutinising existing moral arguments that are already informing significant political work. This requires academic philosophers to relinquish any remaining notion that moral and political philosophy exists solely in the work we produce, and, until very recently, have almost exclusively studied. Instead, when proponents of widely used philosophical arguments choose not to acknowledge or take

ownership of them, those arguments may be hidden or exist in inchoate forms. Often, these arguments will come in language and formats considered 'unconventional' by academic philosophy's standards, such as pamphlets, manifestos, texts that appear as historical narratives or propaganda, social media posts, and so on.

Academic philosophers should also train future citizens in this exercise. When philosophy students learn to identify moral arguments in a text like Nkrumah's *Neocolonialism* and to subject these arguments to scrutiny, they may be better at recognising and scrutinising real-world philosophical arguments. When students are taught that the ideas of justice put forth in Fanon's *The Wretched of the Earth* are, just like any other ideas of justice in a philosophical text, up for debate, they may be better at critically engaging with the discourse of those whose political positions they may otherwise agree with. Rather than treating these texts as mere propaganda or situated narratives that should not be questioned, our young citizens would be alert to the public philosophy that is embedded in our social life. *And* they would be sensitive to the sophisticated moral-philosophical arguments presented by those who are too-often dismissed as 'just politically motivated'.

I've argued in this essay that philosophy, especially moral and political philosophy, has always played indispensable roles in public life and in real-world change. A central task of the academic philosopher is to use her position and skills to engage in critical dialogue with public philosophers, and to train future citizens to do the same. Doing so does not require denying the worth of academic philosophy, nor does it require setting up stark distinctions between us, the professional theorists, and them, the public who needs to be persuaded by our theories of justice. Instead, it simply requires an openness and curiosity toward the philosophy and practices that have shaped our world, and continue to do so, for better or for worse.

Public philosophy for gremlins

Dr Lorna Finlayson, University of Essex

A friend of my dad's used to have a saying which we may call, in his honour, the 'Julius Principle'. The Julius Principle holds that around 95% of all professionals are utterly hopeless at the thing they are professionals at. Teachers, doctors, plumbers, lawyers, officials of various kinds — you name it: useless bunch, whether because of lack of aptitude or because they are out for themselves, motivated more by their own advancement or convenience than by service to others, or because of some combination of the above. To find that the teacher is a poor communicator and has no affinity with kids, or that the doctor makes the wrong diagnosis, or that the plumber fixes one leak and creates another, should not be the surprising thing. What is surprising is the happy case where the professional actually knows what he or she is doing and does *not* screw it up.

The Julius Principle is somewhat similar to the more famous 'Peter Principle', which holds that every employee in an organisation 'tends to rise to his level of incompetence' (Peter and Hull, 1969). People who are good at their jobs will tend to be promoted based on their competence at that role or level, but this will not necessarily translate into competence at a higher level (for example, a competent educator may not be a good manager). If they *are* good at their new role, they will be promoted again. In the end, anyone who is any good at anything will end up doing something they are bad at. By contrast, the 'Dilbert Principle' proposed by Scott Adams (Adams, 1996) observes that people who are particularly *incompetent* at their jobs are often promoted into management roles, basically to get them out of the way. In the sage words of Dogbert (Dilbert's dog): 'leadership is nature's way of removing morons from the productive flow.' Sadly (as any academic can tell you), this does not remove the power of managers to wreak destruction on the working lives of others.

The Julius Principle differs from these others mainly in that it is stronger and more sweeping: the *vast majority* is condemned as useless. This might seem, at first sight, like a pretty

obnoxious sort of sentiment. Not only is it insulting to all the hard-working professionals out there, we might think, but it's simply not true. It might sometimes *seem* like everyone we encounter is an idiot or a crook and that things are always going wrong, but this is just a kind of Eeyore-ish cognitive bias. Things do go wrong, for sure, but those are just the times we're more likely to remember: we take for granted or do not even notice all the times when they work as they should. After all, if it were really the case that almost everyone is terrible at their job, then society would come to a screeching halt — lorries would crash into each other, surgeons would remove the wrong organs, bartenders would try to pour pints into the wrong end of the glass. Of course, all of the above have been known to happen. But mercifully, they are the exception rather than the rule.

Still, this only refutes an extreme and rather crude version of the Julius Principle. Competence and incompetence are relative concepts, after all (as is corruption or malevolence). This might seem only to raise another problem: relative to what? Almost by definition, most people can't be *below average* at what they do. Thus the suspicion may arise that this is really no more than a bit of snootiness: a way of saying, 'Other people are not as good as *me*.' Whether true in a given case or not, this is not a terribly attractive posture to adopt. In any case, it is merely a statement about the speaker's individual superiority, which (whether real or not) has no obvious wider significance or interest. I often think this about a widespread habit of thought and speech which lays the blame for this or that problem with 'people' *simpliciter*. You hear it all the time: 'people are so stupid', or 'people are selfish' — and *that* is why we can't have nice things. Such statements are ultimately conservative or authoritarian in character, it seems to me. If the problem is simply (other) *people*, then what is the solution? Dissolve the people and elect another?¹

But while we have reason to be wary of the hidden politics behind clichéd statements of casual misanthropy, it's equally important not to fall for a kind of dogmatic 'moderate'-ism that regards any wide-scope or far-reaching criticism as suspect by default. For one thing, maybe the philosopher and 'critical theorist' Theodor Adorno had a point when he said something that would strike most people as obviously false: that only exaggerations are true.² That is, sometimes what is technically an overstatement (and in that sense false) conveys an intended meaning or truth better — and hence more *truly* — than a more caveat-laden, round-the-houses version which would be more strictly accurate. Take a statement like: 'Tory politicians have no empathy for people in poverty.' Probably, at least some of them do, in a way, at least a tiny bit. Maybe they feel a little twinge of humanity now and again, on the rare occasions

when the suffering of the people in question is forced on their attention, for instance when they encounter a suitably sympathetic and presentable poor person (and maybe that's why they prefer to keep such encounters to a minimum). But a revised statement acknowledging all this — for example, 'Most Tory politicians, most of the time, have very little empathy for the poor' — would, I think, do a worse job of conveying the reality. In other words, it would be *less true*. And even if you don't accept that, there is the consideration that apparently extreme or sweeping statements might not always be exaggerated at all. What if the truth *is* extreme?

We can make pretty good pre-theoretical sense, I think, of the difference between seeing the contemporary social world as one in which things basically tick along okay (albeit with a few glitches here and there), and seeing it instead as basically dysfunctional. This dysfunctionality may not necessarily be relative to any other past or present society, but relative to how it *might* or *could* or *should* be, as well as relative to what it *claims* to be — societies invariably find ways to present themselves in a flattering light and to disguise their uglier features. From the perspective of this second position, the Julius Principle makes a great deal of sense. For all sorts of reasons, we might take the view that the way our institutions are designed, the forces to which they are responsive (profit, for instance, rather than the imperatives of human need or public good), and the ways in which their personnel are selected and trained, are not conducive to producing even a narrow or minimal competence. Let alone faring well in a more substantial sense relative to those institutions' stated or widely presumed functions (the Julius Principle as I've formulated it is ambiguous between these somewhat different claims). It might be thought that only the second of these suggestions is respectable. While institutions and their practitioners may not do what they 'say on the tin' (so to speak), surely they are competent in the first, more minimal sense. For example, a journalist will at least know how to write correct sentences, and to deal with editors, and to consult sources, and so on, but may nevertheless be bad at their job in the sense that they are hopeless at thinking critically and holding the powerful to account (the function central to journalism's self-conception). But even this strikes me as an overly optimistic (or what Adorno called 'affirmative') assessment. Moreover, it's not possible cleanly to extricate a thin sense of 'competence' from the more loaded idea of doing a job either in the way we think it should be done, or in a way which lives up to its self-professed standards and rationale. It may be, for example, that journalists often can't write properly and don't bother to check facts because of such factors as the rampant nepotism and elitism of that profession, and the fact that their work is in practice not held to account in the sense of having to be accurate, but only in the sense of flattering or leaving undisturbed the relevant vested interests.³

To put it another way, something like the Julius Principle seems to fit well within — and indeed to be logically dictated by — a 'critical theory' of society, in the sense espoused by thinkers such as Adorno. The conviction of this strand of social theory is (to put it in the crudest terms) that things are both very bad, and bad in ways that are *systematic* as opposed to local and isolable. From that point of view, there would be an incongruity involved in being too 'affirmative' of any portion of social reality. How can you say (as critical theorists of the Frankfurt School do) that society 'as a totality' is the problem — that it is rotten to the core, not just in need of a little reform here and there — and then go on in the same breath to say that what the teacher training colleges do is jolly good and that most teachers do a sterling job? It doesn't really make sense. There is a strong inhibition against noticing this tension, however, which may have a lot to do with a desire not to be seen as elitist or 'anti-worker'. Especially on the political left, it's almost a compulsive tic to say, even when drawing attention to some systemic failing or pervasive abuse, that ('of course') the vast majority of teachers (for example) are excellent and dedicated people doing a vital job. There are some respectable reasons for this, no doubt. You don't want to make blanket statements where they are not true (for example, that all teachers are paedophiles). You don't want to blame individuals for what are really structural problems. But that doesn't remove the fact that individuals are not neatly detachable from structures: individuals are formed by structures and work within them. It doesn't make it true, for instance, that most teachers do a great or even competent job for the kids — nice as this would be. That, in my experience anyway, is not the case (and I don't think I'm alone here: it's telling that when people opine about how great teachers are, they often seem to have in mind some abstract *idea* of teachers, not the actual ones they remember being taught by themselves). The good teachers were the exceptions: 5% sounds about right to me.⁴ Generous, even.

Which brings me, at last, to philosophy. Philosophers — more, I think, than members of many other professions or disciplines, for whom this sort of question either seems misplaced or the answer obvious — periodically ask themselves about the nature of their proper role and distinctive contribution relative to wider society of 'public life'. This, in my view, is a good and necessary thing to ask about. But the ways in which many philosophers ask and try to answer the question of the relationship between philosophy and the wider world, it has always seemed to me, encode a heavy bias towards the status quo. If, for example, we identify this question with that of how philosophers might influence policymakers, we reveal what I would consider an unduly optimistic estimation of the receptiveness of those in positions of power to being influenced in a positive direction (or indeed, to being influenced by argument at all). More fundamentally, to frame the issue in this way takes for granted the

existing structures of power and decision-making which are precisely what need to be called into question.⁵ But I have a still more basic worry that is less often voiced (maybe because voicing it would involve coming uncomfortably close to calling your colleagues a load of idiots or lickspittles — which is never likely to go down well). The worry, bluntly put, is that philosophers might not actually be very fit agents to bring about a positive influence or to make a 'contribution' to public life that amounts to more than an ideological greasing of the wheels of the political and social status quo. What if — as in the well-known comedy sketch — we're the baddies (or the fools, or some combination of the two)?

This might seem like a perverse as well as a rude thing to say. Even if philosophy and philosophers are so terrible at doing what they purport to be doing — that is, contributing clear-eyed and principled understanding that is unafraid to go against power's grain — then isn't that all the more reason to ask how we might do better? In one sense, the answer to this is obviously 'yes'. Setting out and pursuing an idea of how philosophy might be done in a different and better way, one geared towards the end of emancipation from oppression rather than the reproduction of the status quo, is after all central to the project of critical theory.⁶

The source of the confusion here, it seems to me, is that the question of what role philosophy should play relative to politics and public life is crucially ambiguous. On one reading, it is simply equivalent to the question of what is a good way to do philosophy, if you are interested both in interpreting the world and also in changing it. And according to the sort of answer associated with critical theory, at any rate, the way to do that is resoundingly *not* to try to think of ways to be 'helpful', that is, to try to make the system as it exists work better. The social status quo is so deeply and systematically dysfunctional that it cannot be made to work by a few reforms — and is in any case so thoroughly corrupted that for it to 'work' might actually mean something still more horrifying than our social world already is. The only decent response, in that case, is to try to be like a gremlin in the machinery: to aim to intervene in the social world in ways that disrupt its smooth action, that gum up its gears — or, if that is impossible, to stand back from it and refuse to be drawn into the meat-grinder.

On the other reading, by contrast, the question is about what overall role philosophers in general, or academic philosophy as an institution, might feasibly and desirably play in relation to the social world as it currently exists. Though superficially similar, this is a quite different question, one that presupposes a view of either or both the social world in general

and the institution of academic philosophy in particular such that the latter might feasibly come to play a desirable (even a desirably transformative) role vis-à-vis the former. Only on a basically affirmative assumption of this kind is it possible to conflate the two questions.

In other words, many discussions of the relationship between philosophy and public life suffer from a failure of realism when it comes to the way the agent of possible intervention — in this case, philosophy — is thought about. This failure mirrors the way in which people often contemplate the question of 'humanitarian intervention' by liberal states (as if these are entities that can realistically be expected to intervene in the rest of the world in ways that are remotely 'humane'). Even those who cast themselves as radical critics of society are not always so critical when it comes to the institution of philosophy. But the basic problem with the question of what philosophers can do for public life is that philosophers — or at least, about 95% of us, 95% of the time — are awful. You can aspire to be a gremlin in the system, but then you must also expect to be a gremlin in the department.

There are exceptions, of course: philosophers who try to do something other than interpret the world, even if actually changing it seems to be ever further out of reach. But the point is that within the institution of philosophy in society as it is presently constituted, it is no accident that these will be the exceptions rather than the rule. Why is such a pessimistic view of philosophy in order? The long way round in terms of answering that question would be to do a comprehensive critique of the state of philosophy. But the shorter way is just to point out that, if you take a dim enough view of the social world in general, then a rosy view of anything in it — certainly anything as large and as embedded within the 'establishment' as the academic field of philosophy — makes questionable sense. If you really think that society as a whole is the problem, that the rot goes deep, then there is at the very least a strong presumption in favour of the expectation that philosophy — most of it, anyway, most of the time — will be no exception, but more of the same. The most we can do is to strive to be, as far and as often as possible, among the exceptions.

Notes:

- 1 The phrase comes from Brecht's satirical poem of 1953, 'The Solution' (*Die Lösung*); see Thomson (2006: 239).
- 2 Adorno makes versions of this point on a number of occasions; see, e.g., aphorisms 29 and 82 of his *Minima Moralia* (Adorno [1951] 2005).
- 3 Of course, there is also a third possible sense of being 'good' or 'competent' at a job, which is related to what we might take to be the *real* (as opposed to the official or stated) social function of the institution or role in question — for example, we may judge that the real function of journalism is to prop up existing power structures and to 'manufacture consent' for them (cf. Chomsky & Herman 1988), and that most journalists perform this function quite effectively.
- 4 Again, it might be that teachers are quite effective at doing what they are *really* there to do: to contain, subdue and browbeat children into submission until they are ready to be released into the workforce. I leave this question open.
- 5 I have developed this point at greater length elsewhere (Finlayson 2015).
- 6 Cf. Raymond Geuss's (1981: 58) description of the characteristic aspiration of critical theory as to be 'the self-consciousness of a successful process of enlightenment and emancipation.'

References:

- Adams, S. (1996). *The Dilbert Principle*. HarperBusiness.
- Adorno, T. (1951). *Minima Moralia: reflections from damaged life*. Reprint, London: Verso, 2005.
- Chomsky, N and Herman, E S. (1988). *Manufacturing Consent*. New York: Pantheon Books.
- Finlayson, L. (2015). *The Political Is Political: conformity and the illusion of dissent in contemporary political philosophy*. Rowman & Littlefield.
- Geuss, R. (1981). *The Idea of a Critical Theory*. Cambridge University Press.
- Peter, L.J., and Hull, R. (1969). *The Peter Principle*. Reprint, London, Pan Books, 1970.
- Thomson, P. (2006). 'Brecht's Poetry', pp.225-241 in *The Cambridge Companion to Brecht*. Cambridge University Press.

Identity dialogues

Professor Tariq Modood, University of Bristol

Contemporary anglophone political philosophers look to ground liberal politics and social justice in what free, rational individuals would collectively agree to after discussion. The works of John Rawls (1921–2002) is foundational to this approach. The conditions or circumstances in which such dialogue takes place, though, are critical. The discussants should be able to focus on what is good for individuals in general, or to put it differently, what all individuals would want after reflection, not on what individuals like themselves would want. They must think selflessly.

Rawls designs a thought experiment the centrepiece of which is what he calls the 'veil of ignorance' (Rawls, 1971). For the deliberation of individuals to lead to the discovery of social justice or 'fair terms of social cooperation' they must be made ignorant — stripped — of their specific identities such as their gender, class, nationality, culture, religion and so on. So none of the reasoners knows for example whether they are rich or poor, black or white, Christian or Muslim etc. So no one will risk favouring laws and policies that unduly favour a particular class, race or religion, in case — when at the end of the deliberation the veil is lifted and they (re)learn who they are — it turns out they are not of the group they favoured.

Rawls's claim is, then, that the principles of social justice can be only worked out by individuals, intellectuals, lawmakers, benign governments etc to the extent that they approximate to being self-less or identity-less reasoners. That, however, means that dialogue among such individuals is not necessary because, stripped of all their differences, such reasoners are identical.

One reasoner can, in theory, produce the just solution without there having to be a dialogue amongst all the citizens. Moreover, behind the veil of ignorance the debate makes no difference to what is valuable in the product of the debate. The product — the principles that a diverse society should live by — are not influenced by who is or is not included in the debate and so they remain the same however the debate goes. That is to say, the principles are not influenced by the debate and could indeed have been known without any dialogue having taken place. More precisely, they are known by reason not by dialogue or by who participates in the dialogue. In his later work, Rawls acknowledged that in the actual world of liberal democracies

people do not have to have this self-ignorance. He allowed that people are often guided by an ethical philosophy or a religious perspective. But as these were various and were the source of deep and intractable disagreement, they were only permissible as sources for politics if they led to conclusions that would be consistent with the principles of social justice of the kind that his original method was designed to establish. Philosophies and religions that did not lead to liberal egalitarian principles were not acceptable sources to reason from.

Dialogue as used by multiculturalists

This is not, however, how multiculturalists have approached political theory. Dialogue rather than abstract reasoning by a sole reasoner or identical identity-less individuals has motivated multiculturalists, who hold that politics must find ways to work with different perspectives without assuming that all will converge on a single rational point or consensus. Multiculturalists assume that the context for politics is already thoroughly imbued with dominant ways of thinking and doing. That is, imbued with cultural orientations such as national history and language, with religious and/or secular perspectives, with institutional norms etc and that these contextual factors cannot be made abstract to identify a set of culture-free problems. Moreover, the relationship between the relevant parties is likely to involve domination–subordination, inclusion–exclusion and that the weaker or newer party is likely to lack recognition or be misrecognised (Taylor 1994). Dialogue rather than identity-less reasoning will be relevant here for at least three reasons.

First, effort must be put into reaching a cross-cultural understanding to help find a solution to the problem or to establish a problem-solving principle. It is not just a question of taking material interests into account but a matter of (re)designing the shared public space and rules of conduct so that diverse cultural commitments and needs are explicitly addressed. This helps ensure that the public space does not simply reflect the dominant culture but also accommodates and welcomes new or marginalised minorities.

Second, this means that the solution is genuinely open. By this I do not mean that 'anything goes.' Rather, that the solution cannot be predicted in advance — that is, there is no predetermined answer waiting to be discovered, as there would be for a maths problem. The dialogue makes a difference: it contributes to a growth of understanding that is genuinely novel or supplementary, and the quality or character of the dialogue is dependent on the

participants. This is not simply in terms of their power of reasoning but in terms of 'where they are coming from' — that is, different parties would produce a different outcome.

Third, the dialogue is important not just in discovering an outcome but in building a *relationship* of trust, co-operation and belonging. These three reasons make the dialogue quite different to the 'behind the veil' reasoning of identity-less reasoners.

The multiculturalist political theorists I have in mind include Iris Young and her assisting people, who are required to understand themselves as oppressed and to discover themselves in collective identities such as woman, black or gay; and to thus develop a liberatory identity and group politics and to use it to engage with other groups to institute a new form of democratic politics (Young 1990).

James Tully has continually emphasised that cooperation under conditions of deep diversity or 'multiplicity' requires a 'multilogue' (Tully, 1995). Bhikhu Parekh explicitly makes intercultural dialogue central to his conception of multiculturalism. His interventions in relation to *The Satanic Verses* affair are exemplary (Parekh 1989). While recognising that in such cases the majority dominate public discourse — and often in a manner that is not conducive to dialogue or mutual learning — he argues that multiculturalism is not about a relativism of allowing each minority to live as it wishes (Parekh 2006). Rather, it is about ensuring that there is a genuine dialogue, that the minority is allowed to express its point of view. Such dialogues inevitably have a majoritarian or status quo starting point because, even while wanting to express unfamiliar sensibilities and bring in new arguments, minorities are primarily trying to persuade the majority. This often takes the form of a minority claiming that their goals are not so different from those the majority has sought for itself at various points. In making this argument, the minority must justify its position by appealing to, and at times seeking to modify, the existing 'operative public values' that shape public debate and determine what is considered legitimate or reasonable within the current polity (Parekh 2006).

For such multiculturalists the principles of social justice are not known in advance or simply by reason but are arrived at by conflict, *learning*, dialogue and negotiation in circumstances of inequality and minority-claims making. This is not a 'veil of ignorance' activity taking place away from conflict with an isolated reasoner or an assembly of identity-less, self-less reasoners. For the multiculturalists the dialogue is claims-based and contentious, and based on *identity-assertion* (relative to other identities), not identity effacement. The dialogue seeks to get beyond — though

it may never fully reach — the conflict or challenge which provoked it, by urging for acceptance of the excluded, those deemed inferior, and misrecognised, by the formation of new, inclusive, hyphenated and overlapping identities.

The dialogue comes into being because of identity-based claims; it proceeds by recognising identities; and its goal, its teleology, is the construction of new identities and new relationships, which are not reducible to redistribution (Modood, 2013). The success of the dialogue is dependent on the development of overlapping identities. It is not just dependent on, as in Rawls, deriving an ‘overlapping consensus’ from agreed principles of social justice independent of our self-perception as a society trying to achieve a new, respectful inclusion of ‘difference’. I hope it is evident how important this kind of dialogue can be to overcome the spiralling mutual mistrust and polarisation that is a feature of our current politics and public exchange.

Public intellectual engagement

The kind of macro-level dialogue that I am speaking of can also be understood as a form of public intellectual engagement. One of the best-known statements on the nature of public intellectuals is by Edward Said (Said 1996). Said writes ‘of the intellectual as a being set apart’, angry and oppositional, a critic of all worldly powers. Public intellectuals marry the academic’s commitment to intellectual values but combine it with a critique of injustice, which is aimed not just at fellow specialists but as wide a public audience as possible. I can offer my understanding of public intellectual engagement by relating to Said’s idea of a public intellectual, which I find too one-sided and painted too starkly.

An example of this one-sidedness is the detachment from society that Said attributes to public intellectuals. He argues that their aim is to uphold universal ‘standards of truth about human misery and oppression... despite the individual intellectual’s party affiliation, national background, and primeval loyalties’. Of course, this kind of integrity is what one requires not just from public intellectuals but from all professionals such as academics, doctors, judges, engineers and so on. It also does not mean that public intellectuals have to be lesser members of their society; or that they share less understanding and concerns with their co-ethnic, co-religionists or co-nationals; or do not care for the well-being of their groups, including if they protest when they think injustice is being done by their groups. Yet, Said describes public

intellectuals as, and indeed exhorts them to be, 'outsiders and exiles' and admiringly quotes Theodore Adorno: *'It is part of morality not to be at home in one's home.'*

Said notes that 'because the exile sees things both in terms of what has been left behind and what is actual here and now, there is a double perspective that never sees things in isolation.' Thus despite presenting a self-image of the intellectual as standing outside or above the society they are engaging with, Said recognises there must be a commitment to a people(s) or concrete institutions and practices — not just to abstract principles like truth or justice or humanity.

My point is that commitments to groups, people, causes, institutions, one's country and so on are not a nuisance, or incidental to an engaged public intellectual. These commitments are as essential to the public intellectual as the commitment to intellectual integrity. A public intellectual should care about a people, a place or a cause and not just about being an intellectual (demonstrated effectively in relation to George Orwell and Albert Camus in Walzer, 2002). The public intellectual must have a home, but this commitment must not be unaware of or incompatible with an equally strong commitment to intellectuality. Just as, of course, there must not be an unconsidered commitment to certain intellectual points of view and theories, including those which have the prefix of 'critical' (a prefix that seems, to some, to be a badge of adherence rather than something to deconstruct). The public intellectual endeavour is to engage in and lead a society's moral, ethical and political conversation. And while some 'outsider' features can offer some epistemological advantages (and no doubt some deficiencies) one needs to be part of the society that one seeks to engage.

Said cites the African-Americans James Baldwin and Malcolm X as exemplars of public intellectuals. Yet they were individuals who knew which side they were on. They were outsiders to certain structures of power, but they belonged and were committed to the well-being of the groups they represented and believed they were part of. It is most unlikely that they endorsed Said's motto of 'never solidarity before criticism'. Moreover, when it comes to multiculturalist public intellectuals they are likely to belong to more than one group and so unlikely to be either wholly insiders or outsiders. This chimes with Said himself, a Palestinian-US scholar with a lifelong commitment to the Palestinian people.

The public intellectual, then, has to negotiate critical outsidership and epistemological insidership and belonging, solidarity and rootedness. She does not need to give up entirely on

her social roots; indeed, to do so is to risk losing an important understanding and sympathy for her group or society as well as trust and standing with the group and/or society. So, rather than exile, an unconsidered loyalty, or aloofness from one's group, she instead should move towards developing multiple belongings and possibilities of dialogue.

A similar one-sidedness characterises Said's distancing of public intellectuality from 'specialisation' and 'expertise', overlooking that a public intellectual must be part of an intellectual discipline. He argues that 'the particular threat to the intellectual today... [is] an attitude that I will call professionalism', which he describes as treating intellectual work as just a job, on a nine-to-five basis, the demotion of an intellectual vocation for what today is likely to be called 'work-life balance'.

Said also worries about intellectuals seeking acceptance, prestige and honours. I agree that some university institutional cultures — such as those in Britain today — encourage a narrow scholasticism, typified by an appreciation for disciplinary jargon and a disregard for clarity, but Said is too dismissive of professional expertise. He overlooks the fact that scientific and professional expertise improves material living standards, public services and personal well-being. It is about engaging with the pressing needs of individuals and communities, such as seeking a cure for cancer, reducing world poverty or contributing to the advancement of 'the knowledge society' with an objective to improve regional and national productivity and promote technological innovation. We may agree with Said, however, that such activity is not public intellectual engagement. And as for an intellectual seeking honours — while that cannot be the primary motive, there is — and should be — honour and recognition, and social status in public intellectual engagement. And it is odd that Said, who received such acclaim — including being chosen to deliver the prestigious Reith Lectures, in which he presented the views I am discussing — should fail to mention it.

A better understanding of how the professional and the personal interplay — what might be called 'honourable ambition' — is captured in this description of public intellectuals as 'those who live with the tensions generated by the contrasting pulls of specialist focus and peer recognition, on the one hand, and on the other the risks and thrills associated with being known as someone who addresses a much wider range of publics on issues of general concern' (Kenny, 2008: 7). Of course, these elements are not always in harmony, but it is important to recognise the reality of competing motives alongside defining public intellectuals in ascetic and purist terms.

Despite Said's tendency to sometimes express himself in a one-sided way, he also offers a more complex characterisation and is closer to the mark when he does so. For example: 'There is therefore this quite complicated mix between the private and the public worlds, my own history, values, writings and positions as they derive from my experiences, on the one hand, and on the other hand, how these enter into the social world where people debate and make decisions about war and freedom and justice'. There is however one issue on which I do not simply think Said has a preference for one-sidedness but where our views collide: where Said's combination of intellectual elitism and wanting to be on 'the same side with the weak and unrepresented' becomes unstuck.

When I previously referred to Parekh as a multiculturalist public intellectual I evidenced his interventions in relation to the crisis around Salman Rushdie's novel, *The Satanic Verses*. In these he argued for the importance of listening to and conversing with Muslims who are angry and feel they have been humiliated. It happens that Said too refers to this crisis and states that to have failed to have defended this novel is 'to betray the intellectual's calling'. This is because 'uncompromising freedom of opinion and expression is the secular intellectual's main bastion'. It may be that by 'secular' Said does not mean non-religious but as someone who does not have a 'belief in a political god' or 'a total dogmatic system'. If so, I share that view and have warned of 'the danger of ideology' in discussions of multiculturalism (2013: 118-122).

Yet, my view is that we should challenge secularist one-sidedness just as much as religious militancy, avoiding the appropriation of 'truth' or 'light' for any exclusivist perspective. Instead, we should foster a diverse and broad spectrum of both secular and religious intellectual voices. Similarly, a public intellectual must be politically free to be left-wing, right-wing, centrist, religious, secular and so on — and of course to argue for her point of view by attending to other, especially dissenting voices, and respond to objections and critiques. Public intellectual engagement is of course political, not neutral, but it is a dialogue or a multilogue of complementary and contending intellectual-political positions, which together weave together a diverse public sphere.

Dialogue can take more than one form. For example, most of Plato's early dialogues, referred to as eristic, take the form of an interrogation. Socrates is portrayed engaged in a hostile series of questions aimed to show that his interlocutor, often a well-known 'Sophist' or public philosopher, does not know what he is talking about. Socrates sets out to destroy the argument of his interlocutor and to discredit him as a teacher or a learned

person or an authority on wise conduct. These dialogues typically end in a breakdown, with Socrates's opponent alleging that Socrates is constantly twisting his words for his own self-aggrandisement so there is no point carrying on.

The other kind of dialogue, of which *Republic* is the most famous example, is more like an interview and consists of a rational cooperation to discover the truth. It is most important that the first kind of dialogue does not dominate the public sphere but is subordinated to the second, as in Plato. This duality must not be reduced to oppositionalism, deconstruction or the merely 'critical'. A public intellectual must be creative and constructive as well as critical. In this regard, Andrew Gamble's description of an eminent British multiculturalist effectively captures the genre:

'The political theorists of multiculturalism such as Bhikhu Parekh... have been active participants in politics in the sense that they seek to advance the political education of citizens, by articulating choices, framing questions, offering alternatives, and challenging orthodoxies and entrenched attitudes. They address themselves to the public, not to [just] coteries of experts, or office holders. They are essential builders of the public sphere...' (Gamble, 2015: 297).

References:

- Uberoi and T. Modood (eds), *Multiculturalism Rethought*, Edinburgh: Edinburgh University Press, pp. 273-299.
- Kenny, M. (2008). Britain's Anti-Intellectual Intellectuals: Thoughts on Stefan Collini's *Absent Minds*. *Political Studies Review*, 6(1), 1-13.
- Modood, T. (2013) *Multiculturalism: A Civic Idea*, Second Edition, Polity Press.
- Parekh, B. (1989) 'Between holy text and moral void', *New Statesman and Society*, 24 March, London: pp. 29-33.
- Parekh, B. (2006). *Rethinking multiculturalism: cultural diversity and political theory*. Second Edition. Macmillan.
- Rawls, J. (1971). *A Theory of Justice*. Harvard University.
- Taylor, C. (1994). *The Politics of Recognition*. *Multiculturalism: Examining the Politics of Recognition*. (ed) A. Gutmann. Princeton: Princeton University Press.
- Tully, J. (1995). *Strange multiplicity: Constitutionalism in an age of diversity*. Cambridge University Press.
- Walzer, M. (2002). *The company of critics: Social criticism and political commitment in the twentieth century*. Basic Books.
- Young, I. M. (1990) *Justice and the Politics of Difference*. Princeton: Princeton University Press.

Philosophy matters

Professor Debra Satz, Stanford University

I find the question 'does philosophy matter to public life?' puzzling. On one hand, I think, how could anyone believe that ideas and methods from philosophy do not matter to our individual and collective lives? Our world has been rocked by claims about values — from the claims of human equality to those involved in the birth of liberalism to debates about the appropriate role of religion in public life. Philosophical reflection addresses questions about values and has helped guide human culture as societies (and individuals) have grappled with such issues.

On the other hand, the charge that 'the philosophers have only interpreted the world, in various ways; the point, however, is to change it' goes back to Marx (1845) and, indeed, much further back than that. Today's political realists have added to that charge, claiming that abstract philosophical ideas, by themselves, change nothing (Geuss, 2008). And in a world where too often it is those in power who claim a monopoly on what is right, my students worry that philosophy is only talk.

Many people, therefore, do not find the question puzzling. Some reject the kinds of questions contemporary philosophers ask; some bemoan the professionalisation and specialisation which turned many people away from asking big questions; some think the focus on 'ideal norms' and 'abstract arguments is pointless; and some think that without an army, philosophy departments might as well devote themselves to the self-address of the eternal (for all it matters). These are not baseless concerns. But they do not state that different philosophical ideas about values — about what matters, about what we ought to do, and about how we should live together — have not played a large role in public life. Here, I will argue that such ideas are not mere talk but have given shape to human aspirations and channelled individuals' and societies' energies in particular directions.

I'll point to two broad examples to make my case about the importance of philosophical ideas for public life; list a shorter selection of specific examples from contemporary philosophy; and in the light of this, revisit the question.

The birth of liberalism

Modern liberalism emerged in the seventeenth and eighteenth century in the aftermath of the religious wars in Europe and amid challenges to the divine right of kings. Rather than privileging the idea of a natural corporate social order based on status, liberals argued for the equality and freedom of individuals. But as soon as they emerged, liberal ideas were subjected to at least three pointed attacks.

The first line of attack argued that liberal individualism was incompatible with social order. Early modern European societies were held together by bonds of status which assigned differential rights and duties to groups, and by widely shared religious doctrines. In the absence of such shared beliefs, and without a given social hierarchy, how could a society of individuals be stable? Why would equal and free individuals obey the law?

Starting with Thomas Hobbes — who argued against the view that individuals were the king’s property and therefore must give allegiance — and running through such thinkers as Locke, Rousseau and Rawls, philosophers developed a set of answers to this challenge. They argued that under conditions of scarcity, rational individuals will find it in their interest to form a society with rules and to empower a sovereign — the Leviathan — to enforce those rules. Of course, these thinkers differed on the extent of the powers of the sovereign, the nature of the rules that would be agreed to, and the set of alternatives that should be considered. But they each argued that a political state is justified and will be stable to the extent that it improves individuals’ well-being against some set of alternatives while preserving their freedom and equality. They emphasised the importance of liberalism to rule of law, decentralised market exchange, and widely distributed property.

The second line of attack argued that liberalism was incoherent. Liberalism asks people with strong personal convictions to refrain from using state institutions to further those convictions. It does so out of a commitment to a principle of neutrality, the idea that the state should be neutral between different conceptions of the good. But isn’t liberalism itself such a conception of the good? If so, then why should non-liberals accept liberalism? Many philosophers during liberalism’s rise responded to this line of attack.

And John Rawls more recently provided a strong argument (Rawls, 1971). He argues that because under conditions of freedom, people will not converge on a shared view about

what is a good life, we should not use such conceptions to allocate social resources such as money, power and opportunity. We should instead derive principles of justice that treat all individuals fairly. Liberalism should be neutral between individual views of value and meaning. Of course, liberalism cannot be neutral to its own presuppositions, in particular, its claim that all individuals are free and equal. However, Rawls also argued that even those who do not support such neutrality towards conceptions of the good life on liberal grounds — because they believe there is one truth given to us by God — may subscribe to such neutrality on other bases. For example, appealing to pragmatic grounds, or if the conclusions of such neutral reasoning are compatible with their religious precepts, etc.

The third line of attack was that inequalities between individuals are naturally produced and are inevitable in a free society. Here too, political philosophers explored the ways that equality and liberty could be reconciled. From Locke's view that all are entitled to 'equal freedom', to Rousseau's idea that natural differences by themselves do not entitle some to more than others, philosophers hammered out different views of what freedom and equality actually amount to. It is critical to remember that in pre-Lockean Europe natural hierarchy was the 'common sense' of the time (although there were critics) and it took a concerted philosophical attack to articulate and defend a coherent alternative.

Liberal philosophers alone, of course, did not make liberalism. But they strengthened the alternative to the 'common sense' of the old social order and connected with popular aspirations and social movements. Just think about some of the policies that came out of liberal thought: the extension of the franchise to women, propertyless men and former slaves; the rise of a social safety net to protect citizens from certain kinds of risk; and the creation of a private sphere of life.

Neoliberalism

Philosophical ideas also shaped the doctrine of neoliberalism which rose to prominence in the decades after the second world war. The term neoliberalism, of course, has been used in many ways. It is sometimes described as an empirical doctrine, positing a relationship between markets and growth and contrasting that to the inferior relationship between government and growth. But its early champions were quite explicit about the *philosophical* commitments of their theory. Indeed, the success of neoliberalism arguably depended on

the integration of its economic model with a moral framework (Bowles and Carlin, 2021). That framework privileged 'freedom from' government coercion along with a view of human beings as primarily, although not only, self-interested. This latter assumption of self-interest applied not only to the owners of companies but also to government representatives and functionaries. Moreover, according to neoliberalism, these government agents held much greater centralised power than mere capitalists.

Given these two moral/political commitments it is not surprising that neoliberals sought to reduce the scope of government and empower individuals. Indeed, neoliberals were suspicious of most forms of collective action. They thought that governments should play a much smaller role in capitalist economies than it had during the mid-twentieth century and that markets offered far better means of achieving not only economic growth but individual freedom. They also sought to articulate and defend a theory of the proper role of government.

Hayek's *The Road to Serfdom* (1944) was written during the second world war, at a time when fascism and a totalitarian socialism seemed likely to push out liberal democracies. Hayek was especially worried about centralised government regulation and planning which he believed was not only inefficient but a threat to freedom. Governments do not have access to the essential information that individuals have about what they want and what is good for them — although Hayek and Friedman worried that all of us are remarkably ignorant about what is good for us. Instead, experimentation, choice and decentralisation are the best means we have of learning. Hayek launched a major attack on socialist thought that shaped much mid-twentieth century thinking.

Although Hayek was trained as an economist, his writings are also works of political theory; the same is true of James Buchanan and, to some extent, of Milton Friedman. And the doctrine of neoliberalism succeeded not simply because it centred on an economic theory ('free' markets), but because it articulated a plausible theory of human behaviour, a view of the proper role of government, a set of emblematic policies, and an attractive (if controversial) view of individual freedom (Burgin, 2012). Friedman was especially good at popularising this vision of freedom and helped successfully deploy it to convince Richard Nixon to end the military draft and institute a volunteer army (Appelbaum, 2019). Friedman later recalled his role in Nixon's decision as his 'most important' policy achievement (Friedman, 1995, p. 38).

Both liberalism and neoliberalism, in my view, give testimony to the power of ideas. Moreover, these ideas were debated, tested and transformed in ways that are familiar to philosophers.

No philosophy in public life?

Of course, there are a host of other areas where philosophers and legal scholars have recently played important roles in shaping public policy debates: animal welfare; disability accommodation; drug policy; punishment; effective altruism; universal basic income; tax policy; freedom of speech; healthcare ethics; and campaign finance law. But to say they have played important roles does not mean that their views prevailed or that they won over most of the public. Or that their views were even correct.

Conclusion

Given the outsized roles of philosophical ideas and argumentation in the rise of liberalism and in the articulation of neoliberalism, and the examples of other areas where philosophers have made contributions, why is there so much scepticism? Why is the role of philosophy in public life even a question? I can think of three reasons.

First, it is undoubtedly true that it is not by the public’s reading of philosophical arguments in professional journals that philosophers make their mark. Instead, philosophical ideas are spun out in classrooms, law schools, policy arenas and places where people get together and try to reason about what to do. There is a continuum between what everyone does and what philosophers do — we try and consider and make sense of our options and evaluate them. Where philosophers excel is in the business of argument and counterargument, logical inference, drawing connections, and making distinctions.

Second, while I have suggested that ideas influence society, it is of course true that society and social conditions influence ideas. But the claim that ideas don’t matter at all — that the causality goes in only one direction is implausible: societies with roughly similar social conditions have taken markedly different trajectories driven in part by ideas — including religious ideas about the world. Max Weber’s famous formulation states: ‘Not ideas, but material and ideal interests directly govern man’s conduct. Yet frequently the “world images”

which have been created by "ideas" have, like switchmen, determined the tracks along which action has been pushed by the dynamic of interests.' (in Gerth and Mills, 1958, p. 280). Weber's key example was the rise of capitalism in the Protestant west and not in the more advanced civilization of China.

Third, although ideas matter, how best to implement ideas is not the core provenance of philosophy. You cannot make one move from Kantian or utilitarian theory to conclude how best to regulate the internet, or how healthcare should be allocated or whether to legalise marijuana. Philosophers have something to contribute to the assessment of alternatives and can help clarify the multiple and sometimes conflicting values at stake in our options. But we need to work with others with other kinds of expertise if we want to contribute to policy — and narrow graduate training and professionalisation don't make such collaboration easy. Such collaboration is, in my view, quite important, if we are to help develop a compelling alternative to neoliberalism.

To clarify — I don't think the only two ways of contributing to policy are top down or bottom up. Even bottom-up approaches will find it helpful to sometimes employ abstract reasoning to clarify a conflict and to help guide us. And there is a place for 'ideal' theorising that stretches the limits of our sense of possibility. Who thought, after all, in the late eighteenth century that slavery was wrong and should be abolished everywhere? As Adam Hochschild describes it in his wonderful book *Bury the Chains: Prophets and Rebels in the Fight to Free an Empire's Slaves* (Hochschild, 2006), the idea seemed 'totally utopian, crackpot, wildly too optimistic'. So too did the equality of women when Mary Wollstonecraft argued the case in 1792. But such ideas have helped shape the arc of the moral universe. And reflection on these and other ideas is also part of philosophy.

References:

- Applebaum, B. (2019) *The Economist's Hour: False Prophets, Free Markets, and the Fracture of Society*, Little Brown and Company.
- Bowles, S and Carlin, W. (2021) 'Rethinking Economics' *IMF Finance and Development Report*.
- Burgin, A. (2012) *The Great Persuasion: Reinventing Free Markets Since the Depression*, Harvard University Press.
- Friedman, M. (1995), 'Best of Both Worlds: An Interview with Milton Friedman' *Reason Magazine*, June 1995.
- Gerth, H.H. and Mills, C. (1958) *From Max Weber: Essays in Sociology*, Oxford University Press.
- Geuss, R. (2008) *Philosophy and Real Politics*, Princeton University Press..
- Hayek, F. (1944) *The Road to Serfdom*, University of Chicago Press.
- Marx, K. (1978, 'Theses on Feuerbach' in *The Marx-Engels Reader*, Second Edition, Norton and Company.
- Hochschild, A. (2006) *Bury the Chains: Prophets and Rebels in the Fight to Free an Empire's Slaves*. Boston: Houghton Mifflin.
- Rawls, J. (1971), *A Theory of Justice*, Harvard University Press.
- Wollstonecraft, M. (1792) *A Vindication of the Rights of Woman*. J. Johnson.

Philosophy as democratic underlabour

Professor Adam Swift, University College London

To the best of my knowledge it was Stuart White who first suggested that political philosophy can be seen as a kind of 'democratic underlabouring' (White, 2003, p.29). White was riffing on John Locke, for whom it was 'master-builders' such as Boyle and Newton whose 'mighty designs, in advancing the sciences, will leave lasting monuments to the admiration of posterity' (Locke, 1690). For philosophers like himself 'it is ambition enough to be employed as an under-labourer in clearing the ground a little, and removing some of the rubbish that lies in the way to knowledge'. To regard philosophy as democratic underlabour retains Locke's insistence on its modest, ancillary role but changes the goal. Rather than preparing the path for knowledge, philosophical groundwork is necessary for our democratic processes to work well.

On the one hand, democratic decision-making is constitutive of people's public lives: we live our public lives partly as citizens, deliberating and making decisions together. By enriching the quality of our deliberation, philosophers can help us flourish in our role as citizens engaged in the process of collective self-rule, improving the quality of our democratic processes. On the other hand, the outcomes of those processes determine the content of our public lives: what we are collectively deliberating *about* are precisely the shared rules that we are to live by, and impose upon one another. By clarifying the normative issues at stake, and offering their considered views about what those shared rules should be, philosophers can also help to improve our decisions. Democratic underlabour may be done in service to their fellow citizens, but philosophers need not be modest about the importance of their contribution.

There are two kinds of philosophical work involved. The first is purely analytical. Philosophers can clarify concepts in a way that enables citizens, and those seeking their votes, to understand the normative issues at stake in political debate. Politicians sometimes invoke grand but vague ideals like 'justice', 'liberty', 'democracy'. Take 'justice' for example. There are different conceptions of that concept. Are they referring to some idea of fairness, of giving people what they deserve, or of people getting what they are entitled to? These three diverge in important ways.

Or consider something less thoroughly value-laden, like 'inequality'? On hearing that word, philosophers will immediately want to know whether we're talking about inequality in distributive or relational terms. Is the concern that people are unequal in the distribution of something, or is it rather that they do not relate to one another as equals? If distributive, are we supposed to be thinking about inequalities in outcomes or in opportunities? Is the speaker really interested in inequality at all? It often turns out on closer inspection that they are actually talking about something else, such as the plight of those who have least. 'Conceptual analysis' may sound fancy but it's simply the obviously important business of working out what people mean when they say things.

Sometimes the imprecision is deliberate. The vaguer the description of something that sounds nice (or nasty), the wider the support one can garner for (or against) it. Still, even where those seeking the power to make decisions may find it strategic to present their views in elusive feelgood (or feelbad) slogans, they have no good reason for thinking that way. And in any case, citizens choosing between them are empowered when they can distinguish between — and so are in a position to interrogate seriously — the various offers on the table. The democratic process is improved when citizens can see beyond the rhetoric to pin down where exactly politicians are disagreeing. Are they aiming at different goals or do they have different views about how best to achieve them? If the latter, is the disagreement principled or does it rest on different views about what is feasible?

Philosophical analysis can also contribute to our collective understanding of what is at stake when it comes to decisions about specific policies. It's rare for policy disagreements to be entirely technical — the kind that can be resolved merely by looking at the relevant evidence about 'what works'. Whether policy wonks realise it or not, policy choices nearly always depend on value judgements. Sometimes these are judgements about what outcomes one is trying to achieve. Sometimes they are about which means of achieving them are permissible.

Some democratic underlabour involves exposing the different moral considerations that underlie policy decisions, so that citizens and their representatives can better understand what it is that they are choosing between.

Consider the domain of education policy. People disagree about what state schools should teach, about how they should select their pupils, about whether parents should be permitted to send their children to elite private schools, or to schools that will instruct them in a particular religious worldview, and so on. Advocates of different positions on these issues typically endorse different value judgements. Some of these judgements are consequentialist: they concern the outcomes that educational policy should be aiming for. Disagreement here will turn on views about, for example: what constitutes a good childhood; the mixture of educational goods — valuable knowledge, skills, attitudes and dispositions — that we should be looking to schools to produce in their students; how opportunities for those goods (and all the others to which education is instrumental) should be distributed; and so on. Some of the judgements are non-consequentialist: they concern not the value of outcomes but people’s entitlements to be treated, or not to be treated, in certain ways — that is, whether the methods of producing such outcomes are acceptable. Perhaps parents’ rights to spend their money on their children’s education, or to send their children to an instructive religious school — or children’s rights not to be sent to such a school — should be respected whatever the consequences.

In *How To Think About Religious Schools: Principles and Policies* (Clayton et al., 2024) my co-authors and I suggest an analogy between philosophical and chemical analysis. Just as a chemist analysing a compound will identify its constituent elements, and be familiar with the various ways in which those elements tend to combine to form more complex units, so the philosopher can analyse people’s political views, identifying the particular combinations of normative judgements that motivate them. Where the chemist’s analytical framework is given by the periodic table, which identifies and organises the discrete and basic constituents of the material world, the framework we offer identifies and organises the various moral values and principles that combine to form people’s views about religious schools. This type of analytical, clarificatory work is one kind of democratic underlabour — one way that philosophy can prepare the ground for democratic deliberation, improving both the process and the product by helping people to understand more clearly their own and others’ beliefs, and the nature of their disagreements.

But philosophers can contribute more than clarification of the normative judgements that underpin political views, important though that is. They can also offer arguments for and against substantive conclusions, giving reasons why their fellow citizens — and their representatives — should endorse some views and reject others. Such arguments may operate at different levels of abstraction, and be of less or more immediate relevance to decisions currently on the political agenda. At the more abstract level, they can be about, for example: the best way to understand 'justice', 'freedom' or 'equality'; the complex ways in which property rights both protect and interfere with people's freedom; the moral basis of the family and its implications for parents' rights; and so on. Where values conflict, the arguments can be about which balance between competing values — which mixtures of elements in the philosophical periodic table — are better than others. At the more concrete level, they can be about the merits and demerits of specific policy proposals, whether concerning immigration, assisted dying, the regulation of social media, or whatever. Our book about religious schools begins with an analytical framework but it concludes with a concrete set of policy proposals for the regulation of religious schools in England, those proposals inevitably reflecting our substantive — and reasoned — judgements about the proper balance between the various value considerations at stake.

Although more than clarificatory, this second kind of philosophical work still counts as 'analytical' — as long as the arguments are presented in a suitably clear, precise and rigorous way. Does it really qualify as 'democratic underlabour'? Its aim is surely less modest. The goal here is to improve the quality of democratic decisions directly, by persuading others that some decisions are better than others. Advocating particular outcomes is more presumptuous than seeking only to improve the process. But the label still applies. What philosophers are offering when they engage in this kind of persuasive work remains precisely that, an offering. It is their contribution to democratic deliberation, their reasoned justifications offered as recommendations to those with whom they are deliberating. Those recommendations may be considered 'expert' in so far as they are the product of professional training, of the sheer amount of time spent worrying at getting the arguments right, of the improving filter of their intellectual community. But it is up to their fellow citizens — or their representatives — to decide whether to accept what they are offered. In this respect they are no different from economists and other social scientists, whose contributions typically take the form of the empirical evidence, and predictions based on that evidence, that are also needed to guide decisions. The Brexit referendum showed that they too can do no more than hope that their fellow citizens will take heed.

There is nothing novel or, I think, controversial in the story so far. So let me complicate the discussion by raising, necessarily briefly, three further issues. The first concerns how, on this conception of philosophy's contribution to public life, we should understand the significance of the philosophical input. Exactly what difference might it make? In my opening paragraph I claimed vaguely that philosophical underlabour is necessary for our democratic processes to 'work well'. In my second I distinguished between its contribution to the quality of processes or procedures, on the one hand, and decisions or outcomes, on the other. Philosophy can help us both to a better way of making decisions and to a way of making better decisions. However one conceives the relation between those two aspects — and much democratic theory is devoted to that question — it's not hard to see why we might value them both.

Less obvious, perhaps, is the thought that, in making its contribution, philosophy may be not merely improving the process or product, but contributing to the legitimacy of the political system. Political decisions are coercively imposed on all of us, including those who voted against them. The question of legitimacy, crudely, is the question of what justifies the winners' forcing the losers to comply, and it may be that part of the answer concerns the kind of deliberation involved. Where citizens contribute to collective decisions with little understanding of what they are voting for, or why they are voting for it, then it's questionable whether the winning outcome can legitimately be imposed on those who disagree. To be sure, one might think that legitimacy derives primarily from respect for the kind of equality that is embodied simply in majority rule, without regard to the quality of the deliberation (if any) that precedes it. But the procedural demands may be higher: for example, it may matter whether those participating in the practice of collective self-rule have an adequate sense of the values at stake. If so, philosophy can improve our democratic practice by making it more legitimate.

What happens, though, when a philosopher's reasoned view amounts to the claim that some matter is just too important to be left to the democratic process? This is the second complicating issue. My happy picture — philosophers have no special standing in public life but are simply underlabouring on behalf of their fellow citizens — fails in those cases where their judgements about how to balance the various values at stake lead them to the conclusion that something should not be up for democratic grabs in the first place. The plausibility of that conclusion is likely to depend, in part, on the extent to which real-world democratic processes approximate the philosophical ideal. In societies like ours, where the

opportunity to influence political discussion is so unequal, and where many of those with greatest influence show so little interest in reasoned argument, it isn't hard to be sceptical about the legitimising force of (so-called) 'democracy'. But even in the ideal political system, there will be cases — typically conceived as matters of grave injustice or violations of basic rights — where philosophers may think that there is too much at stake for the matter to be decided by the collective will of their fellow citizens. In such cases, it is incoherent for them to regard themselves simply as one voice among others. They are committed to denying that modest view of their contribution.

In practice, however, things are complicated. Sometimes the very question of whether an issue should be subject to democratic decision is itself decided democratically. It can be through our (less or more) democratic political processes that we collectively decide who gets to decide what. Consider, for example, how recent prospective leaders of the Conservative party have campaigned on a ticket of withdrawing from the European Convention of Human Rights. In that kind of case, the happy picture still applies: philosophers can straightforwardly see themselves as participating in the process by trying to persuade their fellow citizens about what view they should take. Sometimes, though, it is other bodies that make decisions. These may range from grand judicial institutions like the Supreme Court, to bodies like the National Institute for Health and Care Excellence (NICE) or the Human Fertilisation and Embryology Authority (HFEA).

Whether philosophers addressing decision-makers in such organisations can claim to be performing democratic underlabour depends on how we think about the democratic credentials of those organisations. Where the organisations operate with devolved authority — given the task, by their fellow citizens (or their representatives), of making decisions within a particular domain of expertise — then that description remains apt. Where they are better understood not as part of but different from — and indeed sometimes antithetical to — the democratic parts of our decision-making system, then philosophers contributing to their decision-making processes must accept that they are making a different kind of contribution. That's not a problem. There are proper limits on democracy and philosophy can contribute to valuable non-democratic elements in our decision-making arrangements. But we should be clear that, in such cases, we are no longer talking about philosophy as distinctively *democratic* underlabour.

This last scenario raises the third complication. When they are addressing their fellow citizens directly, philosophers should feel free to offer their own views on the matter in hand, however controversial or personal those views may be within the philosophical community. Given their expertise, they have a responsibility — we might think of it as a role obligation — to put forward their best arguments for their conclusions, eschewing the slogans and rhetoric that so often characterize political debate. It's their distinctive task to offer their contributions as philosophers. (I'm talking here about philosophers acting in their capacity as philosophers. With other hats on they can sloganise as much as they want.) But what they are arguing *for* can be as idiosyncratic as they like. Things are different when they are not addressing the public as a whole but are rather given special standing as advisers or consultants to bodies — like NICE or HFEA. Then they are more like representatives of the philosophical community than individual citizens who happen to be philosophers. Here their contribution seems properly limited to conveying the collective — though rarely consensual — wisdom of that community. That doesn't limit them to clarification or analysis, especially if considerations judged important by philosophers are not being given proper attention, but they should not see their position as an opportunity to push their own particular line from a position of influence.

How might philosophers better discharge the role I've described? It's hardly original to suggest that they might present their views in ways that make them accessible. Philosophers are incentivised to produce research that makes new, sophisticated or fundamental moves in specialised professional debates. That's not altogether a bad thing. I have nothing against philosophers talking to other philosophers about things that — and/or in ways that — only they can understand. I do some of that myself, and nothing I've said should be taken to belittle that kind of work. In time, and with the help of those skilled in rendering scholarly work accessible to wider audiences, it may even qualify as democratic underlabour. But, although individual philosophers may not always be well suited to the task, philosophy as a discipline best fulfils its promise in that role when it speaks to its audience of fellow citizens — or their representatives, or those with devolved decision-making authority — in ways that are intelligible to them.

Relevance matters too. Philosophers collectively might do more to connect their research to issues that are currently, or at least foreseeably, up for democratic decision. These are the matters where their input is most pressing and where, in allowing their fellow citizens to set the agenda, their contribution can most clearly be seen as democratic underlabour. Urging

this is quite compatible with acknowledging the potential long-term significance of work that is not directly oriented towards concrete, current concerns. It is perfectly possible to do both. But applying philosophy to the here and now requires recognition that useful guidance typically depends on combining abstract values — if not ideal theory then at least theory of ideals (Hamlin and Stemplowska, 2012) — with context-specific empirical evidence. My final suggestion — also happily familiar — is that philosophers do more interdisciplinary work, and especially that they engage more with relevant social science.

References:

- Clayton, M., Mason, A., Swift, A. and Wareham, R. (2024) *How To Think About Religious Schools: Principles and Policies*. Oxford, Oxford University Press.
- Hamlin, A. and Stemplowska, Z. (2012) ‘Theory, Ideal Theory and the Theory of Ideals’, *Political Studies Review* 10(1), pp.48-62.
- Locke, J. (1690) *An Essay Concerning Human Understanding*.
- White, S. (2003) *The Civic Minimum: On the Rights and Obligations of Economic Citizenship*. Oxford, Oxford University Press.

The marketplace of ideas: who's buying?

Professor James Wilson, University College London

Much philosophical writing is in principle relevant to decision-making in the public sphere, but in practice has absolutely no effect on the day-to-day concerns of policymakers. Only a tiny percentage of philosophical writing is explicitly addressed to practical problems of a kind that policymakers need to solve, and even where it is, it is often written in a way that is too technical for those without graduate-level training in philosophy to understand. And when philosophers aim to say something helpful about live policy problems, they often misunderstand relevant contextual features, and recommend simple solutions that, if adopted, might make the real-world problem worse.

Why is this, and what should be done about it? The deepest difference between philosophers and policymakers is in their prevailing assumptions about the role and usefulness of theory. Policymaking is fast-paced and focused on making improvements in a context in which many factors are beyond the policymaker's control. As a result, policymakers are most interested in what will work here and now, where 'working' involves only making an improvement according to some baseline, rather than completely solving the problem.

Philosophy works at a rather slower pace. Philosophers usually take their discipline's problems to be abstract and highly idealised. In approaching problems, they often deliberately ignore many factors such as context, history and how the problem interacts in practice with others. This idealisation shapes philosophers' sense of what a good theory would look like. Philosophers tend to assume that a theory cannot be correct if someone can produce a successful counterexample to it, and in exploring potential counterexamples,

they are drawn to conceptually interesting edge cases that often involve outlandish scenarios. While focusing most philosophers’ analytical firepower on highly contrived cases might provide the best way of determining whether any philosophical theories can be defended against all possible counterexamples, it has the unfortunate result that philosophers often largely ignore the common cases which are the bread and butter of public policy.

It is easy to point to celebrated historical examples of philosophers’ influence on policymaking — such as Mary Warnock’s official inquiry and report that led to the establishment of the Human Fertilisation and Embryology Act 1990 — but it is rather more difficult to identify similar contemporary examples of policy influence. While a generation ago it seemed natural for politicians to look to prominent philosophers like Mary Warnock, Anthony Quinton, Bernard Williams or Stewart Sutherland to provide wise counsel, they now seek advice elsewhere. Twenty-five years ago, there were four philosophers in the House of Lords, but Onora O’Neill is now the sole remaining philosopher, and will retire soon.

Growing numbers of philosophers now see it as a problem that the discipline punches below its weight in its contributions to public policy, and this has coincided with increased financial incentives from research funders to demonstrate the impact of research. Recent years have seen increased emphasis by philosophers on demonstrating how philosophical expertise can improve the conceptualisation, framing, understanding, and weighing of options around matters of public concern. This has included writing academic articles, newspaper commentaries, media interviews, writing commissioned reports, and working with public bodies to help solve problems as part of a multidisciplinary team. The REF 2021 Impact Case database (REF 2021, 2023) provides a good sense of the range of activity.

What has been missing so far is a rigorous and realistic account of how policymakers view the involvement of philosophers in the conceptualisation, framing, analysis and weighing of policy options. Do they think that there is something significant missing from democratic decision-making without the input of philosophers, or do they think that philosophers’ usefulness is rather more marginal? And if policymakers do not find philosophy useful for their purposes, what should philosophers do about this?

There are sharp differences in how policymakers use scientific and philosophical insights. Scientific knowledge is descriptive rather than normative, and facilitates new or different means of attaining whatever goals a community has. Therefore, the same body of scientific

knowledge can be useful to individuals or governments with widely differing political goals. While scientific and technological advances sometimes lead to profound social shifts (such as the smartphone) and these may also reshape social values, the desirability of such social changes is a matter for general political discussion, rather than something on which scientists are uniquely qualified to pronounce. Conversely, philosophical literature in ethics and political philosophy is overwhelmingly normative, rather than descriptive. Philosophers' aims are not usually to provide tools that can be used better to pursue whatever goals a political community already has, but to make arguments about the principles and the goals that determine what societies ought to do.

The defining feature of democracy is that it is an open-ended project in which the people, and their elected representatives, set its direction. Such fundamental commitments are difficult to reconcile with the idea that philosophers (or any other group) have a special kind of insight into core ethical values, and should be deferred to on this basis. As a result, not even philosophers themselves claim that they are best placed to dictate the values that should inform public policy. Philosophers are of course just as entitled as any other citizen to intervene in public debates, and to contribute to a public conversation, whether via newspaper articles, media appearances or political activity, but it is not the case that others should defer to them because they are philosophers on matters relating to values.

How can philosophers influence public policy in a way that is not only responsible but also effective? A basic quest for philosophers is to ascertain what other sources of policy advice they are in competition with. Are the contributions of philosophers in competition only with those of other philosophers, or with a wide range of other potential sources of advice? As philosophers, it is tempting for us to think that what we can provide is unique and can only be provided by those with philosophical expertise. However, this is to mistake the nature of the policy environment.

On each particular policy issue, there are many interested parties who want to shape the policy landscape and the detail of what is implemented. These include other government departments, executive agencies, arms length bodies, industry bodies, think tanks, trade unions, individual businesses, non-governmental organisations, pressure groups and private citizens. Though some may attempt to steer government decision-making in a way that seems regressive or self-serving, most will draw on values and considerations that have wide currency within society.

Policymakers' central challenge is how to sift through this surfeit of information, advice and lobbying about what should be done. Policy advice is thus a market, in which what philosophers offer may be in competition with a wide range of non-philosophical offerings. How to categorise philosophically informed policy advice, and how much to value it in comparison to other sources of advice, will be determined by policymakers rather than philosophers. As in any market, what matters is not whether the seller thinks that there is something uniquely valuable about their product, but whether buyers agree.

We can think of the market for philosophically informed policy-relevant advice from both the supply side and the demand side. From the supply side, questions we can ask are: What sorts of public-policy-relevant philosophy is being produced, on what topics, and how much of it? How much effort is it for policymakers to transform philosophical outputs into something that is useful for their purposes? And how does this differ according to the kind of output that the philosopher produces? For example, a complex journal article will be much more difficult to assimilate than a well-targeted briefing. On the demand side, basic questions are: How many policymakers want philosophically informed advice? How will they use it? How much do they want it? And what other goods that, from the perspective of the policymaker, will substitute for it?

Reflecting on these supply and demand features encourages a more realistic view of philosophy's competitors as sources of advice to policymakers, and helps us understand the costs that the policy system will need to be willing to pay if it is to make different kinds of philosophical insights usable for its purposes. These costs include time to research relevant philosophical material; resource to transform materials into practical products; and effort to change plans or structures in response to learnings and insights.

As has been thoroughly explored within innovation research, generating new knowledge through university research makes little difference to the economy unless governments or businesses can transform the research into practical tools and resources. Cohen and Levinthal called this quality 'absorptive capacity', and defined it as the 'ability to recognize the value of new information, assimilate it, and apply it' (Cohen and Levinthal, 1990).

A government department may wish to learn from scientific findings, but lack the time or expertise to be able to assimilate the research, and explore its relevance and how it could improve responses to policy problems. Scientists' ideas are more likely to be implemented if

they are accessible, for example if they provide a summary paper and a well-written report with realistic recommendations. What philosophers think policymakers need, in terms of philosophical input, may not align with what is required. Philosophers need to produce philosophy that maps on to policymakers' problems — and that aligns with policymakers' absorptive capacity.

Suppose I want a fish finger sandwich. There are several ways of achieving this, depending on my culinary skill and the time I am willing to invest. I could make my own bread from scratch, and fashion fillets of fish into breadcrumb covered fingers; I could buy bread and read-made fish fingers, and assemble the meal myself; or I could order the sandwich from a local takeaway. If philosophers are in effect selling sourdough starter kits and fresh fish, while others are selling the complete sandwich, it is easy to see why their potential contributions are often overlooked by busy policymakers with tight deadlines.

The absorptive capacity for philosophical research in the UK is weak. While there is significant government infrastructure devoted to making science, technology, engineering and mathematics (STEM) useful for government policy, there is hardly any for the humanities, and none specifically for philosophy. To give just a few examples, the UK government currently has 25 Chief Scientific Advisers (GOV.UK, 2016), each based in different government departments and agencies, including the Government Chief Scientific Adviser who regularly briefs the prime minister. No government departments have a Chief Humanities Adviser (let alone a Chief Philosopher). There is a Government Office for Science, but not one for humanities. The Council on Science and Technology advises the prime minister, but there is no corresponding committee for humanities. There are of course individuals who studied philosophy (especially philosophy, politics and economics) at university who are employed as policymakers, but their role is not to facilitate the two-way translation of ideas between philosophy and government or the wider economy. It is reasonable, therefore, to assume that absorptive capacity for philosophy is currently very slight.

Philosophy's lack of influence on public policy in the UK is unsurprising, given the lack of absorptive capacity within government. As a philosopher, it is natural to think that this lack of absorptive capacity is a problem not just for philosophers, but for society more broadly. But it is not straightforward to articulate either what the government is missing as a result, or how to fix the problem. One reason is that it is not clear that philosophy generates knowledge of a

kind that could, or should, be taken up via ‘knowledge transfer’, even if politicians wanted to. Given the level of profound and sustained disagreement among professional philosophers, it is rarely possible for policymakers to determine that one established philosophical view is more substantively correct than another, though policymakers may of course find some approaches more congenial than others. Therefore, the problem of weak absorptive capacity for new philosophical insights is, understandably, not treated with the same urgency as, for example, weak absorptive capacity for novel insights in AI or renewable energy.

Value inquiry — improving the ways a society pursues and reconciles its existing ends, and how it comes to change its ends — is crucial for well-functioning democracies. However, value inquiry is something to which all citizens need to contribute, rather than just an elite. While some philosophers may plausibly be described as experts in some aspects of value inquiry, it does not follow that philosophers should wield inflated influence over a democracy’s goals and guiding principles.

The idea of the public interest provides important insights about what the role of philosophers in public policy should be, and about the ideal absorptive capacity for philosophy within government. Everyone has a right to have a say in the public sphere, but it is fundamental to a well-functioning democracy to distinguish between public and private roles. What may be done when someone is acting in a public role is constrained by a set of norms, which aim to ensure that such actions are guided by public rather than private interest. Documents, such as the UK government’s Nolan Principles, Civil Service code, and Ministerial Code, articulate some of the responsibilities associated with different public roles. Thus, it is legitimate for a politician to aim to benefit their friends and family in their private life, but it will amount to corruption in public office if they use their public office to do so. Civil servants may disagree with government policy, or think it unethical, but their role requires them to help implement such policies, and to follow proper channels in reporting serious concerns they have about any particular policy. More broadly, public servants and elected officials should take account of arguments and representations only in so far as they are relevant to the pursuit of the public interest.

To be relevant for policymakers, experts — whether philosophers or those from any other discipline — need to frame their contributions in line with the requirements of policymaking. In particular, they need to articulate an understanding and analysis of the policy problem — and any related recommendations — from the perspective of the public interest rather

than their private perspective. In recent work with colleagues, we defined the role of the ‘critical friend’ as providing the best way of combining success in influencing government policy with democratic legitimacy. A critical friend provides advice and challenge to a public institution based on an accurate and sympathetic understanding of the kinds of constraints that an institution faces (Wilson et al., 2023). Acting as a critical friend requires humility: a willingness to work as an under-labourer within a broader process of democratic deliberation; accepting the broad framing of the problem as set by public servants; and if required, challenging this framing to ensure it is working in the public interest.

Philosophers who argue for policy positions based on ethical theories usually think of themselves as acting in a principled rather than self-interested manner. They might expect that what they are doing will fall squarely in line with the requirements of policymakers. But, unless a philosopher articulates their argument as a response to the problem, and recognises the relevant institutional constraints, their intervention will fall into the same category as those of lobbyists or advocates. If philosophers’ attempts to influence policy — without properly considering the public interest — are ineffective, this may be a sign that the system is working well rather than badly. That is, if the supply does not meet the demand, it should be rejected.

The most pressing question for further work, which should involve both philosophers and policymakers, is how philosophical expertise can help to improve the quality of a public deliberation and the design of value frameworks to ensure they effectively support government decision-making. I doubt that policymakers would say that what they value about philosophers is their expertise in uncovering what the uniquely correct normative principles are, or what should be done given these principles. What they are more likely to find useful is philosophers’ application of ethical theories to help frame and reframe policy problems, their capacity to articulate novel concepts that bring underrepresented experiences into focus, and their ability to spot and help resolve ambiguities and inconsistencies. Building absorptive capacity for these kinds of philosophical expertise should be a priority not just for philosophers, but for anyone interested in improving the functioning of our democracy.

References:

- Cabinet Office (2024). *Ministerial Code*. [online] GOV.UK.
Available at: <https://www.gov.uk/government/publications/ministerial-code/ministerial-code>.
- Civil Service (2015). *The Civil Service Code*. [online] GOV.UK.
Available at: <https://www.gov.uk/government/publications/civil-service-code/the-civil-service-code>.
- Cohen, W.M. and Levinthal, D. (1990) *Absorptive capacity: A new perspective on learning and innovation*. Philadelphia, PA: Reginald H. Jones Center, Wharton School, University of Pennsylvania.
- Committee on Standards in Public Life (1995). The 7 Principles of Public Life. [online] Gov.uk. Available at: <https://www.gov.uk/government/publications/the-7-principles-of-public-life/the-7-principles-of-public-life--2>.
- GOV.UK. (2016). *Chief Scientific Advisers*. [online]
Available at: <https://www.gov.uk/government/groups/chief-scientific-advisers>.
- REF 2021 (2023). *Impact database: Results and submissions: REF 2021*. [online]
Available at: <https://results2021.ref.ac.uk/impact>. (Accessed 7 February 2025)
- Wilson, J., Hume, J., Cian O'Donovan and Smallman, M. (2023). *Providing ethics advice in a pandemic, in theory and in practice: A taxonomy of ethics advice*. Bioethics. doi: <https://doi.org/10.1111/bioe.13208>.

The point is to change it

Professor Jonathan Wolff, University of Oxford

Many political philosophers happily recite *Thesis Eleven* of Marx's *Theses on Feuerbach* (inscribed on his gravestone): 'Philosophers have only interpreted the world in various ways; the point, however, is to change it.' This is an admirable sentiment. The motivation for many of us who work in moral and political philosophy is to help right the wrongs we perceive in the world. Yet we often find that the norms of professional academia encourage us to concentrate on more abstract matters. Our energies become devoted to such things as clarifying our positions in fine detail and showing how they differ in subtle ways from the existing literature, and much less on finding common cause with others to bring about change in the world. Hence a renewed concentration on making change seems very welcome. And now especially, when we suffer the twin threats of the arc of the moral universe (to use Martin Luther King Jr's evocative phrase) bending away from — rather than towards — justice; and the general disparagement of the humanities — philosophy among them — as marginal or irrelevant to what is 'important' in the world.

Yet if we are to encourage philosophers to think about how they can contribute to making change, we also need to pay attention to the practicalities and processes of doing so. In other words, what is, or should be, our theory of change? How do we propose that change should be made, or at least how we can contribute to that process? Marx had a theory of change: proletariat revolution, and a conception of his own role as the theorist of that movement. But what theory of change is appropriate in current circumstances? In this article I will consider the potential for philosophy and philosophers to contribute to beneficial policy change.

Marx was, of course, by no means the first of the socialist thinkers to think about how to bring about a new society. Charles Fourier, for example, set out the details of utopian 'phalanxes', and while he made little progress in his native France, he inspired experimental, if short-lived, communities in the US. Robert Owen was able to try out his ideas in practice in the mill he

managed in New Lanark in Scotland, and is reputed, for example, to have introduced the first nursery school, among many other innovations. Ultimately disillusioned with the limits of New Lanark, which could only remain embedded within a voracious capitalist economy, Owen too tried to set up utopian communities in the US, although they never took root. Fourier and Owen can be understood as attempting to prompt social change through leadership by example. However, in terms of changing social structures, they did not get past what is, in effect, the pilot stage, albeit often very valuable in themselves, and in Owen's case inspiring lasting innovations.

This is not the age of revolution or small-scale socialist experiment — although some radicals still favour the former, and some religious groups have adopted something similar to the latter. Contemporary political philosophers naturally assume that social change must be consistent with the rule of law, and hence made through either the ballot box, or through legal challenge in courts of human rights. Within the democratic process two important vectors of influence are to appeal to voters to help determine who gets elected; or to appeal to politicians directly to try to influence their decision making. Much of what we do as philosophers — make arguments, write reports, assist think tanks — probably has the latter form, hoping to influence the decisions of policymakers simply by the quality of and the rationale for our ideas. This, of course, is where our expertise tends to lie. But in reserving our activities for such spaces we are making some perhaps naive assumptions about how politics works: as if it is a Habermasian ideal speech situation — that is, a hypothetical situation when people have equal opportunity to participate in rational discourse without coercion or manipulation. This picture leaves out constraints of political pressure, crisis points, deal-making, receding timelines, and so forth. In other words, we shy away from realpolitik. Perhaps that is as it should be, and philosophers should stay in their lane. But it leaves us with a nagging question: what can we do to (help) to bring about change when democratic politics is too slow or not listening, or captured by bad-faith actors? I do not think that there is a single, special way, and that broadly we are doing the right things, given who we are and what we are capable of. But it is also possible to come to a better understanding of best practices, and in doing so also see how to focus our efforts more effectively.

Inspiring social change

The idea of wanting to bring about social change by influencing those in power has probably always been with us, and it is worth considering how this has been attempted by successful social movements. To return to Martin Luther King Jr as a leading example, we find the

suggestion that non-violent civil rights struggle is essentially a matter of progressively improving the oppressed group's negotiating position until the battle is won. He proposed that each wave of protest should go through four stages: collection of facts; negotiation; self-purification, in the sense of preparing for action; and direct action, to create crisis and tension to improve the negotiating position for the next round. The dignity of the struggle is essential to its success. It should be public, peaceful, and, to use John Rawls's language, addressed to the sense of justice of the majority. Protestors should be willing to accept legal punishment for any laws they have broken.

All elements of this picture can, of course, be challenged. James C Scott discussed the practices of informal, essentially private protest: sabotage, foot-dragging, evasion, false compliance, pilfering, feigned ignorance, arson, dissimulation and slander. Yet these are methods of individual self-protection rather than social change. At the other end of the scale lies public violence. Some examples, as in the UK poll tax riots of 1990, or the more recent Black Lives Matter protests, which sometimes turned violent, are associated with the acceptance of the need for change, but in such cases typically take part alongside many other actions calling for the same measures. Violent protest without a scaffolding of non-violent activism seems much less likely to bring about change in modern democracies. Reason, reflection and argument seem unavoidable.

The potential of philosophy

And, of course, academics are naturally drawn to reason, reflection and argument, rather than protest or violence. Yet it is far from clear how these lead to change. One barrier is that academic philosophy is normally a solitary activity. Papers and books are typically published in the name of a single author (albeit often with wide acknowledgements) and although co-authored work is becoming more common it is still very much a minority activity and rarely extends to more than two or three collaborators, except in the case of commissioned reports. But social activism is a collective activity, requiring a powerful consensus and pragmatic compromise rather than the scrutiny and approval (or rejection) of every detail, as we are used to in academic scholarship. How then can we have any chance of influencing policy? At worst we seem to have assumed a theory of change that anticipates our papers in academic journals will somehow influence people in power. In more adventurous moments we supplement these with blog posts, podcasts or newspaper columns, but we typically fall far short of concerted effort.

But before we consider how to address this gap, it is worth reflecting on what we think our influence as philosophers should be. We are not policy experts, and even if we have strong arguments — for example, to reduce inequality, address poverty or welcome refugees — the practical application of these arguments requires knowledge and experience that we, as philosophers, simply do not have. What, then, should our aim be? It is all very well to say that we want to bring about change, but are we confident that we know, in detail, what changes are desirable and feasible for our societies?

These practical questions have not been an obstacle to the development of ideas of justice.

As I have mentioned, we tend to think that this is where our expertise lies, but as many have pointed out, the world is already rich in theories, and some have asked whether there is really any space or need for more. Indeed, this was a point made 150 years ago by Mikhail Bakunin: ‘During the last nine years more than enough ideas for the salvation of the world have been developed in the International (if the world can be saved by ideas) and I defy anyone to come up with a new one. This is the time not for ideas but for action, for deeds.’

While Bakunin is right that a huge array of theories and variations existed, even 150 years ago, it may be too pessimistic to assume that there is no room for any more. In recent decades, for example, the theory of universal basic income, especially as promoted in writings of Philippe van Parijs and colleagues, has gone from an eccentric academic fantasy to part of regular, mainstream discussion. Some will reply that this is not so much a new idea as a repackaging of an older one, and van Parijs himself is keen to identify earlier thinkers with similar ideas. This response, however, leads me to an observation made by political philosopher Margaret MacDonald in 1940. Often what a political philosopher does, even under the self-image of presenting an entirely new theory, is simply to draw attention to a value or idea that is currently relatively neglected. In other words, even if Bakunin is right that there is a vast storehouse of theories, it is necessary to apply effort to resurface these theories and propose them as solutions to current problems. Presenting a theory as new can contribute to its power.

Furthermore, even when new ideas are introduced, it is a mistake to think that the only form of innovation is confined to theories of justice. Here I draw on the distinction between ideal theory and non-ideal (or ‘real-world’) political philosophy. In most basic terms, ideal theory attempts

to produce theories of the just society, whereas real-world political philosophy recognises our actual, unjust society, and draws attention to existing injustices which can be unknown, underplayed or taken for granted as part of the natural order.

One exponent of this method is Amartya Sen who is in a unique position as an empirical and theoretical economist as well as a political philosopher. In one memorable example, Sen noted that the everyday policies of gender prejudice around the world mean that vast numbers of women are ‘missing’, a reference to the avoidable deaths of women and girls, especially in childhood through neglect of health or nutrition. Each case is treated as a tragic misfortune but adds up to a global pattern of extreme gender inequality. Another of Sen’s central contributions to non-ideal political philosophy is to argue that famines are a political failure that can be avoided under the right political conditions, rather than inevitable natural disasters. These arguments, of course, draw on his work as a social scientist, but when combined with his normative sensitivity present a powerful and persuasive case that those with power and influence need to think and act differently.

Sen did not, in these examples, present a new philosophical theory, but shifted perceptions, using existing language. But he is also responsible for conceptional innovation through his invention of the language of capabilities to describe human well-being and potential. Capability theory, either in Sen’s version or Martha Nussbaum’s development, has been taken up in a wide range of other fields to help develop measures and methods. Introduction of new vocabulary can itself be surprisingly powerful. Iris Marion Young, for example, has at least two contributions that have had wide traction. In *Justice and the Politics of Differences* she explains what she calls ‘five faces of oppression’: exploitation, powerlessness, marginalisation, violence, and cultural imperialism. And in *Responsibility for Injustice*, she puts the idea of structural injustice at the centre of analysis, and provides a powerful framework for others to make sense of and to address injustices in different spheres.

These examples of Sen and Young show a way of doing political philosophy which is distinct from the development of detailed theories. It is a matter of creating a vocabulary — naming things to bring them to consciousness — which can then be taken up by others. How this can work in detail is very well expressed in a passage by feminist political philosopher Alison Jaggar:

As a young woman ... I was unable to articulate many vague and confused feelings and perceptions because the language necessary to do so had not yet been invented. The vocabulary I needed included such terms as ‘gender’... ‘sex role’, ‘sexism’, ‘sexual harassment’, ‘the double day’, ‘sexual objectification’, ‘heterosexism’, ‘the male gaze’, ‘marital, acquaintance, and date rape’, ‘emotional work’, ‘stalking’, ‘hostile environment’, ‘displaced home-maker’ and ‘double standard of ageing’.

More recently a particularly striking example of the introduction of a new term has been Miranda Fricker’s concept of ‘epistemic injustice’. In one way what Fricker points to is nothing new. It has long been known that there are gender, class and racial biases regarding who will be listened to, taken seriously, or treated as an authority, often without any justification. But it was often taken for granted as part of a type of natural order of things, and those who complained were typically regarded as a nuisance or attention seeker. However, once we have the vocabulary of epistemic injustice — just as in Jaggar’s examples — there are at least three types of advantage. First, it can be studied as a systematic phenomenon, with wide application and several forms of variation. Second, it allows people to understand their own situation as an example of injustice. And third, it makes clear to other parties that it is not something we should simply accept as an unchangeable reality. Hence the mere naming creates a potential for change, in which the philosopher is part of a broader eco-system of activism.

The philosopher’s theory of change

Although some academics may have a double life as scholar and activist, this is not a route that many of us will feel equipped to take. We are reasoners. But the forms reason can take are many. Here I’ve suggested that among the ways in which moral and political philosophers can contribute to debates in public life include: the development of new theories; the promotion of theories that have been relatively neglected; the identification of ignored injustices; and the development of new vocabularies. Often these will go together: new vocabulary is a type of mini-theory, and it could bring prominence to something previously known but not given due weight, and it can help identify injustices that have simply seemed part of the moral furniture.

But our question is how to contribute to a process of change. Earlier, I rather unfairly suggested that many philosophers think they can somehow contribute to change by publishing their research behind a paywall, but of course there are already many alternative

routes to sharing ideas. We try to find ways to make our work available to the public and to policymakers. But I would suggest that another — indirect — model has in practice been more successful: the model of change by diffusion. The work of Sen, Nussbaum, Young and now Fricker, has been taken up by academics and thinkers in other fields including some who are themselves often more immediately involved in the policy process, and are looking for new ways of conceptualising problems and solutions. It does not happen in isolation, and requires us to overcome our diffidence about promoting and publicising our work, and working with others. It requires us to publish the same ideas in different venues, without worrying about 'self-plagiarism'. It requires us to collaborate and to take the expertise of others — academics in other fields, activists, policymakers, journalists, civil servants — seriously, and to learn to see things from their perspective, as well as our own. Sometimes it requires us to tolerate simplification, other times to insist on rigour. It is a matter of doing our best work, looking for, and taking, opportunities, and welcoming wide engagement. For most of us, little if anything will change in the world as a result, however hard we try. Furthermore, each of us will feel much more comfortable at some stages of this process than others. Not every philosopher can be a philosopher-activist, and in fact very few have done so successfully. Even those with a reputation for activism have rarely done more than chair public committees or write for newspapers and periodicals. Many will feel that they can contribute best to the background work of simply trying to develop the best theoretical and conceptual material. But we should all remain heartened by the fact that we are part of a collective endeavour that has contributed — and will continue to contribute — to identifying and sometimes even mitigating suffering and injustice.

Section 2: Reflections

Did that answer your question?

Professor Clare Chambers, University of Cambridge

What have philosophers ever done for us?

Well, setting aside grandiose things like the Socratic method, the principles of logic, and the categorical imperative, we actually have evidence to answer this question. Since 2014, universities have had to show that their academics' research has impact outside the sector, through the Research Excellence Framework (REF). According to the impact case studies published for REF 2021, philosophy has had a wide range of influence, including: securing millions of dollars in charitable donations; shifting practice around end-of-life care; increasing public awareness regarding the risks of advanced AI; informing journalists and politicians on media ethics; affecting the design of smartphones; changing the way prisoners are educated; evidencing the interpretation of historical items in stately homes; inspiring artistic works on migration; shifting public perspectives on beauty; protecting cultural heritage in conflict situations; facilitating flood insurance; and encouraging computer game designers to pay more attention to historical accuracy (REF 2021, 2023). And this is just a small sample!

Philosophers can be found in the highest level of government and policy work, too: in the House of Lords, at the head of policy organisations and public bodies, chairing government-commissioned reports, and advising parliament. Our skills seem to be in demand.

Why, then, do we need to discuss the relevance of philosophy to public life? Isn't it self-evident?

Well, no. Despite all this public engagement and demonstrated impact, philosophy and the wider humanities are under threat. Government policy over recent decades has been to emphasise STEM (science, technology, engineering, and mathematics) subjects and insist

that all education, at school level and beyond, should be focused on career success. A career in philosophy somehow doesn't seem to count.

Perhaps that's not a niche view. When my son was about nine years old he had a philosophy lesson at school. My son told his school friends that his mum was a philosopher. They were not impressed. "She can't be!" one of them replied. "Philosophy isn't a real job!"

Philosophy is a real job, of course, and in this piece I'll show how at least some philosophy plays a vital role in public life. I write from a position of familiarity with the interplay between philosophers and policymakers, as my own work addresses issues of direct political importance. I'll give two examples.

First, my book *Against Marriage: An Egalitarian Defence of the Marriage-Free State* argues that state-recognised marriage is a violation of both liberty and equality. In place of marriage, I develop an alternative model for regulation of intimate relationships which protects the vulnerable. This work involved considerable engagement with academics and practitioners in law and politics as well as philosophy. My analysis contributed to the successful campaign and Supreme Court case for equal civil partnerships led by the trailblazers Rebecca Steinfeld and Charlie Keidan. The result is that different-sex couples may now have their relationships recognised as civil partnerships, without the gendered and religious associations of marriage.

Second, much of my work concerns a major public health issue: the overwhelming pressure to modify our bodies in pursuit of an ever-expanding expectation of physical perfection. I've developed the concept of 'the unmodified body', my term for *a body that is allowed to be good enough just as it is*. The unmodified body is a political idea and, I argue, a key principle of equality. Through examples ranging from bodybuilding to breast implants, make-up to male circumcision, I analyse the power structures and oppressive forces that demand we alter our bodies. Instead, I offer a vision of the human body that is equal without expectation: an unmodified body that is not an image of perfection or a goal to be attained, but a valued end in itself.

This work has had impact in several ways. Through publication of my book *Intact: A Defence of the Unmodified Body* (Chambers, 2022), and an extensive programme of public engagements, I work to raise public awareness and make a provocative intervention into public debate. Moreover, working with the Nuffield Council on Bioethics, I influenced

significant new legislation: the Botulinum Toxin and Cosmetic Fillers (Children) Act 2021 (UK), which makes cosmetic Botox and fillers unlawful for children under 18. In addition, through lectures to members of the Royal Society of Medicine and the British Association of Plastic, Reconstructive and Aesthetic Surgeons, I prompted many cosmetic surgeons to revise their practices. These are concrete ways in which work in philosophy can directly affect policy and practice.

But why should policymakers listen to philosophers specifically? Consider some perennial questions of political philosophy. Does freedom mean making your own choices, whatever they are, or can we sometimes be made more free by being guided towards more rational options? Does equality mean that everyone should be treated the same, or might it sometimes require treating people differently? Should people be held responsible for their choices even if that leaves them very badly-off? Does justice require taxing the rich to give to the poor? What is the difference between women and men, and what does that imply for how they should be treated?

These questions of political philosophy all make a considerable difference. Answering them one way or another has serious implications for many aspects of public life: what our laws should be, how we should punish criminals, who we should tax and how much, whether we should permit abortion or secure a right to maternity leave. But, a sceptic might ask, why should philosophers' answers to these questions be any better than anyone else's? Philosophers seem to lack both the technical knowledge of the scientific disciplines and the practical know-how of hands-on professions. Many philosophers devise new terminology and complex formalisations, but these hardly seem necessary for thinking about matters like fairness and the state. Political philosophers debate matters on which almost everyone has an opinion. We cannot claim a monopoly on thinking about questions of justice, and we are seldom awarded deference.

Philosophy matters to public life, though, because philosophers apply themselves to matters of profound significance using distinctive tools and for distinctive purposes. Our contribution lies in our ability to apply philosophical method in the pursuit of truth for its own sake. At our best, philosophers have a luxury denied to politicians: we can reason about politics without having to attend to politics. What I mean by this is that we do not fear the ballot-box; we are not constrained by being 'on the record'; we are accountable only as reasoners and not as representatives.

All of this means that philosophers do not have the power to implement our recommendations. We rely on our work inspiring others. That’s not a bug, it’s a feature. It allows us to consider political matters disinterestedly, without fear or favour. We can step back and view the bigger picture, zoom in and identify the founding principles, scan the horizon and identify contradictions and implications, and pass on the results to those with the democratic mandate to act.

What, then, is philosophical method? I’ll explore two types of philosophical methodology: reasoned argument; and careful attention to objections. Put this way, the principles don’t sound particularly distinctive to philosophy. But philosophy makes use of these methods in a way that politicians seldom can, with distinctive results that are a real contribution to public life.

Reasoned argument

Reasoned argument is the hallmark of philosophy. René Descartes employed this method to the extreme in his *Meditations on First Philosophy*, where he wrote:

Some years ago I was struck by the large number of falsehoods that I had accepted as true in my childhood, and by the highly doubtful nature of the whole edifice that I had subsequently based on them. I realized that it was necessary, once in the course of my life, to demolish everything completely and start again right from the foundations if I wanted to establish anything at all in the sciences that was stable and likely to last (Descartes, 1641).

Descartes attempted to derive philosophy from the ground up, using reason alone. He concluded, famously, that the one thing he could be certain of was that he himself existed: ‘I think, therefore I am.’

This attempt to follow the argument where it leads is a hallmark of philosophy. Some philosophers go too far — if the argument leads you somewhere terrible, perhaps you’ve taken a wrong turn along the way — but the basic idea is sound: philosophers start without being sure of their destination. My own approach to philosophy takes this method of reason as a motivation. I choose topics that seem to have real-world importance — this much is essential to me — yet are deeply puzzling. Sometimes I philosophise when I’m not sure what I think but I

know the issue matters. Alternatively, I turn to philosophy when I know what I think but I'm not sure if my view stands up to questioning.

An example of the first type is my work on marriage. I began that work because I knew that opposition to traditional marriage was a core tenet of feminism, but I was unsure whether marriage could be redeemed by reforms such as equal spousal rights and the recognition of same-sex marriage. Through philosophy — reading, thinking, talking, and writing — I identified a puzzle. How could marriage be both oppressive to women and liberatory to lesbian and gay people? Why was it bad both to be married and to be denied marriage? This question became a decade of research, culminating in *Against Marriage*.

Politicians and policymakers cannot spend ten years reading, thinking, talking, and writing about a question such as this; nor can they do so in a way that follows reason rather than re-election. This is not to disparage members of parliament (MPs), many of whom develop significant expertise and discharge their various committee and legislative responsibilities with care and clarity. But even the most dedicated MPs cannot dedicate themselves to a question of policy over multiple years. They must follow the issues of the day, develop a position on a wide range of policy areas, and follow party discipline. Perhaps this is why philosophers are sometimes found in the House of Lords, one part of our democratic process not shaped by electoral pressures. We want our law and public policy to be based on rigorous, careful thinking — the combination of research and innovation. Philosophers do this work in a way that MPs cannot.

My second motivation for philosophy is when I know what I think about an issue but I'm not sure if my belief is justified. This is how my work on *Intact* began. I wrote that book because I had long had an intuition that there was something good about leaving your body alone: letting it be just as it is, without unnecessary interference. I have no piercings, not even in my ears; no tattoos; and I have dyed my hair only once, in my teens. When I started the project, many modification practices were deeply unappealing to me in a way that felt ethically significant.

But why did I feel that way? Was it just a preference, even a prejudice? Or was it the identification of something significant? Could it be resistance to a culture that corrupts and distorts our bodies, layering them with shame? I wanted to know if my unexamined feelings were justified, and the only way to do that was to examine them. I am no Descartes but, like him,

I embarked on an attempt to think about body modification that did not rely on my established, unexamined beliefs.

This time, the research process took only three years — short in philosophical terms, an inconceivable age in politics. I was fortunate to receive funding from the Leverhulme Trust, which gave me three years of research time uninterrupted by teaching and administrative duties. I could not have written that book without relief from the considerable workload that is now standard in academia: though we are expected to produce world-leading research, we are not routinely given the time to do it. During those three years I read countless books and articles, learned from clinicians and beauticians, went to bodybuilding competitions and cosmetic surgery conferences, and talked to many people about their feelings about their bodies. The result was *Intact*, an impassioned manifesto for the political relevance of the unmodified body and an indictment of a culture that tells all of us, all the time, that our bodies are never good enough. I did not know where *Intact* would take me when I started thinking about it; I only knew that the philosophical method would get me somewhere worthwhile.

Attention to objections

Philosophy is the pursuit of truth via the most difficult path possible.

Philosophy is never finished because we can never be sure that we have found the right answer. As before, this is not a bug; it's a feature. In philosophy we can never be sure we have found the right answer because we must always be searching for the strongest possible objections to our view. We must test our conclusions against our opponents. We must consider for ourselves how our opponents can demolish us, and we should regard that demolition not as mere destruction but as the necessary precursor to construction.

Another way of putting this is to say that philosophers employ the principle of charity. It is not philosophy if you assert without considering objections, and it is not philosophy if you consider only the weakest version of your opponent's views. Instead, the principle of charity instructs us to interpret our opponents in the strongest possible light and to create the most damning objections to our own positions.

We address objections to our own position either by hearing them from others or by creating them ourselves. I once attended a seminar at University College London in which the renowned political theorist Cass Sunstein presented a paper. One student asked a question which, as always in philosophy, was framed as a critique of Sunstein's argument. (Philosophers don't ask the speaker for more information, no matter how eminent they are — we tell them they're wrong. You know you've given a good presentation when there are tons of questions and every one of them is an attempt to prove your error.)

The student's question was not particularly challenging, and Sunstein had an easy answer. But he did not end the exchange there. 'Now,' Sunstein said, 'if you change your question slightly, so that you ask me this instead' — here he set out a connected but different objection — 'oh wow! Now I'm really stuck! Now you've got me! Great question.' Sunstein helped the student to be a more effective critic, not merely as an act of pedagogy and kindness, but because philosophers like objections more than praise, and we like difficult objections most of all.¹

Unlike a commitment to reason, which is shared to some extent by all scholarly disciplines, the method and extent of seeking objections is quite distinctive in philosophy. We have a disciplinary reputation for toughness. At best this toughness is our strength; at worst it can descend into machismo. Other academics can interpret philosophical debate as aggression, as I learned when the anthropologist chair of a discussion at an interdisciplinary conference on marriage asked me to rein in my line of questioning because 'things were getting heated'. From my point of view — and, I'm pretty sure, that of the lawyers I was asking — the discussion was just starting to get interesting.

The urge to subject one's work to objections is in stark contrast to debate in political life, where the ultimate task of a politician often seems to be to prove that she is right and her opponent wrong — preferably by painting the opponent in as poor a light as possible. Adversarial politics and the electoral system mean that point scoring and misrepresentation are often rewarded. Point scoring for the sake of it can happen in philosophy too, of course, but when it does I regard it as a failure of the discipline; as having moved from philosophy to something else.

Another key distinction between philosophical combat and adversarial politics is the role of listening and answering. In philosophy it is essential to listen carefully to the challenge being put, because only then can one answer it precisely. In politics, in contrast, listening seems unimportant because politicians seldom attempt to answer questions. Instead, they seek

to deflect them. Politics elevates winning above being right; cares little about compromise if conquest is an option; subordinates truth to power. In a good philosophy seminar, the speaker asks the questioner 'did that answer your question?' If not, or if the answer brings up new problems, the questioner has another go. Public life could benefit enormously if its participants paid greater attention to the questions they are asked and were more careful to ensure they answered them.

The mark of a philosopher, then is that she *always* asks herself 'What if I'm wrong?' More demanding, she asks herself '*How* might I be wrong?' A question I routinely ask my students, and I apply it to my own work, is this: 'What is the strongest objection to your argument?' Next question: 'How would you answer that objection?' Next question: 'And what is the strongest argument against your answer?' When we go through this process we see not only that we have developed a stronger position of our own, but also we are aware of its limitations. We see when nuance lies. We see where ambiguity rests. We see how others could think differently, and we see what opponents might have in common. We understand that, although in philosophy nothing is ever settled, emphasising uncertainty brings us closer to truth.

So, what can philosophers bring to public life? We bring time and space to address the most complex questions, including political questions, away from political imperatives. We bring a commitment to the method of reason as a contrast to received opinion or ideological loyalty. And we bring a deep tolerance of disagreement and difficulty. In philosophy, conflict does not mean animosity, nor does it mean that one side is disreputable and must be silenced. It means that the question is a good one, and we should keep trying to answer it.

Notes:

- 1 At least, objections are what we like as philosophers. As humans we like praise, too.

References:

- Chambers, C. (2019) *Against Marriage: An Egalitarian Defence of the Marriage-Free State*. Oxford University Press.
- Chambers, C. (2022) *Intact: A defence of the unmodified body*. London: Penguin Books.
- Descartes, R. (1641). *Meditations on First Philosophy*. Cambridge University Press, 1986.
- REF 2021 (2023). *Impact database : Results and submissions: REF 2021*. [online]
Available at: <https://results2021.ref.ac.uk/impact>. (Accessed 7 February 2025)

Breaking bread with the enemy

Dr Anca Gheaus, Central European University

Towards the end of winter 2023, a minor scandal erupted on social media about an event that had taken place in Romania, my country of origin. Someone had filmed a man getting out of his car in a peripheral neighbourhood in Bucharest with two cats in his arms, which he then carefully placed over a low fence of a garden belonging to a block of flats. He then sought to get back into his car, in a rush. The woman who filmed this understood immediately that the man was abandoning the cats. She knew that their fate would likely be grim: they were about to become homeless in bad weather; they would be at risk of attack by stray dogs; and they would also have to scavenge for food, or to be at the mercy of people feeding them erratically. The greatest hazard for stray cats in Bucharest is other stray animals, especially dogs, themselves hungry.

The woman followed the man, continuing to film, and trying to convince him to change his mind, with a mixture of aggressive cajoling and outright shaming and blaming. The man attempted to explain he couldn't do otherwise. The exchange escalated to her calling him names; he fled. As Facebook users learned from this ill-fated dialogue, there was more to the story. The cats had been adopted from the streets as kittens, years before, by the same man. In fact, they were now being returned to the very same garden where he had found them. Not having led a stray life, they weren't streetwise: they didn't know how to chase mice or avoid the dogs that would chase them; they were worse off than their fellow street cats who had never known the protection of a home. It all seemed particularly cruel, and a bit absurd: why rescue cats, house them for years and then callously get rid of them? Many people were outraged, shared the news, and proposed bitter sanctions for the man. One of these people picked up the cats and, I think, eventually re-homed them.

No surprise that the story was made public by animal rights activists and their supporters. I learned about it reading one of their posts on Facebook, alongside plenty of outraged

comments on behalf of the cats. Perhaps I wouldn't have followed it up had I not come across a comment which, like the rest, was condemning the man, and in this case, also stating that his autistic child was no excuse for his actions. I was intrigued and went on to watch the footage myself. In it, the man spoke of the excessive burdens of his parental duties; he had no help outside of the family and was simply not coping. None of the people he asked wanted to take the cats, and there was no room for them in shelters — in short, he had no other solution. I don't remember many commentators willing to cut him any slack for his predicament. I wondered if the story had also been reported in activist circles concerned about the situation of neurodivergent people. I imagined readers coming down fully in the man's defence and expressing rage at those who put animals' fate above that of people.

This little anecdote sounds tamer than it looked on film; where I come from, people don't hold back from showing emotions, at least the unpleasant ones. I felt sympathy with the animal activists' outrage, and then I also felt sympathy for the hypothetical crowd who would probably have been outraged on behalf of the man and the rough treatment that *he* received.

This is the sort of situation that practical philosophers describe in the classroom when we teach ethics — in all its guises, including political philosophy. We try to explain how both the claims of the cats and those of the man matter. But, when something like this actually takes place, it is also the kind of occasion where one wants — where I would have wanted — a philosopher ready to step into a public role. Someone able to articulate the case against abandoning animals, and the case against demanding that people with difficult caring responsibilities push themselves beyond their limits for the sake of their pets. Maybe — I thought — if someone were to give the most persuasive case for both sides, the outpouring of anger could have been pre-empted. With less anger, and more receptiveness to understanding the man's motivations, there would have been less blame. People would have surely understood his reasons. And, who knows, it may have led to better prospects for improving the cats' lot. Not that philosophers have a unique power to throw light on the moral complexity of difficult choices. But, in virtue of our analytical training, and of the oft-repeated desideratum of uncovering and addressing the strongest objections to our views, we should be particularly well placed to do so.

I'm not, of course, suggesting that philosophical analysis could have done any good on the spot. Learning about philosophy in school, and earlier, could have helped: those who teach philosophy to children think that we are most receptive to it before puberty. An

understanding of philosophy in mass media and social media could have helped too. You might find it naive to suggest that philosophy can help those with conflicting interests break bread with each other; or that it can let us see that, often, the issues engaging our passions have no fully good answers, at least not in the short term. But I'm optimistic because I have myself been in those people's shoes. I'm not an animal welfare activist, nor have I struggled to improve the lot of neurodivergent people and their family members — however, I've witnessed many other social issues pitted against other perfectly legitimate concerns.

Because I spent time in activist circles, I know that philosophy can lead us out of our bubble of beliefs and commitments. It has the power to remind us that there are limits to what we should do, and expect of others, for the sake of our most dear ethical causes. It occasionally dispels the self-righteous rage which precluded any dialogue between the man and the woman in my story. While doing this, it also sometimes uncovers the source of the problem: in this case, that state institutions should have, but did not, put aside enough resources to avoid conflicts of existential interests between cats, people with autism, and their parents. Even when philosophy seems helpless in terms of guiding actions, there is something hopeful about it: it casts in a more truthful, better light those who would have otherwise been seen as pure villains. The habit of looking for reasons against one's own views, even against one's moral instincts, makes it impossible to depict as evil a man who cannot look after his child with additional needs and the family pets at the same time. It makes it equally impossible to dismiss animal welfare activists' concerns about vulnerable, cognitively complex animals being at risk of hypothermia, starvation and attack by dogs and other animals.

Russian writer and Soviet dissident Alexander Solzhenitsyn wrote: '...the line separating good and evil passes [...] right through every human heart' (Solzhenitsyn, 1973). Engaging with each other in what are, essentially, exchanges that are informed by philosophical candour can move that line, inching towards the good.

My favourite example is not a professional philosopher, but the North American Black jazz musician Daryl Davis, who, over several decades, managed to talk over 200 people out of their membership in the Ku Klux Klan (KKK) (Davis, 1998). He didn't do this by reason alone: he befriended many of these people, often through music and music-related conversations. Yet, the first step was almost always to knock on their doors, introduce himself and ask them about their reasons for being white supremacists. He studied the ideology of the Klan, to understand the racist beliefs that animated its members and meet

them with arguments that went to the core of those beliefs. He did this with the aim of eventually exposing the inconsistency between their racist views and other — better — ethical beliefs these people held.

Davis's accomplishment required him to retain hope that two people committed to radically opposed causes can nevertheless agree on some fundamental principles, and, eventually, on what they should do. He couldn't have succeeded without trust in his interlocutors' desire to justify their ethical commitments to others. That trust was repaid, and so his stories provide some optimism about the prospects of public philosophy, whether or not practised by professional philosophers.

Public philosophy can burst our social bubbles when people practising it assume that those with whom they radically disagree are responsive to reason and might get something right. This might be a controversial view, and, in any case, it invites several objections: that such attitudes are psychologically difficult to sustain, that they are dangerous, that they are perhaps even morally objectionable. There are people — philosophers included — who react with irritation when I tell them about Daryl Davis. Some say that his actions shouldn't be taken as a blueprint for what others can and should do. He is, they think, a privileged person, unlike others in so many ways that it's offensive to suggest that other victims of racial injustice should try and emulate him. Others go further, stating that Davis's willingness to put aside his feelings and talk with KKK members, giving them the benefit of the doubt, is hardly compatible with self-respect. And that engaging in this way is possibly unjust, as it treats the KKK members with undeserved good will. Worse even, it risks empowering them further.

The view of philosophy entailed by these objections is very different to my own. It understands the role of philosophy, when practiced professionally, as providing comprehensive criticism of the wrongful ideologies that pervade contemporary societies, including the mechanisms through which they reproduce, and crafting tools for undoing those ideologies. *Public* philosophy is conceived of as part of the pushback against pernicious ideologies, identifying allies for good causes and ameliorating the world by helping to drive those ideologies into extinction. On this account the line that divides good and evil primarily cuts through the ideological field, pitting harmful and desirable ideologies against each other. The issue of whether and how the line also cuts through the heart of every human being is less important.

One problem with this model is practical: contemporary societies contain multiple, complex and only partly overlapping ideological divides. Therefore, once we broaden the scope of our concern to include all the causes worth fighting for, few alliances can be stable enough to be effective. Worse still is the problem of principle. Many conflicts of interest don't have (politically feasible) solutions which fully satisfy all the legitimate claims. For instance, parents of autistic children cannot always continue to look after domestic animals who need their protection and who may have gotten attached to that family in particular; it's either one kind of suffering for one party, or another kind for the other, or maybe suffering of a lesser magnitude for all involved. And stories like this raise questions about public policy: for example, under what circumstances should people be legally able to abandon their pets? Without principled rejection of an 'us versus them' attitude, it is easy to become polarised about these matters and fail to do justice to all parties.

Anyone interested in ideology critique will of course note that the conflicts of interests in this case occurs because the background institutions that allocate resources do not work as they should. Our societies are rich enough to enable considerable luxury for a large minority, and should make it a priority to address non-luxury issues: for example, to ensure that children with special needs and their parents don't find themselves in practical conundrums like the one above, and that animals are shielded from hunger, cold and violence. But, until and unless institutions provide adequate resources, — that is, the collective "we" — do their job, we will continue to be caught in similar conundrums. It is important how we react to these conundrums, and how the public reacts to those who make the choices.

Besides, it is not possible to pre-empt all conflicts of interests through better collective action and more just institutions. The difficulty can be inherent to the problem: should we, say, impose high risk of ill mental health and serious developmental problems on young children for the sake of protecting the lives of others, and in particular the elderly? Should we do this even if the younger generation has worse life prospects than those of their grandparents? Readers will recognize the inter-generational conundrum raised by the recent pandemic. And in some cases, the conflict can involve different but similarly weighty interests of the same party. For instance, is it more important to prevent suffering so extreme that it might end in suicide in pre-pubescent children with gender dysphoria by providing access to puberty blockers; or ought these children to be protected from embarking on a momentous process whose bearing on their lives they are too young to evaluate? Unlike situations like the man abandoning his cats, these issues tend to polarise not only public

opinion, but also many professional philosophers. It is a mistake to dismiss some of the arguments on each side in these and similar dilemmas; in the most charitable interpretation of issues such as these, both sides are animated by reasonable concerns.

I submit that an important role of philosophy — whether done in conference rooms and journal articles, in classrooms, or in more public venues — is to reduce polarisation. Well-executed public philosophy is needed to show how the deepest disagreements of our age might involve people who are all motivated by the same fundamental principles and values, yet focus their attention on individual pieces of the practical puzzles at the expense of the general picture. To gain such understanding, philosophers must assume that the puzzles really are puzzling, that they lack obvious solutions. People are already engaged in this enterprise of public philosophy, which is proof of its possibility, though there is no denying the psychological burden involved. But is the alternative of not doing it less burdensome? According to one understanding of self-respect (Bird, 2010), its pinnacle is reached when one can withstand others' mistreatment without doubting one's own moral worth. If so, it's worth taking the moral risk of genuine dialogue with KKK members. And the fact that so many of them responded to Davis's influence and abandoned the Klan suggests they deserved his generosity.

Public philosophy is necessary for figuring out some of the most pressing problems of the day, from economic policies to cultural wars, without vilifying the supporters of different solutions. This little essay, then, turns out to be as much about the role that philosophy should have in public life as about what professional philosophers would have to do if our work is to play that role. It is not really a piece about what philosophers-as-they-are can do to improve things, but about what philosophers-as-they-could-be would be especially well-placed to do. The problem confronted by Davis is philosophically easy; many others aren't. This is why we should marshal not only our analytical tools but also our experience with the complexity of ethical problems to reject glib solutions; and, while doing so, reject the sentiment expressed in Elizabeth Anscombe's verdict: 'I do not want to argue with him; he shows a corrupt mind' (Anscombe, 1981).

References:

- Anscombe, G. (1981) *Modern Moral Philosophy: Ethics, Religion and Politics*, Oxford, Blackwell.
- Bird, C (2010). 'Self-respect and the Respect of Others', *European Journal of Philosophy* 18(1): 17–40.
- Davis, D (1998). *Klan-destine relationships: A Black man's odyssey in the Ku Klux Klan*. New Jersey. New Horizon Press.
- Solzhenitsyn, A (1973). *The Gulag Archipelago*. London. Vintage Classics.

Of weasels and women, or, what is public philosophy anyway?

Dr. Katharine Jenkins, Glasgow University

When I was asked if I would like to write something about public philosophy for this volume, my initial thought was: *OK, yes, I've done enough public philosophy to be in a reasonable position to say something about that.* Implicit in this thought was the idea that I had a fairly good handle on what ‘public philosophy’ was — good enough, at any rate, to be confident that I had done some. Like the hapless dinner guests hanging out with Socrates in one of Plato’s dialogues, however, I swiftly came to realise that it wasn’t this simple. Not only could I not come up with an explanation of what I meant by “public philosophy”, but when I tried to think about the different occasions on which I have “done some public philosophy”, I immediately ran into difficulties figuring out which things should count, and which should not.

For example, I specialise in the philosophy of gender, and in 2018 there was a public consultation running on reforming the Gender Recognition Act 2004, the legislation that enables some trans people to change the gender marker on their birth certificate. Together with two other philosophers, Lorna Finlayson and Rosie Worsdale, I wrote an essay aimed at the public that ran on the blog of the left-wing publisher Verso (Verso, 2018). We argued in favour of reforming the act to enable people to change the gender marker on their birth certificate based on a simple declaration of their self-identified gender, rather than having to submit ‘evidence’ to a panel of ‘experts’. With the support of the Nottingham Institute for Policy and Engagement, I ran a session in Westminster for civil

servants, MPs, members of the House of Lords, and their staff, and co-wrote an associated briefing document with Ruth Pearce, a sociologist and an expert in trans health (Jenkins and Pearce, 2019).

I also met individually with civil servants from the Government Equalities Office to discuss the consultation. A key philosophical point that I made on each of these occasions, as well as in a subsequent book, is that we should not think that questions about how society should handle gender recognition can be settled by working out what a woman ‘really is’, and thus which people ‘really are’ women (Jenkins, 2023). Rather, it should be approached as a pragmatic and political question: which of the arrangements around gender recognition that we could adopt at present best promote people’s safety and dignity? This question, I have argued, is much simpler to answer, and points firmly towards reform.

Before thinking about it too much, I would have said that each of these activities should count as public philosophy. Yet in terms of their format, they were certainly very different. Some were in-person activities; some were written material; some were on the internet for anyone to read; some were open to a specific group of people; and some were literally me and two other individuals in a room. If we think that ‘public philosophy’ is philosophical activity that somehow involves ‘the general public’, then most of these activities wouldn’t count.

Perhaps, though, public philosophy is not about *reaching* the public directly — that does seem a bit too literal — but about something like ‘influencing public life’. That would include things like meeting privately with civil servants, because the work of the civil service is surely part of public life, if anything is.

But there’s a hitch here too — in fact, there are several. For one thing, some philosophy that has had a very great influence on public life doesn’t seem appropriately described as public philosophy. Especially with quite technical areas of philosophy, it is possible for philosophical work to influence people’s everyday lives significantly, but in an indirect way. Think of Alan Turing’s philosophical work on computation, which has had a far-reaching effect on so many aspects of society via its impact on computers (Wikipedia says he is ‘widely considered to be the father of theoretical computer science’, which sounds about right). I’m not an expert on his work by any means, but I’ve read some of it, and it’s hardly something most of us would consider to be public philosophy.

In the other direction, what about cases where a philosopher doesn't manage to change the direction of aspects of public life as intended? This is, for the most part, the case with the activities mentioned here. My impression — doubtless fallible — is that the responses to the public consultation around the Gender Recognition Act pointed towards reforming it in the direction of self-identification, which is also the outcome that my interventions sought to promote. This reform then didn't happen because it wasn't politically expedient for the government of the day — trans people are a convenient football in the 'culture wars' — and, unfortunately in my view, subsequent political dynamics have shifted us, if anything, even further away from reform.

If we want to think of public philosophy in terms of influencing public life, then, we will end up including more than perhaps we bargained for, and we will need to be happy with a pretty homeopathic understanding of 'influencing' for many of the things that we usually think of as public philosophy to count. Either that, or we will need to appeal, not to the philosophy's effect in the world, but on the intention of its originator. That, however, seems much too cheap.

For example, a scholar penning a convoluted and pedantic treatise that only a handful of their fellow academics are likely to read, let alone understand, might be convinced that a proper understanding of their views on justice would change the course of global politics, and might fervently hope that this will come to pass. But it doesn't seem right to say that they are therefore doing public philosophy.

Equally, someone might blithely jot down an argument about something like land rights or just war theory that, while fairly abstract on the face of it, is seized on by those who seek to dispossess or even exterminate a particular group of people, and used to lend an impression of legitimacy to their genocidal campaign. The philosopher here might be able to quite sincerely say that they didn't mean anything to *come* of it; they were simply working through a philosophical view. Now, suppose that this genocide was a well-known risk at the time the philosopher wrote, and that the philosopher was either aware of the potential connection to their work, or would have been aware if they had taken the time to think about it. In cases such as this, talking about the philosopher's *intentions* seems rather beside the point.

The philosopher's protestations might seem both more relevant and more reasonable if we imagine that the risks of wrongs to the targeted group of people were not knowable at

the time the philosopher was writing, but only emerged later on. Still, even in this scenario, the philosopher's work *did* influence events in the world at large, and it seems wrong to completely discount this fact when it comes to thinking about what kind of philosophy it is.

Overall, then, one's philosophical work can fail to have the effects one intended, and can have effects one did not intend, or even foresee. So there seems to be something odd about treating my intentions about my philosophical work's effects as decisive when thinking about its status as public philosophy or not. In fact, this seems like giving philosophers licence to be out-and-out weasels, sneakily disclaiming responsibility for the predictable consequences of what they say and write on the grounds that they didn't *mean* for those things to happen.

A similar worry also occurs if we think, not about a work of philosophy's intended *influence*, but, perhaps more simply, about its intended *audience*. Going back to the activities listed earlier, one thing they have in common is that they were not intended to just reach other academics. By contrast, when writing an academic journal article, one is usually anticipating that its only readers (if indeed there are any at all) will be other academics — and one writes accordingly. Similarly, when writing something like a textbook, one is mostly anticipating that it will be read by students. So maybe that's a simple, and fairly modest, way to characterise public philosophy: it's philosophy that's aiming to reach *at least some people who are not academics or students*. But of course, the mere fact that I wrote something with a wider readership *in mind* does not guarantee that what I write is at all *suitable* for this purpose or *likely* to indeed be widely read.

Might we try once more to get away from intentions? We could simply say that public philosophy is philosophy that *is in fact engaged* with by people outside the academy/higher education. As before, this opens up a potentially troubling amount of uncertainty. In my own case, for example, of all the philosophical activities I've engaged in, the one that attracted the most public attention was something that I hadn't been thinking of as public philosophy at all. In 2022, I organised, along with the philosopher Alexis Davin, a small online conference on trans philosophy. Although it was open to all — anyone could submit an abstract or sign up to attend — we felt it was unlikely there would be very many participants who were not academics or graduate students, and we certainly didn't advertise it much beyond the usual channels that such people tend to keep an eye on. However, a former philosophy academic made disparaging remarks about it on social

media, and this somehow became a news story in a national broadsheet newspaper (the substance of which really was as minimal as: an event is happening, and a person thinks it is silly).

I'm not sure what to make of this — was the conference not public philosophy to begin with, but then became public philosophy after the news story ran? That would pre-suppose that organising an event along the lines of a conference is the sort of thing that could in theory count as public philosophy. That is, we're not restricting the idea just to instances when one engages in an act of philosophising, such as giving a talk oneself. But even if we grant this assumption,¹ it seems peculiar to say that the publicity itself made a difference to the type of thing the event *was*, given that the conference itself proceeded in pretty much the same way it would have done if the newspaper story hadn't run (zoom security was little tighter, due to risks of trolling, but that hardly seems like the right kind of difference).

So it's a complicated question — perhaps even philosophical question — what counts as 'public' in the context of public philosophy? Does it mean addressing some philosophical work to (large audiences of?) non-academics? Or to people with some kind of public role, such as policymakers? In either of these cases, at least you would normally know whether you're doing it. But what if it is more to do with having influence? Or actually reaching a wide audience? In either of *those* senses, you could do public philosophy without knowing that you were doing it — at least, not until some (possibly quite distant) later date — and you could also find yourself doing public philosophy of quite a different kind than you would ever have intended (as when your work is used to support some policy or action that you in fact deplore). And all of this is without our even beginning to wrestle with the question of what makes something — some writing or activity — *philosophy* in the first place.

At this point, you might be wondering whether the difficulty of defining public philosophy is an *interesting* difficulty. It is fairly simple to point to something — anything really — and insist that we do not have an adequate definition of it. In fact, this might seem like the pet hobby of the philosopher with too much time on their hands. *So what, you might think, if we cannot give a definition of public philosophy? We know it when we see it, and that is quite enough to be getting on with.*

It's a good question to ask; sometimes, things don't admit of a clean definition, and pointing out that we can't find one doesn't really have any practical implications: a niche pastime

at best, and a waste of time at worst. It might even seem like its own brand of weaselling: *Making a difference in the world? No thanks, I'll just stay in my comfy armchair and think about what public philosophy really is.* I don't think this is one of those cases, though. As I've tried to demonstrate, we don't always know public philosophy when we see it; I don't, at any rate.

More importantly, if we find in the end (as I suspect we will) that no neat account can be given of public philosophy as some cleanly separated type of philosophy, then I think this *would* show something interesting, namely, that all philosophy is at least *potentially* public philosophy. This stance aligns to at least some extent both with philosophers known as critical theorists, who think that the job of the philosopher is to theorise the society they are in, which includes them and their philosophical activities; and with feminist philosophers, who have been sceptical of a sharp distinction between 'the public' — often associated with men — and 'the private' — often associated with women (Jenkins, 2024).

So the trouble we face in defining public philosophy should, I think, lead us to the view that it is wrong to think that philosophy happens in a bubble separating it from society, or that the ways in which our philosophy connects with society are totally under our control. If philosophers are thinking otherwise, we're really just kidding ourselves. And that way lies weaselling, for this self-deception may well be a convenient one, in that it seems to exempt us from the messy and often dispiriting business of thinking about how what we're doing fits into the world around us. Only it doesn't exempt us; nothing can.

If this is right, then it's *not* trivial or cheap to point out the difficulty of defining public philosophy, because it leads us to realise that it is not some distinct thing that can be neatly hived off from philosophy in general. Avoiding this mistake should help to keep us honest about the philosophy we do — *all* the philosophy we do.

What does it look like, then, to do philosophy in the absence of a fairy-story of separateness? This is a complicated question, and deserves a more thorough treatment, but here's the best I can offer right now. Although we may at times deliberately seek to contribute to public discussions we consider important, we should also understand all our philosophical work as at least potentially interacting with public life (whatever exactly that is), even if we feel this is unlikely in a particular case. And yes, I really do mean *all* philosophy, on any topic, however esoteric. Even if there is some philosophical research that in fact has no chance whatsoever

of impacting public life — and I think this is likely — my point has been that we are not in a good enough position to reliably identify it in advance.

Where we can discern points of connection between our philosophical investigations and our social context, or suspect there might be some, we should take the time and trouble to inform ourselves about this context, considering this to be part of our core task as a philosopher. We should write to be understood as widely as possible, on the grounds that we cannot predict who it will be relevant to. And we should remember that what happens next really isn't up to us.

Notes:

- 1 It seems like a pretty reasonable assumption to me, but if you don't agree, then think about one of the talks that was given at the conference, and transpose my questions to that instead of the conference as a whole.

References:

- Jenkins, K. and Pearce, R. (2019). *The Gender Recognition Act: A trans-inclusive feminist approach*. Available at: <https://www.nottingham.ac.uk/policy-and-engagement/documents/policy-briefs/briefing-gender-recognition-act.pdf>
- Jenkins, K. (2023). *Ontology and Oppression: Race, Gender, and Social Reality*. New York: Oxford University Press, pp. 200-239.
- Jenkins, K. (2024). *Feminist Philosophy: A Very Short Introduction*. New York: Oxford University Press, pp. 42-65; 66-90.
- Verso (2018). 'I'm not transphobic, but...': A feminist case against the feminist case. [online] Available at: <https://www.versobooks.com/en-gb/blogs/news/4090-i-m-not-transphobic-but-a-feminist-case-against-the-feminist-case-against-trans-inclusivity>.

On the inequity of ethics

Dr Winnie Ma, Juris Doctor Candidate at UC Berkeley
School of Law

1. Introduction: The apparent paradoxical inequity of ethics

Prevalent today is the sense that academic philosophy is irrelevant to the real world and to public life; that philosophy is the paradigmatic 'ivory tower' discipline and profession, perhaps by definition concerned with the abstract and ideal and unconcerned with messy and complicated practical realities. Indeed, concern that there is a kernel of truth to this impression, and that as a philosopher, I personally needed, first, more hands-on experience with the subjects of my research, which include AI, profiling, stereotyping and discrimination; and second, a more direct means of applying my research to create positive societal change through influencing policy, law, public opinion etc. It was this desire to make a practical impact that led me — perhaps only temporarily — to abandon academic philosophy to pursue a career in law.

This chapter will focus on a particular factor contributing to the widely held sense that academic philosophy, and in particular ethics, is irrelevant to the real world and to the messy matters of public life: what I call the apparent paradox of the inequity of ethics.

(a) Ethical imperative #1: Make AI/tech ethical.

We might, for example, take big tech companies like Google, Meta and Microsoft. Only enormous tech companies like these with hundreds of billions of dollars in yearly revenue have engaged teams of experts to reflect on and implement ethical guidelines to govern their behaviour. Startups with much more limited financial resources very rarely engage such ethics teams. But this is not necessarily because super-wealthy big tech companies are 'more ethical' than smaller tech-companies. Rather, it is often because only they can

afford to engage these ethics teams, and implement the ethical guidelines that these teams develop.

And we should of course be attentive to the efficacy of and real motivations behind these initiatives by big tech companies. For example, many critics have questioned the sincerity of big tech's initiatives, accusing them of ‘ethics-washing’, or the practice of insincerely and often ostentatiously presenting oneself to the public as being concerned with ethics, without, however, doing much of substance to pursue ethical objectives. We could add the further nuance here that it might be that only big tech companies and other financially privileged actors can afford to make themselves appear ethical, which might come apart from and perhaps be more cost-effective than actually acting ethically. Other critics have noted that ethics teams have tended to be under-resourced and under-supported even in big tech companies. Finally, the fact that such ethics teams are often last hired (in robust economic periods, and after such companies have reached a high threshold level of financial security) and first fired (for example, during economic downturns) may also be indicative of big tech companies’ actual priorities regarding ethics versus, for example, profit (Belanger, 2023; Ars Technica, 2023; Criddle and Murgia, 2023; Criddle and Murgia, 2023; Field and Vanian, 2023).

(b) Ethical Imperative #2: Eat ‘ethical foods’.

As another example of this paradoxical inequity, consider the ethical food movement, and more generally the idea that it is more ethical to purchase and consume foods certified or labelled ‘fairtrade’, ‘organic’, ‘sustainably-sourced’, and so on, as well as foods that are considered healthy, as opposed to foods without the organic or other ‘ethical food’ label, or ‘junk foods’ like McDonald’s or Burger King. You might recall instances in which, at the grocery store, you were faced with a decision between purchasing a certified organic or fairtrade or sustainable food item, and a conventional, non-certified alternative that cost less. Many times — often with a pang of guilt — I’ve chosen the cheaper, ‘less ethical’ alternative because of financial constraints. Perhaps too, if you’ve made similar food-related decisions for yourself or others in which you’ve prioritised budget over ‘ethics’, you’ve experienced a similar pang of guilt and/or others’ spoken or unspoken negative judgement.

It is at least on the face of it paradoxical, however, that abilities to engage in the purportedly ‘more ethical’ behaviour of purchasing and consuming ‘ethical foods’ are inequitably distributed in favour of the financially better-off and those with fewer constraints on their

time and other resources. It is also apparently paradoxical that the capacities and resources required to competently reflect on how to make ethical food decisions are similarly inequitably distributed in favour of those with money, time and energy. For myself, I only attained a very rudimentary understanding of the ethical foods landscape after watching the lectures of a free Coursera course titled 'Food Ethics' (Chignell, 2024). My understanding was rudimentary in the sense that, through watching these lectures, I have only begun to understand how complex and unclear the answers are to questions about which foods are ethical and why.

(c)Ethical imperative #3: Don't be a sellout; be a public interest lawyer.

Another example of the apparent paradoxical inequity of ethics is the acute tension felt by many law students from low-income backgrounds between their need for financial security — typically gained by working for a corporate or "Big Law" firm — and their desire to become public interest lawyers — typically working for non-profit organisations or the public sector, for significantly lower pay.

Emblazoned on a highly ranked law school's public interest careers page is a quote by US Supreme Court Justice Sonia Sotomayor: 'We educated, privileged lawyers have a professional and moral duty to represent the underrepresented in our society, to ensure that justice exists for all, both legal and economic justice.' While I think that the vast majority of my peers certainly agree in spirit with this, many also struggle to balance this sense of duty with the practical burdens of massive student debt and feelings of responsibility to provide for their families. And for many students, I think this also involves a hard choice between pursuing their passions (for example, of working in the public interest sector) and financial security. And oftentimes the very idea of figuring out what their passion is feels like a privilege that has been and will be kept out of reach until after achieving financial solvency and security.

Such students are often faced with a choice between taking the supposedly more ethics-driven path and working in public interest law, versus working in large, private law firms where salaries are generally significantly higher. Some choose to first do their time in big law firms, where the hours tend to be long and arduous, to gain experience and pay off student loans much more quickly than they otherwise could. And, after they've attained a level of financial security, they permit themselves to turn their attention to public interest work or to figuring out what kind of work they're really passionate about. For all of us who will be making similar career choices in law and in other domains, however, I think it's clear that privilege is a hugely

significant factor — and one that can sometimes press us in conflicting directions. That must be taken into account when considering what it means to make ethical choices.

2. On equity, diversity, and inclusivity in academic philosophy

Related to the above discussion about the relationship between privilege and public interest work in the legal domain, we also need to consider the persisting fact that, although progress has been made in recent years, in the US and the UK those studying and working in academic philosophy — including those working in ethics — still disproportionately come from privileged as opposed to marginalised backgrounds with respect to socioeconomic status and, relatedly, ethnicity, gender, (dis-)ability, etc (Schwitzgebel, 2020).

For example, results from various surveys — from the UK, Australia and the US — show that philosophy undergraduates, postgraduates, faculty staff and professors are disproportionately white and male.

One reason I left academic philosophy for law school was because of increasing financial responsibilities that had become incompatible with the relatively low salary levels of philosophy professors, especially early career ones. More generally — given the lower salaries in academic philosophy compared to fields such as law, and science, technology, engineering, and mathematics (STEM); the limited available funding for graduate study in philosophy; and the time required to train as a professional philosopher — the study of philosophy, especially at the graduate level, often becomes accessible primarily to the financially privileged.

Many have already written about the implications of the skewed demography of academic philosophy (Schwitzgebel et al., 2021; Steward, 2023; De Cruz, 2018). This skew in philosophy, and here in particular in ethics, can mean analogously skewed ethical perspectives. An ethical perspective worked out by group X, for example, will often tend to focus on the particular challenges, concerns and preferences of that group, often taking away focus from the particular challenges, concerns, and preferences of other groups Y, Z, and so on. In the case of academic philosophy, given the greater average degree of financial privilege of our demographically skewed cohort of academic ethicists, there's the danger of mainstream ethics being unduly insensitive to the concerns of the less financially privileged.

Thus, these skewed ethical perspectives likely contribute to the paradoxical inequity of ethics today. That is, mainstream ethical systems, as formulated and propounded by a demographically skewed philosophical academic establishment, seems to be guilty of, among other things, being inequitable both in terms of their practicability and their theoretical accessibility. Ethical imperatives — such as defining and embodying transparency, responsibility and fairness (especially in tech companies); determining which foods are ‘ethical’ foods and consuming them accordingly; opting for public interest law over a higher-paying job at a Big Law firm; or even studying philosophical ethics to better understand and practice ethics — seem to be more easily fulfilled by those who are financially and otherwise privileged. Such ethical imperatives are also likely, justifiably, to irritate those on whom they are unsolicitedly pressed, especially if there are asymmetrical power relationships at play between the presser and pressed.

But, if we accept that there is such a paradox — that ethics today is inequitable in that those with various kinds of privilege can more easily learn how to engage in competent ethical reflection and behave ethically due to their privilege — doesn’t that undermine the legitimacy of today’s ethics? Put another way, how can an inequitable ethics be legitimate?

3. More questions (than answers?)

I’m not sure I have definitive — or even particularly good — answers to this or other questions that naturally follow from it. Such follow-up questions include: If our ethics is/are inequitable, and if all ethics are formulated by individuals with different positionalities, challenges, concerns and preferences, can we escape inequity in ethics? Or must we accept it as an intrinsic feature of any human-formulated ethical system? Moreover — and this question verges into what philosophers call ‘metaethics’ — have we assumed, and should we assume, that all ethical systems are human-made? And if we accept that all ethical systems are human-made and therefore always at least susceptible to inequity, are there still ways for us to mitigate this inequity, and in doing so, enhance the legitimacy of the ethical systems we follow?

Another related string of questions may include: to what extent, if at all, do we believe that formal (philosophical) training is necessary for engaging in the kind of competent ethical reflection that ought to shape the ethical imperatives and norms guiding our societies? Conversely, to what extent should ethical reflection be (re-)conceptualised as a democratic

activity — one that anyone can competently engage in, regardless of their personal background or level of philosophical training, including formal ethics training? Should the ethical reflections of any member of the public be deemed as just as credible as those of an ‘ethical expert’, such as a professor of philosophical ethics? And, if we (re-)conceptualise ethical reflection as the kind of ‘democratic’ activity discussed above, how do we mitigate the problem of the marginalisation of minority voices, that is, minority ethical perspectives — a problem that plagues all democratic systems?

Again, I don’t think I have answers to many of these questions. However, to respond to a few, intertwined lines of inquiry here, I can venture one minimal suggestion, which has of course already been put forth by many others: we should work to promote greater inclusivity in the composition of ethicists who shape the ethical frameworks that guide our societies. This could (and, I think, should) mean making the relevant kinds of philosophical/ethical training and employment (for example, as philosophical ethics professors) more accessible to all members of society. But this could also mean rethinking what kind of training and/or expertise competent ethical reflection requires.

4. Conclusion

To conclude, this chapter was written partly to combat the widely held view that philosophy is irrelevant to, unconcerned with, or potentially even harmful to public life, and especially to the lives of its most marginalised members. It was also written in part to grapple with personal concerns about the practical utility of philosophy and the many kinds of privilege associated with philosophy and philosophers today. These concerns are in part what motivated my move to the profession of law. In law, I hope to apply my previous philosophical work on topics like the ethics of belief, stereotyping, profiling, and algorithmic bias and discrimination through practical, legal routes to make a tangible impact in the world.

Partway through my legal studies, I am still optimistic, although by now I’ve seen that law certainly faces its own, perhaps even greater, challenges around inequity and privilege. We are left with the questions posed above, and I ask for your help in answering them. Because, at the very least, when it comes to the question of who should decide who engages in competent ethical reflection on behalf of our societies, it seems as though everyone — philosophical expert or not, lest we beg the question — deserves to have an equal say.

References:

- Beebee, H. and Saul, J. (2021). *Women in Philosophy in the UK*.
<https://bpa.ac.uk/wp-content/uploads/2021/11/2021-BPA-SWIP-Report-Women-in-Philosophy-in-the-UK.pdf>
- Belanger, A. (2023). *Report: Microsoft cut a key AI ethics team*. Ars Technica. <https://arstechnica.com/tech-policy/2023/03/amid-bing-chat-controversy-microsoft-cut-an-ai-ethics-team-report-says/>
- Chignell, A. (2024). 'Food Ethics'. Coursera. Princeton University.
<https://www.coursera.org/learn/foodethics/home/info>.
- Criddle, C. and Murgia, M. (2023). Big tech companies cut AI ethics staff, raising safety concerns. *Financial Times*. [online] 29 Mar. Available at: <https://www.ft.com/content/26372287-6fb3-457b-9e9c-f722027f36b3>.
- De Cruz, H. (2018). Prestige bias: An obstacle to a just academic philosophy. *Ergo*, 5 (10).
<https://doi.org/10.3998/ergo.12405314.0005.010>; [to add more citations].
- Field, H. and Vanian, J. (2023). *Tech layoffs ravage the teams that fight online misinformation and hate speech*. [online] CNBC. Available at: <https://www.cnbc.com/2023/05/26/tech-companies-are-laying-off-their-ethics-and-safety-teams-.html>.
- Goddard, E. (2008). *Improving the participation of women in the philosophy profession*. Australasian Association of Philosophy: Committee of Senior Academics Addressing the Status of Women in the Philosophy Profession. Available at: http://www.aap.org.au/Resources/Documents/publications/IPWPP/IPWPP_ReportA_Staff.pdf
- National Association of Colleges and Employers (NACE). (2022). *NACE Salary Survey: Starting salary projections for Class of 2022 new college graduates*. https://careers.unc.edu/wp-content/uploads/2022/02/NACE-Salary-Survey_Winter-2022.pdf
- National Science Foundation, National Center for Science and Engineering Statistics (2019). *Doctorate Recipients from U.S. Universities: 2018*. Special Report NSF 20-301. Alexandria, VA. Available at <https://nces.nsf.gov/pubs/nsf20301/>.
- Peterson, G. and, & Hustoft, Z. (2021). *Second Wave Survey of Women's Representation in U.S. Philosophy Graduate Programs*. Harvard Dataverse, V1 <https://doi.org/10.7910/DVN/IJZW8Y>.
- Ryan, J. and Sackrey, C. (1999). *Strangers in Paradise: Academics from the Working Class*. South End Press.
- Schwitzgebel, E. and Jennings, CD. (2017). *Women in Philosophy: Quantitative Analyses of Specialization, Prevalence, Visibility, and Generational Change*. *Public Affairs Quarterly* 31:83-105.
- Schwitzgebel, E. (2020). *Diversity in Philosophy Departments: Introduction*. Blog of the APA; Blackwell Publishing Ltd. <https://blog.apaonline.org/2020/06/11/diversity-in-philosophy-departments-introduction/>.
- Schwitzgebel, E., Bright, L., Jennings, C., Thompson, M., and Winsberg, E. (2021). *The Racial, Ethnic, and Gender Diversity of Philosophy Students and Faculty in the United States: Recent Data from Several Sources* [online] Available at: <https://faculty.ucr.edu/~eschwitz/SchwitzAbs/DiversityUS.htm>.
- Stewart, S. (2023). *Philosophy Needs More Women*. The Atlantic. <https://www.theatlantic.com/books/archive/2023/04/how-to-think-like-a-woman-regan-penaluna-book-review/673728/>;

Philosophy in the flow of political life: realism, moralism and community wealth building

Professor Martin O'Neill, University of York

Debates about the relation between political philosophy and real-world political life can sometimes tend towards extreme positions. A more careful attention to the place of values in political life can help us towards a genuinely 'realistic' view of the relation of theory and practice, which avoids the lure of both a naive moralism *and* of an equally unsatisfactory anti-moralism that takes things too far in the opposite direction. Moreover, an embedded approach to 'philosophy in the flow of political life' gives us a more productive engagement with our core political values, and a better understanding of the demands of those values in their full significance and complexity.

Philosophy and the real world: 'Enactment' and the spectre of moralism

In his influential critique of 'moralism' as a way of thinking about the role of political philosophy in relation to political practice, Bernard Williams attacked what he called '*the enactment model*', whereby 'political theory formulates principles, concepts, ideals and values; and politics (so far as it does what the theory wants) seeks to express these in political action,

through persuasion, the use of power, and so forth' (Williams, 2005). On Williams's view, the enactment model posited an imagined division of labour that might seem rather attractive to philosophers (or at least to *some* philosophers); a division of labour in which most of the important substantive content of politics is seen as settled already in the philosopher's study, or the seminar room, with the work of practical politics then relegated to a technocratic activity of implementation (performed more or less faithfully, and more or less skilfully). Fundamental philosophical work on political values and principles is imagined here as done in isolation from the practical domain of politics, even if there is some more applied work of translation and application that then has to be done in light of an assessment of circumstances. The political domain is then viewed as in some sense subservient to the domain of philosophy — a zone of translation and implementation, to be judged primarily with regard to how well it manages to bring reality into line with what philosophy would demand of it.

On this kind of 'moralist' view that he was seeking to criticise, Williams diagnosed the central error in this approach as seeing politics as a mere instrument for realising the demands that are made of the world by philosophy. According to Williams, any view that builds in the priority of morality to politics, or which sees politics as 'something like applied morality' has made a deep mistake: it has failed to register the autonomy of politics as a distinctive domain of human life. Any view that sees politics as no more than a transmission mechanism for the enactment of ideas that have already been fully worked out in the philosopher's study can thereby be seen as both implausible and hubristic. The *enactment model* fails to register the significance of politics as a distinct domain of human interaction, while also placing an unsustainable weight of significance on what Williams called the philosopher's 'panoptical view', through which the philosopher's perspective on the world is that of a semi-detached observer 'surveying it to see how it may be made better'.

Williams's paradigm case of the *enactment model* was utilitarianism; and while it is not the aim of this essay to prosecute the case against that particular way in which utilitarian philosophers often approach questions of policy, one might well take the view that there are abundant recent examples of how utilitarianism — especially in its 'longtermist' variant — has clearly put itself four-square in the way of Williams's critique. Any honest assessment of Williams's attack on the *enactment model* would have to allow that in many cases philosophers' engagement with issues in public life can display both a tin-ear for distinctively political questions, and an almost comical hubris in their abundant faith in their own panoptic vision.

Williams himself moves from a critique of 'moralism' to defending a more minimalist conception of the role of philosophy in political life, according to which view ideas of basic legitimacy — what Williams calls the 'basic legitimation demand' — should be kept at the centre of philosophical engagement with the political domain. My aim here is not to assess Williams's brand of 'realism' in detail, but instead to make the case for a broader role that can and does exist for normative political philosophy, and for philosophical thinking about values and principles, seen as *part of* the public political life of our societies. My argument is in favour of a clear-eyed appreciation of the real potential for philosophy within politics, and in defence of the philosophical articulation of political values as an internal rather than external aspect of our shared political life itself. We can accept something like Williams's critique of 'moralism', as applying to various forms of badly conceived and badly executed philosophical engagements with political reality — as epitomised by the cruder instances of the 'enactment model' — but this does not mean that a genuine appreciation of the complexities of political reality need involve too sweeping a rejection of a role for normative political philosophy. Our values and principles are a central part of our politics, and philosophy is indispensable in the development and articulation of those values.

In making this case, I want to draw a contrast between the kind of 'enactment view' that Williams describes (and rightly critiques) and the kind of work that many of us try to do when we bring normative political philosophy to bear on real-world issues and problems. Paying closer attention to the role of political philosophy in the real world, and moving away from the kind of schematic model with which Williams begins, allows us to vindicate a substantive and non-minimalist role within politics for philosophical work on normative values and normative principles. We can have a philosophical focus on political values, viewed as a subset of a broader set of moral values, without falling into any kind of pernicious 'moralism'. To put things polemically, Williams's contrast between 'moralism' and 'realism' is a false dichotomy: a genuine realism would not quickly dismiss the role of normative values in political life, and indeed it is a particularly egregious failure of engagement with politics as it is really practised, if we think of the political domain as somehow autonomous from the domain of normative values.

Community wealth building: Values, philosophy and local economic development

This so far no doubt sounds rather abstract, so let me make things more concrete with the aid of an example. One part of my own work in recent years has been articulating and defending an approach to local development known as ‘community wealth building’. This is an approach to local economic policy in which city or regional governments, as well as locally based ‘anchor institutions’ such as universities, colleges, museums, hospitals, charities and so on, seek to use their social and economic power (as employers, purchasers of goods and services, stewards of land and assets etc.) to influence the development of their local economies in particular directions — in particular to make those local economies more sustainable, inclusive and collaborative.

This approach to local economic development has gained an increasing array of adherents in different parts of the world over the past 20 years, having been pioneered by the Washington DC based ‘think-and-do tank’, The Democracy Collaborative (Democracy Collaborative, 2024) which first put these ideas into practice in developing the ‘Cleveland Model’ of economic development in Cleveland, Ohio (Democracy Collaborative, 2014). My own engagement with this work has itself been thoroughly collaborative, working alongside my colleague Joe Guinan, including co-writing our book, *The Case for Community Wealth Building* (Guinan and O’Neill, 2020). Guinan is involved in these policies in a practical way as President of the Democracy Collaborative, with a long history of practical engagement and activism in cause of establishing alternative economic models, and developing a more democratic economy.

A joke that my co-author and I shared when we set out on writing about *The Case for Community Wealth Building* was that, while these policies certainly worked *in practice*, our job was to try to make them work *in theory* as well. We saw what we were doing, in writing on this approach in a way that combined philosophical and empirical analysis, as starting from a set of practices that already existed, rather than starting with a blank sheet of paper and stipulating — on the basis of some prior philosophical theory — the way in which policy should be pursued. Clearly this was an approach that was a long way from what Williams conceived as ‘the enactment model’.

Nevertheless, we took it that the task in which we were engaged had a real philosophical dimension, and specifically that it involved the elaboration and defence of an account of community wealth building (CWB) in terms of an understanding of the values on which it rested, and which it sought to promote. On our analysis, the relevant values are the values of *equality and democracy*. Our view was that, despite the fact that CWB could involve a wide variety of different kinds of policies, pursued in different contexts, what united the different elements of the approach was a concern to move towards an economic settlement that is both *more democratic* and *more egalitarian*. Moreover, on our account, those values are not merely contingently related within this approach, but should instead be seen as mutually supportive and mutually reinforcing. We see the creation of an economic model that devolves more power to individuals, at the local level, as helping to realise more egalitarian social relations, of the kind that both express a conception of citizens' status as democratic citizens, and which in turn creates conducive conditions for a richer democratic culture.

One thing to emphasise here then is that this philosophical account of the foundations of CWB is not plucked from the air, but has been a product of immersion in the real-world practices under discussion. It has been developed after countless interactions and discussions with many individuals, in different parts of the world, who have been involved with CWB as a matter of practical political agency. Moreover, it is not that we looked to offer an interpretation of some separate realm of practice, but rather that discussions about values and principles have always been an important part of the overall practice of developing CWB. Those who are involved in CWB in a day-to-day way do not see philosophical discussion of the values that drive their work as something outside or alien, but simply as part of the constructive reflection on our own aims and activities that is involved with consciously pursuing any set of goals over time, and which is perhaps especially important in the political domain. Philosophical reflection and practical political activity here are closely interrelated: they are not two wholly separable domains; rather, they stand in relations of mutual support and dependence.

Clearly this kind of approach to thinking about philosophy and public political life is not a version of *enactment* on Bernard Williams's model. But neither is it about mere, inert 'normative description' of a set of political activities that would carry on in their own way anyway. Rather, there is a kind of reflexivity between theory and practice. Part of what makes this possible is that the political agents involved are themselves also concerned with political

values, and their own political activities are, in part, driven (in this case) by their affiliation with values of equality and democracy.

There is a certain self-consciously hard-headed 'realist' line that would chide the 'moralist' for a naive fixation on values, and for a purported failure to see that politics is in general about the exercise of *power* and the battle between different *interests*. No doubt such naivety can exist, of course; but there is also an opposite kind of mistake involved in seeing politics as about *nothing but* power and the battle of interests, which can lead to a refusal to understand or appreciate the way in which values and principles are often at the heart of what political agents are motivated to do. There is a further, related error, which falls close to constituting a patronising attitude towards those involved in practical politics, whereby no credit is given to political agents' ability to reflect critically on their own values and principles, or to change course in light of reasoned argument. Our experience in writing in a philosophical idiom about CWB, and about analysing its normative foundations, is that this is an activity that is of great interest to the political agents who are themselves involved in the practice of CWB. Critical reflection on values and principles is something in which they themselves are eager to engage; and this kind of normative engagement is, indeed, inseparable from what it is for anyone to engage seriously with any kind of political project over time.

A clearer understanding of the normative basis of a policy approach such as CWB can also help in its practical defence. A common criticism of CWB, made by publications such as *The Economist*, was that the emphasis on buying local amounted to no more than a collectively irrational form of 'municipal protectionism', where each locality engaged in 'beggar-thy-neighbour' policies that would be collectively self-defeating (The Economist, 2017). But much of the bite of that kind of criticism can be drawn away when one points out that the prioritising of 'local' spend under CWB strategies is justified not in terms of localism per se, but in virtue of the fact that the democratic and egalitarian aims that would be met by helping to develop (for example) new cooperatives or social enterprises can best be pursued with regard to local firms. Conversely, the decision by local councils or anchor institutions to move procurement away from firms with aggressive labour practices, or which engage in devious forms of tax avoidance, will often involve doing less business with certain multinational firms, in favour of working with smaller and more local suppliers. A normatively satisfying justification of the policies in question, in terms of the values at the centre of their underlying rationale, thereby provides the basis for rejecting lines of criticism that might otherwise appear persuasive.

The benefits of engaged philosophy for philosophy itself

I want to end by returning to part of what is of value in Bernard Williams’s critique of moralism, and to say something about the way in which a more situated and engaged approach to political philosophy can help to avoid some of the pitfalls that Williams identifies. On the *enactment model*, we imagine a situation where we see our values and principles as calling for the world to be made some particular way — and we see the role of politics then being to bring that into reality. On such a view, we imagine that our interpretation of our political values involves seeing them calling on us simply to bring about particular kinds of outcomes, whether characterised in terms of the overall level of utility, or the pattern of distribution of resources, or whatever else. My suggestion, though, is that if we start in a more embedded way, engaging with real political practices, and the work of real political agents, we are likely to be led away from this more general and abstract account of what we take our values to demand of us.

Thinking about values of democracy and equality with regard to CWB led us, in *The Case for Community Wealth Building*, to give an account of the values in question here in terms of the way in which citizens relate to each other within economic life, and in terms of their own experiences of both their own status and their relations to their fellow workers and fellow citizens. This gives us not an instrumental account of the value of democracy simply as a means towards certain favoured outcomes, but an account that focuses on the standing and status of democratic citizens, and the value to those citizens of having more power and control over the direction of their lives in the economic and political domains. Similarly, this account involves a way of thinking about equality not in terms of the intrinsic value of some distributive pattern, but in terms of the value of citizens being able to relate as equals. This naturally leads on to considering how institutions can be shaped so that they could develop what we might think of as the ‘infrastructure of social equality’. These more social or relational ways of thinking about our core normative political values have a close and plausible fit with the real-world political projects and practices that we were investigating.

Attending carefully to real-world politics is a good way of avoiding what we could think of as ‘the lure of excessive abstraction’. There can be a tendency in philosophy to be dragged towards excessively abstract or austere interpretations of our values, and to lose some of the sense of reality that we retain when we stick more closely to the phenomena of political and social life. In particular, considering our values through the lens of this more embedded and engaged approach makes it much less likely that we will interpret our normative commitments

as merely striving to bring about some simply described, general overall outcome — whether that is some high-level pattern of distribution, or the maximisation of aggregate utility, or whatever it might be.

In short, objecting to Williams’s ‘enactment model’, at least in its starker versions, is justifiable not only for methodological reasons, but also on substantive and normative grounds. If we care about citizens’ status as democratic citizens, and the character of their social and economic relations, then we need to attend to the functioning of myriad institutions, at different levels and in different contexts, and not just demand that the world should enact some simple overall outcome. We can reach this conclusion, then, by the opposite route to the one taken by Williams: not by downplaying the centrality of normative values in politics, but instead by keeping those values at the centre of our attention, seeing them not as *outside* but as *within* and *essential* to the flow of our political lives. Taking our political values seriously in the full specificity of what they demand of us, and seeing those values as themselves an essential aspect of politics, can lead us to a plausible view that avoids the excesses of the starker varieties of both realism and moralism.

References:

- Democracy Collaborative (2014). *The Cleveland Model: How the Evergreen Cooperatives Build Community Wealth*. [online] Available at: https://www.youtube.com/watch?v=s_kLye_6VBc.
- Democracy Collaborative (2024). The Democracy Collaborative. [online] The Democracy Collaborative. Available at: <https://www.democracycollaborative.org/>.
- Guinan, J. and O’Neill, M. (2020). *The case for community wealth building*. Cambridge, UK: Polity Press.
- Williams, B (2005). *In the Beginning Was the Deed: Realism and Moralism in Political Argument*. Princeton, NJ: Princeton University Press.

To the shoemakers and the ship-builders

John Tasioulas, University of Oxford

What is the contribution that the arts and humanities, including philosophy, can make to public engagement with the increasingly pervasive technology of artificial intelligence (AI)?

Choice

Perhaps the most fundamental contribution of the arts and humanities is to make vivid the fact that the development and deployment of AI technology is not a matter of destiny, but instead involves successive waves of highly consequential human choices. It is important to identify the choices, to frame them in the right way, and to raise the question: who gets to make them and how?

This is important because AI, and digital technology generally, has become the latest focus of the myth that social evolution is determined by independent variables over which we, as individuals or societies, are able to exert little control. A recent book by a leading industry figure employs the metaphor of a dehumanised force of nature to characterize the AI revolution — ‘a coming wave’ obeying a ‘a single, seemingly immutable law’, according to which a new technology ‘gets cheaper and easier to use, and ultimately it proliferates, far and wide.’ (Suleyman, 2023).

The humanities are vital to combatting this techno-determinism, which is profoundly disempowering for individuals and democratic publics alike. They can do so by reminding us of other technological developments that arose the day before yesterday — such as the

harnessing of steam or nuclear power — and how their development and deployment for good or ill were always contingent on human choices, and therefore hostage to systems of value and to power structures that could have been otherwise (Acemoglu and Johnson, 2023).

Ethics

Given the necessity for choice, the second contribution the arts and humanities can make is to emphasise the inescapability of ethics in articulating and making these choices.

Ethics is inescapable because it concerns the ultimate values in which our choices are anchored, whether we realise it or not. These are values that define what it is to have a good life, and what we morally owe to others, including non-human animals and to nature. The arts and humanities in general, and not just philosophy, engage directly with ethics — the ultimate ends of human life. And, in the context of AI, it is vital for them to resist various worrying contractions and distortions that the notion of ethics is apt to undergo. Thanks in part to the incursion of big tech into the AI ethics space, ‘ethics’ is often interpreted in unduly limited ways. For example, as a form of self-regulation lacking legal enforceability. It would be a hugely damaging capitulation to the distortions wrought by big tech to adopt its anaemic understanding of ethics as — at best — essentially self-regulation, or — at worst — corporate PR.

All forms of ‘regulation’ that might be proposed for AI, whether one’s self-regulation in deciding whether to use a social robot to keep one’s aged mother company, or the content of the social and legal norms that should govern the use of such robots, ultimately implicate choices that reflect ethical judgements about salient values and their prioritisation. Reclaiming a broad and foundational understanding of ethics in the AI domain, with radical implications for the re-ordering of social power, is an important task of the arts and humanities.

The dominant approach

The next question we might ask is: what is the shape of the ethical self-understanding that the arts and humanities can help to generate? The starting-point, I think, is to recognise that there is already a dominant approach in this area, that it has grave deficiencies, and that a key task for the humanities is to help us elaborate a more robust ethical outlook. I take the dominant approach to be that which is found most congenial by the powerful scientific, economic and governmental actors in this field.

Like anyone else, AI scientists are prone to the illusion that the intellectual tools at their disposal have a far greater problem-solving purchase than is actually warranted. This is a phenomenon that Plato diagnosed long ago with respect to the technical experts of his day, such as shoemakers and ship-builders. The mindset of scientists working in AI tends to be data-driven, it places great emphasis on optimisation as the core operation of rationality, and it prioritises formal and quantitative techniques.

Given that intellectual framework, it is little wonder that a leading AI scientist like Stuart Russell finds himself drawn to a preference-based utilitarianism as his overarching framework of ethics (Russell, 2019). According to Russell, the morally right thing to do is that which will maximise the fulfilment of human preferences. So, ethics is reduced to a technical exercise in prediction and optimisation — deciding which act or policy is likely to lead to the optimal fulfilment of human preferences.

But this view of ethics is, of course, open to serious challenge. Its concern with aggregating preferences threatens to override important rights that erect strong barriers to what can be done to individuals. And that’s even before we observe that human preferences may themselves be infected with racist, sexist or other prejudices. Ethics operates in the crucial space of reflection on what our preferences should be, a vital consideration that makes a belated appearance in the last few pages of Russell’s book. It does not take those preferences as ultimate determinants of value.

This crude, preference-based utilitarianism also exerts considerable power as an ideology among leading economic and governmental actors. This is less easy to see, because the doctrine has been modified by positing wealth-maximisation as the more readily measurable proxy for preference-satisfaction. Hence the tendency of GDP to hijack

governmental decision-making around economically consequential technologies such as AI, with the consequent sidelining of values that are not readily captured by market demand. Hence, also, the legitimisation of profit-maximisation by corporations as the most effective institutional means to societal wealth-maximisation.

The three Ps: pluralism, processes and participation

The kind of ethics we should hope the arts and humanities steer us towards is one that ameliorates and transcends the limitations of this dominant paradigm. This humanistic ethics would have at least the following three features: the three Ps.

Pluralism

This humanistic ethics would emphasise the irreducible plurality of values, both in terms of the elements of human well-being and the core components of morality. This pluralism calls into question the availability of some optimising function for determining what is all-things-considered the right thing to do. It also undermines the facile assumption that the key to the ethics of AI will be found in one single master-concept, whether that be safety, trustworthiness, human rights or something else. How could human rights be the overarching framework for AI ethics when, for example, AI has a serious environmental impact that cannot be exclusively cashed out in terms of its bearing on anthropocentric concerns? And what about those human values to which we do not think of ourselves as having a right but which are nonetheless important, such as mercy, kindness, or solidarity? Nor can trustworthiness be the master value. It is at best parasitic on compliance with more basic values, hence it cannot displace the need to investigate those values. Equally, safety is an impoverished comprehensive rubric for AI ethics, since the concerns of the latter go well beyond protecting life and limb, even if highly speculative worries about the 'existential risk' to humanity posed by AI systems have tended to draw attention away from this fact.

Acknowledging the existence of a plurality of values, with their nuanced relations and messy conflicts, heightens the need for choice, and accentuates the question of whose decision will prevail. This sensitive exploration of a plurality of values and their interactions is what the arts and humanities, at their best, offer us. I say at their best because, of course, they often fail in this task. My own discipline, philosophy, has itself in recent years often propagated the

highly systematising and formal approach to ethics that I have condemned. Philosophers have much to learn from closer engagement with other humanities disciplines, like classics and history, and with the arts, especially fiction, which often penetrate to the heart of issues like the significance of personal relations, or the nature of human emotion, in ways that the more discursive methods of philosophy cannot. Exemplary here, of course, has been the work of philosophers like Martha Nussbaum, Charles Taylor, and Bernard Williams.

Processes not just outcomes

Of course, we want AI to achieve valuable social goals, such as improving access to education, justice and health care, and to do so in an effective and efficient way. The COVID-19 pandemic cast into sharp relief the question of which outcomes AI is being used to pursue. Did it help us, for example, reduce the need for our fellow citizens to undertake hazardous labour in the delivery of vital services, or is it primarily engaged in profit-generating activities, like vacuuming up people’s attention online and encouraging crass consumerism, that have little redeeming social value? The second feature of a humanistic approach to ethics is to underline that what we rightly care about is not just the value of the outcomes that AI can deliver, but the processes through which it does so.

Compare the use of AI in cancer diagnosis and its use in the sentencing of criminals. When it comes to cancer, what may be all-important is getting the most accurate diagnosis, and it is largely a matter of indifference whether this comes through the use of an AI diagnostic tool or the exercise of human judgement. In criminal sentencing, however, there is a powerful intuition that being sentenced by the robot judge — even if the sentence is likely to be less biased or more consistent than one rendered by a human counterpart — means sacrificing important values, such as accountability and empathy, relating to the process of decision.

This point is familiar, of course, in relation to such process values as transparency, procedural fairness and explainability. But it goes even deeper, because of the dread many understandably feel when contemplating a dehumanised world in which decisions that affect our deepest interests have, at least as their proximate makers, automated systems that can neither be held directly accountable nor empathise with the plight of the humans subject to their decisions (Campbell, 2020; Tasioulas, 2023).

Participation

The third feature relates to the importance of participation, whether as an individual or as part of a group of self-governing democratic citizens, in realising our ethical values. At the level of individual well-being, this takes the focus away from theories that equate human well-being with some end-state, such as pleasure or preference-satisfaction, that could in principle be secured by a process in which the person who enjoys them is entirely passive, for example, by pumping vast quantities of an antidepressant drug in the water supply. Contrary to this view, the exercise of our distinctive human faculties for reason, communication and social connection lies at the core of human wellbeing. This has immense relevance for assessing the significance of the increased deployment of AI systems in place of human endeavour.

One of the most important sites of participation in constructing a good life, in modern societies, is the workplace. According to a McKinsey study, around 30% of all work activities in 60% of occupations are capable of being automated (McKinsey Global Institute, 2017). Should we accept the idea that the large-scale elimination of job opportunities, due to automation, can be compensated for by the benefits that automation brings? The answer depends on whether the participatory self-fulfilment of work can, any time soon, be feasibly replaced by other activities, such as art, friendship, play or religion. If it cannot, addressing the problem with a mechanism like universal basic income, which involves the passive receipt of a benefit, will not be enough (Tasioulas, 2024).

Similarly, we value citizen participation as part of democratic self-government. And, arguably, we do so not just because of the instrumental benefits of democratic decision-making in reaching better decisions (‘the wisdom of crowds’ factor), including better decisions regarding the fair distribution of the benefits of new technologies. The need for democratic control, and the diffusion of power more generally, to ensure the fair distribution of the wealth resultant from new technologies is an important theme in Acemoglu and Johnson’s *Power and Progress*. But we also value participatory decision-making processes because they affirm the status of citizens as free and equal members of the community. This is an essential plank in the defence against the tendency of AI to be co-opted by technocratic modes of decision-making. These modes can erode democratic values by seeking to convert matters of political judgement into questions of purely technical

expertise. In this process the rhetoric of safety, and existential risk, has played a significant role in constraining democratic control over AI by encouraging an expert-based discourse that sidelines the experience and perspectives of ordinary citizens.

At present, much of the culture in which AI is embedded is distinctly technocratic, with decisions regarding the ‘values’ encoded in AI applications being taken by elites within the corporate or bureaucratic sectors, often largely shielded from vigorous democratic control. Indeed, a small group of tech giants accounts for the lion’s share of investment in AI research, dictating its overall direction. Meanwhile, AI-enabled social media poses risks to the quality of public deliberation by promoting the spread of disinformation, aggravating political polarisation, and so on. Similarly, the use of AI as part of corporate and governmental attempts to monitor and manipulate individuals undermines privacy and threatens the exercise of basic liberties, effectively discouraging citizen participation in democratic politics. On the positive side, we need to think seriously about how AI and digital technology more generally can enable, rather than hinder and distort, democratic participation, not just in the sphere of formal law-making, but also in corporate governance and beyond (Landemore, 2024).

Democratising the arts and humanities — The Lyceum Project

If the arts and humanities are to advance a humanistic ethical agenda in AI of the kind I have described, they must themselves be democratised. In a democracy, it is not enough to give people a vote while effectively excluding them from meaningful deliberation; and if they are to deliberate as equals, they must have access to the key sites where basic ideas about justice and the good are worked out.

The arts and humanities are prominent among those sites. Hence the wisdom of Article 27 of the Universal Declaration of Human Rights, which includes a right to participation in science and culture. We can see manifestations of this right, enabled by digital technology, in the resurgent citizen science movement (Vayena and Tasioulas, 2015). But we must also address the exclusion of our fellow citizens from the domains of artistic creativity and humanistic enquiry. This means that work on AI within the arts and humanities should not merely be accessible to a wider public, nor should it merely model civil and rational debate. It should also afford our fellow citizens the opportunity to articulate and critically assess their own views in dialogue with others.

Lest all this sound unduly abstract, I should add something about one of the more notable attempts I have made, as Director of the University of Oxford's Institute for Ethics in AI, to practice what I am preaching.

On 20 June 2024, in partnership with Stanford University, 'Demokritos', Greece's National Centre for Scientific Research, and the World Human Forum, a Greek-based think tank, the institute held the Lyceum Project, a one-day public event: 'AI ethics with Aristotle'. The event took place at the Athens Conservatory, a stone's throw from the original site of Aristotle's school, the Lyceum. Around 500 people attended, from different walks of life, to listen to talks and panel discussions featuring both early-career and established philosophers, entrepreneurs, policymakers, computer scientists, and a speech by the Greek prime minister, Kiriakos Mitsotakis. As part of the event, the ancient historian Josiah Ober and I co-wrote a white paper on how an Aristotelian framework affords the best approach to the ethics of AI, illuminating such matters as work, democracy and global tech regulation (Ober and Tasioulas, 2024). The evening ended with a youth dialogue on the hallowed site of the ancient Lyceum that was followed by an inspiring performance of Sasha Waltz's 'In C' by the Conservatory's graduating class of dancers. These two concluding events served as powerful reminders that the future is in the hands of the younger generation and that humans have capacities for grace, beauty, and emotion far beyond any automated system (ox.ac.uk, 2024).

This event exemplified, for me, how philosophers might devise formats that foster inclusive and rigorous engagement with the largest questions confronting democratic publics. And it was gratifying to see that the Greek prime minister's High-Level Advisory Committee on Artificial Intelligence, on which I was privileged to serve alongside another philosopher, Vasso Kindi, subsequently recommended an annual global AI ethics summit to be held in Athens, one drawing on the rich traditions of philosophical inquiry (Special Secretariat of Foresight, 2024). In a world in which a small number of powerful actors increasingly dominate the public discourse on questions of great human significance, it is important for philosophers to seize such opportunities for different voices to be heard.

This chapter is a shortened and updated version of the following blog post: John Tasioulas, 'The role of the arts and humanities in thinking about artificial intelligence (AI)', <https://www.adalovelaceinstitute.org/blog/role-arts-humanities-thinking-artificial-intelligence-ai/>

Notes:

- 1 Mustafa Suyleman, *The Coming Wave: AI, Power and the 21st Century's Greatest Dilemma* (Bodley Head, 2023), p.6.

References:

- Acemoglu, D. and Johnson, S. (2023). *Power and Progress: Our Thousand-Year Struggle over Technology and Prosperity*. London: Basic Books.
- Campbell, J. (2020). *Causation in Psychology*. Harvard University Press, ch.3.
- Landmore, H. (2024). ‘Can Artificial Intelligence Bring Deliberation to the Masses?’, in R. Chang and A. Srinivasan (eds), *Conversations in Philosophy, Law, and Politics*. Oxford: Oxford University Press.
- McKinsey Global Institute (2017). *A future that works: Automation, employment, and productivity* | McKinsey. [online] [www.mckinsey.com](https://www.mckinsey.com/featured-insights/digital-disruption/harnessing-automation-for-a-future-that-works/de-de). Available at: <https://www.mckinsey.com/featured-insights/digital-disruption/harnessing-automation-for-a-future-that-works/de-de>.
- Ober, J. and Tasioulas, J. (2024). ‘The Lyceum Project: AI Ethics with Aristotle’. *SSRN Electronic Journal*. doi: <https://doi.org/10.2139/ssrn.4879572>.
- Ox.ac.uk. (2024). *The Lyceum Project: AI Ethics with Aristotle* | Ethics in AI. [online] Available at: <https://www.oxford-aiethics.ox.ac.uk/lyceum-project-ai-ethics-aristotle-O>.
- Russell, S. (2019). *Human Compatible: AI and the Problem of Control*. London: Allen Lane.
- Special Secretariat of Foresight. (2024). A Blueprint for Greece's AI Transformation - Special Secretariat of Foresight. [online] Available at: <https://foresight.gov.gr/en/studies/A-Blueprint-for-Greece-s-AI-Transformation/>.
- Suyleman, M. (2023). *The Coming Wave: AI, Power and the 21st Century's Greatest Dilemma*. London: The Bodley Head.
- Tasioulas, J. (2023). ‘The Rule of Algorithm and the Rule of Law’, in C. Bezemek, M. Potacs, A. Somek (eds), *Vienna Lectures on Legal Philosophy vol 3: Legal Reasoning*. Oxford: Hart Publishing, pp.17-39.
- Tasioulas, J. (2024). ‘Work and Play in the Shadow of AI’, in D. Edmonds (ed.), *AI Morality*. Oxford: Oxford University Press.
- Vayena, E. and Tasioulas, J. (2015). ‘We the Scientists’: a Human Right to Citizen Science. *Philosophy & Technology* 28, pp. 479-485.

Section 3: Interventions

Philosophy protects the climate

Professor John Broome, University of Oxford

The influence of philosophy

Philosophy has an immense influence on public life. As an example, think of the influence of liberalism in politics. Liberalism is largely the creation of philosophers, including particularly John Locke and John Stuart Mill, and it is now the guiding principle of many countries' political systems. However, philosophy generally works its influence slowly, as ideas percolate from philosophical writing out into society at large.

It also sometimes happens that important public issues call directly for philosophical analysis. They cannot wait for the slow percolation of thought. For instance, in the 1980s advances in the technology of human reproduction demanded regulation. It was recognised that moral philosophy had an important contribution to make in determining the regulations. The philosopher Mary Warnock oversaw an official inquiry and report which led to the establishment of new UK legislation (Warnock, 1984). Her work was also influential elsewhere in the world.

Another example is climate change. Governments must decide how to respond to it. The need for action is urgent. Despite international efforts that have continued for more than three decades, emissions of greenhouse gas are still increasing. We are on track for warming of about three degrees by the end of this century, which is expected to be catastrophic. This is an emergency. Decisions need urgently to be made about how to drive down emissions. And good decision-making about climate change demands a philosophical contribution. Philosophy can help combat this most serious threat to humanity's future.

What has philosophy to do with climate change? One answer is that climate change itself is immoral in one respect. Some people in the world — mostly in rich countries — cause greenhouse gas to be emitted for their own benefit, and this does harm to other people — mostly in poor countries — who themselves contribute little to climate change. This is an injustice done by the rich to the poor, which is a sort of moral wrong. Wherever morality is at issue, there is a role for moral philosophy. Philosophical questions arise about the injustice of climate change. For example, how far is a country responsible for recompensing those who suffer from the emissions it released in the distant past?

I shall concentrate on a different contribution that philosophy can make. Decision-making about climate change involves judgements of value, and value is within the domain of moral philosophy. These judgements are generally made in practice using the tools of economics, by means of cost–benefit analysis broadly construed. For example, the *Stern Review* (Stern, 2007) — a major UK government report on the economics of climate change — estimated the benefits to be gained in 200 years from limiting climate change, and compared them with the present costs of limiting it. Cost–benefit analysis is a matter of valuation: the value — goodness — of benefits is weighed against the value — badness — of costs.

Philosophy allied with economics

Climate change affects all the world’s population over centuries. It requires methods of analysis that can cope with such numbers and such a long time. Economics possesses those methods, whereas moral philosophy has traditionally been concerned with the actions of individuals and relations between small numbers of people. But moral philosophy does supply underlying principles of valuation. Economics can apply these principles to the enormous problem of climate change.

The branch of economics that is concerned with value is known, oddly, as ‘welfare economics’. It encompasses cost–benefit analysis. Welfare economics is in effect applied ethics: moral philosophy applied to economic matters.

True, some economists believe their discipline to be independent of ethics. This belief was revealed in the reactions of some economists to the *Stern Review*. The *Stern Review* explicitly recognised that welfare economics rests on ethical assumptions. William Nordhaus

and Martin Weitzman — the two most distinguished US economists of climate change — each wrote a review of the *Stern Review*, and each denied that economics rests on ethics (Nordhaus and Weitzman, 2007). These reviewers insisted that values should be taken, not from ethics, but from the preferences of the public as they are revealed in markets.

This view about the source of value is an ethical theory. Some English speakers think that they speak their language without an accent, and only other speakers have an accent. They are so deeply imbued in their own accent that they do not recognise it as an accent at all. Nordhaus and Weitzman are so deeply imbued in their own ethical theory that they do not recognise it as an ethical theory at all. They accept it without thought, and suppose that anyone who takes the trouble to think about ethics is arrogantly trying to impose her own view on other people. Nordhaus wrote that the *Stern Review* ‘takes the lofty vantage point of the world social planner, perhaps stoking the dying embers of the British Empire’. Weitzman accused the *Review* of ‘relying mostly on a priori philosopher-king ethical judgements about the immorality of treating future generations differently from the current generation’.

These authors’ own arrogant position would be more defensible if their ethical theory were a good one. But it is not, at least in the context of their two reviews. This context is a discussion of ‘discounting’, which is the practice of giving less value to future goods than to present goods. Nordhaus and Weitzman recommended deriving the rate at which future goods are discounted from the rate of interest in present financial markets. But the preferences of future generations are not represented in present markets. Even if it were right in general to derive values from people’s preferences, it would still not be right to derive them from the preferences of only some of the affected people — the present generation — and ignore the preferences of the rest.

The idea that economics is independent of ethics must be rejected. It remains prevalent, especially in the US, but many economists side with Nicholas Stern in rejecting it. Philosophers who wish to influence the world’s response to climate change would do well to ally themselves with these economists. Economics has traditionally been influential in policymaking. Governments rarely call for advice from philosophers, but they regularly call on their economic advisers. Moreover, economic theory offers crucial insights that must be recognised by policymakers if climate change is to be overcome. If philosophers can ensure that governments’ economic advisers give good advice, properly based on principles of ethics as well as on economic theory, they will have made a major contribution. An alliance of philosophers and economists can together deliver good advice to governments.

An intellectual alliance already exists. The interface between economics and ethics is a well-populated field. Economists and moral philosophers have many interests in common. They include foundational questions about value: What is the source of value? Is it preferences? How do different values combine? Should different people's well-being combine additively as utilitarians suppose? Should the worse off have priority? And so on. They also include questions about particular values: What is bad about inequality, if anything? How should future goods be weighed against present goods? What value should be attached to changes in the world's population? How bad is a person's death? And so on. Many of these questions are directly relevant to climate policy.

In trying to answer them, philosophers and economists have always interacted. Indeed, they have sometimes been one and the same person; remember Adam Smith, John Stuart Mill and Henry Sidgwick. When the interaction breaks down, things go badly. When philosophers began to make the distinction between egalitarianism and prioritarianism in the 1980s and 1990s, they did not realize that the same distinction had been investigated by economists in the 1960s and 1970s, using different terminology. As a result, philosophers wasted effort in rediscovering what was already known.

The philosophers' ignorance arose partly because many of them could not read economists' writings. They did not understand their pervasive mathematical notation. Interacting with economists is demanding for a philosopher; it requires an investment in understanding their technical language and methods. But in value theory the investment is well worthwhile. Indeed, in the more quantitative areas of value theory it is essential.

The value of human life

Bad economics is a constant threat to good public decision making. The valuing of human life is an excellent example. This is a crucial issue for climate change, since it is widely agreed that the greatest harm climate change will do is to shorten people's lives. The US Environmental Protection Agency (US EPA) has recently produced a new estimate of the 'social cost of carbon', which is supposed to measure the harm done by emissions of greenhouse gas (US EPA, 2023). It estimates 'mortality costs' — the harm of killing people — as more than half of all the harm.

Economists have a way of setting a value on human life that conforms to their general ethical assumption that value should be based on people's preferences. They base the value of life on what people are willing to pay to reduce their risk of dying. Poor people are willing to pay less to reduce their risk than rich people are. Consequently, many economists value the lives of people who live in poor countries much less than the lives of those who live in rich countries. The US EPA does the same.

This is foolish. The amount of money someone is willing to pay to reduce risk to her life depends on two things: first, on the value she attaches to her life, and, second, on the value she attaches to money. Poor people attach a greater value to money than do rich people because they have a greater need for the things money can buy. This is the reason poor people are willing to pay less money for reducing risk to their lives. It is not because their lives have less value to them.

Money is not a good measure of value unless adjustments are made to correct for the differing values money has to different people. These adjustments are called 'distributional weights'. They can be made, but bad economics does not make them. The US EPA's valuation of life is bad economics.

Indeed, it is incompetent. Measuring value by means of people's willingness to pay, unadjusted, is known as the 'Kaldor-Hicks criterion' (Kaldor, 1939; Hicks 1939), and the US EPA explicitly uses this criterion. But it has been known for more than 70 years that the Kaldor-Hicks criterion leads to contradictory implications (de Scitovszky, 1941; Gorman, 1955; Blackorby and Donaldson, 1990; Broome, 2024). It is refuted by *reductio ad absurdum*, that is to say.

It is also immoral. In forming its response to climate change, the US EPA values the lives of Americans 35 times higher than the lives of Bangladeshis. Yet Americans on average emit 11 times as much greenhouse gas as Bangladeshis do. It is a licence for the rich to kill the poor.

Philosophy is not required to identify this fault in the US EPA's economics. All economists should have learned about it during their education. But actually, welfare economics is little taught in economics departments these days. The Kaldor-Hicks criterion, though refuted 70 years ago, is still regularly applied in cost-benefit analysis for policymaking. Contact with philosophy might help economists recognise the logic

of *reductio ad absurdum*. But the real role for philosophy in improving the economics of valuing lives is to think about what, actually, a person loses when she dies.

This has been a topic within philosophy since antiquity. The answer cannot safely be left to the preferences about risk that people happen to have, since those preferences are rarely thought out carefully. There is a group of economists who reject the Kaldor-Hicks criterion, and take the view that what a person loses when she dies is the rest of her life. They value this on the basis of how long and how good the rest of her life would have been. This value is codified by health economists in the form of quality-adjusted life years or qalys. The World Health Organization uses disability-adjusted life years or dalys. Practical measures of value such as these are very amenable to philosophical analysis.

A generalisation of qalys and dalys is to value a person's life by the total of well-being the person enjoys during her life, which is to say the well-being she has at each time, added up over her life. This could be called the total of 'well-being-adjusted life years'. Adding up this value across people implies that the harm done by climate change is the total of well-being it takes away from the world's population. This is a utilitarian measure of value.

Population ethics

Measuring value this way raises a problem that economists of climate change have so far entirely missed. They recognise that mitigating climate change saves lives. They forget it also affects births: it alters the number of people who are born and also the identities of the people who are born. This is because action to mitigate climate change alters people's lives significantly. Consequently, it affects how many children people choose to have, when they choose to have them, and whom they choose to have them with.

An action that mitigates climate change extends the lives of some people. It also causes some people to live who otherwise would not have lived at all. If value is measured by the total of people's well-being, extending lives counts as a benefit equal to the total of well-being it adds to existing lives. Creating new lives counts as a benefit equal to the total of well-being contained in the added lives. Action to mitigate climate change also causes some people not to live who otherwise would have lived, and this counts as a harm equal to the total of well-being those people would have enjoyed had they lived.

This is intuitively odd. Well-being can be added to the world in two ways. One is by extending the life of an existing person; the other by adding a person to the population, who then enjoys well-being. Measuring value by the total of well-being treats either of these ways as equally good. But they are not intuitively equally good. Indeed, many people's intuition is that adding people to the population is not good at all, but ethically neutral, even if their lives will be good. The philosopher Jan Narveson wrote: 'We are in favour of making people happy, but neutral about making happy people' (Narveson, 1973).

This issue has taken us into the realm of population ethics. It is not surprising that the question of valuing lives should force us to consider the value of creating lives as well as the value of extending lives. Since climate change will change the world's population, good decision making about climate change must take population ethics into account. This is a place where philosophy can make a vital contribution. Once more, it is most likely to have influence if it works through economics. Indeed, some of the leading contributors to population ethics are the economists Charles Blackorby, David Donaldson and Walter Bossert.

The IPCC

My own experience of working with economists has been good. It helps that I was a professor of economics before I became a professor of philosophy. I left economics in 1995 but was drawn back to working with economists again a decade later when I returned to work on the ethics, and hence the economics, of climate change. I was a lead author of the 2014 assessment report of the Intergovernmental Panel on Climate Change (IPCC). I believe that my colleague Lukas Meyer and I are the only philosophers who have ever been lead authors for the IPCC.

Each IPCC report consists of three huge volumes — one from each of three working groups — together with several summaries. Each volume has a 'Summary for Policymakers', which is the most widely read and widely influential part of it. Meyer and I were in Working Group 3, which was concerned with how climate change can be mitigated. We joined about a dozen economists in writing a chapter entitled 'Social, Economic and Ethical Concepts and Methods' (IPCC, 2014). I found the economists much more tolerant of philosophy than I expected. Welfare economics has been little taught since the 1970s, but these economists at least recognised the contribution philosophy could make. So did Ottmar Edenhofer, the chair of the working group.

In the working group’s volume, I was given the space to describe aspects of value theory that underlie economic valuations of the harm done by climate change, and of the benefits that can be achieved by policies that limit climate change. I was able to discuss non-human values and cultural values, as well as human well-being, and how human well-being can be aggregated across time and across different people. I mentioned problems of population ethics and considered how lives should be valued. Meyer wrote about issues of justice including the question of how far a country should be held responsible for its emissions from long ago.

The Summary for Policymakers of this volume was written by a small subset of the authors, including me (IPCC, 2014). At the meetings of this group, I was with a few economists among many other social scientists and some lawyers. An affinity emerged between economics and philosophy as the more analytic disciplines among the rest. We wanted to write in clear, sharp sentences, stating our meaning exactly. Other contributors preferred to write long sentences containing many caveats and conditions, whose meaning was often not entirely determinate. I later saw the benefit of this vagueness. When we came to present the report to governments, it gave them less to disagree with and made it easier to cobble together a consensus. Still, it was striking how economics and philosophy converged on the aim of precision.

The Summary for Policymakers is highly compressed. Nonetheless it had enough space to state explicitly that climate change raises issues of justice and ethics. It states that climate policy can be aided by ethical analysis, and can take into account values of different sorts, including non-human values. It also makes it explicit that distributional weights should be applied to monetary measures of benefits and harm, to take account of the differing values money has to different people. These are not radical statements, but their presence is significant because each sentence in the Summary for Policymakers has been individually considered and approved by consensus at a meeting of the IPCC’s members. These members include the governments of every country in the world.

So, for instance, every country, including the US, has agreed that distributional weights should be applied to measures of benefits and harms. Consequently, when the US EPA declined to apply distributional weights in measuring the social cost of carbon, it failed to honour a national commitment. As it happens, the Office of Management and Budget — another arm of the US government — has recently given its permission for US agencies to use distributional weights (Office of Management and Budget, 2023).

Did all this work in cooperation with economists achieve any practical influence for philosophy? The IPCC's 2014 report provided the scientific background for the meeting of the UN Framework Convention on Climate Change that took place in Paris in 2015. That meeting led to the Paris Agreement, which is now the governing document of the international community's fight against climate change. It is the framework regulating each country's contribution. So the IPCC's report as a whole could scarcely have been more influential in international politics. Philosophy's small contribution to it is a foot in the door.

References:

- Blackorby, C. and Donaldson, D. (1990). A Review Article: The Case against the Use of the Sum of Compensating Variations in Cost-Benefit Analysis. *Canadian Journal of Economics*, [online] 23(3), pp.471–494. Available at: <https://ideas.repec.org/a/cje/issued/v23y1990i3p471-94.html>.
- Broome, J. (2024). 'The Value of Life in the Social Cost of Carbon: A Critique and a Proposal'. *Journal of Benefit-Cost Analysis*, pp.1–17. doi: <https://doi.org/10.1017/bca.2024.21>.
- de Scitovszky, T. (1941). 'A Note on Welfare Propositions in Economics'. *The Review of Economic Studies*, 9(1), pp.77–88. doi: <https://doi.org/10.2307/2967640>.
- Gorman, W.M. (1955). 'The Intransitivity of Certain Criteria Used in Welfare Economics'. *Oxford Economic Papers*, [online] 7(1), pp.25–35. doi: <https://doi.org/10.2307/2662000>.
- Hicks, J.R. (1939). The Foundations of Welfare Economics. *The Economic Journal*, [online] 49(196), pp.696–712. doi: <https://doi.org/10.2307/2225023>.
- IPCC (2014). *AR5 Climate Change 2014: Mitigation of Climate Change*. [online] [ipcc.ch](https://www.ipcc.ch/report/ar5/wg3/). Available at: <https://www.ipcc.ch/report/ar5/wg3/>.
- Kaldor, N. (1939). 'Welfare Propositions of Economics and Interpersonal Comparisons of Utility'. *The Economic Journal*, 49(195), p.549–552. doi: <https://doi.org/10.2307/2224835>.
- Narveson, J. (1973). 'Moral Problems of Population'. *The Monist*, [online] 57(1), pp.62–86. Available at: <https://www.jstor.org/stable/27902295>.
- Nordhaus, W.D. and Weitzman, M. (2007). 'A review of the Stern Review on the economics of climate change'. *Journal of Economic Literature*, 45, pp.686–702; 55, pp.703–24.
- Office of Management and Budget (2023). 'Circular A4: regulatory analysis', pp.65–7. Available at: <https://bidenwhitehouse.archives.gov/wp-content/uploads/2023/11/CircularA-4.pdf>.
- Stern, N. (2007). *The Economics of Climate Change*. Cambridge University Press. [online] doi: <https://doi.org/10.1017/cbo9780511817434>.
- US Environmental Protection Agency (2023). *Report on the Social Cost of Greenhouse Gases: Estimates Incorporating Recent Scientific Advances*. Available at: https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NCEE&dirEntryId=356357.
- Warnock, D.M. (1984). *Report of the Committee of Enquiry into Human Fertilisation and Embryology*. London, HMSO.

The pathology of the prison

Dr. Sarah Bufkin, University of Birmingham

In HMP Wandsworth, 80% of the men are locked into overcrowded cells, most for more than 22 hours a day. The wings are rife with broken furniture, insects and rat faeces. Some cells reach over 35 degrees Celsius in the summer. Staff absences mean that many men struggle to regularly access the showers, attend medical appointments, or make it to classes or the library. In May, Chief Inspector of Prisons Charlie Taylor noted that he had encountered in Wandsworth ‘a degree of despondency that I have not come across in my time as Chief Inspector’. Seven inmates have taken their own lives in the past year alone.

Wandsworth is by no means an outlier. Similar problems are rife across the UK’s overcrowded prison system, where Taylor has issued seven ‘urgent notifications’ to highlight substantial failures since November 2022. Education, work and resettlement programmes were cut to the bone during the pandemic and have since been neglected. More than half of men and a staggering four-fifths of women in state custody reported experiencing mental health issues in 2022–2023, and yet mental healthcare provision is profoundly inadequate, with people in crisis spending too long in poor conditions before transfer to hospital. In turn, violence is rife inside, with many prison staff failing to turn on their body cameras when using batons or wielding chemical PAVA spray. Assaults are two-thirds higher in men’s prisons than they were a decade ago, and rates of self-harm among incarcerated women are eight times that of incarcerated men, reaching record highs last year.

From reading media reports over the so-called ‘prison crisis’, it can be all too easy to assume that the UK prison and probation system is caught up in an exceptional moment — a perfect storm of post-pandemic, post-austerity conditions that have deformed what would otherwise be the normal and humane functioning of the carceral state. Renovate the Victorian buildings, crack down on drug use, build new prison beds, and His Majesty’s Prisons will be back to their stated purposes as a tool of criminal justice and rehabilitation.

Within these debates, the failures are presumed to lie at the level of performance, not with the institution of the prison itself or with the practice of incarceration as punishment.

Yet this reading of the prison crisis only provides a narrow window into the brutalities that characterise contemporary carceral state. By carceral state, I refer to the system of institutions, agencies, technologies, laws, and discourses that use repressive state power to manage and control 'dangerous' or 'disorderly' populations. Such a state apparatus tends to encompass both the sites that immobilise people (such as jails, prisons, detention centres, asylums, etc) and the broader systems that decide who is to be surveyed, detained, charged, and locked up (such as the police, prosecutors, the courts, border control forces, and even welfare state agencies). This chapter contends that when we treat prisons like Wandsworth as mere aberrations from broader norms of criminal justice, we miss the systemic role that the prison plays within the carceral state and its broader project of racialised and class-based population control. Such an idealised account of the prison can indict and condemn poor prison policy, but it can't explain or diagnose why we continue to incarcerate greater and greater numbers of people, even as crime rates have fallen. Nor does it have much to say about who is most likely to be intensively surveyed by police, arrested, remanded to prison, and sentenced to custodial sentences — and whose lawbreaking rarely brings them into contact with the carceral state.

This is where critical philosophy takes a different path from what we might call normative political philosophy. Instead of first identifying abstract values or principles and then applying them to various concrete situations, critical theorists embrace a practice of social critique — they work from actual instances of oppression to diagnose how power asymmetries produce these harms and to gesture towards the sorts of emancipatory practices that might overcome them. In this respect, social critique is problem-focused rather than principle-focused; it aims to understand the why and how of oppression, rather than simply provide a normative objection. And critical theorists acknowledge that building a more just future is rarely a straightforward task of simply revealing to people that their practices don't align with their values (or that they ought to be valuing differently). It is not that the world cannot be changed, but that the work of doing so is never a simple project of asking people to become better moral agents or states to become better stewards. Existing power hierarchies place real constraints, both material and ideological, on what we can do and become, even as we might work to move beyond those limits.

The practice of social criticism, of course, is not the sole provenance of philosophers or political theorists. Nor is it an exercise exclusively carried out within the institutional remit of the academy. Within the discipline of philosophy, however, this model of social critique is often associated with the Frankfurt School, which took shape in and around the Institute for Social Research in Frankfurt during the 1920s. Under Max Horkheimer, the institute adopted a problem-oriented, interdisciplinary research agenda that sought to disclose the distortions and dysfunctions that capitalism produced — including the rise of fascism in 1930s Europe and in the Fordist US. As Horkheimer wrote in a famous essay on what distinguished critical from traditional theorizing, 'every part of the theory presupposes the critique of the existing order and the struggle against it along lines determined by the theory itself'.

For contemporary Frankfurt School thinkers like Axel Honneth, what makes this critical approach distinctive is its commitment to the diagnosis of social pathologies. Such a method is not uncontroversial. After all, the language of 'pathology' has all too often been used to blame racially and sexually minoritised communities for their own oppression. But many critical theorists argue that it is worth holding on to social pathology as a concept for two reasons. First, it provides a broader and more robust ethical toolkit through which to understand what can go wrong in a social formation. Instead of only focusing on individual wrongdoing or on structural inequalities, social pathology critique indicts specific forms of life and social practices as unethical, alienating, inhuman or otherwise dysfunctional. It reminds us that there are certain worlds that we shouldn't accept, that corrode or otherwise distort what we are capable of as human animals and planetary co-inhabitants. And second, this mode of critique focuses on the ways in which some societal problems tend to reproduce themselves. A pathological social formation degenerates; its problems become more entrenched and harder to dislodge. (One could think here of the continuing demand for intensifying fossil fuel extraction in the face of a climate catastrophe that is already upon us.) Like many illnesses or malignancies, social pathologies can only be successfully ameliorated by being tackled at their roots.

It is my contention, within this short and polemical essay, that the prison system today is best understood today as a kind of social pathology. Rather than view prisons as the 'legitimate' response to 'diseased' or otherwise 'deviant' individuals, I argue that the modern prison regime itself constitutes a pathological social development that not only fails to solve, but further entrenches broader problems of inequality, racialised and engendered violence, precarious employment, state neglect, and predatory practices

of dispossession. In making this claim, however, I take only indirect inspiration from the Frankfurt School, whose thinkers did little to problematise the prison or the carceral state more broadly. (The notable exception here is Otto Kirchheimer and Georg Rusche’s 1939 book, *Punishment and Social Structure*.)

Instead, I look to US-based abolitionists like Angela Davis, Ruth Wilson Gilmore, and Mariame Kaba, who ask different questions about the police and the prison. Rather than query the idealised ends a prison ought to serve, they investigate what prisons actually do to the people incarcerated in them and to our communities. Much of this abolitionist activism and scholarship has focused on challenging the racialised rise of mass incarceration in the US. But in this essay, I argue that such a conversation can also be productively re-deployed to tackle the expansion of the UK prison system. For all that we must recognise Britain’s disparate historical trajectory, I contend that those committed to sentencing and penal reform in the UK could learn from US abolitionists. Scholar-activists like Davis and Kaba remind us to ask which communities are more likely to end up behind bars and why — and to dare us to imagine a society that doesn’t see the police and the prison as the solution for a wide range of social problems.

Abolitionist social critique first seeks to denaturalise the prison — to show that it is not a ‘normal’ and natural part of a society’s response to crime, but instead a malignant development that produces more harm than good. This requires breaking the presumed link between crime and punishment. In an ideal society, we might hold that a person who breaks the law should be proportionately punished for that crime as a form of restitution. But over the last few decades, both the US and the UK have locked up many more people, for longer and harsher sentences, even as crime rates have fallen. In turn, most people who break the law won’t come into contact with the criminal justice system. It is increased state surveillance that produces convicts, not the act of offending itself. Nor do prisons today function as sites of rehabilitation, notwithstanding government rhetoric. Deep cuts to education, skills and work programmes; the horrifying decline in prison conditions; the decimation of offender resettlement services; and a lack of support for the formerly incarcerated once they are back on the outside — all these shifts reveal that the prison has become primarily a tool of incapacitation, as Ruth Wilson Gilmore puts it. We cage people to punish them and to remove them from society, not to transform or empower them.

Critics of the carceral state also challenge the idea that prisons make us safer. After all, moving people into state custody does little to solve the problems that cause law breaking in the first place, such as poverty, insecure housing, structural unemployment, drug addiction, mental health crises, gendered and sexualised violence, and racist pogroms. In theory, the prison should serve as a deterrent, especially to those who have suffered through it before. And yet recidivism rates remain incredibly high. In fact, evidence suggests that locking people up not only doesn't prevent them from reoffending, but that it increases their chances of breaking the law. Such studies should come as no surprise when we consider what prisons do to people. Incarcerating someone is to separate them from their family and their community, to expose them to routine forms of violence and indignity, to deny them healthcare and addiction treatment, to remove all sense of autonomy and purpose. Being incarcerated often means that people lose their housing and go into debt. When they leave, the state provides little support for re-entry and employers and landlords often discriminate against those with a criminal conviction. In turn, incarcerating too many people can scar the places they come from, cutting children off from their parents and breaking the social bonds of trust and support so vital to healthy communities. Prisons don't control crime. They help to produce it.

This forces us to ask, why have the numbers of the incarcerated, in both the US and the UK, grown so significantly over the past 50 years? Who benefits and who suffers from this aggressive expansion of state power? Just whose safety is the state concerned with — and who does it construe as the danger to be surveyed, policed, contained, and disposed of?

Here genealogical critiques can make a productive intervention. A genealogical critique problematises the status quo by revealing the contingent relations of force that produced a given institution, practice or discourse. Such historical reconstructions reveal that societies could have been organised otherwise and gesture towards the social functions that a given institution or practice serves. For example, scholars have shown how the development of the police and the prison system in the US has been intimately linked with forms of racialised, gendered, and class-based population management. Angela Davis has written extensively about how the afterlives of racial slavery and empire have shaped the modern carceral state, from the chain gang after abolition through to the mass incarceration crisis of the past 50 years. Abolitionists have re-conceptualised the late 20th century US carceral state as a 'prison-industrial complex', drawing on Dwight Eisenhower's warning about the growth of a 'military-industrial complex' in the wake of the

Vietnam War. In doing so, they draw strategic links between militarism, punishment and capitalism, arguing that a 'punishment industry' has taken root that profits from expanding and supplying the state's expansive custodial estate.

Perhaps no one has offered as clear a diagnosis of this punitive (and profoundly racialised) drive to lock people up in the late 20th century US than Ruth Wilson Gilmore, whose *Golden Gulag* explores why liberal California built upwards of 23 new major prisons, in addition to a range of other custodial institutions, between 1984 and 2005. Contrary to readings that attribute the mass incarceration crisis to the demand for cheap prison labour or the profit incentives of the private prison industry, Gilmore shows how it was the state, with the support of finance capital, large landowners, and de-industrialising rural areas, that drove this prison-building project. Ultimately, Gilmore argues that the growth of the US prison regime has been inseparable from the broader neoliberal project of 'organised abandonment' that stripped the welfare state for parts and left racialised, poor communities to fend for themselves as capital moved production abroad. For thinkers like Davis and Gilmore, then, it is not that today's prisons are failing to operate as they should and are therefore in need of reform, but that they are functioning as they are designed to — namely, as racialised, class-based, and gendered technologies of containment. They have become the primary means by which the state deals with racialised and class-based 'surplus populations' whose labour is both not needed and may never be needed under conditions of structural unemployment.

This critique of the US prison system is well trodden terrain. Less remarked upon is the ballooning numbers of people living behind bars in the UK. But I contend that the British prison regime can also be understood as a pathological state formation. It too functions as a form of surplus state power that has taken on a life of its own, sweeping up more targets and entrenching itself deeper within the institutional fabric regardless of whether it is necessary or useful. Today, the UK locks up a higher proportion of its people than any country in Western Europe. About 6,700 people in Britain are serving life sentences; by comparison, the remaining 51 countries in the Council of Europe (excluding Turkey) have a total of 9,002 people on life sentences combined. And the British state's carceral footprint has more than doubled in the past four decades, as the number of people in state custody went from around 40,000 in 1991 up to more than 88,500 people in early September of this year. The Ministry of Justice predicts that Britain could have as many as 114,800 people behind bars by March 2028, if current trends continue. Significant drivers of these skyrocketing

incarceration rates have been ill-thought-out and overly punitive sentencing laws, including the much-maligned Imprisonment for Public Protection sentences and mandatory minimums for certain crimes. Sentences have got longer, more people are spending time in prison for petty crimes, and the state has criminalised more and more activity, including forms of 'disruptive' protest.

This explosive growth in the prison regime, however, cannot be adequately diagnosed without reckoning with the long *durée* of Britain's imperial past or the inequities of its postcolonial present. In the wake of white backlash to postcolonial migration from the Caribbean, South Asia, and Africa in the 1960s and 70s, British authorities adopted many of these repressive techniques of population control it had developed in colonial counterinsurgency campaigns in Malaysia, Kenya, and Cyprus to discipline growing racially minoritised communities at home.

As Stuart Hall and his colleagues presciently showed in *Policing the Crisis*, the state and the media fed a racialised moral panic over young Black men and 'mugging' in the 1970s and, in doing so, built a popular consensus around the need for harsh 'law-and-order' measures, including long custodial sentences. Such a crisis was 'lived' and explained through race, Hall and his Centre for Contemporary Cultural Studies colleagues argued, but it had its roots in the broader economic crises of postwar British capitalism and the fast pace of social and cultural change that had left many British people feeling displaced and insecure.

Over the interceding decades, a revolving cast of racialised scares has driven the public appetite for tougher sentencing laws, from concerns over 'gang' activity and knife crime to Islamophobic worries over the 'Muslim terrorist' to the recent preoccupation with racially minoritised migrants and asylum seekers. As five senior former judges pointed out in a September briefing paper denouncing Britain's 'sentencing inflation', many of today's punitive sentencing requirements can be linked back to 'single-issue campaigns' that were mobilised by 'emotive media attention' and called for mandatory minimums and harsher state responses, thereby removing judicial discretion.

Structural racism, overpolicing, educational inequities, substandard housing, changing migration laws, and increased susceptibility to precarious under- and unemployment, in turn, mean that racially minoritised communities disproportionately suffer the repressive heel of the British state. Racially minoritised people, especially Black men, are much more

likely to be arrested, and they are more likely to be sent to prison than white people for the same offences. It should come as no surprise, then, that the racially minoritised are overrepresented in British prisons (27% of those incarcerated versus 18% of the general population). There are almost three times more Muslims in prison today than there were two decades ago. As the senior justices warned, 'Without urgent remedial action to address sentence inflation, this country will soon experience US-style mass incarceration.' The new Labour government, however, seems more invested in 'fixes' to the prison crisis that expand the footprint of the carceral state, rather than in radical solutions that would reduce the number of people we lock up and the communities we abandon.

Perhaps most importantly, abolitionists like Davis advance a fierce ethical indictment of incarceration as a deeply inhuman and inhumane practice. This is where critical theory diverges from applied analytical jurisprudence. Within analytical jurisprudence, philosophers of law debate how to best to understand legal concepts and theories of punishment, to ascertain how our practices might be better brought into alignment with those theories. Such a focused analytic exercise, of course, is important and can bring real insight into what we should (and should not) expect from the law. But critical theorists want to offer a more radical critique of punishment in modern societies than what is available within jurisprudence alone. On one hand, they want to diagnostically situate the rise of repressive state power within a broader picture of neoliberal capitalism and the making-surplus of whole communities. And on the other, critical theory mobilises a more robust ethical critique, one in which existing practices of penal 'justice' are indexed against a theory of human flourishing and found wanting. Whereas sociologists or criminologists might offer data on the (in)effectiveness of specific carceral policies, practitioners of social critique are unafraid to indict a broader form of life that sees the prison as a normal and natural response to socially produced insecurity and deprivation.

Social critique therefore raises the question of whether we want to live in a society that locks people in cages for years at a time, sometimes for decades, and calls it justice. That abandons people to overcrowded cells where they are trapped for 22 hours a day without purpose or hope. That exposes them to routine violence, medical neglect, unsanitary conditions, and the daily humiliations of prison life. That sanctions routine sexual coercion and the intensive state disciplining of gendered bodies. And social criticism calls many of us out for supporting policies that have abandoned entire communities to violence, poverty and neglect while pathologising them as 'deviant' or 'dangerous' and therefore making them

disposable. In doing so, critical philosophy of this kind dares us to envision a world without prisons, without militarised policing and immigration detention centres.

Such a political horizon might seem utopian, especially in a social formation in which global capitalism sets the terms of possibility. But abolitionists aren't as worried by the charge of utopianism. After all, while everyday political change tends to take place through incremental reforms, mobilising people often requires a radical and emancipatory vision — one committed to the idea that other worlds are possible. And there are concrete decarceral strategies that can move us further away from these repressive and dysfunctional responses to inequality, poverty, drug use, migration and crime without opening the doors on all the prisons tomorrow. As Davis points out, abolition isn't just about the negative work of dismantling repressive institutions and state practices; it is also about building new institutions that can make the prison obsolete. This vision of 'abolition democracy' involves making a radical investment of state resources into communities — from healthcare to education to housing to childcare to public spaces — and doing so on terms that give communities autonomy and input over that support. Social critique certainly doesn't have all the answers about how we get to this decarceral future. But it does the important work of denaturalising what Pierre Bourdieu called the 'inert violence in the order of things'. Abolitionists force us to ask what kind of human and humane world we might want to live in together, if we just had the conviction and the courage to work for it.

Call the midwife

Professor Roger Crisp, The Uehiro Oxford Institute

Professor Katrien Devolder, The Uehiro Oxford Institute

Dr Jonathan Pugh, The Uehiro Oxford Institute

Who was the first philosopher? Most people would say Socrates, and there is no doubt that he, and the character named after him in the dialogues of Plato, have played a huge role in the development of philosophy over the last two-and-a-half millennia. Socrates, then, in a sense gave birth to philosophy. But how could that be, given his frequent claims to knowing nothing except his own ignorance? Partly because those claims can easily be read as ironic, but also because his dialectical way of doing philosophy enabled others to put forward and develop their own ideas. Indeed, one of Socrates's most famous images is that of himself as a midwife. In this paper, we shall suggest that modern philosophers might be seen, and see themselves, in the same way, using the resources of philosophy to help others give birth to their own ideas.

At least as he appears in the evidence available to us, Socrates was a truly public philosopher. He believed he had a divine mission to interrogate everyone he encountered on fundamental moral issues, and through exposing their ignorance put them in a position to live more virtuous lives. His methods contrast with those of modern academic philosophers. He sought to discuss philosophy with all, not only his students or fellow philosophers. He did not see philosophy as a profession and indeed would take no payment from his interlocutors. Nor did he publish, believing that the only way to make progress in philosophy is to discuss it with real people. He argued that the written word cannot defend itself or be questioned and cannot teach the truth effectively.

Academic philosophy as we know it, then, might be said to have begun with the work of Socrates's most famous discussant, Plato, who founded a famous school for philosophy in Athens (the Academy), wrote down his views, and disseminated them. And when we consider the history of philosophy from the time of Socrates and Plato until now, it is hard to think that Plato's decision to record his thoughts was a mistake.

But someone might suggest academic debate in philosophy is pointless since there has been so little agreement, even on fundamental issues such as truth, the criteria of personhood, or ethical principles. Here it is worth remembering that philosophy is, probably, still at an early stage of its development: we often forget that we may well be living at the very beginning of an extraordinarily long intellectual history. Further, philosophy, including philosophical ethics, has made a good deal of progress. By the end of the 20th century, for example, the differences between various consequentialist and deontological positions were significantly clearer than they had been at the end of the previous century. And, through successful public engagement on a large scale, several philosophers have in recent years greatly advanced the understanding of ethics among the general public. These include, for example, Kwame Anthony Appiah in his *New York Times* column ‘The Ethicist’; Mary Warnock in her work on bioethical issues; Michael Sandel, in many television programmes and online courses; and Toby Ord, in his work on existential threats.

In addition, philosophers have, partly through their publishing, contributed to important ethical advances. For example: the abolition of slavery — Adam Smith; equality for women — Mary Wollstonecraft and J.S. Mill; education reform — John Dewey; the decriminalisation of homosexuality — H.L.A. Hart; the global distribution of healthcare resources — Larry Temkin; attitudes to global poverty and the treatment of non-human animals — Peter Singer; and human rights and economic development — Martha Nussbaum and Amartya Sen.

Philosophy, then, can change things, and work in practical ethics can sometimes have quite revisionary implications for the world of policymaking. This should not surprise us. Professional moral philosophy champions the rigorous scrutiny of even the most widely and deeply held beliefs and values, and some of its best work can substantially change our understanding of these beliefs and values, or what they might entail.

This willingness to challenge existing precepts and frameworks is in many ways a virtue of the field, and this sort of work in philosophy can of course sometimes have real world effects on policymaking. Yet this willingness can also sometimes be a limitation when it comes to thinking about the potential of philosophy to impact real-world policy. It is one thing to convince an academic colleague of a philosophical argument to change the world and change their current understanding of a practical problem; it is quite another thing to argue that it would be politically feasible or expedient for those in the policymaking sphere to follow suit.

At any rate, philosophy *can* assist with progress. Further, this is something we should expect (both expect to happen and *expect* from philosophers). But there is no doubt that a lot more could be done. One problem is that the explosion in professional applied or practical ethics happened during the 1970s, when publication had already become the primary mode of assessment of philosophical reputation, and hence the driver for academic activity. In the UK, for example, there have been eight major national research assessments since 1986, and only in 2014 was 'impact' considered (and then only in the form of somewhat artificial 'impact cases', rather than the direct engagement of individual academics with the public through publication, the media, committee work and so on).

Philosophers themselves have tended not to object to the focus on academic output, primarily because they like doing philosophy, and having their ideas taken seriously by their peers in the discipline. The result is that it would be difficult to significantly lengthen the list of 'engaged' philosophers given above. But philosophers should not be blamed for this, nor should members of the public for their ignorance of what philosophy is and how it might play a much larger part in public life. Both sides are the products of their times. If a finger is to be pointed anywhere, it might perhaps be towards Kenneth Baker, the education secretary in the Thatcher government, for not including philosophy in the 1988 national curriculum. Had the public become more aware of philosophy, there would have been more opportunities to benefit from interaction with professional philosophers, an interaction which would almost certainly have benefited professional philosophy as well as individual philosophers.

Let us operate on the assumption that clarity, understanding of arguments and their logic, and perspective are, in general, valuable in public life. Philosophers, then, because of the peculiar focus on argument in their training, have a huge amount to contribute, but they are not called upon as often as they might be. The current situation is in some ways analogous to a programme in which medical researchers are hired to make progress on treatments for various illnesses, and merely to publish their papers in journals, which are known to be read mainly by other medics. (Medics, of course, do get kudos from the success of their research, though the drivers behind their success often arise from the interest of drug companies. Philosophy is equally in the public interest, but not marketised.)

All, or at least most, philosophers, then, should be required by their employers to take part in public life. (This would of course require a national strategy, imposed by the Department of

Education.) This includes non-ethicists: they have much to contribute merely through their philosophical expertise. Further, for ethics in particular, 'ethical consultancy' programmes should be instituted, analogous in some ways to existing business ethics consultancies. Though they should, wherever possible, be non-profit, and their focus should not be restricted to the culture of the organisation that uses their services. Their function should not be seen as in any way moralistic or protreptic. The aim would be to help others to think through the issues they are facing with a more self-conscious focus on clarity and rigour.

It might be objected that these changes will slow down progress in academic philosophy. But, one might respond, the current situation may be inefficient, and fewer papers would not necessarily lead to slower progress in philosophy. And, through greater interaction, the philosophy that is published, especially on practical issues, would be more informed and helpful. Philosophy does move forward slowly, that is, carefully, and maybe things are better that way. And, even if the objection is correct, the advantages to the public of greater engagement by philosophers outweigh the drawbacks. Nor need we expect every philosopher (even the very best) to be required to spend the same amount of time on public engagement. Further, public engagement should be understood broadly. Some philosophers are better at, or more inclined to, working with the media, while others may prefer to serve on committees, publish accessible papers, or to work with organisations such as the Public Philosophy Network.

Earlier, we mentioned the potential for philosophy to lead to radical or revisionary conclusions. But it needn't do that, and in fact moral philosophy can also play a quite different, and arguably more powerful role in this sphere. Rather than using the tools of moral philosophy just to develop theoretical arguments supporting change (as do Appiah and the 'public intellectuals' listed above), the philosopher in the public sphere can also seek to use these tools to enhance understanding of the practical options that are on the policymaking table.

This is where philosophical obstetrics comes in. Let us return to ethical consultancy. The University of Oxford and some other leading universities have focused on 'research-led' teaching, that is, the teaching of undergraduates, in particular, who then go on to make a great contribution to their country (in most cases — politics may be an exception ...). Ethical consultancy is analogous to that, but it extends the net to include the public at large. As we noted, this already happens in certain businesses, and when it is done properly, rather than as a form of 'ethics-washing', it can improve corporate governance as well as the lives of

employees and other stakeholders. Here at the Uehiro Oxford Institute, we and many of our colleagues have been involved in work with government, both national and local, the NHS, schools, the police and prison service, as well as a wide range of other institutions through our Master of Studies (MSt) in Practical Ethics (University of Oxford, 2025), Bitesize Ethics (University of Oxford, 2024) series, and other programmes.

The consultancy role is perhaps of most use in areas of public debate in which there are a variety of conflicting moral considerations supporting mutually exclusive policy choices. This is particularly so when the public debate on the topic has become politicised in a manner that precludes different sides of the issue from acknowledging any weakness in their own position, or any strength in their opponent's. In such circumstances, progress can be extremely difficult, and the stalemate in the public and political debate may even lead to a kind of policy paralysis.

This is one important area in which the skills of the philosopher — not just the moral philosopher, but also the logician, the metaphysician, the epistemologist, if they have the right communication skills — can help to enable progress. By clarifying concepts, analysing the structure of the arguments that are employed to support different policy options, and pointing out where their different strengths and weaknesses may be, the public philosopher can help the policymaker to chart and navigate the different value trade-offs that they will unavoidably have to make in developing coherent policy on highly contested issues.

The benefit of philosophical input here is not that the philosopher can ‘solve’ the moral problem facing policymakers; unfortunately such neat solutions are rarely available in either politics or philosophy. Instead, it is to help ensure that the policy response to an issue with moral dimensions is well reasoned and systematically considers a wide range of morally relevant factors. Ultimately, then, philosophical input can help to ensure that policy provides moral justification to those affected by the policy. After all, to retain their legitimacy, it is crucial that policymakers can offer this group a reasonable answer to the question, ‘Why did you decide that this was the best course of action?’

It might be objected that this is all well and good, but that philosophy as an academic discipline will suffer because its practitioners are spending time on public engagement rather than academic research. This objection cannot be dismissed out of hand. But it is worth noting, first, that a great deal of philosophical research is funded by the public, who might not unreasonably expect philosophers to be prepared (indeed required),

like real midwives trained at public expense, to directly contribute to the public good. Further, to some extent, the value of research diminishes marginally, and it may be also that philosophers’ ability to think and to express their own views will be advanced through communication with non-specialists, with a focus on accessibility, avoiding jargon, and so on.

One example of putting this kind of vision of philosophy’s role in public life into practice is the Uehiro Oxford Institute’s Decision Aid for the Restitution of Cultural Artefacts (DARCA). The question of whether cultural institutions (like museums) ought to return certain items within their collections has been subject to increasing debate in recent years. There are two broad competing moral obligations that largely shape the debate in this sphere. On one hand, it is sometimes claimed that cultural institutions have a moral obligation to the publics they serve to preserve and showcase valuable cultural artefacts, because these items manifest important parts of humankind’s shared cultural heritage. Conversely, however, these same institutions are sometimes understood to be under a moral obligation to return certain artefacts because another party has a plausible moral claim to that item, perhaps because its return would constitute an appropriate form of reparation for a historical injustice involved in the object’s transfer of ownership. In view of this conflict, DARCA was developed to enable those facing the practical policy challenges of cultural restitution to adopt coherent and justifiable policy responses on a case-by-case basis.

DARCA was developed in close partnership with an interdisciplinary group of stakeholders. Across two workshops, and following a review of existing sector guidance, law, policy, and academic moral philosophy, the DARCA team identified a range of moral considerations that can be understood to affect the relative strength of these two moral obligations. Following iterated consultation with experts in the field, the team aimed to synthesise this body of work into an approachable and accessible tool that can enable individuals to think systematically through various considerations relevant to ethical questions about cultural restitution. To do so, the team developed a series of questions that can be used to assess how these factors might arise and interact in a particular case of cultural restitution. These questions form the theoretical basis of DARCA, and users answering these questions are invited to reflect on some of the philosophical complexities of each of the questions raised with carefully curated guidance notes. Upon completing DARCA, the user is provided with a summary of their answers, any further written justification they may have given for those answers, an overall assessment of the comparative strength of the argument for restitution, and an explanation of that assessment.

DARCA cannot solve the moral issue of cultural restitution, and that is not its intended goal. Indeed, it is intended to guide and supplement the user’s own moral reflection on this complex area of policy, rather than to replace it. However, after using DARCA, the user should have a better sense of where the moral arguments appear to lead in their case, why they do so, and what moral factors have made a difference. The role of philosophers in developing DARCA, then, has not been to provide policymakers with the ‘right’ answer; it has instead been to help them reach their own answer in a manner that is systematic, well-informed, and philosophically grounded. The role of the engaged practical philosopher, then, is primarily not to persuade, but to provide assistance to others in reaching their own, reasoned conclusions.

References:

- University of Oxford (2024). *Bitesize ethics 2024*. [online] Available at: <https://www.practicaethics.ox.ac.uk/event/bitesize-ethics-summer-programme> [Accessed 21 Feb. 2025].
- University of Oxford (2025). *MSt in Practical Ethics*. [online] Available at: <https://www.practicaethics.ox.ac.uk/mst-in-practical-ethics> [Accessed 21 Feb. 2025].

Ours to question why

Professor Cécile, Fabre University of Oxford

As I write this in November 2024, the war between Ukraine and Russia appears to have no end in sight, and neither does Israel's conflict with Hamas in Gaza. The UK is supplying military assistance to Ukraine and Israel. While the UK is not supplying 'boots on the grounds', it is contributing to the deaths of thousands of people. In living memory, its forces have directly killed individuals who had been declared the country's enemies, by order of the elected government of the day.

Of all the things that we do, killing another person stands in the greatest need of moral and legal justification: in the absence of such justification, it can constitute manslaughter or murder. A decision to resort to military force, regardless of who makes it and of its rationale, is in effect a decision to authorise soldiers — on our behalf — to kill enemy combatants and civilians, whether through deliberate acts or collateral consequences of military action. Thus, when parliament votes in favour of such action, it authorises soldiers to kill, not merely on our behalf but also, in some important sense, at our behest. When the government agrees to supply weapons to its allies, it provides the means for the act of killing. If we are to take the presumption against killing seriously, we — citizens, soldiers, elected representatives and government ministers — must also take seriously the task of morally justifying war.

Some view this task as hopelessly naive, futile, or both. It is naive, because states go to war if and when it serves their interests, and whatever philosophers have to say about the right and wrong of doing so has little purchase. It is futile, because what matters the most, surely, is not whether states are morally allowed to go to war and whether soldiers, so authorised, are morally allowed to kill enemy combatants and civilians: what matters is whether those acts are lawful.

And yet, a striking feature of recent parliamentary debates on military action is the highly moralised language employed by members of parliament (MPs) for and against war. In March

2003, the House of Commons voted in favour of the government's motion to wage war against Iraq, as part of a US-led coalition. Thus, arguing against the deployment of forces, Douglas Hogg MP, stated: ‘I do not think that any of the usual characteristics of a just war have been satisfied. If we were dealing with a situation in which Iraq had attacked another country or had mustered troops on the frontiers of another country, or if there were compelling evidence that Iraq was delivering to terrorists weapons of mass destruction with which they could attack another country, I would vote for war, but none of those circumstances exists.’

Or consider Peter Bradley's impassioned speech in support of intervention: ‘I believe in just wars. I believe that they are commissioned in defence of freedom, and against oppression. I also believe that, for them to be just wars, they must be the last resort. Diplomacy must come first, but if we are to prevail in defence of what we believe to be right, there must also be a limit to diplomacy. [...] If we are to set aside our prejudices and accept that doing nothing is not an option, if we accept that diplomacy is at an end and that Saddam continues to defy and threaten us, what is the alternative to war?’ There is none, Bradley concluded, siding with the government, though noting the difficulties of establishing, there and then in 2003, whether the conditions for a just war were met.

Ten years later, in August 2013, the House of Commons voted against the deployment of a British intervention force in Syria. MPs framed the decision whether to intervene in Syria in the language of rights and duties. When debating whether intervention in either Iraq or Syria was lawful, there seemed to have been little doubt in their mind about the moral legitimacy of the international legal framework within which Britain had to operate. As a constitutional matter, the decision to go to war is a royal prerogative: it does not require the approval of parliament, and any such approval (or disapproval for that matter) is merely advisory. Thus, in 2018, the then-prime minister Theresa May authorised military strikes against President Assad's regime (which had used chemical weapons against rebel forces and civilians) without seeking the consent of parliament. Nevertheless, as Hansard records, she addressed the House of Commons at length in similar terms to her two predecessors: ‘it was not just morally right but legally right to take military action [...] to alleviate further human suffering’.

Turning to the present, public debates, both inside and outside parliament, on the wars between Russia and Ukraine, and between Israel and Hamas, are suffused with moral language: the claim that Ukraine is morally justified in defending itself from Russia's unwarranted aggression is seen as beyond dispute; likewise, the claim that Hamas committed atrocities

against Israeli civilians on 7 October 2023 is widely endorsed (including by proponents of an independent Palestine); so is the view, at least in this country, that the humanitarian costs of Israel's retaliatory war in Palestine since then are disproportionate.

Both within parliament and, indeed, in the press, there is a remarkable degree of consensus on the conditions which a war must meet to be just and lawful. It must have a just cause — such as national defence against a military aggression, stopping a humanitarian disaster, or upholding the moral and legal prohibition on the use of weapons of mass destruction — and reliable evidence that there is such a cause. It must aim to bring about those just ends. It must be a proportionate and necessary means of achieving those ends, including in the way it is fought. More specifically, minimising harm to civilians is of paramount importance. Disagreements focus on whether those conditions apply in the specific cases under consideration.

Without explicitly acknowledging it, our MPs and prime ministers are engaging in philosophical arguments: they do what philosophers and theologians have done for millennia in both western and non-western traditions. Anecdotally, from my own experience as a philosopher of war, so do many soldiers. The question, then, is not so much whether foreign policy actors should engage with the philosophy of war, but rather whether they do so as well as they could. No less crucially, it is also what philosophers of war can learn from those public debates: if philosophy has a role to play in public life, so does public life in philosophy.

These conditions of a morally just war are familiar, intuitively powerful, and, at first sight, provide secure foundations for the laws of war. On closer inspection, and as philosophers of war have noted, they are much less straightforward than they appear. Let me highlight two serious difficulties. First, they raise thorny interpretative problems which, once brought into view, weaken prospects for consensus. Consider for example the claim that unwarranted military aggression is a just cause for war. In most cases, and paradigmatically so for Ukraine in February 2022 or Poland in September 1939, this may well be true. Suppose however that the aggressor-state invades the defender-state without shedding blood — as Russia did when it annexed Crimea in February 2014. Or that it seizes a remote and uninhabited part of the defender-state's territory. The aggression takes place without violating anyone's right not to be killed. The question, then, is whether and when violations of the comparatively less important right to territorial sovereignty warrant the use of lethal force against enemy combatants who have played a minimal role in those violations, as well as the collateral deaths of civilians.

Consider next the claim that war must be a proportionate means to achieving justified ends; that is, the bads it does must not be excessive in relation to the goods it brings about. To ascertain whether the war would be proportionate, thus, we must decide which goods, relative to which evils, are relevant. For example, suppose we go to war against an authoritarian state which has invaded us. There may be an opportunity to remove the authoritarian state's leaders from power and help its population install a democratic regime. It may seem that these are goods which ought to count against the bads our forces will inflict. Yet, it is doubtful that regime change is itself a just cause for war. If it is not justified to wage war to bring about this particular end, why should the outcome of a successful regime change be a factor when considering proportionality? And, if we do go to war against our unjust aggressor, the result could be that the aggressor's increased military spending leaves its civilians entirely destitute. Should those bads count against our decision to go to war?

Assume, next, that we can determine the relevant goods and bads. As the war goes on, policymakers must decide whether to carry on with the war or whether to sue for peace and withdraw their troops. This includes deciding whether the costs incurred so far count in the proportionality calculus, or whether they can be discounted as sunk costs. Suppose our elected representatives authorise the government to defend our homeland and anticipate a proportionate and thus acceptable cost of, say, 15,000 combatant fatalities. However, suppose the war goes badly: six months into the conflict, there are 10,000 combatant fatalities, for much smaller territorial gains than anticipated. Our military leaders estimate that a final push would deliver victory, but at a cost of 8,000 lives. Assume that they are correct: victory, then, in the end, would cost 18,000 lives.

There are two ways of thinking about the 10,000 lives already lost. On the one hand, they can be thought of as 'sunk costs'. Given that they can never be recovered, our leaders need not take them into account as they consider what to do next: rather, they must focus on the additional lives that will be lost if they proceed. As 8,000 is well below the 15,000 threshold, continuing with the war is not disproportionate. On the other hand, we can deem those 10,000 lives not to be 'sunk' and take the view that those costs are to be carried forward. If so, the war must stop, as it would fail to meet the proportionality condition. What should our leaders do?

What about the requirement to minimise harm to civilians? On the face of it, it is clear civilians should be protected from harm wherever possible. Matters are not so simple, however. For the problem is that some civilians contribute to unwarranted wars of aggression — for example,

those who work in the arms industry. There are others: civilian members of intelligence services; journalists and so-called 'influencers' who feed the war propaganda machine; and also civilian leaders, including (why not, after all?) civilian elected representatives who, in regimes which are or claim to be democratic, vote to authorise and facilitate the war.

Perhaps, then, the requirement to minimise harm to civilians only applies to innocent civilians — who do not participate in those ways in the war, let alone take up arms. They clearly are not legitimate targets; and whatever harm is done to them, as a collateral effect of our war, must be kept to a minimum. So far, so simple. Suppose, however, that complying with the requirement would entail a serious loss of life on our side. To whom must our leaders give priority, and on what basis? Answers to this most-difficult question are likely to depend on the kind of war we fight: whether it is a war of self-defence, or a war of intervention — and on whether the fatalities are soldiers (and whether these are conscripts or fight voluntarily) or fellow citizens.

There are no easy answers to the questions I have raised. The main lesson to draw from my remarks, if there is one, is that, in the light of a well-developed body of philosophical work on the morality of war, the appearance of consensus is likely to be deceptive — and indeed deceptively dangerous. The costs of being mistaken about, for example, what is a just cause and what is proportionate are measured in lives, not just in financial costs, votes, or reputation.

The second difficulty is that the morality of war is not as clearly aligned to the laws of war as is often supposed by many outside philosophy (and also, by quite a few within the just war tradition). The laws of war sharply distinguish between the legal status of a decision to resort to war, and the legal status of acts of killings within the war. Thus, take for granted that Russia unlawfully invaded Ukraine on 24 February 2022, and that Ukraine had a lawful cause for responding by force. Yet, although Russia's war of aggression is unlawful, the individual acts of killing carried out by Russian soldiers against Ukrainian soldiers in pursuit of their regime's unlawful ends are not, in themselves, unlawful. Indeed, they are not any less lawful than the acts of killing carried out by Ukrainian soldiers against Russian soldiers in defence of their country.

Morally speaking, however, this cannot be right. For if it is morally wrong to wage a war of aggression, then by implication it is morally wrong to kill in pursuit of that war. To be sure, it may well be that soldiers on the 'unjust' side have no choice but to fight — perhaps they are conscripted under life-threatening duress, or perhaps they are unavoidably unaware of their regime's true ends. It is worth noting, however, that those reasons would not be

considered as exculpatory of the deliberate killing of innocent civilians. Yet, even if, for example, Russian soldiers are not blameworthy for killing Ukrainian soldiers, it remains the case that they are not morally permitted to do so, whereas (by implication) Ukrainian soldiers are morally permitted to kill in defence of their homeland.

Some readers might have no difficulty granting the point in this case — so deep is moral outrage, in this country at least, at this war. But it is worth highlighting what it implies: if, as some readers might think, the UK, the US and their coalition partners lacked a just cause for invading Iraq in 2003, then by implication and whatever the laws of war may say, soldiers of the coalition force were not morally permitted to kill Iraqi soldiers who were resisting the invasion. Theorists of the just war are familiar with such a view — indeed many explicitly endorse it. Outside academic philosophy, however, I wager that it is rather controversial — so much so, perhaps, as to cast doubt on the usefulness of philosophy to public policy, at least in this area.

Here is what non-philosophers may well be tempted to say. What follows is a rough summary of remarks I have often heard over the years:

'Well, many of our soldiers believed that they were doing the right thing, in fact the MPs who voted in favour of the war in 2003 believed they were doing the right thing. All those people acted in good faith. The invasion of Ukraine, or the Second World War for that matter, which philosophers of war keep invoking, are easy cases — too easy to serve as reliable guides given the complexities of most conflicts. Besides, suppose you're right and that our soldiers were not morally permitted to kill Iraqi soldiers who were defending their country. So what? Do we really want to revise the laws of war, so that killing enemy combatants in an unjust aggression is a war crime? Surely not: who would stand in judgement? How could we possibly enforce those laws? Besides, let us not pretend that Iraqi soldiers were saints: as well as defending their country, they were also defending a murderous regime so if the war was unjust, it was unjust on both sides, and it is really not clear what to make of that.'

Philosophers of war can learn at least four lessons from these sorts of remarks. First, they need to pay greater attention to the conditions under which policymakers — members of parliament, cabinet ministers, defence officials, military personnel — are required to make those decisions that will inevitably lead to deaths. This requires that they seriously

consider empirical evidence about war and political decision making, as well as records of parliamentary proceedings and publicly available policy briefs.

Second, they need to be more attentive than they have been so far to the fact that, according to their own principles, no war is fully just or unjust. This obvious observation, which one often hears from non-philosophers, has profound implications for our moral assessment of decisions to go to war in the first place, and of individual military missions with that war.

Third, they need to contribute more than they have done so far to debates about the relationship between the morality and the laws of war, and to discussions about the dilemmas faced by public officials whose moral judgements are in tension with the law. 'Resign!', we might say. But what if the law offers the best moral compromise we can hope for under the circumstances? What about the costs of keeping one's hands clean?

Finally, philosophers of war need to do much more work than they have done so far on the rights and wrongs of decision-making under conditions of uncertainty. By uncertainty, I mean both uncertainty about the facts of the case (for example, does a satellite image of troop movements along a border constitute evidence that our neighbour is about to attack us?), and uncertainty about what is the right thing to do (for example, does the loss of 10,000 soldiers constitute a sunk cost or not?).

Moral uncertainty is particularly troubling. There is no settled consensus amongst philosophers on many of the issues discussed here. Our leaders do not have time to wait until we develop a plausible view on sunk costs before deciding whether, for example, to increase military support for Ukraine. But what they ought to expect from philosophers is some guidance on the moral implications of both views.

To be sure, philosophers have addressed those issues — just as policy actors are aware of the moral complexities of seemingly straightforward principles. All too often, however, they fail, it seems, to communicate their views outside academia. To make progress, what is needed from both philosophers and policymakers is greater reciprocal engagement. This could take many forms, from compulsory ethics teaching in military academies, to setting up more systematic and regular exchanges between philosophers of war, our elected representatives, and members of the armed forces. In our current geopolitical moment, ethical reasoning that is both rigorous and attuned to the realities of armed conflicts is more important than ever.

On the new demise of ethics

Dr. Elke Schwarz, Queen Mary University of London

This essay is about the disappearance of ethics in an era which is characterised by the capacity to inflict large-scale violence at machine-speed. In what follows, I argue that in a technologically saturated military environment, where artificial intelligence (AI) is fast becoming a foundational infrastructure, ethics has taken a leave of absence. This should worry us, because the erosion of ethics in the domain of war and conflict might also be a harbinger for the demise of ethical thinking more broadly in public life.

At the time of writing, the world is witness to two aggressive international wars that have been raging on with breathtaking levels of human and material destruction. These two wars, in Ukraine and Gaza, have accelerated the development and rollout of new AI technologies that strive for greater and faster delivery of force, and reports of AI-enabled autonomous weapons systems being fielded are beginning to proliferate. The systems of greatest concern are those in which the human decision-maker plays an increasingly marginal role. In such times, which stand under the tragic sign of expanding violence, escalating conflict, and rising geopolitical tensions, the evaporation of what it means to give ethical consideration to other human beings should give us pause for thought.

At a 2024 AI Expo conference in Washington DC, a panel of experts with military, government, and defence industry backgrounds contemplated the wars of the future. The panel's core message was that future wars will be both devastating and incredibly lethal, but they will also be precise and humane because of new AI-enabled weapons. With this, the experts seemed to suggest that responsible powers must acquire AI weapons to deter irresponsible powers from using such weapons, and if responsible powers must use such weapons, they will do so, of course, responsibly. The ethos of this peculiar justification is perhaps best summed up by the remarks of one of the panellists — the CEO of an AI company that specialises in offensive targeting — who proclaimed with a remarkable

degree of enthusiasm: 'The peace activists are actually the pro-war activists, and we're the peace activists!' — proposing the argument that, if you don't want war, somehow the responsible action is to gear up for one. You would be forgiven for not following this logic. It makes a somewhat muddled allusion to the mechanisms of deterrence theory, a Cold War game theory-inspired way to justify stockpiling that most absurd of all weapons, the nuclear bomb. It is not a coherent discourse. But it is one in which the erosion of the status of ethics is patently demonstrated. This erosion has taken place gradually, concurrent with the ascendancy of computational rationality across various domains of human affairs.

In his 1954 book, *The Human Use of Human Beings* (Wiener, 1954), the grandfather of cybernetics Norbert Wiener makes some incisive observations about the co-constitution of humans and autonomous machines. He saw the perils of cybernetic technology for society as twofold. First, its inner logic might be wielded by one group of humans as a form of power and control over others. Second, this logic might also lead humans to believe it would be better to outsource complex and difficult moral decisions to machine authority. As an eminent scientist of technology and mathematical thinker, Wiener was ill at ease with the way in which technological thinking was fast becoming a dominant mode of philosophical and social reasoning. An insight we appear not to be heeding today.

Indeed, cybernetics, especially in the form of computational processes, has captured the collective imagination of the west since the 1950s and 60s, prompting many to understand the world and everything in it through feedback logics and probabilistic reasoning. Cybernetics is, first and foremost, a theory of processes and mechanisms. It is a theory of functionality in which inputs are translated into outputs. In cybernetics, the world is conceived as a collection of information that is to be transmitted. Information is understood as calculable data and even humans are conceived as objects of information, rendering them functional elements in a system determined by process. Cybernetics is unique in that it enables material and mental phenomena to be conceived in the same computational register. This makes its technological logic especially alluring.

It is unsurprising, then, that certain, prominent strands of moral philosophy were seduced into this schema. Derek Parfit's approach to applied ethics, for example, was thoroughly informed by the probabilistic logics of expected utility and game theory, and various modes of analytic moral reasoning about justifications for harm — from the Trolley Problem onwards — are characterised, in form, by approaching ethics as an algorithmic

puzzle. All this fits hand-in-glove with modern approaches to warfare, which are themselves more and more dependent on computational reason.

Every epoch has its own paradigmatic ways of conceptualising wars and of justifying the harms these paradigms produce. Ours stands under the sway of workflow efficiency. Current military doctrine and practice increasingly mirrors the logic of computational optimisation, prioritising cost-benefit calculations in the interest of speed and scale. All ethically salient considerations are swept up in this logic. The notion of permissible civilian harm, for example, is rerouted through the process of calculating 'collateral damage' — the percentage of civilians (men, women, children) that might be harmed if a situation necessitates it. A notoriously controversial calculus that clearly echoes the parameters of the Trolley Problem.

The Trolley Problem and its many variations are early iterations of algorithmic rationality. It is also fundamentally a problem about justifying harm. The basic scenario is fixed around the same dilemma: there is a runaway trolley, five people are tied down on one track, one person — either on another track or a footbridge or in some other perilous situation — will die. A decision must be made — will one person be sacrificed or shall five die? Each variation sets out to test to what degree different variables in the scenario matter to the moral outcome. New variables are introduced to change the moral calculus — a lever to divert the trolley; a tractor to rescue the five but which will then crush a bystander; the option to push someone off a bridge to stop the trolley. Technological instruments are always posited to make a difference to the justification of the lethal action. The role technologies play in shaping moral choices is almost never acknowledged. But crucially, this mode of reasoning implicitly constitutes an objectification of all that is human. Every subject becomes object in a world of computable things.

To be fair, Philippa Foot's original version of the problem was not intended to be used in such a formulaic manner, nor did Foot seek to spawn an entire 'Trolley-ology' industry. One might argue that utility-focused approaches to ethical reasoning originated precisely from a moral concern with social reform, refracted through a then radically new mode of accounting for egalitarian and universal values. (Although, it must be said, even in Bentham's era, the idea of smoothing the plurality of human life into a single calculus was controversial.) But the point at issue is this: ethical reasoning does not happen in a vacuum. It develops and changes alongside ways of envisioning and practising human life. In our time, these visions and practices stand under the sign of computation. Consequently, our forms of ethical reasoning are increasingly inflected by Bayesian statistics and game theory.

Nowhere is this more clearly felt than in the context of warfare. As military weapons become more computationally sophisticated, ethics is pressed further into a schema of technicity, to the point where ethics itself becomes technical, an exercise in which systematic process matters more than substance. Analytic approaches to ethics facilitate this. We can see this within abstracted conceptions of 'just war' reasoning, where the language of calculable utility dominates and the human is objectified, quite literally, in one of two ways: he or she is either the recipient of a moral decision and becomes a set of data points within a wider configuration of variables; or he or she is tasked with administering the moral calculation about who may die and in this way becomes a functional element in the wider data infrastructure. In their most abstract forms, analytic modes of just war reasoning become indistinguishable from technical discourse. They abound with mathematical syntax, taking on the pure form of algorithmic rationality itself: if X condition applies, then, *ceteris paribus*, Y must logically follow.

Algorithmic rationality seems tailor-made for thinking through ethical problems associated with autonomous, potentially lethal machines. The MIT Moral Machine project, for example, wants to identify the most agreeable ethical principles with which to programme self-driving cars, so that these cars can make the 'right' kill decision autonomously. Other research efforts by well-known technology companies are underway to design 'AI morality', so that humans can receive 'ethical guidance' on issues such as who should receive a kidney. When humans are rendered as objects of information, it is only logical that the most complex ethical calculations should be left to machines.

Abstraction marginalises the ethos of ethics altogether, especially in and for warfare. The aforementioned deterrence theory, popularised during the Cold War, is one such form of abstract reasoning about lethal technologies. It assumes that all actors are purely rational; that all actions and consequences of these actions are, to some degree, predictable; and that, based on this, the best outcome can be ascertained by drawing on expected utility and decision theory. This is ethics as economic science, as a technique for solving optimisation problems. In the context of nuclear deterrence, it overlooks entirely that the rational grounding of this approach rests on a threat of civilisational annihilation. In the contemporary context, with AI-enabled weapons, moral reasoning fits effortlessly into a workflow process, and it equally becomes ever more remote from recognising the ethos of the ethical problem itself, which is almost always an intractable social problem, resistant to optimisation efforts.

The abstraction is not accidental, but precisely the point. Abstraction decouples empathy from action. It decouples the consequences of actions from the moral weight of taking responsibility. It objectifies all those embroiled in the morally relevant calculus as data. It removes the most fundamental element of ethics — that of understanding one's responsibility in relation to a decision which affects others. It also works to render the 'rational' so fantastically technological that surely, only savvy experts can be trusted with engineering the difficult choices. This directs our gaze onto 'solvable' ethical problems and away from those that do not sit so comfortably within a techno-economic logic.

However, our morality, and thus our moral decision-making, is anchored in a shared history of human plurality and human social relations. It is this condition that makes us not just functional actors, but moral actors. Moral agency means the capacity not simply to have, but also to take moral responsibility in relation to others. And that means having the capacity and space to reflect on one's position and that of others in any given configuration of moral relevance, particularly in warfare. The ever-present potential for open-endedness in moral decision-making and responsibility-taking does not lend itself to the parameters of systems that function best with and within closures, and at speed and scale, as is the case with AI.

Analytic approaches to ethics are, of course, only one way of thinking about ethics as it relates to war and technologies of war. Instead of an algorithmic programme, ethics might be understood as a relational practice. Ethics as a practice is open-ended, deliberative, and ultimately concerned with broader questions concerning human relations, not technical processes. It rests on a foundation of human relations, anchored in both vulnerability and trust. As pivotal parameters for social and political life change, ethical guidelines, their associated practices, and our human relationship with ethical values and beliefs become subject to new inquiry. Ethical thought should inform the laws, guidelines, and rules that govern social and political life. But it must also serve to prompt critical reflection about the ways in which morality might become fossilised in these laws and rules, to avoid the risk that the form supersedes the content. I am reminded of Hannah Arendt's observation about morality and mores: 'What people get used to is less the content of the rules [than] the possession of the rules under which to subsume particulars.'

Having a set of codified ethical principles in place does not guarantee ethics in practice. This is tragically confirmed in our present moment when the law appears to be inadequate to curb the worst excesses of the ongoing conflicts, let alone stymie their expansion. Neither just

war principles nor international humanitarian law were intended as a checklist that could be satisfied, or optimised, in purely procedural ways. Rather, the intent with early just war thinking was to grapple with and understand that actions and circumstances need to be considered carefully in morally charged environments. That there are always broader ethical questions to consider in any war, questions about how we wish to live with one another, about the possibility for peace once fighting ends.

There is no easy answer to the question of how many innocent people may be killed to achieve a goal in warfare, nor is it sufficient to think of the ethics of war in mere numerical terms, counting only the dead. Those that make a decision to kill must wrestle with the weight of responsibility attached to such actions. For this, we must remain humans, not become objects ourselves, and we must tolerate the discomfort of an irresolvable ethical challenge, despite being embedded within a network of machine functionality. Herein resides moral responsibility.

The ethos of international humanitarian law or the just war tradition is still intuitive to most of us: those blameless should not have to suffer, violence should not be a first resort, and there should be a mandate for restraint to violence because otherwise peace cannot be restored. Where ethics becomes automated, this ethos is exchanged for variables that can be adjusted based on preference, interest and aim. In other words, it becomes subject to power, which ultimately is a feature of nihilism. But that is where we are today. The accelerated rollout of AI-enabled technologies of warfare, their imperative to make war faster and more lethal, and their logics of process and efficiency leave neither space nor reason for ethics to thrive, not even as a calculation. No wonder, then, that ethics is on its way out as an uncomfortable necessity, with the many intractable, irresolvable, difficult challenges that arise in war. Ethics is about meaning; AI is about function. The two are not easy bedfellows.

Two points are worth stressing in closing. First, AI's emphasis on speed is utterly incompatible with ethical deliberation. I have written elsewhere on how the iterative logic of failing and learning through accelerated processing power is both the foundation of AI as well as the technology industry itself. The motto 'move fast and break things' is not just a PR slogan; it is an ideology. Increasingly too, it seems, for the defence domain. With this shift, ethics is necessarily sidelined. The second point is that computational logics always objectify their environments, and that means objectifying humans. When an AI system identifies a human as a target-object, that human is immediately objectified, and thus

dehumanised. And from the history of mass violence, we know that dehumanisation plays a significant role in facilitating an erosion of moral restraint.

And algorithmic ethics cannot address either of these issues. Quite the contrary. Taking the ethos of ethics seriously would mean, first and foremost, not succumbing to the speed and scale mandate of killing as a systematic process. It would also mean that we ask more of those involved in AI-enabled war, not less; that we should consider maximally restrictive ideas about the use of violence rather than follow the logic of machine functionality. This means that those involved in the business of war should become habituated into thinking about humans as humans, not objects, depicted on an interface.

The ethical stakes are patently high, in warfare and otherwise. By functionalising the notion of agency and by marginalising the human as a decision maker, we diminish our ability to understand what moral agency, and responsibility, means in practice. To be sure, moral agency can be theorised and defined in different ways, and I accept that perhaps our idea of what constitutes agency should be revised in an AI-saturated environment. But agency and *moral* agency are not the same. Moral relations do not arise *ex nihilo*. As humans, we understand and are able to judge the contextual nuances of human relations in a way that technological artefacts are simply not able to. Technological artefacts can comprehend neither the complexities and incongruities of human life nor its meaning.

AI-enabled weapon systems are such technological artefacts. And they are also technologies of violence. Instruments of violence are designed to inflict damage on human bodies and human property. Too often we discuss these systems as though they were somehow separate from this violence against human life; as though war is primarily an engineering challenge and as though they somehow carry within themselves their own justification for the violence they enable. Algorithmic ethics facilitates this configuration.

Rethinking ethics not as systematic process, but as practice rooted in human relations, is an urgent task. If we let ethics disappear, and with that the intrinsic, mutual value of human life, we are readying ourselves for a future of violence, not peace.

References:

- Wiener, N. (1954). *The human use of human beings: Cybernetics and society*. New York: Da Capo Press.

The world and his wife

Dr Musab Younis, University of Oxford

Our current political moment is characterised by such extreme and shocking violence that any measured reflection on the role of philosophy in public life can seem impossible. The elevated ideas that accompanied the rise of international organisations in the 20th century — peace, harmony, order, cooperation — appear as distant from reality as at any moment since the end of the second world war. The events of the year 2024, and in particular Israel's merciless genocide of Palestinians in Gaza, resolutely backed by the UK, the US and other western powers, were described by observers as a 'breaking point' and 'death blow' for international order.¹ And yet if brute force still governs the world, the language of force does not. Principles that we might call philosophical in nature are still being articulated constantly in public life. Across the world, in both popular and official discourse, we find references to progress, fairness, efficiency, justice, development and peace.

I want to suggest in this essay that philosophy has a role to play in politics even if its actual operation in the world is not easy to discern. Philosophy, understood capaciously — I leave aside in this essay debates about the differences between political theory and political philosophy — can challenge the common assumptions and limits that frame political discourse. By asking questions that go beyond the usual bounds of public debate, it can illuminate the unexamined and ahistorical assumptions that are to be found everywhere in public life. This type of philosophy is not the exclusive preserve of scholars working in universities. Anyone who has sought to name and describe a form of domination they experience, with the aim of resisting it, is — in my sense of the term — a philosopher. Many of those who lived under colonial rule, for example, developed forms of anticolonial thought that radically challenged the legitimacy of empire. Their arguments are expressions of political philosophy that show why imperial formations like Britain cannot be understood as isolated and singular national units.

Since the wave of decolonisation in the 20th century created a world of sovereign and supposedly equivalent states, many academic political theorists have been interested in the question of borders. Do national borders have any moral value? Do we owe more to fellow citizens of our own states than to the citizens of other states? But these questions are not often found in public discourse. What predominates in the mainstream media and among the largest political parties is a kind of certainty that sees the central actor of politics as a nation, distinct from and locked into endless competition with other nations. In a competitive world, ‘we’ must not lose out to ‘them’ — the vaguely hostile multitudes living beyond our protective borders but threatening constantly to traverse them.

This position corresponds broadly to a philosophical approach known as communitarianism. It emphasises the moral value of communities, which in the modern world are typically understood as sovereign states. It finds the borders that protect those states to be morally just. And the aim of foreign policy is basically to maintain or improve a country’s position vis-à-vis the positions of other states. Here is Winston Churchill explaining in 1914 that ‘we’, the British: ‘... are not a young people with an innocent record and a scanty inheritance ... We have engrossed to ourselves an altogether disproportionate share of the wealth and traffic of the world. We have got all we want in territory, and our claim to be left in the unmolested enjoyment of vast and splendid possessions, mainly acquired by violence, largely maintained by force, often seems less reasonable to others than to us.’ (Darwin, 2009). Churchill’s remarks were not intended to be self-critical. As a politician, he saw his role as helping to maintain Britain’s ‘disproportionate share’ rather than bring it into question.

Communitarian attitudes dominate conversations about immigration in Britain. If public opinion is divided on the question — polls show complicated and shifting attitudes, depending on how the question is framed — there is much less complexity in electoral politics. Politicians generally compete in their expressions of hostility to new arrivals. After winning the 2024 general election, the Labour party swiftly announced ‘new measures to cut historically high levels of net migration.’ At the Conservative party conference in October 2024, the contenders for party leadership all stressed their anti-migrant attitudes. ‘The world and his wife and their extended family...’ cannot come to Britain, said Robert Jenrick. ‘The age of mass migration must end.’ On X, he promised to ‘deport foreign criminals, get terrorists off our streets, and end illegal migration’. Immigrants who bring ‘foreign conflicts’ with them should not come to Britain, said Kemi Badenoch, who

eventually won the leadership contest. She added that 'not all cultures are equally valid' and that some immigrants bring with them 'ancestral ethnic hostilities.' James Cleverly, a former government minister, said that 'dozens' of asylum seekers would now be in Rwanda if he and his colleagues had remained in power. All these are strongly communitarian arguments. They animate much of the public and media discussion on migration through the use of terms like 'secure borders', 'small boats', 'asylum seekers', 'illegal immigration', 'foreign criminals', and 'deportations' — all of which stress the idea of an embattled political community. I am talking about Britain, but the story is a similar one in other Western countries like the US, France, Germany and Australia.

And yet political theory and philosophy, by which I mean critical and sustained reflection about politics that is not confined to the logic of policymaking, shows many alternatives to the communitarian position. Some argue straightforwardly for a radical cosmopolitanism that treats all people across the world as moral equals. Others focus on how nations have historically come to be imagined, and what is included and excluded from those acts of imagination. This latter view is, I suggest, a particularly useful one for thinking about the question of borders and immigration. With respect to Britain, it shows how the state was constituted through — and cannot be separated from — a form of hierarchical engagement with the world in the form of imperialism. A public discourse might paint Britain as an embattled political community. But this elides the fact that, as Churchill suggested, Britain is, on balance, the beneficiary of a global order that it constructed in its own interests. And this makes the idea of Britain as a bounded entity unconvincing in both historical and moral terms.

Such arguments have long been articulated by colonised peoples. 'It was the sweat of the black man's brow which laid the foundation for the present day opulence of Britain and White America,' wrote an anonymous Sierra Leonean editorialist in 1938. The editorialist compared the relationship of Britain and Sierra Leone to that of 'the cheat' with 'his victim', and insisted the peoples of West Africa had a right to reparations after suffering for centuries from enslavement, colonisation, and forced transportation ("Rambling Talks," *Sierra Leone Weekly News*, July 2, 1938.). Similar arguments across anticolonial writing saw Britain as a country whose present existence cannot be separated from its history as the centre of what was once the world's most expansive empire. As Kojo Koram wrote recently: 'Rather than saying Britain had an empire, it would be more accurate to say that the empire had Britain.' (Koram, 2024).

A historically informed reflection on national identity opens up ways of thinking about politics beyond the reflexive defensiveness that characterises so much of contemporary discourse. It also throws into question the boundaries of the body that is imagined as the subject of politics. The categories through which our current political language is articulated depend on the idea of a separation between the domestic and international. Yet anticolonialists have long challenged that separation. In its place, they have insisted on the co-constitution of what came to be known as the global north and south. As Fanon put it: 'Europe is literally the creation of the Third World' (Fanon, 1963). Since the encompassing of the earth by western empires, national groups have always been imagined as elements of world-spanning projects. This is true for anticolonial nationalisms no less than colonial ones.

There are of course those who will bristle at any attempt to challenge the deeply rooted idea of national interest. 'Some people love to talk our country down,' said James Cleverly during the Conservative leadership contest. But 'this country has given so much to so many people.' It is easy to imagine the retort to this claim. The left is more patriotic than the right because its vision of 'our country' is more egalitarian and less beholden to elite interests. Yet each side of this debate, which dates back at least to the French Revolution, is circumscribed by the same spatial demarcation that refuses to accept how 'we' are connected to 'them', those beyond our borders who are excluded from the conversation even as they so often constitute its obsessive subject.

In the popular imagination, political theory is written by academics who are mostly based in universities in Europe and the US. But there are many traditions of serious thought about politics — including those that emerge from anticolonial traditions, social movements and radical political formations — that are not included in this definition of political theory. Those traditions are vital in understanding the injustices of the world and the necessity for its reorganisation and reconstruction. The events of recent years have acted as a sharp reminder of the rigid hierarchies that still govern international politics and the extreme difficulty for the vast majority of the world's peoples of resisting imperial power even at its most shocking and rapacious. Since the advent of European socialism, some people have criticised these hierarchies from the perspective of working populations of the imperial core. They have suggested, for example, that workers in Europe and North America would benefit from diverting funding from military uses to hospitals, schools and other social services. These arguments are convincing, of course. And yet — for all their well-

intentioned nature — they reproduce the idea that the people who really count are those who constitute part of ‘our’ collective national body.

When we consider the ways in which Britain was historically constituted as an empire, and is still constituted today as part of an imperial formation, it is much less easy to justify any such political demarcation that separates Britain’s population from the subjects of its ex-colonies. One of Edward Said’s key contributions was to show how national identities, especially those forged in a colonial context, implicate other peoples in ways that cannot be forgotten or ignored. This means that there is simply no realistic concept of what Britain is that does not take into account its expansion across so much of Africa, Asia, and the Caribbean, not to speak of North America, Australia and New Zealand, in a centuries-long process that created not just a single nation or state but many different states locked into a global — and extremely unequal — system of production and exchange. Palestine is often described as relevant to Britain because of Britain’s history of colonisation of the region. This is true. Yet we also have to insist — looking to Gaza, Sudan, the Congo, Haiti — that the history of colonialism still lives with us in the present, in a singular world that our prevailing discourse wrongly depicts as compartmentalised.

Notes:

- 1 An interview with seven leading international experts on genocide found that in relation to Israel’s destruction of Gaza “the question is not controversial—even for those who previously rejected the label” (NRC, May 14, 2025).

References:

- Darwin, J. (2009). *The Empire Project*. Cambridge: Cambridge University Press.
- Fanon, F. (1963). *The Wretched of the Earth*. New York: Grove Press.
- Koram, K (2024). Kemi Badenoch should read some Edmund Burke. [online] *New Statesman*. Available at: <https://www.newstatesman.com/ideas/2024/04/kemi-badenoch-should-read-edmund-burke> [Accessed 20 Feb. 2025].

Author biographies

Author biographies

David Archard

David Archard is Emeritus Professor of Philosophy at Queen's University Belfast, having previously taught at the Universities of Ulster, St Andrews and Lancaster. He has published extensively in applied ethics, political philosophy, and philosophy of law, especially on the topics of children, the family, sexual consent, and the public role of philosophy. His book *Children: Rights and Childhood*, now in its 3rd edition, was the first book philosophically to analyse the moral and political status of children.

He has been Honorary Chair of the Society for Applied Philosophy and is its Vice-President. For twelve years he was a Member of the Human Fertilization and Embryology Authority, and latterly its Deputy Chair. He was Chair of the Nuffield Council on Bioethics, the United Kingdom's de facto national ethics committee (2016-2023), is Chair of the Ethics Advisory Group of the 'Every Story Matters' exercise of the UK's COVID Public Inquiry, and a Member of the Clinical Ethics Committee of Great Ormond Street Hospital.

Shuk Ying Chan

Shuk Ying Chan is a Lecturer in Political Theory in the Department of Political Science at University College London. In her recent work, she has been interested in questions of colonialism and decolonization, empire and race, and reparative justice for historical injustices. Her monograph, *Postcolonial Global Justice* (Princeton University Press), was published in August 2025.

Prior to joining UCL in 2023, she was a Postdoctoral Prize Research Fellow at Nuffield College and the Department of Politics and International Relations at Oxford University. She completed her PhD in political theory at Princeton University. Her dissertation was awarded the 2022 Leo Strauss Award (for best dissertation in political philosophy) from the American Political Science Association. Originally from Hong Kong, she holds an MPhil in Political Theory and a BA in History and Politics from the University of Hong Kong.

Lorna Finlayson

Lorna Finlayson is a lecturer in philosophy at the University of Essex, and was previously a junior research fellow at King's College, Cambridge. She has published on a range of topics in political philosophy and its methodology, including feminism, political realism, and theories of ideology. She has also written for popular outlets such as the *London Review of Books*, the *New Left Review*, *The Guardian* and *The New Statesman*, on topics including universities, the Labour Party, free speech, animal rights, schools, and the NHS. She is the author of two books, *The Political Is Political* (London: Rowman and Littlefield International 2015) and *An Introduction to Feminism* (Cambridge: Cambridge University Press 2016). A third book, on the politics and philosophy of childhood, is under contract with Penguin Press.

Tariq Modood

Tariq Modood is Professor of Sociology, Politics, and Public Policy at the University of Bristol, and founding Director of the Bristol University Research Centre for the Study of Ethnicity and Citizenship. He has held over 40 grants and consultancies (UK, European and US), authored and edited over 35 books and reports, and published over 250 articles or chapters in political philosophy, sociology and public policy. He is the co-founding editor of the international journal, *Ethnicities*, and has served on the Commission on the Future of Multi-Ethnic Britain, the National Equality Panel, and the Commission on Religion and Belief in British Public Life. He was elected a Fellow of the British Academy in 2017.

Debra Satz

Debra Satz is the Vernon R. and Lysbeth Anderson Dean of the School of Humanities and Sciences at Stanford University, the Marta Sutton Weeks Professor Philosophy, and, by courtesy, Political Science. Her research focuses on the ethical limits of markets, objections to inequality, and issues in practical ethics. In 2004, Satz received the Walter J. Gores Award, Stanford's highest teaching honor.

Among her publications are *Why Some Things Should Not Be for Sale: The Moral Limits of Markets* (Oxford University Press, 2010); *Economic Analysis, Moral Philosophy and Public Policy* (Cambridge University Press, 2016) (with Dan Hausman and Michael McPherson); *Ideas That Matter: Democracy, Justice, Rights* (Oxford University Press, 2019) (with Annabelle Lever) and was a coauthor of the Institute for Fiscal Studies *Deaton Review: Dimensions of Inequality*, published by Oxford Open Series in 2024. A co-authored follow up to the *Deaton Review, Changing Inequalities: Causes, Concerns and Consequences* is forthcoming in 2026. She is a member of the American Academy of Arts and Sciences.

Adam Swift

Adam Swift is Professor of Political Theory /Philosophy in the Department of Political Science at University College London. He previously taught at Balliol College, Oxford — where he founded the Centre for the Study of Social Justice — and the University of Warwick. His book *Political Philosophy: A Beginners' Guide for Students and Politicians*, now in its 4th edition, has been translated into nine languages.

He has an M.Phil in Sociology from Nuffield College and is especially interested in interdisciplinary and collaborative work, often combining normative philosophy with empirical social science. He has published on a wide range of topics but has come to specialize in questions about social mobility, social justice, education and the family. His books include *How Not To Be A Hypocrite: School Choice for the Morally Perplexed Parent*, *Family Values: The Ethics of Parent-Child Relationships*, *Educational Goods: Values, Evidence and Decision-Making* and *How To Think About Religious Schools: Principles and Policies*.

James Wilson

James Wilson is Professor of Philosophy at UCL, and co-Director of the UCL Health Humanities Centre. He has pioneered a socially engaged and reflexive approach to philosophy, which not only brings existing philosophical research to bear on practical problems, but also examines how detailed understanding of practical problems reveals gaps and conceptual problems in existing philosophical theories. He has published extensively in applied and engaged ethics, including influential articles on health inequalities, healthcare resource allocation, paternalism, methodology in normative thinking, intellectual property, and the right to privacy. He is author of *Philosophy for Public Health and Public Policy: Beyond the Neglectful State* (OUP, 2021).

Wilson is joint editor of the philosophy journal MIND, and is a trustee of the Royal Institute of Philosophy. He has been a member of the National Data Guardian's Panel of adviser since 2016, and Deputy Chair of the Metropolitan Police Service's research ethics committee since 2020.

Jonathan Wolff

Jonathan Wolff is Emeritus Alfred Landecker Professor of Values and Public Policy, Blavatnik School of Government, University of Oxford. Previously he was Dean of Arts and Humanities and Professor of Philosophy at UCL. His books include *Robert Nozick* (1991), *An Introduction to Political Philosophy* (1996, 4th edition 2023), *Why Read Marx Today?* (2003), *Disadvantage* (with Avner de-Shalit 2007), *Ethics and Public Policy* (2011, 2nd edition 2019), *The Human Right to Health* (2012), *An Introduction to Moral Philosophy* (2018, 3rd edition 2025) and *City of Equals* (with Avner de-Shalit 2023). He is particularly interested in the relation between values and public policy, and has worked with The Home Office, Department of Culture Media and Sport, Railway Safety and Standards Board, The Nuffield Council on Bioethics, and the World Health Organisation. He is a Fellow of the British Academy and President of The Royal Institute of Philosophy.

Clare Chambers

Clare Chambers is Professor of Political Philosophy and a Fellow and Dean of Jesus College, University of Cambridge. She is the author of *Freedom & Equality: Essays in Liberalism and Feminism* (OUP, 2024); *Intact: A Defence of the Unmodified Body* (Penguin, 2022); *Against Marriage: An Egalitarian Defence of the Marriage-Free State* (OUP, 2017), which won the 2018 David Easton Prize of the American Political Science Association; *Sex, Culture, and Justice: The Limits of Choice* (Penn State Press, 2008); *Teach Yourself Political Philosophy* (with Phil Parvin, Hodder, 2012); and numerous articles and chapters on political philosophy, gender, and bioethics. Clare Chambers is also the editor of *The Routledge Handbook of Philosophy of Sex and Sexuality* (with Brian D. Earp and Lori Watson, Routledge, 2022) and co-Editor-in-Chief of *Res Publica*, the journal of legal, moral, and social philosophy (with Sune Laegaard).

Clare Chambers regularly appears on BBC Radio, and her research has featured in print and online media as diverse as *The Guardian*, *The New Statesman*, *El País*, *la Repubblica*, *iNews*, *The*

Times of India, the *Times Literary Supplement*, *The American Conservative*, *The New Humanist*, *Metro*, *Slate*, *Philosophy* 24/7, *Philosophy Bites*, *Daily Nous*, and *Aeon*. She was a Council Member of the Nuffield Council on Bioethics from 2020-25.

Anca Gheaus

Anca Gheaus is an Associate Professor in the Political Science Department at the Central European University in Vienna. She is a political philosopher interested in justice and the normative significance of personal relationships. She co-authored the book *Debating Surrogacy*, which came out at the Oxford University Press in 2024, and co-edited *The Routledge Handbook of the Philosophy of Childhood and Children*, published in 2018 by Routledge. She is also the author of numerous journal articles and book chapters, primarily on issues concerning childrearing, gender justice, non-ideal theory and methodological issues in political philosophy. At the moment she is working on a monograph on child-centred childrearing, under contract with Oxford University Press.

Katharine Jenkins

Katharine Jenkins is Professor of Philosophy at the University of Glasgow. She previously taught at the University of Nottingham, and held a Junior Research Fellowship at Jesus College, Cambridge. She has published widely in social philosophy, especially feminist philosophy and social ontology, and is the author of a monograph, *Ontology and Oppression: Race, Gender and Social Reality*, which explores how social categories such as races and genders exist and how these categories are bound up with systematic injustices. She has also published a short book for a general audience, *Feminist Philosophy: A Very Short Introduction*. She is the co-director of the Society for Women in Philosophy UK, and (for the period of 2024-2028) a Co-Editor-in-Chief of the academic journal *Hypatia: A Journal of Feminist Philosophy*.

Winnie Ma

Winnie Ma is a J.D. candidate at the University of California, Berkeley, School of Law, with particular interests in civil rights and ethical AI innovation. As a law student, she has worked at UC Berkeley's Samuelson Law, Technology & Public Policy Clinic, and has clerked at the Youth Law Center, whose mission is to transform the United States' juvenile justice and child welfare systems. Formerly, she was Lecturer in Philosophy at King's College London, specializing in the epistemology and ethics of AI, including medical algorithms, and the ethics of belief of stereotyping more generally. She has taught courses ranging from epistemology to the politics of health and medicine at King's College London in the Departments of Philosophy and of Global Health & Social Medicine, as well as the medical school. And she was formerly Research Associate and Project Manager at the Sowerby Philosophy & Medicine Project, working to introduce philosophy into the curricula that train clinicians and encourage interdisciplinary dialogue and collaborative research.

Martin O'Neill

Martin O'Neill is Professor of Political Philosophy at the University of York, having previously taught at the Universities of Cambridge and Manchester. He has published widely on the theory and practice of social and economic justice, in journals including *Philosophy & Public Affairs* and *Political Philosophy*. Among his books is *The Case for Community Wealth Building* (Polity Press, 2019), co-authored with Joe Guinan. He currently serves on the executive committee of the British Philosophical Association, and is a member of the Council of the Royal Institute of Philosophy.

Martin is a member of the Trustee Board of the Democracy Collaborative (a 'think-and-do tank' based in Washington DC), and previously served as a member of the UK Labour Party's Community Wealth Building Unit. He has consulted for international organisations including the International Labour Organization (ILO) and the World Economic Forum (WEF), and his work has been supported by funders

including the Institute for New Economic Thinking (INET), the Joseph Rowntree Foundation, the British Academy, the Leverhulme Trust, and the Interdisciplinary Social Research Foundation (ISRF).

John Tasioulas

John Tasioulas is Professor of Ethics and Legal Philosophy at the University of Oxford and inaugural Director of Oxford's Institute for Ethics in AI. He has previously taught at Glasgow, King's College London, and University College London, and held visiting appointments at the University of Chicago, Harvard University, and the University of Notre Dame. He has published extensively in moral, political, and legal philosophy, with a focus on human rights, punishment, international law, and the ethics of AI and digital technology. He is the author of *On Justice and Mercy: Essays in Moral and Legal Philosophy* (forthcoming, OUP) and the editor of *The Cambridge Companion to the Philosophy of Law* (CUP, 2020).

He has acted as a consultant on human rights to the World Bank and served as a member of the International Advisory Board, Panel for the Future of Science and Technology (STOA), European Parliament and as a member of the Prime Minister of Greece's High-Level Advisory Committee on AI. He is a Senior Fellow in Schmidt Sciences's AI2050 program and is currently engaged in an AI2050 project on humanistic AI ethics together with Professor Hélène Landemore (Yale University).

John Broome

John Broome is Emeritus White's Professor of Moral Philosophy at the University of Oxford, Honorary Professor at the Australian National University, and an Emeritus Fellow of Corpus Christi College, Oxford. He was previously Professor of Philosophy at the University of St Andrews and Professor of Economics at the University of Bristol.

He works on the philosophy of normativity and reasoning, and on the philosophy of climate change. He was a Lead Author of the Fifth Assessment report of the Intergovernmental Panel on Climate

Change. His most recent books are: *Climate Matters* (2012), *Rationality Through Reasoning* (2013), and *Normativity, Rationality and Reasoning* (2021).

Broome is a Fellow of the British Academy, the Australian Academy of the Humanities, and the Royal Society of Edinburgh. He is a Member of the Swedish Royal Academy of Sciences and the American Academy of Arts and Sciences. He is an Honorary Fellow of Trinity Hall, Cambridge, and holds an Honorary Doctorate from Lund University.

Sarah Bufkin

Sarah Bufkin is Assistant Professor of Political Theory at the University of Birmingham. She works on Black Atlantic political thought and Critical Theory, with a specific focus on racism and racialization, imperial techniques of government, and racial capitalism. Prior to Birmingham, she was an Examination Fellow at All Souls College, Oxford, where she completed her doctorate in political theory. Her writing has been published in *Cultural Studies*, *Philosophy & Social Criticism*, and *Identities*, among other journals. Sarah is a book review editor at *Perspectives on Politics*, an editor at *Contemporary Political Theory*, and a member of the editorial advisory group at *Soundings*. She is currently writing a book theorizing Frantz Fanon for the afterlife of empire. More broadly, she is interested in anticolonial thought, feminisms, critical social theory, continental philosophies of self and Other, and critiques of the carceral state.

Roger Crisp

Roger Crisp is Director of the Uehiro Oxford Institute, and Professor of Moral Philosophy at the University of Oxford. He has long-standing interests in the nature of ethics, ethical theories, and real-life ethical problems. With Tony Hope and others, he helped to establish in 1995 one of the first UK Clinical Ethics Committees, in the Oxford Radcliffe Hospitals Trust. He is the author of *Mill on Utilitarianism*, *Reasons and the Good*, *The Cosmos of Duty: Henry Sidgwick's Methods of Ethics*, and *Sacrifice Regained: Morality*

and Self-interest in British Moral Philosophy from Hobbes to Bentham. He edited the *Oxford Handbook of the History of Ethics*, and translated Aristotle's *Nicomachean Ethics* for Cambridge University Press. He is currently translating and commenting on three of the Platonic dialogues concerned with the death of Socrates, as well as thinking, when time allows, about Buddhism and personal identity.

Katrien Devolder

Katrien Devolder is Professor of Applied Ethics and Director of Public Philosophy at the Uehiro Oxford Institute and GB Fellow at Reuben College, University of Oxford. She has published two monographs: one on compromise positions in the embryonic stem cell debate (OUP, 2015) and one on the ethics of human cloning (Leuven University Press, 2001). She has also published numerous papers on ethical issues pertaining to gene editing, gamete donation and genetic parenthood, life extension, compromise in bioethical debate, moral complicity, animal ethics, conscientious objection, and euthanasia in prisoners. Her most recent work focuses on the concept and ethics of laziness. She also produces (conducts, films and edits) interviews with academics to make complex ethical debates accessible to a wide audience which can be viewed on YouTube's 'The Practical Ethics Channel' or listened to on Apple Podcasts ('Thinking Out Loud').

Jonathan Pugh

Jonathan Pugh is a Senior Research Fellow at the Uehiro Oxford Institute. He is also the Theme Lead for Values and Society at Reuben College, University of Oxford. His research interests lie primarily in issues concerning personal autonomy in practical ethics, particularly topics pertaining to informed consent. He was previously a member of the UK Pandemic Ethics Accelerator, and the Parfit-Radcliffe Richards Fellow at the University of Oxford. He is the author of *Autonomy, Rationality and Contemporary Bioethics* (OUP, 2020), and

he has also published work in philosophical and medical journals on a number of topics in neuroethics, research ethics, public health ethics, the ethics of psychiatry, and clinical ethics. He is currently the Principal Investigator on the project ‘Predictive Technologies, Ethics, and the Future of Insurance’, and the academic lead on the ‘Decision Aid for the Restitution of Cultural Artefacts’ project.

Cécile Fabre

Cécile Fabre

Cécile Fabre is Senior Research Fellow in Politics at All Souls College, Oxford. She holds degrees from La Sorbonne, the University of York, and the University of Oxford, and had held positions at the LSE and the University of Edinburgh. Her research interests include theories of distributive justice, the rights we have over our own body, and the ethics of foreign policy. She has published papers in (amongst others) *Ethics*, *Law and Philosophy*, *The Journal of Political Philosophy* and the *British Journal of Political Science*. Her books include *Social Rights under the Constitution* (OUP 2000), *Whose Body is it Anyway?* (OUP 2006), *Cosmopolitan War* (OUP 2012), *Cosmopolitan Peace* (OUP 2016), *Economic Statecraft* (Harvard UP 2018), and *Spying Through a Glass Darkly* (OUP 2022). She delivered the Tanner Lectures on Human Values at Stanford in 2022 on the ethics of preserving cultural heritage, and is expanding the lectures into a monograph. She is a member of Academia Europaea and a Fellow of the British Academy.

Elke Schwarz

Elke Schwarz is Professor of Political Theory at Queen Mary University London, UK. She holds a PhD from the London School of Economics and Political Science (LSE), an MA in Conflict Studies from the War Studies Department at King’s College London (KCL) and a Bachelor in Business Studies from Belmont University (USA). She has published extensively on the ethics of technology and warfare, with a specific emphasis on new and emerging military technologies, including military Artificial

Intelligence (AI), autonomous weapon systems, drones and robots. She is the author of *Death Machines: The Ethics of Violent Technologies* (Manchester University Press). Elke is currently Vice-Chair of the International Committee of Robot Arms Control (ICRAC) and also an Associate with the Imperial War Museum (IWM) and an RSA Fellow. She held a 2022/23 Fellowship at the Center for Apocalyptic Studies (CAPAS) and was a 2024 Leverhulme Research Fellow.

Musab Younis

Musab Younis is Associate Professor of Political Theory and Jarvis Doctorow Tutorial Fellow in Politics at St Edmund Hall, University of Oxford. Previously, he worked at Queen Mary University of London (QMUL) as Senior Lecturer in the School of Politics and International Relations. He completed his doctorate at St Catherine’s College, Oxford, and has held visiting positions at the European University Institute and the University of London Institute in Paris.

His research explores political theory in relation to race and empire, with a focus on the history of anticolonial thought. He is the author of *On the Scale of the World: The Formation of Black Anticolonial Thought* (University of California Press, 2022), which won the Sussex International Theory Prize and was named as one of the best academic books of 2023 by *The New Statesman* magazine. In addition to his academic writing, he is a regular contributor to the *London Review of Books*. His essays have also appeared in *n+1*, *Prospect*, *Baffler*, and *The Guardian*.

Acknowledgements

This volume was co-edited by Professor David Archard and Matt Davies.

Ben Brown, Harry Farmer, Mahi Hardalupas, Martin Davies, and Victoria Briggs provided substantive contributions and editorial support. Copyediting was carried out by Rebecca Ghani and design work by Eureka! Design Consultants. The production of this volume would not have been possible without extensive project management support from Ben Brown, and important contributions from other Nuffield Foundation colleagues including Raf Bogan, Michelle Wuisan, Nicola Webb and Victoria Briggs. The editors are grateful to all of the above for their work on this project.

Tim Gardam, former CEO of the Nuffield Foundation, provided the initial impetus for this project. The editors would like to thank Tim and his successor as Nuffield CEO, Gavin Kelly, for their role in making this publication a reality.

