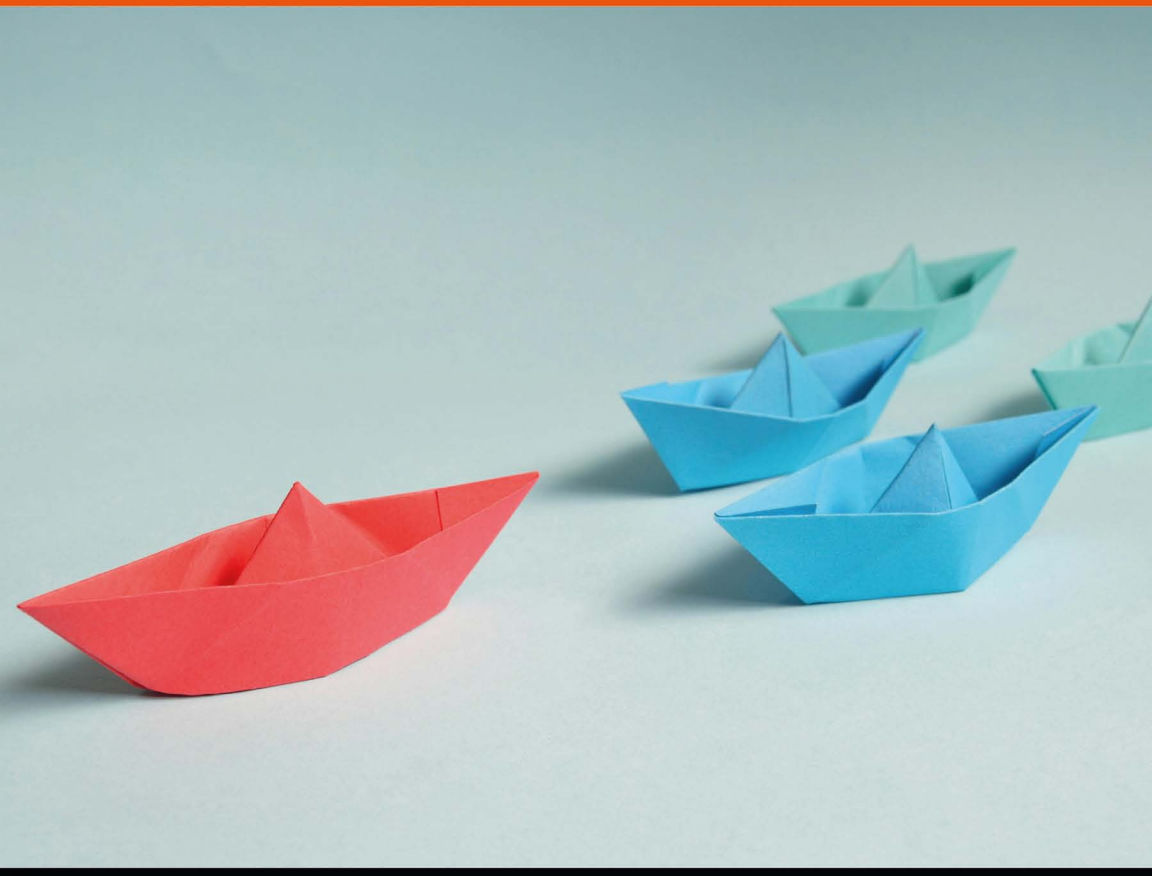

Political Beliefs

A Philosophical Introduction

Oliver Traldi

ROUTLEDGE

POLITICAL BELIEFS

Anyone who's had an argument about politics with a friend may walk away wondering how this friend could possibly hold the beliefs they do. A few self-reflective people might even wonder about their own political beliefs after such an argument. This book is about the reasons that people have, and could have, for political beliefs: the evidence they might draw on, the psychological sources of their views, and the question of how we ought to form our political beliefs if we want to be rational.

The book's twenty-four chapters are divided into four larger parts, which cover the following: (1) the differences between political and other types of beliefs, (2) theories of political belief formation, (3) sources of our political beliefs and how we might evaluate them, and (4) contemporary phenomena – like polarization, fake news, and conspiracy theories – related to political beliefs.

Along the way, the book addresses questions that will arise naturally for many readers, like:

1. Does the news you choose to watch and your own social media leave you stuck in an "information bubble"?
2. Are you committed to a certain ideology because of the history of your society?
3. Are people who believe "fake news" always acting irrationally?
4. Does democracy do a good job of figuring out what's true?
5. Are some political beliefs good and some evil?

As the book investigates these and other questions, it delves into technical, philosophical topics like epistemic normativity, the connection between belief and action, pragmatic encroachment, debunking arguments, and ideology critique. Chapter summaries and discussion questions will help students and all interested readers better grasp this new, important area on the border of politics and philosophy.

Oliver Traldi is a John and Daria Barry Postdoctoral Research Fellow at the James Madison Program at Princeton University. He received a PhD in philosophy from the University of Notre Dame.

“Gives a broad overview of a range of topics, identifies lots of interesting new ideas, questions, and avenues for further research, and contains a wide array of helpful references to follow up. . . . I think this would work well for philosophy majors but also for PPE students or even philosophy courses for political science majors, communication science majors, or public administration majors. Parts of it could also be a great addition to political philosophy classes.”

– **Jeroen de Ridder**, *Vrije Universiteit Amsterdam*

POLITICAL BELIEFS

A Philosophical Introduction

Oliver Traldi

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

Designed cover image: Miguel Á. Padriñán/Pexels

First published 2024
by Routledge
605 Third Avenue, New York, NY 10158

and by Routledge
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

Library of Congress Cataloging-in-Publication Data
Names: Traldi, Oliver, author.
Title: Political beliefs : a philosophical introduction / Oliver Traldi.
Description: New York, NY : Routledge, 2024. | Includes bibliographical references and index.
Identifiers: LCCN 2023055644 (print) | LCCN 2023055645 (ebook) | ISBN 9781032409122 (hardback) | ISBN 9781032409108 (paperback) | ISBN 9781003355274 (ebook)
Subjects: LCSH: Political psychology. | Political culture. | Political participation. | Political science—Philosophy.
Classification: LCC JA74.5 .T73 2024 (print) | LCC JA74.5 (ebook) | DDC 320.01—dc23/eng/20240304
LC record available at https://lccn.loc.gov/2023055644
LC ebook record available at https://lccn.loc.gov/2023055645

ISBN: 978-1-032-40912-2 (hbk)
ISBN: 978-1-032-40910-8 (pbk)
ISBN: 978-1-003-35527-4 (ebk)

DOI: 10.4324/9781003355274

Typeset in Times New Roman
by Apex CoVantage, LLC

In memory of Jeffrey Friedman

CONTENTS

PREFACE

Two friends – maybe it's easier to imagine them as lovers – run laughing through the hills and valleys of a great green land. They play a game: when they reach a pair of hills, each stands on one and explains why the other should come to their side. Eventually one is convinced, or playacts as though they are, and they join hands and run again together under a clear sky.

They make their way to a certain pair of hills and take their places as usual. But as they each start to speak to the other, a wind picks up, carrying their words off with it. They try to shout, but the wind only grows louder. Suddenly on each hilltop materializes a huge collection of other people, one surrounding one friend, another surrounding the other. A rope appears, standing taut in midair between the two hills, each group of people pulling as hard as they can from their side.

The friends are left with a dilemma. They can continue their game in a new form, participating in the tug of war instead of the mock debate they'd chosen. Or they can abandon their positions on the hills and go off to do something else. What's not open to them is to go on as they had been.

Different features of this story will for different readers be attractive illustrations of the difference between political disputes and other kinds of disputes. For some, political disputes fall short of our desired standards of rationality. They are ultimately just matters of tugging, not of offering one's own reasoning and listening to other people's. For others, political disputes are what happen when a disagreement is no longer a game. Becoming political is just what happens when participants actually start to care about how the dispute is resolved – when there are "stakes" to it. Yet others might emphasize that political disputes can happen "anywhere" in our belief space – a hill of science, a field of morality, a river of

physical force. It's almost arbitrary whether or not a dispute is "politicized." I don't think any of these views is quite right. But I think the story does capture something important.

You've probably heard many people express their political beliefs. Maybe you've expressed some of your own. In particular, you've probably heard many people have political disagreements – disputes with each other, in words, over their political beliefs. Such disputes occur in debates, in dorm rooms, over the dinner table. They occur between competing candidates, between classmates, between comrades. Maybe you're the instigator of such disagreements. Or maybe you join unwillingly, pulled in by the emotional force of the conversation or the certainty that someone else has said something incorrect or even evil. Or maybe you're the one quietly asking: "Why did you have to make it all political?" Whichever role you play, this book is for you.

As you've probably gathered from the title, this book is a philosophical introduction to political beliefs. It's a book in the branch of philosophy we call epistemology. In epistemology, we study belief and associated concepts like knowledge, certainty, and rationality. The way I think about it, epistemology is the study of how to evaluate our beliefs. So this book is an introduction to debates about how to evaluate our political beliefs. If you like, you can think of it as an introduction to disputes about how to resolve our political disputes.

Because I think the notion of a political belief is a kind of moving target, as I'll explain soon, there are a lot of different questions we could have about our political beliefs. But I'm going to limit my scope a bit to make this project manageable both for myself and for you. To form political beliefs in a rational or reliable way probably requires that we do basic things like perceive a real, actually existing world, learn from our experiences, and understand basic arithmetic. Philosophers often worry about skeptical challenges to foundational forms of knowledge like the perceptual, inductive, and mathematical. I won't worry about that in this book, though. I'll be concerned with problems that are relatively characteristic of the political realm. At the same time, I won't get too concrete. This book won't tell you how to form political beliefs about any particular issue, like abortion or climate change.

This middle level of abstraction puts us in an area of philosophy which has a very long pedigree but has also seen a very recent rise in interest: social epistemology. Social epistemology asks how we should form our beliefs, and how we should evaluate the ones we have, given that we live in a world with other people. Should we trust other people? What should we do when they disagree with us? Political epistemology, the topic of this book, is much like social epistemology applied to politics, but it also includes a few extra dimensions. First, we don't just live in a world with other individuals; those individuals form groups, and we ourselves are likely members of certain kinds of groups as well. Just what it means for individuals to form a group is difficult to pin down, but many

political theorists take the idea of a group to be part of the essence of politics. Second, political beliefs are often beliefs about what we ought to do, and those sorts of beliefs, what I'll call moral beliefs or values, introduce their own epistemological puzzles. Third, to the extent that politics is a distinct field of inquiry, it may present challenges of its own, like complexity.

The sort of philosophy I do is focused around theories and arguments. Those are two kinds of structures that put together a bunch of different ideas or claims. Sometimes I'll call an idea or claim a *proposition*. A theory is just a bunch of propositions put together. An argument is a bunch of propositions put together with one picked out as a conclusion.

Classical philosophical puzzles of "global skepticism" target all of the knowledge we might think we have. How can we be certain that we're not dreaming? How can we be certain that we're not being tricked by some evil demon? How can we be certain that we're not brains in vats hooked up to some illusion-generating machine? If any of those skeptical scenarios obtain, it could mean that none of our beliefs are true, and that possibility might suggest that we don't truly know anything at all. But such arguments rely on these odd scenarios being saliently possible, and philosophers have resisted global skepticism by resisting either the claim of possibility or the claim of salience in that regard.

When it comes to politics, however, we might worry that we are actually in the odd scenario, which would justify a local skepticism about political beliefs. There are a lot of potential arguments for political skepticism. Some argue that political cognition is subject to too many biases for us to rationally trust our own judgment when it comes to politics or that political cognition is necessarily subject to distorting ideologies. Others argue that there is so much disinformation in the current media environment that it is irrational to trust anyone else. Still others say that politics in the modern world is necessarily too complex to license rational political beliefs. Finally, some say that the mere prevalence of disagreement about politics should be enough to make us unsure about our own political beliefs. These arguments will be seen throughout the book, but it's good to start thinking about them now.

The skeptical perspective is not the only approach people take to political epistemology. Some start with a set of political beliefs they take to be obviously false and then try to characterize what goes wrong when people believe those things. Still others avoid both political skepticism and political dogmatism. They think that there are reliable ways of getting at the truth in social epistemology – those we'll consider include listening to the experts and deferring to the majority or social consensus – and that there's no reason these shouldn't remain reliable when it comes to politics. More generally, some might wonder whether politics deserves its own epistemology at all; what makes it so special?

The plan for the book is as follows. In the first part of the book, I'll go through some preliminaries: the surprisingly rare question of what counts as *political*, the

surprisingly contentious question of what counts as a *belief*, and then the first effort, to my knowledge, of characterizing what counts as *political belief*. Then I'll talk a bit about the nature of *political conflict*, of which various theories have been propounded, and about ways to interpret *political disagreement*, including the principle of charity, which unfortunately does not, in light of the various possible causes of conflict, always give us clear guidance.

In the second part, I'll go through some major *theories of political beliefs*, with some input from philosophers but just as much from social scientists, especially psychologists. What I mean by a theory of political belief here is a scientific, causal account of what gives different people their different beliefs. Some varieties of theories we'll consider: theories based on *personality type*, theories based on *ideology*, theories based on *group membership*, theories based on *social location*, theories based on *cognitive heuristics*, and *minimalist*, *expressivist*, *and eliminativist* theories, which suggest that people might not have political beliefs at all.

In the third part, I'll go through some sources of evidence or justification for our political beliefs. First will be two types of skeptical arguments for the conclusion that we should doubt or even abandon our political beliefs. A *debunking argument* takes one of the etiologies of political belief and suggests that, if we think our political beliefs were caused in that way, we should abandon them. The argument from *widespread disagreement* suggests that political beliefs are in such controversy that we should abandon many of them. A natural next step is to ask: What about when there is a clear majority on one side or another of a political belief? That's the question of the *epistemology of democracy*. If we don't trust majority votes, we might trust other *decentralized* mechanisms for aggregating the viewpoints of many different people. A central alternative to the epistemology of democracy is the idea that we should *trust the experts*, the subject of the following chapter. We'll consider the possibility that we should determine our political beliefs based on which beliefs would be *morally right or wrong* to hold. Finally, we'll take a close look at specific epistemological issues related to the political theory of *liberalism*.

The fourth part will be similar to the third, except that I'll focus more on real-world political phenomena that have been matters of public debate. I'll talk about *polarization*; "polarization" really is used to name (at least) two different kinds of phenomena, *sorting* and *extremism*. I'll talk about *conspiracy theories* and the difficulty of explaining what's irrational about them without condemning a lot of our normal thinking and theorizing. I'll talk about *propaganda and dehumanization*, tools attributed to powerful people and groups by the ideology theories considered earlier, and about *fake news and political rumors*, which cast doubt on how much of other people's testimony we can believe. I'll talk about political *narratives* and how they affect people's snap judgments about news items. Then I'll close with a relatively underexplored topic, the relationship between

our political beliefs and the *philosophy of history*, especially our narratives about historical progress or historical decline.

Don't get the wrong idea: even in the parts where I'm not providing my own original ideas, my perspective has influenced the way I've written this book, from choices about what to include to the organization of the material to little sidebars and snippets about which arguments I think are good and which arguments I think are bad. So I want to urge you: don't take my word for it. Do your own thinking about everything that comes up in this text; treat it like it's under dispute, just as you would a text about politics itself. To my mind, that's one of the core insights of epistemology, including political epistemology. At some level, at some point, you're on your own. You have to figure it out for yourself.

ACKNOWLEDGMENTS

Many people have helped me with this project, both with the book itself and with the underlying ideas. It would be impossible to list everyone with whom I've discussed the topic of political beliefs, including academics in philosophy, political science, psychology, and economics; policy wonks and political operatives; and "civilians" with no professional link to politics. I wrote much of this book while traveling; so I'll start by thanking the people who have put me up: Mom and Dad, Matthew and Lesley, Rebecca and Johan, Jeff and Tina, Sara and Jan, Larisa, Joey and Annie, Alice, and any number of Airbnb hosts. They have all talked through my ideas with me as well – well, not the Airbnb hosts, but the others.

The three philosophers who have read the most of my work, and especially the most of my work on political beliefs, are my advisor Blake Roeber, Dan Greco, and Liam Bright, who calls himself an expert on "Oliver Traldi thought." Everyone at Notre Dame has been really helpful to me while I've been there, and I want to thank the undergraduates in my fall 2021 class, which was basically a dry run for this book, in particular, as well as David Eil and Béatrice Leydier, who listened to a different version of it on Zoom. I've had great and encouraging conversations and online "exchanges" with spectacular philosophers of political belief like Rima Basu, Spencer Case, Gabriele Contessa, Kevin Dorst, Michael Hannon, Hrishikesh Joshi, Enzo Rossi, Jason Stanley, Kevin Vallier, Brandon Warmke, and Alex Worsnip, as well as non-philosophers who work on similar topics like Musa al-Gharbi, Paul Bloom, Jonathan Haidt, Dmitri Halikias, Robin Hanson, Philip Tetlock, and Adrian Vermeule. Thanks to Adam Gibbons in particular for many talks over the years and for giving an early draft a thorough read. He made many helpful suggestions which were integrated into the text, as

did two anonymous referees at Routledge. Becca Rothfeld and Jane Cooper also read some short early excerpts and gave me helpful feedback.

Andy Beck, my editor at Routledge, has been great through this whole process, and I'd also like to thank several anonymous reviewers for really helpful comments. I made all sorts of additions based on those comments, added helpful sources and qualified questionable claims I'd made, and fleshed out plenty of sections which were sketchier when I'd submitted them. In particular, I spent September of 2023 integrating nearly every suggestion made by an incredibly helpful review of the entire book manuscript. I talked about these issues constantly at the Philosophy, Politics, and Economics Society conferences in New Orleans, and virtually everyone gave me good ideas and feedback. Some of the ideas in the book's final chapter were initially written as a paper for now-Supreme Court Justice Amy Coney Barrett's course on modern constitutional theory in the spring of 2019, which I audited; many thanks to her and to my classmates. The philosopher groups I'm in on Slack and Discord have also helped me relax and formulate my thoughts, and I've gotten pretty much constant feedback from many people on Twitter (now X), much of which is very helpful. I've published versions of some of these ideas in magazines and newspapers, so thanks to all of my editors as well. In particular, this text cites an article of mine in the *Bulwark*, where my editor was Adam Keiper, and a book review of mine from the *Hedgehog Review*, where my editor was B. D. McClay.

Special thanks are due to the Institute for Humane Studies for a very generous grant which helped me focus on this project while other things threatened to distract me from it. This publication was also supported by the Princeton University Library Open Access Fund.

Finally, thanks are very much due to Jeffrey Friedman, whose political epistemology workshop I attended in Berkeley in the summer of 2018. We argued a lot about what good writing is, what good arguments are, whether there's any use to analytic philosophy, whether historicism is mere pedantry, and so on. He passed away in December 2022 and will be missed for his kindness, his scholarship, and his incisive thinking. Though I don't know if he would be any more fond of my approach now than he was then, his ideas are all over this book.

PART I

Preliminaries

1

WHAT IS POLITICAL?

Just which beliefs are political? In these first few chapters, we'll start by thinking about what it means for something to be political and go on to thinking about what it means for something to be a belief. Hopefully from those two inquiries we should be able to say at least a little bit about what it might mean for something to be a political belief, a topic about which very little has to date been written. I'll offer my own theory by the end of the third chapter, but like most theories in philosophy, it's probably wrong. I hope that, rather than internalizing or memorizing it, the reader will apply the same scrutiny to my theories as I do to other people's throughout this text.

The notion we want to define in these first few chapters is "political belief." In the first chapter, we'll ask which things are political, and in the second chapter, we'll ask which things are beliefs. In the third chapter, I'll give my own view of which beliefs are political. I suspect you'll be able to find ways in which it is overinclusive and underinclusive. The quest for real definitions rarely finds treasure.

Power, conflict, and order

In an unpublished paper, political scientist T. J. Donahue surveys fifteen different analyses of the concept "politics," finding them all ultimately wanting – that is to say, either overinclusive or underinclusive – and then offering his own. Donahue writes[1] that it is "[a]n oddity of political philosophy . . . that it spends so little time answering this question" – that is, the question "what is politics?" This is no doubt correct, and it is true of fields other than political philosophy. For instance, the Routledge textbook *Introduction to Political Psychology*, in an introductory

DOI: 10.4324/9781003355274-2

section called "What is Political Psychology?",[2] does not undertake any sort of outlining of what counts as politics, or "political behavior," the object of study of political psychology, the textbook's topic. This is especially odd because politics itself includes a great deal of debate about what politics consists of.

Rather than covering all the definitions Donahue considers, I think it's better to note that they fall roughly into three groups. To some theorists,[i] the dividing line between the political and the nonpolitical has something to do with *power*. To other theorists,[ii] this dividing line has something to do with *conflict*. And to yet more, the line has to do with *setting things in order*.[iii] Donahue offers quick counterexamples to all the theories he surveys, and quick counterexamples to his own theory are also available. But we should try to say something more general about each of these approaches.

Let's start with a note about all of the approaches together. A characteristic situation that can be unobjectionably called "political" might involve two or more different "sides" jockeying against each other in one way or another to try to make sure that things are done in a certain way in society. Note that this situation has elements of all three theses: the power, the conflict, and the eventual setting down of some sort of social order, whether by legislation or by some other means. One method for figuring out which definition seems best to you might be to start trying to take away pieces of this situation and see if it still seems "political" to you. For instance, you might think that if no conflict existed, it would be hard to understand what the jockeying for power even involved.

Is "political" really just one thing? Maybe we use the word in multiple senses. Or maybe it doesn't mean anything: maybe it's just something we throw around to try to affect how others see things. Or "political" could be what Ludwig Wittgenstein[3] called a *family resemblance concept* or *cluster concept*. Whether or not a cluster concept is properly applied in some case depends on a variety of factors, and it is possible that none of them will be either sufficient or necessary. The full range of instances of such concepts thus exhibits only a "family resemblance" among its members. An interesting feature of such concepts is that they can come in degrees based on how many factors in the cluster obtain in some specific case. If some activity or belief can be more easily said to be *more* political or *less* political than some other activity or belief rather than being said to be political or not political full stop, then that might be good reason to think that "political" is a cluster concept.

i Donahue cites Raymond Geuss, Adam Swift, Peter Nicholson, and Max Weber.

ii Donahue cites David Miller, John Dunn, Bernard Crick, Carl Schmitt, Mark Warren, and J. D. B. Miller.

iii Donahue cites Michael Oakeshott, Chantal Mouffe, and Bertrand de Jouvenel. His own theory also fits here. (Note that some of the theories he considers could arguably fall into more than one of my categories.)

On to power. The basic problem with theories that define politics in terms of power is that "power" itself is a poorly defined term. If we take power to be any sort of influence, then anyone convincing anyone else to do anything at all would count as political; Donahue gives the example of one person convincing another to give them a ride when the first person's car has run out of gas.[4] But if we narrow the sort of power we're talking about, we end up excluding from the scope of "political" activities that should definitely count as politics. For instance, a definition of politics in terms of *state* power ignores all the sorts of political activity that might occur in the absence of formally organized states;[5] a definition of politics in terms of *force* ignores all the sorts of political activity that might involve attempts at rational persuasion.[6]

Now consider a possible way of protecting theses about power. A power theorist might say: "You're right that the term 'power' is too general. What we really mean is *political* power." This response would neglect the task at hand. The task at hand is precisely to say what it means for something to be political. To qualify a term in a proposed analysis of politics by restricting it to political phenomena is to propose no analysis at all, for the same reason that we are taught in school not to include a word in its own definition: it relies on what it provides prior to the provision.

There is a more concrete problem with the idea of "power," too. We can distinguish between the power *over* others and the power *to* perform certain kinds of actions.[7] An analysis of politics that includes all instances of both types of power is clearly overinclusive. We gain the power to do new things when, for instance, we learn to play an instrument or to speak a language, but those are generally not political activities. However, to entirely exclude power-to from the analysis of politics would seem underinclusive. The fact that society is, at least according to some, set up in a way that determines that members of certain (privileged) groups have a wider scope of action than members of other (disprivileged) groups is arguably a central case of a political fact. Even including all instances of power-over in the analysis of politics might be a mistake. People can gain power over other people by being beautiful, charming, or talented. But it seems clearly overinclusive to think that all performances of beauty, charm, and talent are political activities.

Arguments against defining politics in terms of conflict[8] are similar in form to the arguments against defining politics in terms of power. There are plenty of conflicts that aren't political (say, a physical fight between two people who both want the last piece of a delicious cake). Are there political situations that don't involve conflict? Say an emergency situation arises in some country. Everyone agrees on what the government should do about it, and effecting this outcome requires that the legislature pass a certain bill, which indeed does pass unanimously. Was the event of that bill's passage not a political event? The possibility of bipartisan political action might mean that defining politics in terms of conflict

is underinclusive. However, we might still save the definition if bipartisan political action is political only when undertaken in institutions which are ordinarily conflictual.

I have a little argument which favors seeing politics as a matter of conflict. When people append the word "politics" to the name of some other activity, as in the phrases "academic politics" and "office politics," they are referring to an aspect of that activity that involves some sort of conflict. So it is natural to think that the "political" side of an activity will be the conflictual one. Though I think it's cute, this argument should not move the needle very much. This use of phrases like "office politics" would be comprehensible even if conflict were merely a frequent feature of politics rather than an essential one. And "politics" could indicate something more general than conflict, like unpleasantness.

Donahue argues against most theories of politics as a matter of setting things in order because they exclude cases of purely negative political expression – political actions which seek only to break a group's current arrangements, radicals setting themselves on fire to protest the passage of a statute, and so forth.[9] This includes theories of politics that essentially involve concepts like governing and ruling. To my mind, these theories, including Donahue's, also suffer from a problem we saw that plagued theories of politics as power: just as there are intuitively political and nonpolitical instances of power, there are intuitively political and nonpolitical instances of order. Take a case in which four friends have decided to play a board game together every few weeks and must pick a time and a place for those regular meetings. These friends will set their affairs in order as a group and make general arrangements which are relevant to their goals. (They might even overcome some conflict.) However, it doesn't seem natural to me to say that the friends are doing politics. There are aspects of the activity that kind of seem political, but something seems to be missing, too. Readers, of course, might disagree.

A natural thought is that we might be able to find a definition of "politics" or "political" that includes power, conflict, *and* order. But I think there are still obviously conflicts over power which involve setting affairs in order which aren't political. For instance, imagine that two friends each want to be the one to choose where the group goes to dinner tonight, and one of the two friends very nastily unplugs the other friend's phone, so that they can't call a restaurant to make a reservation, leaving the first friend with the power to do so. This involves power (the power to make the reservation), conflict (the conflict over where to eat), and setting things in order (deciding where the group will eat), but I think it's intuitively not a political event or a political struggle. Just as before, we want to say: this isn't a matter of *political* power; this isn't a matter of *political* conflict; this isn't setting things in order in a *political* sense. That this response is natural indicates that there's some other sense of "political" lurking behind all these definitions.

Is everything political?

In an introduction to the volume *What is Politics?* Adrian Leftwich writes of politics that "debates about its proper definition and the scope of its subject matter are themselves political."[10] But it's very hard to make sense of this sentence if "political" is left undefined. Indeed, that it may be a "political" matter what counts as "political" leads some into odd inferences. After all, though we may all agree that it is a political matter, we may not know what that entails until we *resolve* that political matter: until we figure out what "political" means. Some extremists about the nature of politics may take the view that there is no truth in political matters – that they are inevitably resolved by force, for instance, and that there is no ultimate "fact of the matter" where they're concerned. But then there is no genuine inference to be made from the fact that it is a political matter what counts as political, since there is no ultimate "fact of the matter" about what inferences can be made from something counting as political. Even worse, some people argue from the premise that it's a political matter what counts as political to the conclusion that *everything* is political. This recently popular catchphrase is rationally unsupportable, and indeed if it were true it would be hard to make sense of sentences including words like "politics" and "political" or to understand phrases like "political debate" or "political science."

Bad arguments are sometimes worth careful study; seeing how they went wrong can help us do better when we come up with our own arguments. One sometimes sees other bad arguments for the conclusion that everything is political.

A first such argument starts from the observation that whenever we spend our time doing something other than politics (or that seems to be other than politics), or whenever we introduce concerns into our decision-making that aren't political (or that seem not to be political), we are giving political considerations less weight than we might, which itself is a political choice. Thus, the argument seems to go, *every* choice is a political choice, insofar as every choice involves making some sort of decision about how to weigh political considerations. This argument can be defeated by an analogous parody. Consider the idea that the choice not to engage in musical activity is a musical choice. After all, it involves weighing musical goals against other goals and deciding that they are not important enough to guide a choice. By this logic, we could argue that every *token* choice is in part every *type of* choice: that everything we do is political, musical, economical, romantic, gustatory, artistic, recreational, medical, and so on all at once. This is obviously an implausible conclusion. Perhaps there is something different about political choices which makes this kind of premise more plausible in the political case than in other cases. But in fact none of the analyses of the concept of politics we examined led to such a conclusion. The argument for an overly broad conception of politics assumes an overly broad conception of politics to begin with.

Another such argument starts from the observation that many individuals, objects, and situations have been affected by politics in some way, at some point in what philosophers might call their "causal history." Politics might be part of the historical explanation of why your clothes are made of a certain material, why your meal is seasoned with a certain spice, why certain courses are offered at your college, and so on. Such historical facts can of course be very interesting. But first, not *every* object has been affected by politics: take celestial objects, for instance. And second, it is not in general the case that every part of an object's causal history inheres in that object. A novelist having written a book while listening to classical music would not make that book a piece of music, let alone a piece of classical music; and this is the case whether or not the piece of music caused the novelist to write the book. Objects are forgetful in this way, and perhaps this justifies a certain amount of envy toward them. Again, there could be something unique about politics such that an object's having a causal history bound up with politics justifies calling that object "political," but it's hard to see what that would be.

Inferring political conclusions from the analysis of politics

Occasionally, people talk as though we can infer conclusions about what political actors *ought to do* from an analysis of politics. In particular, commentators who adhere to the slogan "politics is about power" will often say that this means that political actors *ought to* focus on accumulating and deploying political power rather than thinking about what's right or wrong or trying not to fall afoul of political norms.[11] Apart from the problems with the power-based analysis of politics that we saw earlier, this sort of conclusion can never really be justified from a premise about what politics consists in. For the mere fact that politics is "about" some kind of activity or goal cannot explain why we *ought to* engage in that activity or work toward that goal. There is a big academic debate about this which concerns the thesis of "political realism," which has it, roughly, that moral concerns are out of place in politics. In a recent article,[12] Jonathan Leader Maynard and Alex Worsnip consider five arguments in favor of this kind of conclusion, finding all of them wanting. To me, the biggest problem with this sort of idea is that it undermines the very political appeals its supporters tend to make. If you tell me that politics isn't concerned with what we ought to do, then I'll respond: "So much the worse for politics. Let's not do it anymore." Charitably, I think what's actually at work a lot of the time when people say "politics is about power" is not necessarily a denial that political actors should do what's right but an attempt to emphasize that to do what's right sometimes requires background conditions, like the possession of political power. This is obviously true. If you want to get elected or pass a bill, you need the votes. If you want to prevent criminal organizations from terrorizing a neighborhood or prevent evil dictators

from brutalizing their subjects, you need a police force or an army. So there's a kind of soft reading of the notion that politics is about power rather than morality which interprets it to mean instead that being morally right isn't enough and that one must plan carefully and gather resources to be able to execute one's plans, including moral ones. But we'll see more about this kind of idea in the chapter on political conflict.

Conclusion

We haven't come to much of a conclusion about what politics is. Power, conflict, and order all seem to have something to do with politics, without there being an easy way to make them individually or jointly into a real definition of politics. But I think we may see a way to make use of what we have done. For now, though, let's go on to consider what beliefs might be.

Discussion questions

1. Say two people are arguing over whether something is political. What do you take them to be disagreeing about? See if you can state the disagreement without using words like "political" or "politics." Just what is the importance of the category of "political"?
2. What are some political activities you engage in or that the people around you engage in? What sets them apart as political?
3. In 2005, former President George H. W. Bush said of helping raise money for recovery efforts after a tsunami in Japan: "This is bigger than politics. This is about saving lives." In 2017, the then-House Minority Leader Nancy Pelosi was quoted as saying of the 2016 presidential election: "The integrity of our elections, this is bigger than politics, bigger than Democrats and Republicans. This is about our country." In 2019, describing a vote to impeach Donald Trump, Representative Elissa Slotkin was quoted as saying: "This is bigger than politics." In 2021, an adviser for the campaign to recall the then-Governor of California Gavin Newsom was quoted as saying about the recall: "This is not a Republican recall; this is a group of concerned citizens. This is bigger than politics."[13] What do you think these people mean when they say something is bigger than politics? What do these instances make you think about the word "political"?

2

WHAT ARE BELIEFS?

We've thought about what politics might be, so our next step in figuring out what political beliefs might be is to figure out what beliefs might be. Since philosophers use the word "belief" in a specialized way, it is important to clear up a few potential misunderstandings first. When nonphilosophers use words like "belief" and "opinion," they sometimes mean to contrast them with words like "truth" or "fact," with the intended contrast being that beliefs and opinions are uncertain, sort of up for grabs, whereas truths and facts are certain. This is not the contrast that philosophers usually draw between these ideas, though. Philosophers generally use words like "belief" and "opinion" to refer to things that, so to speak, have bearers, what we'll sometimes call *epistemic agents*. These epistemic agents are individuals who *hold* those beliefs and opinions – who, in verb form, *believe* or *opine* in the relevant way. Truths and facts, on the other hand, don't have bearers. They are part of the world. However, the way philosophers use these terms, an epistemic agent can believe a truth, even a completely certain one, just as an epistemic agent can believe an uncertain truth or a falsehood. A retort like "that's not my belief; it's a fact" thus doesn't make much sense the way we use these words, although I think it is sensible in everyday language.

Epistemologists sometimes prefer to talk not about beliefs but about closely linked attitudes called *credences*. Roughly, a credence is something like a level of confidence in a proposition. Thinking there's a 60 percent chance that it will rain is something like having a credence of .6 in the proposition that it will rain. The relationship between belief and credence is contested (see Jackson 2020 for some of the leading theories and arguments about those theories). Some epistemologists think credences are just beliefs about likelihoods while others think beliefs are just credences that meet a certain threshold.[1] Some epistemologists

DOI: 10.4324/9781003355274-3

eschew talk of credences while others eschew talk about beliefs.[2] I would hate to take a stand on those sorts of deep theoretical issues in a text of this nature. Instead, I'll sometimes talk about credences but usually talk about beliefs, and the reader should generally take my intent to be that comments about one kind of attitude will be translatable into comments about the other attitude in a straightforward way.

Two theories of belief (or credence)

So what are beliefs (or credences, if we prefer)? Two kinds of views are popular.[3] One *representationalist* view has it that beliefs are kinds of pictures or markers in the mind. In my head somewhere is the belief that Paris is the capital of France; it's part, in some sense, of my inner life, with a relationship to my outer life that might be rather complicated. Another *dispositionalist* view has it that the relationship is not so complicated as all that; according to this view, beliefs just are dispositions to behave in a certain way, so that my belief that Paris is the capital of France is just the disposition to say that Paris is the capital of France, to act as though Paris is the capital of France, to be surprised when people say that other cities might be the capital of France, and so on. Dispositionalism fit well within a period in intellectual history during which philosophers were very reluctant to posit mental entities and had ambitions to reduce our typology of the mental to fully observable physical phenomena. It also seems to explain some of our habits of attributing beliefs to others based on actions.

But there is a big problem for at least naive theories of dispositionalism, which we can see if we think a bit more about how we attribute beliefs based on actions. Philosophers have often thought that belief has a special relationship to action, especially action that is *intentional*. Joseph Raz wrote that "[a]cting with an intention or a purpose is acting (as things appear to one) for a reason,"[4] and Donald Davidson wrote that

> [w]henever someone does something for a reason, . . . he can be characterized as (a) having some sort of pro attitude toward actions of a certain kind, and (b) believing (or knowing, perceiving, noticing, remembering) that his action is of that kind.[5]

Similarly, Robert Audi writes that "actions [are] doings that have a description under which they are intentional (this holds with at most a few exceptions)" and "intentional action is explainable by appeal to a set of beliefs and desires."[6] Philosophers call *theoretical reasoning* the process of coming to conclusions about what it is rational to *believe*. Philosophers call *practical reasoning* the process of coming to conclusions about what it is rational to *do*. A rational process of coming to conclusions begins with *premises*, which form the rational basis for

the conclusions. But for me to reason from a premise to some conclusion, and for me to go on to accept the conclusion as rationally compelling when it comes to the question of what I ought to believe or do, it seems that I must believe that premise.

If this sounds a bit technical, a few examples should simplify it. If I desire a chocolate bar, and I believe that I can get a chocolate bar by checking the candy bowl in the department office, then that explains and justifies as rational my action to check the candy bowl in the department office. What if I didn't desire a chocolate bar or didn't believe that there would be one in the candy bowl? Then the action wouldn't make so much sense. Now, in some cases, I'll have multiple desires, and they might conflict. I might want to go to the party because someone with whom I am enamored will be present; I might want to stay at home because it is sleeting outside. Thus I have reasons to do two incompatible things. Philosophers sometimes say these are *pro tanto* reasons, or reasons to a certain extent, and that the conclusions they can provide about what is rational for me to do are *ceteris paribus* conclusions, or conclusions about what is rational "all else being equal."

The multiplicity of possible desires and the fact that we never have just one belief at a time are problematic for dispositionalism. If I believe that you are innocent, then I might be disposed to speak out in your favor to help exonerate you; then again, my desire might be to spite you, so that I will keep quiet about my belief. If I don't like the rain, then I might be disposed to get an umbrella; but if I do, then I might be disposed to act as though it isn't raining at all. And adding another belief can remove the apparent disposition to act from a first belief. If I believe (erroneously, of course) that chocolate is poisonous, I might not be disposed to check for it in the department candy bowl, but I might still believe that it's there. In addition, not all beliefs eventuate in actions; Eric Schwitzgebel gives the example of "an American homebody's belief that there is at least one church in Nice."[7] On the other hand, we will see that people may be at least a bit less reliable when it comes to such inactive beliefs. Dispositionalism also seems to be less intuitive than representationalism when it comes to the sort of project we're undertaking in this book: evaluating our belief-forming processes normatively with regard to how rational they are.[8] The idea of representation seems to have built into it a correctness condition: a representation can be faithful or unfaithful, accurate or inaccurate. This explains something of our sense that beliefs go wrong when they are false – that they "aim at truth," in a philosophical slogan. This connection between belief and truth, mediated by normative notions like rationality, justification, and knowledge, is at the core of epistemology, and so we might think that our theory of belief ought to give a good account of it.

We might also think that although dispositionalism connects belief to action, it somehow reverses the connection in doing so by defining belief in terms of action rather than vice versa. Our experience of belief and action is that we are

often spurred to action by what we believe. The belief seems to come first and the action later. As Galen Strawson has written, the view that "I don't act on some information because I believe it to be true. I believe it to be true because I act on it, or am disposed to" might seem to us to be "the wrong way round."[9] Bad news for dispositionalism if so. (We will consider the question of whether believing itself is an action, and what sorts of actions can eventuate in beliefs, in a later chapter on the ethics of belief.)

Two kinds of belief (or credence)

Throughout the rest of the book, it'll be pretty important to separate two kinds of belief: moral and, well, otherwise. It's hard to find a coinage that captures nonmoral beliefs in a way that doesn't take a side in the debate over whether there are "moral facts," but I'll sometimes say "factual beliefs" to mean nonmoral beliefs. Moral beliefs are beliefs about moral propositions, and moral propositions are propositions about what's good or bad, or what's right or wrong; they plausibly also include "thick" claims that include both moral and nonmoral elements, like claims about what's cruel, what's unjust, and what's oppressive, but also what's "lewd, rude, and nasty," as put by the clever title of a recent book on the subject.[10] Moral statements, which express moral propositions, are things like "It's wrong to murder," "It's wrong to eat meat," "It's good to maintain your friendships," "It's unjust to deny victims a voice," and so on. Nonmoral statements are anything else: "It's raining outside," and so on. Moral beliefs are a subclass of normative beliefs, which might include beliefs about what is in some person's best interest ("You ought to take the second job rather than the first; it pays more"), what is one's responsibility as a result of some role they occupy (say: parent, teacher, or doctor), or what one ought to do in a context with some obvious nonmoral goals (like in chess: "White ought to capture the rook"). What we say about moral beliefs will largely generalize to other normative beliefs.

Moral beliefs are a bit like desires in that they can be combined with nonmoral beliefs to produce justifying explanations of our actions; a moral belief is another kind of "pro attitude," in the words of Davidson above. If I believe that it's right to prevent suffering, and I also believe that feeding the stray cat will prevent suffering, then those two beliefs together fully explain, and explain as a rational action, why I am choosing to feed the cat. That said, most philosophers think that we don't always act in line with our moral beliefs. When we fail to, it's called moral incontinence, weakness of will, or *akrasia*, from the ancient Greek.

Famously, David Hume argued that we cannot come to conclusions about what we ought to do merely from claims about what there is, without premises about what we ought to do; this is called the *is-ought gap*. Because of the is-ought gap, moral beliefs face some epistemological challenges that might not be present for other kinds of beliefs. Just what sorts of facts would moral facts

be? Where are they? How can we test them? Why do moral beliefs vary so much from time period to time period or culture to culture, and why is there so much disagreement even within cultures and times about what's right and wrong? These sorts of questions cause some people to think that, even if there are facts about such matters as right and wrong, good and evil, justice and injustice, and so on, we might not be able to *know* such facts. For our purposes, what's important is that this kind of skepticism would also entail skepticism about some, but not all, political beliefs.

Hilary Putnam gave an argument against the distinction between moral and nonmoral propositions, or between values and facts, from the existence of thick concepts (discussed earlier).[11] Putnam thought that the moral and nonmoral parts of thick concepts (like *lewd*, *rude*, and *nasty*) couldn't be disentangled from one another and that this meant that there wasn't a real, final distinction between facts and values. But this anti-disentanglement thesis, sometimes called *inseparability*, is controversial; I don't see why, with enough work, we couldn't figure out just what the moral and nonmoral aspects of thick concepts are. (Consider the following claims: "Downloading music is stealing," "Downloading music isn't stealing," "Downloading music is stealing, but some stealing is morally acceptable," and "Downloading music isn't stealing, but it's morally unacceptable anyway." Can't we always remove the evaluative element from a supposedly thick concept?) And even if these parts can't be disentangled, it's hard to see why this would mean that there's no distinction between facts and values. To be honest, this wasn't one of Putnam's best arguments.

The role of moral beliefs in politics is also disputed. We'll talk about this a bit more soon, but there are different camps when it comes to the question of just how moral beliefs relate to political beliefs and how political disputes relate to morality. On one extreme side are those who think that every political dispute is about morality, and on the other extreme are those, the so-called political realists, who think that moral concerns have little weight in politics.

Conclusion

I'm disposed to come down on the side of representationalism, to think that there is a distinction between moral and nonmoral beliefs, to think that there are moral facts, and to think that moral beliefs matter for our political beliefs. But some disagree about these things. Regardless, the connection between belief and action, which is so important to dispositionalism, is the major point that we'll take forward into forming our account of political belief.

Discussion questions

1. Think about your day so far, and the actions, whether ordinary or extraordinary, you've taken. What sorts of beliefs and desires would an observer infer from those actions? Would any of those inferences fail to capture your real beliefs and desires?
2. Think about some of the choices you or other people have made recently based on a combination of your moral beliefs and your nonmoral beliefs. How would you separate out the moral from the nonmoral considerations? Are there ways in which one could easily change that would have made the choice wrong or irrational?
3. How did you come by your moral beliefs? And how did you come by your nonmoral beliefs? Are you more confident in one than the other? What sorts of events, if any, have cast your moral beliefs into doubt?

3

WHAT ARE POLITICAL BELIEFS?

In philosophy it is common to motivate an analytical project by showing that a "naive" view goes wrong, and that's what I'll do first. Consider the following *naive theory*: that a belief is political if and only if it is a belief about politics. This naive theory is both overinclusive and underinclusive. It is overinclusive because it implies that my belief that the American House of Representatives is a political body is a political belief. This is a belief about politics, but it is not a political belief. It is underinclusive because it implies that my belief that global warming is an anthropogenic phenomenon is not a political belief. This is not a belief about politics, but it is a political belief.

Seeing where the naive theory went wrong can help us in constructing our own theory. Intuitively, the reason – or at least *one* reason – why my belief that the American House of Representatives is a political body is not a political belief is that it is not a part of any sort of dispute. We might expect, then, that at least in general, political beliefs will be in some sort of dispute: that there is some threshold number of people who hold the opposite belief – that is, if my belief that some proposition p is true is a political belief, then it must be the case that a sufficient number of people believe instead that p is false. Call this claim the *dispute-necessity thesis*:

> *Dispute-Necessity Thesis*: If an epistemic agent's belief in the truth of some proposition p is political, then the truth of p must be in dispute, meaning (something along the lines of) that the number of people who believe that p is false passes a certain threshold.

DOI: 10.4324/9781003355274-4

Note two things about the dispute-necessity thesis. First, it matches up well with some of our theories of the political, which stated that politics was essentially about conflict (but keep in mind that we considered some potential objections to those theories). Second, it is not the same as the following thesis:

> *Dispute-Sufficiency Thesis*: If the truth of some proposition is in dispute, then an epistemic agent's belief in that proposition must be a political belief.

This thesis is in fact false. It is overinclusive: many people disagree about which sports teams are the best, or about which soft drinks are the best, or about which sonatas are the best, but our beliefs about which of such things are the best in their class are not political beliefs. There is another limitation on what sorts of disputes are relevant to whether something is political. It seems to me that *current* disputes are the important ones. If there was some disputed event long in the past that occasioned a war, say, historians might still dispute that event without their beliefs being political; the historians' dispute would likely seem like a scholarly one instead.

One argument we considered against the conflict theory of politics in the first chapter came from bipartisan political action. For instance, a legislative body might pass a law unanimously without any dispute whatsoever, and this still seems like a political action. However, I don't think the fact that this could be a political action means that any beliefs about it must be political beliefs, unless we think that any beliefs about what a political body ought to do are necessarily themselves political beliefs.

Why, in turn, do the beliefs about the sports teams, soft drinks, sonatas, and scholarly puzzles fail to be political? We have ruled out the possibility that they fail to be political because they aren't *about* politics. However, I think it is correct to say that these beliefs fail to be political because they aren't sufficiently *connected to* politics. First, let's set this idea out and name it:

> *Connection-Necessity Thesis*: If an epistemic agent's belief in the truth of some proposition *p* is political, then *p* must have a certain kind of connection to politics.

When it comes to the sports teams, soft drinks, and sonatas, both the beliefs themselves and the disputes regarding them lack the kind of relationship to politics that would render them political.

What kind of relationship is that? We'll think about the other theories of politics here. Some of them were about making decisions collectively or setting things in order. And some specific ones which we didn't look at in detail have

something to do with creating groups which are set against each other. In particular, Donahue attributes to Carl Schmitt what he calls *the friend-enemy thesis*, that politics "is the production of friend-enemy groupings."[1] This has never seemed like an analysis of all politics to me, but it does seem like it captures something about politics and about political beliefs in particular. Some of our beliefs are political not because they affect what we ought to do but because they mark the splits among groups. In an age of high polarization, in which everything from sports fandom to coffeeshop consumption seems to be split on party lines, this seems to be even more the case (see later chapter on polarization). This suggests two potential sufficient conditions for the sort of connection we're looking for:

> *Action-Connection Thesis*: It is sufficient for an agent's belief having the right sort of connection to politics, as required by the Connection-Necessity Thesis, that that belief figure in the right way in practical reasoning about specifically political action.

And:

> *Aggregate-Connection Thesis*: It is sufficient for an agent's belief having the right sort of connection to politics, as required by the Connection-Necessity Thesis, that that agent's having that belief be an important factor in that agent's inclusion in or exclusion from at least one political aggregate or group.

Note that it is *not* illicit to describe actions or aggregates as "political" in these definitions, because it is no longer the term "political" we are looking to analyze but specifically the notion of a political belief.

These theses leave some questions open. For instance: must an epistemic agent be *aware* of their belief's connection to politics for it to count as a political belief? To deny this might lead to some strange consequences. For instance, it might mean that some belief of mine could go from a political belief to a nonpolitical belief, or vice versa, without anything in my own head changing. However, this result will not look *too* strange to contemporary philosophers. Many important (though, as always, controversial) claims in contemporary philosophy concern different kinds of *externalism* about beliefs. One claim, which is pretty widely accepted and is usually called *semantic externalism*, has it that the *content* of our beliefs doesn't depend only on what's in our own heads. Another, which is also accepted by many philosophers (though not by me) and which we might call *externalism about justification*, has it that whether or not our beliefs are *justified* doesn't depend only on what's in our own heads. "Externalism" about politics doesn't seem so radical in light of these.

Another issue is more worrying. There seem to be ways that a belief might be disputed and connected to practical reasoning about political action without

being a political belief. Say that it is *not* in dispute that we ought to take some collective political action. Say that some people, call them Group Q, believe that the right basis for taking that action can be represented by the proposition *q*, and say that other people, call them Group R, believe that the right basis for taking that action can be represented by the proposition *r*. Say further that people in Group Q believe that *q* is true and *r* is false and that people in Group R believe that *q* is false and *r* is true. Then under the theses we've provided thus far, *q* and *r* are political beliefs for the people in Group Q and Group R. However, this might seem like the wrong result. Given that members of both groups are committed to the same political action, just which syllogism of practical reasoning is theoretically appropriate to justify that action might seem like a philosophical question and not at all a political one. (*q* and *r might* justify *other* different political actions, but there's no guarantee of this.) So maybe we should modify the Action-Connection Thesis:

> *Action-Connection Thesis (Improved)*: It is sufficient for an agent's belief having the right sort of connection to politics, as required by the Connection-Necessity Thesis, that that belief figure in the right way in practical reasoning about specifically political action, provided the conclusion of that reasoning is itself a matter of political dispute.

We also might wonder whether it makes sense to focus on political *beliefs* rather than political *propositions*. Wouldn't it be simpler to say that some propositions (statements, ideas, etc.) are political and then say that political beliefs are beliefs regarding the truth or falsity of those statements? It would, but philosophers generally think of propositions as abstract objects, which don't undergo change from one time to another. However, as we will soon discuss, beliefs can go from being political to being nonpolitical or vice versa. An epistemic agent's belief that a candidate for local office is corrupt is a political belief. Two hundred years later, an archivist's belief that that same candidate was corrupt will likely have become a matter of mere historical curiosity. Say that a political group in the archivist's time begins taking up this candidate as some sort of figurehead. Then the archivist's belief may well become political – it may well gain the right sort of connection to inclusion and exclusion in political aggregates, for instance.

So my theory is something like the following: that a belief is political requires that it be in dispute and that it have the right kind of connection to politics; two ways of having the right kind of connection to politics include figuring in the right way in practical reasoning about specifically political action and being an important factor in an agent's inclusion or exclusion from at least one political aggregate or group. I don't mean to rule out that there are other ways that beliefs can have the right kind of connection to politics; I just haven't thought of any compelling ones. But regardless, I don't think my account will be the final

account of political belief. There are still a few things I'm worried about in it, which you'll see later in the discussion questions.

What is politicization?

One immediate application of our theory of political beliefs is to generate a theory of politicization. A belief that is disputed can be politicized by gaining the right kind of connection to politics; a belief that has the right kind of connection to politics can be politicized by becoming a matter of dispute; and a belief that meets neither of those conditions can be politicized by gaining both. This theory, however, does not quite explain why "politicization" carries quite the negative connotation it sometimes has. I suspect that "politicization" as a negative epithet is more likely to refer to those beliefs that are political in virtue of meeting the aggregate-connection condition rather than the action-connection condition.

Of course, there is some paradox in some uses of "politicization" and even of "political" – especially when those using it are politicians. Politicians sometimes say things like "Let's not make this a political issue" when they feel an issue is very serious. But they say this precisely to accomplish goals like developing a public consensus and passing legislation, which are of course political activities. In the spirit of linking such phraseology to the aggregate-connection condition, I think what such politicians mean might be something like "Let's not make this a *partisan* issue." Alternatively, we might interpret them as meaning that nothing in the issue really ought to be under dispute, which is more or less equivalent to saying that their own position is obviously correct.

I think when we say that other sorts of things, like works of art or music, have been politicized, we sometimes mean something similar: that our attitude toward them has become, in an undesirable way, a criterion for inclusion or exclusion in some political groupings. However, in such cases, we also sometimes mean something more related to the action-connection condition. We might simply think that it is somehow improper to be constructing practical syllogisms whose conclusions concern certain works of art and music at all. These things, we might think, should be outside the scope of politics.

This is something similar to what Julia Driver says about the vice of *moralism*, which is "the illicit introduction of moral considerations" and which involves "seek[ing] to convert others, or apply pressure" on them to change their behaviors.[2] Early on in this chapter we showed what went wrong with some popular arguments for the conclusion that everything is political. Now we may have a name for the bad habit of imposing political considerations on everything or seeking to *make* everything political: *politicism*. However, that politicism is a bad habit does not mean that it will not have many successes. Those who want to, it seems to me, can always find ways to make the world a more political place by increasing the amount of conflict and making membership in political groups

more salient. To me that also seems to make the world a worse place. But that is a case I'll have to make elsewhere.

Discussion questions

1. One consequence of my account of political beliefs is that all fundamental ethical disputes, that is, debates about what's ultimately valuable, are political disputes, since they bear on our collective practical reasoning. Does that seem true or false to you? Can we do ethics without doing politics?
2. My account holds that political beliefs must be controversial. Is this true? Say everyone in a democracy has a high opinion of democracy. Should that belief be called political?
3. Another potential consequence of my account of political beliefs is that it might end up pretty rare for two people A and B in very different times and places to have a political disagreement in the following sense: A and B have a political disagreement just in case there's some proposition *p* such that A believes *p*, B believes not *p*, and both A's belief and B's belief are political beliefs. Why might this be rare? Is this a bad consequence?
4. Imagine two people disagree about moral principles and also disagree about the empirical facts of the world. But imagine further that their disagreements happen to line up in a certain way, so that they actually have all the same opinions on political issues of the present day, vote for the same candidates, and so on. Would we say that they have political disagreements? (Note that my account seems to imply that they do.)[3]
5. Think about some of the political beliefs you hold. Do you think they're political because of something like the action-connection thesis or something like the aggregate-connection thesis? That is, do you think your political beliefs are the sorts of things that have consequences for the political actions we should take, or are they more like the sorts of things that help determine what sorts of political groups you're a member of – groups like political parties, for instance, or social groups where politics is important?

4

POLITICAL CONFLICTS AND POLITICAL DISAGREEMENTS

What is the role of political beliefs in political conflicts and their resolutions? Throughout the history of political philosophy and political theory, different philosophers have argued about the importance of different kinds of political beliefs in generating or resolving political conflict and the nature of political conflict more generally. The theory of political epistemology isn't always combined with the theory of political conflict, but I think it sheds light on the approach we should take to political disagreement and, as we'll see in the next chapter, the different ways of interpreting it.

Let's start with four situations. In each, a polity is trying to decide where to build some sort of power plant: location A or location B.

Situation 1: Half of the polity believes that location A is more environmentally friendly and cost-efficient, and thus they support building the plant in location A. The other half of the polity believes that location B is more environmentally friendly and cost-efficient, and thus they support building the plant in location B.

Situation 2: The whole polity agrees that location A is more environmentally friendly, and the whole polity agrees that location B is more cost-efficient. Half of the polity thinks environmental concerns are more important than financial concerns and thus support building the plant in location A. The other half of the polity thinks financial concerns are more important than environmental concerns and thus support building the plant in location B.

Situation 3: The whole polity agrees that the two locations would be equally environmentally friendly and cost-efficient. The residents of location A support

DOI: 10.4324/9781003355274-5

building the plant in location A, and the residents of location B support building the plant in location B. They may sometimes act as if they think the other location is worse, but this is just a tactic to try to get their way.

In situation 1, there is an empirical dispute – a dispute about the facts of the world. If these facts could be settled, everyone might agree on what to do. In situation 2, there is an ethical dispute – a dispute about what ought to be done given the facts of the world. The conflict emerges from people having different values and priorities. In the language of Chapter 2, in situation 1 there is a dispute involving disagreement over nonmoral beliefs, whereas in situation 2 the disagreement is over moral beliefs.

In situation 3, the dispute is a bit more raw. It is a dispute that emerges from people's interests alone, like the physical altercation that might occur if two starving animals saw a piece of food that could nourish one but not both of them. We might say colloquially that there is still disagreement about what to do, that the parties disagree about where to build the plant. But if we look for some claim that they disagree over – in the clearest case, that one believes to be true and the other believes to be false – we can't find one. This is what I mean by distinguishing a disagreement from a dispute: in what I call a disagreement we can always find such a contested claim; in a conflict or dispute we needn't be able to.

There is a fourth kind of possible situation, too, in which people think they have a dispute of the first or second type, but this impression is actually due to differences in the way they use language. We call this a verbal dispute. Some contemporary writers, as well as a few historical figures like Thomas Hobbes,[1] attribute a great deal of political conflict to verbal disputes. It is a kind of open question in philosophy to what extent, or under what circumstances, we should think of verbal disputes as real disagreements, akin to situations like 1 or 2. We'll talk about this in the next chapter.

Pure political conflict and political realism

Note that in situation 3, nobody disagrees about anything. But there is still a conflict. In the history of political thought there have always been writers, sometimes called "realists" or "materialists" of various stripes, some of whom take the view that all important political conflicts are like situation 3. If they're right, this is a problem for the project of our book. After all, we want to help understand these political disputes people keep getting into by understanding how people can form underlying beliefs that generate disagreements. But if the beliefs don't generate disagreements, they can't generate conflicts. The conflicts come from people's different interests, which they simply try to impose on one another, in a kind of amoral free-for-all.

We will see this kind of view come up a lot, in both general and specific ways. But here is a quick argument against it. In the history of political movements there are many schisms. Members of a political group with common goals will part ways, often viciously or even violently, because of disagreements about the best way to achieve those goals. Of course, some of these cases probably do involve people who think that a schism is the best way to serve their interests and advance themselves within their movement. This is in line with what writer Jon Schwarz calls the Iron Law of Institutions,[2] which is roughly that individual actors will always promote and protect their standing within a political institution or movement, even if it means failure for that institution or movement, rather than weaken their standing within it, even if that would mean success for it. But other debates seem well-intentioned, where both sides simply seem to think that their tactics would work better.

The view that all political situations are like situation 3 is a bold, universal one, and one counterexample is enough to undermine it. A weaker version of the realist or materialist thesis is that political conflicts are not properly evaluated by moral measures; the realms of politics and ethics are distinct. We looked a bit at that notion in the chapter on definitions of the political and it seemed pretty weak to us then. But it might be the case that *most* or *many* political situations are like situation 3, even if not all. As we proceed through the book, we'll encounter many philosophers and psychologists who are skeptical of people's political expressions and who often – though not necessarily always – doubt that we mean what we say when we talk about politics.

Interpreting political disputes

The four types of political conflict give us four possible ways to interpret political disputes. The interpretation of disputes is sometimes thought to be governed, either descriptively (it's how we actually *do* or even *must* interpret disputes) or normatively (it's how we *ought to* interpret disputes), by a *principle of charity*. The principle of charity says roughly that we do, must, and/or should treat our interlocutors as being rational and knowledgeable. It is most associated with Donald Davidson (whose ideas about belief and action we've also encountered), who describes it as "not an option. . . . Charity is forced on us; whether we like it or not, if we want to understand others, we must count them right in most matters."[3] However, popular uses of the phrase do present it as a choice, and there may be enough wiggle room in the phrase "most matters" to reconcile the two conceptions of the principle. Importantly for our purposes, the "counting them right" is done, as philosophers tend to put it, "by our own lights," meaning that when we interpret other people, we don't interpret what they mean and then decide that most of what they believe must be right but rather we figure out what they mean, to begin with, under the assumption that most of what they believe

must be right according to what we ourselves already believe, and any interpretation of their speech that contradicts this assumption must be rejected.

What does it mean for some agent A to interpret people so that as much as possible of what they believe is true according to A's lights when it comes to political disputes? This question amounts to asking this: in which situation, of the four discussed here, is there the most disagreement? But it's a rather difficult question. Interpreting the other side as having no real disagreement with our own, but trying to use some sort of power to get their way, does not seem very "charitable" in the intuitive sense of the word – unless, of course, we ourselves also act that way. Similarly, interpreting a disagreement as being merely a verbal dispute might seem more charitable than interpreting the other party as being wrong about something by our lights,[4] but if the disputed words are central enough and the disputed meaning seems obvious enough to us, assessing a dispute as verbal might be tantamount to assessing the other disputant as an incompetent user of the language, which doesn't seem charitable either. Philosophers sometimes treat verbal disputes as not being real disagreements, but to the extent that we can disagree about the meanings of words, which it seems we must be able to in any language that isn't purely stipulative, verbal disputes are still real disagreements about *something*, just not about what they seem to be about. This gives them a confusing status *vis-a-vis* the principle of charity.

What about the choice between interpreting a dispute as being about facts (i.e., as a disagreement when it comes to nonmoral beliefs) and interpreting a dispute as being about values (i.e., as a disagreement when it comes to moral beliefs)? Nathan Ballantyne and Peter H. Ditto have recently[5] investigated an informal interpretive heuristic sometimes called "Hanlon's Razor," which is quoted as directing us: "Never attribute to malice that which is adequately explained by stupidity." If we take attributing a disagreement to "malice" to mean interpreting an interlocutor as disagreeing with us about values and attributing a disagreement to "stupidity" to mean interpreting an interlocutor as disagreeing with us about facts, then Hanlon's Razor as a whole can be taken to mean that we should, when possible, always interpret disputes to be over facts rather than values. As something of an aside in a famous paper about disagreement, Adam Elga takes a position that might be taken to imply something like this.[6] Elga's view of dealing with disagreement is that we ought to set aside the view about which we disagree, so that if, for instance, A and B have agreed about everything else so far but disagree about some proposition *p*, they should each take the other's view to be as likely as their own and suspend judgment about *p* once they learn about the disagreement.[7] But Elga thinks it's highly unlikely, maybe impossible, that once we find we disagree with someone about important moral and political questions we could assess them to be equally as likely as we are to get those questions right, because there are a lot of issues that are "closely linked" or "allied" with important and moral questions; "in messy real-world cases, the disputed issues

are tangled in clusters of controversy."[8] So, if we follow Elga's reasoning and apply it to the question of interpretive charity, it's more charitable in general to interpret disagreements as being about issues that are easily separable from other issues, and moral beliefs are less easily separable from other beliefs than nonmoral beliefs are, so it might end up being more charitable in general to interpret disagreements as being about nonmoral beliefs than moral beliefs. We'll take a look both at this view of disagreement and at this view of separability in due course.

Another issue is the *metaethical* question about just what the nature of values is: whether our attitudes about morality are beliefs or whether they are emotive or expressive, more like cheers and boos – as though we are spectators at a sporting match yelling "Yay friendship!", "Boo murder!", and so on.[9] We'll see this question rear its head again when it comes to our political beliefs. It is not immediately clear how we can apply the principle of charity to the emotional expressions of the "boo"/"yay" variety. Charity, in Davidson's sense, was a matter of an agent A interpreting other people so as to make their beliefs look as much like A's as possible. It might not seem sensible that this should extend to interpreting other people so as to make their *emotions* look as much like A's as possible. However, there is a philosophical tradition of locating much of what's important in our moral lives in emotion-like *reactive attitudes*,[10] and it is also often argued that emotions have an evaluative component – that they're not just feelings.[11] This may be uneasy ground for someone who thinks that morality is merely emotive or expressive to stand on, though, since this evaluative component of emotion might end up being cashed out in their eyes as just another bunch of emotions.

There is similarly an interesting question about the charitable interpretations of *actions*. As we saw earlier, actions can be explained by reference to beliefs and (something like) desires. So to interpret other people's actions, we likely engage in a complex, dynamical process of negotiating between belief-sets that aren't quite like our own and desire-sets that aren't quite like our own. At any point, we could stop this process and simply assume that the agent whose actions we're trying to interpret has beliefs like our own but desires completely unlike our own. We don't seem to do this, suggesting that we're somehow "charitable" when it comes to interpreting other people's desires as well. This suggests that a broader framework than Davidson's might be necessary to understand the phenomenon of charitable attribution and interpretation.

The interpretation of actions may have more consequences for the interpretation of speech than it might immediately seem. Take the notion of *speech acts*, developed originally due to J. L. Austin.[12] In Jason Stanley's words, speech acts are "way[s] of doing something by using words"[13] – canonical examples include naming, marrying, and ordering.[14] One first-pass way of characterizing when

something is a speech act is that we can always add a kind of preface which describes what a speech act itself accomplishes.[15] The notion of a speech act is clearly important for our understanding of political beliefs and how to interpret spoken political disagreement because, as Stanley puts it, "there are multiple purposes to political speech, only one of which is to assert truths."[16] Indeed, our understanding of disputants' intentions more generally is important: they might be lying about what looks like a disagreement to try to win in some more basic sense, as the supporters of the pure political conflict theory – the political realists and (some) materialists – might suggest. But then there is a question of whether we need to follow the principle of charity in attributing speech acts to those we disagree with and what that charity would look like. There is a sense in which it might be charitable to always assume that those who disagree with you about politics are lying, or trying to waste your time, or trying to shut you up in some way, because that doesn't require that you think that anything *they* think is wrong "by your lights." We'll see more about this effect of charity in the chapter on polarization.

Conclusion

There are many potential sources of political conflict; it's not immediately clear whether there is a rational default interpretation of political disputes among them. What's most likely is that we adjust to every situation of conflict as it arises and explain its source based on its individual characteristics. Political disputes occur when people disagree over facts or values, use words differently, or struggle over finite resources. There is likely no one overarching theory of political conflict, and it is not immediately clear that there are moral or pragmatic reasons, like being charitable, which counsel in favor of adopting such a theory even as a starting point or heuristic.

Discussion questions

1. Think about a recent time when you've had a political conflict or disagreement with someone else. What did the source of the conflict seem to be at first? What did the source of the conflict seem to be after reflection?
2. Have you ever had a disagreement that you felt to be the result of how you and the other party used words, which you found you could resolve simply by using different words?
3. Think about a political conflict or disagreement in the news. What does the source of the conflict seem to be to you?
4. Think about an occasion on which you felt that someone else had misinterpreted you. This could be in the context of a political conflict, but it needn't be.

What was the nature of the misinterpretation? To what did they attribute their differences with you? Is there some other basis on which you wish or wished they had founded their interpretation?

5. Which do you think is more representative of what kind of person you are – your moral beliefs or your nonmoral beliefs? (Or is neither particularly representative?)
6. What do you think is the source of the conflict between or among political parties (say in the country where you live, or the country where you study, or whatever state or country you choose)? Do politicians in the different parties mostly differ when it comes to facts or when it comes to their values? Are any of their conflicts merely verbal?

5

THE POLITICS OF VERBAL DISPUTES

In this chapter we'll look at some issues specific to verbal disputes, since philosophy has a lot to say about them and since many contemporary political disputes take place purely verbally, online or through sound bites, or in a school board meeting or over the dinner table. There is a lot to think about here: talking past one another, changing definitions, whether a verbal dispute is always "mere" semantics or whether it might contain one of the other disputes somehow – whether what looks like a merely verbal dispute might be in fact a dispute over facts or values.

In an influential article, David Chalmers surveyed a few potential characterizations of merely verbal disputes. He liked this one: "A dispute over [some sentence] *S* is (broadly) verbal when, for some expression *T* in *S*, the parties disagree about the meaning of *T*, and the dispute over *S* arises wholly in virtue of this disagreement regarding *T*."[1] In other words, for a dispute over the truth of some sentence to be merely verbal, there must be – and it's enough that there is – a sense in which, if you were to remove the dispute over the meaning of some phrase, you would remove the dispute over the truth of the sentence. However, some of the characterizations he doesn't like as much may actually shed more light on the phenomenon for those unfamiliar with it. Take this one:

> A dispute over [some sentence] *S* is verbal iff *S* expresses distinct propositions *p* and *q* for the two parties, so that one party asserts *p* and the other denies *q*, and the parties agree on the truth of *p* and *q*.[2]

This gives us a kind of test for verbal disputes. Say A thinks a claim is true, but B thinks it's false. But then say that there's another way of phrasing what A takes

DOI: 10.4324/9781003355274-6

the claim to mean, and B thinks that's true; and there's another way of phrasing what B takes the claim to mean, and A thinks that's false. Then they don't really disagree. They just interpret the claim differently.

All this might seem a bit abstract, so here's an example. Say that your friend says that a proposed policy is totalitarian, and you disagree. It might seem to your friend that your disagreement is meant to indicate support for the policy or at least to indicate that you think it's not so bad. But it could easily be that you agree with the sentiment that the proposed policy really is very bad, perhaps that it's an invasion of privacy, or that it tramples on people's rights or something like that. You might simply think that a totalitarian policy must meet very specific conditions – for instance, you might think that to be totalitarian, a policy must undermine democratic rule. Of course, it could be that, in terms of ordinary usage, one of the two of you is *right* about what "totalitarian" means and that the other is *wrong*. But this would be a dispute over the word, not over the policy. Thus, even if we think a merely verbal dispute is still important or worth resolving in some way, understanding it as merely verbal at least redirects our attention somewhere else, toward the *source* of disagreement.

This redirection seems to be an important part of how we interpret political disputes. It's a kind of metaphor, but when we dig deeper into a conflict, we often find out what the real disagreement is *underneath*. When you support a policy and a friend opposes it, those facts aren't just unmoored from your other beliefs and attitudes; you can ask each other *why* you have those attitudes, and the answers – if they're right – will *explain* the disagreement. The question of how to interpret political disagreement is the question of what sorts of explanations of disagreement we look for and accept.

Certain kinds of disputes can't really be only verbal. A dispute in which the two parties can clearly state different plans of action – for instance, if A believes that we ought to do X and B believes that we ought not to do X – may ultimately be rooted in a difference in nonmoral beliefs, but it can never be a purely verbal dispute, because whatever words you replace in their statements, there is still the question about what to do, which they disagree about. Some philosophers think, in part because of the phenomenon of thick concepts which we discussed earlier, that what looks like a verbal dispute may actually be a deep dispute about values. This is sometimes called "metalinguistic negotiation."[3]

In making a case like this, Greg Restall uses the example of the term "marriage":

> On one side, we have the traditionalists who reserve the term for marriage between people of the opposite sex. On the other, we have progressives who take it that some same sex partnerships ought to be recognised as marriages. This is a significant and substantial debate – and no-one thinks that everything is resolved by the philosopher wading into the discussion and magisterially saying something like this: "There are two things you could mean by

> 'marriage'. Let's call opposite-sex marriage 'marriage$_1$' and let's call marriage between partners regardless of sex 'marriage$_2$'. It follows that same sex couples can be married$_2$ but they can't be married$_1$." That wouldn't help at all. Introducing talk of marriage$_1$ and marriage$_2$ doesn't end the debate, because which couples we count as "married" (with no subscript) matters to us, in the law, in our habits, in our conventions, and in so many different ways. Our language is a social phenomenon, with a history and a future.[4]

I think the thrust here is that the idea of marriage is tied up with our sense of morality, our laws and practices and habits – in other words, with our *actions* and which actions we encourage and discourage. So, because "marriage" is a thick concept, which in a sense is partly representative but partly exhortative, we can't simply remove it from our talk. In this example there may also be an added layer: it may be important to some people that certain marriages be recognized *as marriage*, with *no* change to our talk, and that may be part of the political desire of equal recognition or acceptance.

There is a research program in contemporary philosophy which focuses on questioning the definitions of, and sometimes offering new definitions of, precisely these kinds of contested, politically important thick concepts, called *ameliorative analysis*. Rather than trying to come up with thought experiments to figure out how we actually use some term or concept, as we did with "politics," ameliorative analysts start by thinking about what a word *ought* to mean. As outlined by Sally Haslanger in the seminal paper on the subject, it asks of our terms and concepts:

> Are they effective tools to accomplish our (legitimate) purposes; if not, what . . . would serve these purposes better? . . . [W]hat work [do] we want [them] to do for us; why do we need them at all?[5]

One well-known application of the theory of ameliorative analysis is Kate Manne's amelioration of the term "misogyny" in her 2017 book *Down Girl*. In offering a more systemic, political conception of "misogyny," as opposed to a psychologistic, individual definition, Manne cites the "hostile flavor" of the term for the contention that it "plausibly has a certain conceptual role to play in picking out the most hostile and noxious facets of gender-based oppression"; with this in mind, operating under an individualistic conception of misogyny would, Manne thinks,

> threaten to deprive women of a suitable name for a potentially potent problem facing them. This is all the more so in view of the apparent paucity of relevant alternatives. What other English words express similar, and similarly morally weighty, concepts that are particular to gender?[6]

So the idea is roughly that an analysis in this sense – something like a redefinition – of a word like "misogyny" based on its "flavor" or "role" ameliorates social ills by helping people express morally weighty concepts and gives appropriate names to potent social problems. Such amelioration is not the unique province of philosophers, of course. The historian and activist Ibram X. Kendi attempts in his work to redefine "racism" as anything that "produces or sustains racial inequity,"[7] for instance, removing any psychological element from the definition.

Ameliorative analysis of this sort faces a slew of challenges. Since it is philosophical work with a political goal, some of the challenges are philosophical and some of them are political; some sort of combine the two. The most influential challenge, originally posed by P. F. Strawson against Rudolf Carnap's idea of "explication," is that ameliorative analysis "changes the subject" rather than actually addressing some philosophical puzzle about what a term means and is thus irrelevant to our questions involving the terms targeted for analysis.[8] Herman Cappelen has drawn out three different challenges from this one idea; we'll consider two, one of which has two sides to it.[9] The first challenge is that if in ameliorating a term, changing its meaning, we put ourselves in the situation of being able to answer questions about that term or using that term, we shouldn't feel our curiosity is satisfied. The questions are no longer the "old" questions but are instead new ones. For instance, say we are trying to figure out whether cultural appropriation is racist. Then someone comes along with a new definition of "racism" which specifies within it that cultural appropriation is racist. Even if we adopt the new definition, this shouldn't feel satisfying when it comes to our query. The query was about the term under the old definition, not the new one. The second problem has to do with verbal disputes. The first side is that redefining a word can't settle a dispute (just as it couldn't answer a genuinely puzzling question) but rather can merely layer a verbal dispute on top of a substantive one; the second side is that redefining words is apt to cause a great number of verbal disputes, since some people will use the old definitions and some will use the new definitions. Even if ameliorative analysis doesn't "change the subject," it is hard to see how it can answer old questions or settle disputes, since it doesn't generate any actual new evidence that *ought* to affect our beliefs.[10]

Other kinds of critiques are more practical and political. Some people worry that ameliorating the sorts of words that these socially minded philosophers are most often concerned with will dilute or water down their meanings, meaning that they actually *lose* the flavor and moral weight which fueled the ameliorative inquiry to begin with.[11] A related concern is that what is crucial in our thinking about politics is not words but concepts, and when a word changes to match a different concept, we don't stop using the old concept but look for new ways to express it. In this "treadmill" effect, top-down efforts to change language to improve reality by restricting our concepts and our expression of them are

always matched by bottom-up linguistic innovations that enable us to keep talking about what we want to talk about, whether we want to or not. This is a variation on Steven Pinker's idea of the "euphemism treadmill," which is that when we introduce new euphemisms to refer with neutral connotation to traits that are generally taken to be negative, the supposedly neutral neologisms take on negative connotations themselves, occasioning the need for yet more euphemisms, which fall prey to the same process.[12] The overarching principle is that beliefs come first, language only second.

I worry just as much that changing the definitions of words will be politically counterproductive, leading to backlash and also making it harder for people to talk to each other productively. I think it's also sensible to wonder just who will benefit from ameliorative projects and who will be hurt by them. When we come up with a trendy new definition of an important word, it seems like people who are young, well-connected, cognitively flexible, and gifted with a lot of free time are the most able to adapt and promote the new usage, and they also seem to be the most insistent on blaming and shaming those who don't use the word in the new way. For me, this is another nagging *political* source of doubt about ameliorative projects. Finally, ameliorative projects may face political opposition, leading to what Adam Gibbons calls "conceptual conflict."[13]

There are more transparent ways in which verbal political disputes can matter or in which substantive political disputes about facts or values can be worked through in a deceptively verbal way. Take a feminist group that gets into an internal dispute about just what "feminism" means. This group is likely not having a merely verbal dispute, and their interest in the question about what "feminism" means likely isn't a merely academic or semantic interest. Rather, the group is probably debating about what they ought to do. Since the group coheres around feminism to begin with, its members likely take the question of what "feminism" means to be identical to the question of what the group ought to do. This brings us back to the issue of interpreting political disputes, considered in the previous chapter.

Some philosophers have emphasized a sort of reversed way that language can express political disputes. In 1948, on a British radio program, Bertrand Russell introduced the idea of *emotive conjugation*, also called "Russell conjugation," which involves describing the same action or trait in different ways based on who is doing or displaying it; Russell famously gave the example: "I am firm, You are stubborn, He is a pig-headed fool."[14] In political discourse one must always watch for these loaded descriptions. David Beaver and Jason Stanley have used the possibility of something like emotive conjugation (they use the example of "dog" vs. "cur") to argue against the possibility of a neutral debate; because our choices in language, like "firm" versus "stubborn," express our affinities and affiliations, a debate conducted in language can never, for Beaver and Stanley, be a fully neutral examination of the facts.[15] I think that means that Beaver and

Stanley think political disputes tend to be of the third type, with everyone just pushing each other around by connotation, though I'm not sure.

We'll talk more about changes in language over time and their relationship to changes in values in the chapter on political beliefs and the philosophy of history and progress. But the most important takeaway is to think hard about how our ways of using language can interact with political conflict. What looks like a real dispute could be a merely verbal one, whereas what looks like a merely verbal dispute could trace back to real differences of opinion about facts or values.

Discussion question

1. Are there words that you insist upon using in political disagreements, even if their definitions are contested? What sorts of words, if any, are central enough to your political convictions that you can't express those convictions without them? (But before you answer, *try to* express your convictions without them!)

PART II
Theories

6

THEORIES OF TYPE AND PERSONALITY

Just how do people form their political beliefs? In Parts III and IV, we'll consider some specific dynamics, like ideologies, conspiracy theories, and polarization. Also, we'll consider some potentially rational sources of political belief, like experts, news reports, and majority votes. In this part, though, we'll consider some very general theories of political belief formation. We'll make some quick notes about the extent to which they imply that people's political beliefs are rational, but in general they will in fact suggest that people tend to be irrational in forming their political beliefs.

Some theories work from notions of psychological types to explain political beliefs. This has an intuitive appeal: if we feel that political groups divide individuals in a natural and clean way, so that all members of some political persuasion or party have some things in common that members of other political persuasions or parties lack, then the plan of explaining political beliefs in terms of these commonalities is attractive. This idea of political groups as natural and clean divisions is not uncontroversial, and some might prefer to look for similarities between groups rather than differences.[1] We will see some theories later that do so. Lots of different kinds of people have theorized political beliefs in terms of psychological types. Such theories usually seem to involve the value-based vision of political conflict, as is only natural, since that's the potential cause of political conflict most amenable to explanation in terms of psychological types.

DOI: 10.4324/9781003355274-8

General notes on psychological theories of political belief

Theories of political belief formation tend to come in two very different flavors: moderate and radical. The politicization of the study of political belief itself gives a maddeningly self-referential quality to the whole enterprise. The two flavors of political conviction seem to see different phenomena as needing explanation when it comes to political belief. While moderates see political beliefs as subject to a great deal of controversy which stands in need of explanation, radicals see society as easily forming consensus even on false beliefs when it favors the powerful, and this tendency is what stands in need of explanation.[i]

From the moderate perspective, there are (at least) two kinds of approaches to explaining political disagreement. One of them I call the difficulty approach and the other I call the durability approach. The difficulty approach explains political disagreement in terms of some intractability in political questions, meaning that reasonable people might disagree over political questions. The durability approach explains political disagreement as a result of our irrationality: our political beliefs seem to survive contact with facts that should undermine them.

From the radical perspective, there are (at least) three kinds of approaches to explaining political disagreement: the they-tricked-us approach, the we-tricked-ourselves approach, and the nobody-was-tricked approach. The they-tricked-us approach has it that powerful forces pull the wool over the eyes of ordinary people intentionally, maybe through overt means like false propaganda or maybe through more subtle means of suggestion, insinuation, and association. The we-tricked-ourselves approach has it that our own psychologies are somehow built to come up with their own justifications for the dominance of powerful people, institutions, and perspectives. And the nobody-was-tricked approach has it that people don't really have these political beliefs; they're simply acting as though they do, and such actions can be either rational or irrational. (The moderate perspective has its own version of this approach.)

The ease with which a nobody-was-tricked formulation can be thrown together poses a problem. Since virtually any expression of belief can be interpreted as savvy dishonesty, there is pretty much for every theory of what drives political belief a corresponding theory under which the apparent political belief is not really held, and what's driven is merely an agent pretending to have that belief. It might be worth thinking about just how to differentiate between and test two such theories.

Before we look at any of the theories of political belief formation, though, we should consider a kind of null hypothesis: that people form their political

i Thanks to Liam Bright for a discussion which led to the formulation of this distinction. Note that this does not track Bright's distinction between "sexy murder poets" and "basic bureaucrats," which is about romanticism and technocracy.

beliefs the way they form any other beliefs. We've seen that any belief can be politicized, so there's some appeal to this. On the other hand, politicization might undermine our faith in our beliefs' reliability. And how do we form our other beliefs, anyway? We might use observations, intuitions, personal experience, other people's testimony, independent research, news media, expert consensus, and so on. But we've already defined a political belief as one that's involved in some sort of controversy or which splits one group from another. Thus, the question of how we form our political beliefs seems to start with a choice: whether or not to start off from the perspective that one group is better at forming their beliefs than the other.

Moral foundations theory

The most prominent contemporary theory of political belief based on personality is Jonathan Haidt's moral foundations theory, a moderate theory which explains the difference between liberals and conservatives as being caused by conservatives drawing on a larger set of foundational moral principles than liberals do. Though Haidt does not think this expanded moral framework makes conservatives *correct*, he does think it means that conservatives understand liberals better than liberals understand conservatives, a claim we'll evaluate in the chapters on polarization.

Haidt's political psychology is largely based on his prior work in moral psychology,[2] where he advanced an account of moral judgment called *social intuitionism*. Social intuitionism has two parts: the intuitionistic part and the social part; the intuitionistic part is what's most important for us here. Intuitionism about moral judgment is opposed to rationalism about moral judgment, which holds that "the justifying moral reasons we have for our judgments are also the causally effective reasons for why we make those judgments."[3] In a nutshell, intuitionism flips this correlation on its head: it says that the judgments come first, and the apparent reasoning is "a matter of mere *post hoc* rationalizations."[4]

Intuitionism thus reaches back into the classic philosophical dispute between rationalism and empiricism. Haidt takes his work to be updating the great Scottish Enlightenment philosopher David Hume, who wrote in his *Treatise on Human Nature* that "reason is, and ought only to be the slave of the passions"; in the updated version, "reason is the press-secretary of the intuitions, and can pretend to no other office than that of ex-post facto spin doctor."[5] What is it to be a press secretary or spin doctor? Well, such people – who, interestingly for our purposes, play perhaps their most prominent and characteristic roles in politics – take other people's decisions and try to justify them for public consumption. Thus, reason comes *after* judgment and may serve more of a social than an intellectual role – a conclusion also reached by cognitive scientists Dan Sperber and Hugo Mercier in their book *The Enigma of Reason*, which presses an "argumentative theory of reasoning" on which "what reason does . . . is help us justify our

beliefs and actions to others, convince them through argumentation, and evaluate the justifications and arguments that others address to us."[6]

On the intuitionistic view of moral judgment, what comes first is the intuition, which then generates the judgment, and the reasoning comes last; this challenges the rationalist view, on which the reasoning comes first and is followed by a final judgment. Why should we take the intuitionist view and be so pessimistic about moral reasoning? One piece of evidence is the proposed phenomenon of *moral dumbfounding*.[7] Moral dumbfounding occurs in an experimental setting when study participants express a strong moral judgment about a hypothetical scenario but are unable to explain why they have that judgment, despite trying to, which is supposed to show that they did not begin with reasoning but expected it to come after. The hypothetical scenarios depict behaviors like "consensual incest between two adult siblings" and "a woman cooking and eating a piece of flesh from a human cadaver donated for research"; other comparable scenarios are things like "eating one's (already dead) pet dog" and "cleaning one's toilet (in private) with one's national flag."[8] The stories additionally "were carefully written to be harmless" so that some natural harm-based explanations of why wrong had been done (family fallout, disease, social breakdown) were foreclosed.[9] Haidt et al. note what they call "interesting response patterns" among study participants: for one, "participants often directly stated that they were dumbfounded, i.e., they made a statement to the effect that they thought an action was wrong but they could not find the words to explain themselves"; and "participants would start giving an argument but as they were talking they realized that the argument was not going to work and they stopped in the middle of it, without any prompting from the experimenter."[10] They also observed facial movements from participants which they interpreted as expressions of skepticism or self-doubt.

Moral foundations theory builds on social intuitionism by attempting to explain political disagreement in terms of differences in the prerational foundations of moral intuitions. It is probably best seen as a durability theory – one that posits that we're stubbornly irrational in our political beliefs – since it opposes the rationalist model of moral reasoning. Haidt and colleagues argue that liberals form their political beliefs only from "individualizing" moral foundations, which push for people to be treated fairly and for harm to be prevented,[11] and for people to remain free.[12] They call these the "fairness/reciprocity," "harm/care," and "liberty/oppression" foundations. Conservatives, they hold, also have intuitions based on three other "binding" moral foundations, which go beyond individual morality. These include "virtues of loyalty, patriotism, and self-sacrifice for the group, combined with an extreme vigilance for traitors," the so-called ingroup/loyalty foundation; "virtues of subordinates (e.g., obedience and respect for authority) paired with virtues of authorities (such as leadership and protection)," the so-called authority/respect foundation; and "virtues of purity and sanctity that play such a large role in religious laws," the so-called purity/sanctity

foundation.[13] In the main paper on moral foundations theory, they provided four kinds of evidence: differing rates of agreement with explicit statements about moral values; differing judgments about situations (for instance, for "authority/respect," a situation in which a soldier disagrees with an officer's orders); differences in willingness to make tradeoffs and violate taboos; and differences in moral language.[14] Although moral foundations theory is sometimes popularly presented as finding that conservatives somehow "possess" moral foundations that liberals lack, this isn't really what the evidence from Haidt and colleagues shows; rather they find decent-sized but gradual changes in responses depending on self-reported political persuasion, suggesting that liberals possess "binding" foundations as well, just to a significantly lesser degree than conservatives do.

Social intuitionism, moral foundations theory, and some of the inferences their proponents try to make from them about the nature of political disagreement have all been challenged in a variety of ways. Without mentioning moral foundations theory directly, Hrishikesh Joshi has argued that political disagreement cannot merely be a matter of differing values, because contentious empirical premises (e.g. whether or not raising the minimum wage also leads to an increase in unemployment) are needed to reach a belief about public policy.[15] We might add from our definition of political beliefs that some political beliefs have little to do with values at all, like beliefs about the characters of political candidates or whether some candidate committed some or another impropriety; another theory (perhaps one of the ones below) would have to be attached to the moral foundations theory to explain differences in such beliefs.

Hanno Sauer has objected to every aspect of Haidt's theories. First, Sauer suggests that intuition and rationality might not be so easy to separate; we might think that intuitions are kinds of habits of interpreting situations, and it's possible that "we should be inclined to describe the habitualization of conscious deliberation as a rational process."[16] In addition, in certain experimental scenarios in which more time was provided and counterarguments to subjects' intuitions were offered, some study participants changed their viewpoints in response, which seems like an exercise of rationality.[17] Perhaps most interestingly, Sauer notes something strange about how the moral dumbfounding thought experiments were set up. It was precisely "when stories described victim[less] and harmless norm violations" that moral dumbfounding occurred.[18] Thus, the moral foundations concerning harm and victimization may not be subject to dumbfounding. But those are precisely the moral foundations that liberals and conservatives share – meaning that it might be only the conservative-favored moral foundations that social intuitionists should be eager to call irrational.[19] Those are also the foundations that seem to recede for individuals with higher education,[20] which Sauer mentions presumably to buttress the notion that the shared foundations may be the more sensible or accurate moral foundations upon adequate reflection or inspection. There is a more general problem with the idea of liberals

trying to "appreciate" conservative moral foundations: first, social intuitionism seems to suggest those foundations are irrational, rendering such an appeal sort of silly; second, it's not clear in what way one could "appreciate" moral foundations without being motivated by them oneself. (We'll see more on Haidt's and Sauer's arguments in this regard in the chapter on polarization.)

Daniel Jacobson has also objected to the moral dumbfounding experiments which seem to provide support for social intuitionism and thus moral foundations theory, though differently from Sauer. Jacobson asks whether it is really correct to say that the situations presented to study participants are "harmless." That, after all, depends on our prior theory of "harm," itself a substantive issue of ethical theory, and one which it's not clear practitioners of moral psychology should take a stand on in their research. Even in consequentialism, the ethical theory whose conception of harm is perhaps closest to what the social intuitionists need for their arguments, situations like the incest case can still be called harmful for a number of reasons.[21] First, though they might not involve actual harms, they might involve great risks of harm, and consequentialists will generally evaluate an action based on its expected consequences (weighted by their probability) and not the consequences that in fact occur. Second, consequentialists tend to evaluate the harm caused by an action relative to the set of other possible actions the agent might take, so that an apparently harmless action is still harmful insofar as it was taken instead of a more beneficial one. Third, consequentialists will disagree about the nature of ultimate value; some consequentialists might even think that the binding moral foundations concern the sorts of things that are ultimately valuable, so that committing incest could in fact be *inherently* a kind of harm, because it could be disvalued without reference to any other more foundational value. Finally, plenty of ethicists are not consequentialists, and other ethical theories seem sympathetic to the binding moral foundations.

It's worth noting that the interplay of ethical intuitions and ethical theory is a methodological concern in philosophical ethics; perhaps most famously, the moral and political philosopher John Rawls[22] coined the term "reflective equilibrium" to refer to a method of

> working back and forth among our considered judgments . . . about particular instances or cases, the principles or rules that we believe govern them, and the theoretical considerations that we believe bear on accepting these considered judgments, principles, or rules, revising any of these elements wherever necessary in order to achieve an acceptable coherence among them.[23]

This suggests that working ethicists don't really find it problematic to give our intuitions some weight, and indeed what the social intuitionists characterize as post hoc rationalization might sound to them like a reasonable method for working out the best view.

The authoritarian personality

Haidt's theory isn't the only one that divides people up into political types based on their personality and psychology. An older theory of the *authoritarian personality*, still popular among some commentators and academics today, was developed in the wake of the Second World War to explain the rise of fascism and the horrors of the Holocaust. The locus classicus for this theory is a 1950 book called *The Authoritarian Personality* by social theorist Theodor Adorno and others, which begins with the researchers stating that they were

> guided by the following major hypothesis: that the political, economic, and social convictions of an individual often form a broad and coherent pattern, as if bound together by a "mentality" or "spirit," and that this pattern is an expression of deeplying trends in his personality. The major concern was with the *potentially fascistic individual*, one whose structure is such as to render him particularly susceptible to antidemocratic propaganda.[24]

They qualify their view that "personality may be regarded as a *determinant* of ideological preferences" by allowing that "personality evolves under the impact of the social environment and can never be isolated from the social totality in which it occurs."[25] The nature of the authoritarian personality, established by a high score on an F-scale (for "fascism") established by a set of questionnaires, is to love order and hierarchy. In a *Pacific Standard* article, Tom Jacobs has characterized it as "belief in absolute obedience to authority" and "simplistic thinking, intolerance of ambiguity, and racial prejudice."[26]

The 1950 book on the authoritarian personality has lost favor due to its intellectual basis in Freudian psychology and some criticisms of the F-scale and the questionnaires. (The questionnaires asked participants to react to statements like "Although many people may scoff, it may yet be shown that astrology can explain a lot of things," "No insult to our honor should ever go unpunished," "He is, indeed, contemptible who does not feel an undying love, gratitude, and respect for his parents," "It is entirely possible that this series of wars and conflicts will be ended once and for all by a world-destroying earthquake, flood, or other catastrophe," and "Books and movies ought not to deal so much with the sordid and seamy side of life; they ought to concentrate on themes that are entertaining or uplifting."[27]) But the notion of an authoritarian personality has been offered as an improvement on Haidt's theory.[28]

Moreover, similar theories seem to survive in social psychology. A social-psychological theory of group conflict, social dominance theory, posits a *social dominance orientation*, "the extent to which one desires that one's in-group dominate and be superior to out-groups,"[29] which is taken to partially explain political actions and political beliefs, especially the acceptance of various kinds of biased

ideologies. Similarly, the "rigidity of the right" model of political psychology posits a "constellation of psychological attitudes and evocable states – including dogmatism, closed-mindedness, intolerance of ambiguity, preference for order and structure, aversion to novelty and stimulation, valuing of conformity and obedience, and relatively strong concern with threat" to explain conservative political views.[30] This concern with threat has even been traced to differences between the amygdalas of liberals and conservatives[31] – the amygdala being a part of the brain that aids in processing emotions and making decisions. Finally, the most important categorization in the psychology of personality, the Big Five framework that schematizes personality as a combination of openness, conscientiousness, extraversion, agreeableness, and neuroticism (or of their opposites), has been correlated with politics, with liberalism being correlated with openness and conservatism with conscientiousness.[32]

This notion has also been used to explain the political rise of Donald Trump in the United States and the appeal he held for some of his followers. An influential 2016 *Politico* article by Matthew MacWilliams, a political consultant, was titled "The One Weird Trait That Predicts Whether You're a Trump Supporter";[33] the weird trait was the authoritarian personality. Historian Kevin Mattson wrote in *Dissent* in 2018: "One of the crowning works of the Frankfurt School, *The Authoritarian Personality* has much to teach us about the age of Trump."[34] Just a few years earlier, in 2011, political scientist Corey Robin wrote *The Reactionary Mind*, which argued that conservatism consists in "the felt experience of having power, seeing it threatened, and trying to win it back."[35] Such deployments of the theory of the authoritarian personality also call into question the moderate/radical distinction with which I opened this chapter. There is not always going to be a clear delineation between ordinary partisan politics and awful oppression; indeed, partisan politics often involves accusing the other side of awful oppression, and awful oppression often tries to disguise itself as ordinary politics.

The politics of personality research; content overlap

But partisan politics may enter into social psychology research itself, casting doubt on some of these personality models. One way this occurs is "content overlap," in which political content is part of the measurement of the personality trait, or personality traits are part of the political measurement. For instance, a critical article on the rigidity of the right model claims that "scales treated as indicators of conservative vs. liberal ideology often contain content pertaining to religious sentiment, cognitive rigidity, orientation toward authority, and/or intolerance, in addition to (mostly cultural) political content," and that a "dogmatism" scale from 1960 which succeeded Adorno's F-scale directly measured "anti-Communist sentiment, pro-American nationalism, and a hawkish foreign policy posture."[36] When two measures are blended into one another like this, their correlation doesn't prove anything at all. Similar issues arise with the

so-called racial resentment scale, also taken to be part of the explanation for the rise of Donald Trump and contemporary conservative politics.[37] Measures of racial resentment often directly ask questions about contentious political issues and attitudes, meaning that they may directly measure conservative politics rather than finding a correlate of it. And José Duarte has suggested that urban liberal values are built into the measurement of the "openness to experience" construct in the Big Five, saying in an interview:

> They ask things like, "Am I sophisticated in art and literature?" . . . Where I'm from – a rural place, which is also where many conservatives are from – you'd be ashamed to say something like that. . . . We need to ask them things like, "Do you like to sit at night and look at the stars?" "Do you like to learn new things with your kids?" "Do you like to read?"[38]

Indeed, Hrishikesh Joshi suggests the same thing about Haidt's measurements of moral foundations – that his questions about the binding foundations might trigger them in "conservatives in particular."[39] Joshi considers an item on a Moral Foundations Questionnaire used in Haidt's research, which reads "Chastity is an important and valuable virtue":

> There is a worry of circularity here. Since chastity is a value that is prominent among social conservatives, it is no surprise that people who identify as conservative will tend to agree with the prompt. Yet this doesn't establish that liberals do not have a robust Sanctity foundation. For, the foundation may be triggered by other things . . . for instance, having to do with the environment, or with wasting food.

This is the specter of content overlap again.

The theory of the authoritarian personality and its less extreme contemporary variants are probably best thought of as radical theories of political belief formation, probably of the we-tricked-ourselves variety. Focused on the irrationality of political conservatives of one sort or another, from the assumption that such views are deviant, they posit underlying psychological abnormalities to explain them. These psychological abnormalities eventuate in objectionable ideologies and problematic political beliefs, which explain the bad political actions of individuals and conservative political parties and systems more generally.

Much of this research seeks to explain conservatism as an aberration from a set of liberal political and social attitudes assumed to be normal. This tendency, along with the lack of conservatives in social psychology relative to their number in the general population, has led some to suggest that such findings, if not the discipline of social psychology as a whole, may be unreliable on their face.[40] In one very silly incident, multiple papers finding that conservative political attitudes were correlated with psychoticism had to be corrected; their data indicated

in fact that *liberal* political attitudes were the ones correlated with psychoticism.[41] Occasionally, this trend is reversed, for instance in studies on the personality type purportedly behind "political correctness."[42] More often, the sphere of normalcy is shifted from the left to the center, with social psychologists trying to explain political extremism as an aberration from the norm of moderation. For instance, one study argues that "rigidity" is characteristic not of conservatives but of extremists,[43] while another finds that "Dark Triad" personality traits of narcissism, Machiavellianism (manipulation and exploitation of others), and psychopathy are correlated with both "authoritarian political correctness" and "alt-right attitudes."[44] A final reason to take a pessimistic attitude toward this research is the ongoing replication crisis in the social sciences (see the chapter on expertise). Indeed, research on political psychology, highly politicized itself and apt to go viral, is exactly the sort of research that we might expect to be overhyped and to fall apart under scrutiny. Thus, any claims you might hear relating personality types with political attitudes should, I think, be treated with a lot of initial skepticism.

Conclusion

The attempt to explain political beliefs in terms of underlying psychological differences in personality – personality "types" – is a sensible one. It tends to construe political conflict as driven by differences in values, then explains those values by reference to personality types. But these types aren't so clear-cut and neither is the evidence about their existence or their relationship to political beliefs. And in many cases, the attempt seems to begin by taking the correctness of certain political beliefs for granted, which does not seem to me to be a promising research strategy.

Discussion questions

1. What aspects of your personality do you feel might be reflected in your political beliefs? What aspects of other people's personalities do you feel might be reflected in their political beliefs? Are there ways for people with a similar personality to you to reach different political conclusions?
2. What might be some dangers of trying to classify people of a particular political persuasion as being people with a certain personality type or vice versa?
3. What do you think is the more important phenomenon: political disagreement or political consensus? Do you think politics is most fruitfully characterized by its patches of conflict where people might easily find common ground or by its patches of inaction where people might easily take some sort of action?
4. Do you think it is possible to put aside our own political beliefs and biases when studying political beliefs, whether from a philosophical or psychological standpoint? If it's not possible, how should we deal with this fact?

7

THEORIES OF INERTIA AND IDEOLOGY

In general, the personality-types theories see political differences as emerging from human variety. Some other theories, more in line with the radical framework, see the opposite: political sameness emerging from a universal human nature. This posited human nature bears some similarities with the psychological characterization of conservatives and authoritarians some of the personality theorists gave. It is in the nature of humans to avoid complexities, to desire certainty, to trust authority, to react defensively against threats, and to justify the social structures in which they find themselves – so say these theories. This goes, according to those who propound such theories, both for those who are advantaged by those structures and for those who are disadvantaged by them. I called these *inertial theories* since they suggest that political beliefs aid in keeping things the same.

False consciousness, system justification theory, and status quo bias

An early example of an inertial theory is the Marxist theory of *false consciousness*. False consciousness is actually the second half of the Marxist theory of political belief. The first half, which we'll discuss later, is *class consciousness*, an awareness of one's membership in a political group (in the most classically Marxist approach, the political groups are economic strata or "classes") and of that group's political interests. False consciousness is an explanation of why class consciousness does not always obtain and hence why the proletarian revolution is not occurring. False consciousness involves the more powerful class or classes, precisely in virtue of their power, transmitting their own class interests – perhaps

DOI: 10.4324/9781003355274-9

even unknowingly – to the less powerful classes, in the form of political beliefs and perhaps other semi-epistemic phenomena like mental habits or indications about what's salient or deserving of attention. When realized as a sort of comprehensive schema for thinking about society, these beliefs and attitudes are called *ideology*. Knowing transmission that serves a dominant group's interests is often called *propaganda* (which we'll examine near the end of the book). Occasionally, theorists of these concepts take the view that ideology does not just change people's beliefs but constitutes them *as people*.[1] But I find that view incredibly implausible, if not absurd.

Because class consciousness and false consciousness form a theory of political belief together, it's tough to evaluate the notion of false consciousness on its own. It's certainly somewhat natural to think that powerful people can make their interests seem natural or important to others, but I'm skeptical that membership in an economic group or class determines people's political beliefs. And at least in contemporary society, there are plenty of widespread moral and ideological notions that oppose the powerful, even if they don't always lead to effective political change.

An updated form of false consciousness theory, in the language and methods of contemporary psychology, is John T. Jost's theory of system justification. Like many of the personality theorists we saw, Jost seems to begin with a political perspective and take deviations from it as the phenomena that require explanation. Jost writes that his theory is an attempt to answer questions like these:

> Why do some women feel that they are entitled to lower salaries than men, why do people stay in harmful relationships, and why do some African-American children come to believe that white dolls are more attractive and desirable than black dolls? Why do people blame victims of injustice and why do victims of injustice sometimes blame themselves? Why is it so difficult to get people to stand up for themselves, and why do we find personal and social change to be so challenging, even painful? . . . Why do so many poor people oppose the redistribution of wealth? Where is the outrage, even after a succession of worldwide financial crises, meltdowns and bailouts?[2]

Note the divergence between the social-psychological method and our own epistemological method. In epistemology, we allow the possibility that we ourselves are wrong about everything, as Descartes did. Social psychology, however, seems to require a frame of normalcy and deviance, and a researcher's own political beliefs often seem to determine what's normal in the realm of political attitudes.

A useful synonym for "system justification" in Jost's work is "rationalization of the status quo."[3] Things *currently are* a certain way, and people seem to look for arguments in favor of how things are rather than starting from a position of uncertainty – much as, according to Haidt, we look *post hoc* for reasons why

our moral judgments are true rather than reasoning toward those judgments to begin with. Jost offers eighteen theses that constitute system justification theory; we needn't examine all of them, but some notable claims include the following: people rate events as more desirable when they seem more likely;[4] people use stereotypes to rationalize inequality;[5] members of disadvantaged groups may take explanations of their position as legitimate (a "tendency to internalize and perpetuate the system of inequality");[6] favoritism of advantaged groups is stronger when measured as implicit rather than explicit bias;[7] and the legitimacy of the system is linked to higher self-esteem and lower group ambivalence for advantaged groups, and vice versa for disadvantaged groups.[8]

Jost does take a swipe at political conservatism, but as a system or ideology that helps legitimate a bad status quo rather than as a personality trait. He writes:

> Political conservatism is a prototypical system-justifying ideology insofar as it provides moral and intellectual support for the societal status quo by (a) maintaining cultural traditions and resisting social change, and (b) defending the legitimacy of hierarchical arrangements and existing forms of social, economic, and political inequality.[9]

An important element of the false consciousness theory that remains in system justification theory is the focus on political groups; the fact of society-wide injustice is meant to be demonstrated by inequalities between groups – no longer just economic strata but now including races, ethnicities, sexes, genders, and so on – and the predictions of the theory are cashed out in terms of the political attitudes and psychological states of individuals as differentiated by group membership.

According to Jost, "engaging in system justification serves the palliative function of increasing satisfaction with the status quo and addresses underlying epistemic, existential, and relational needs to reduce uncertainty, threat, and social discord."[10] It is therefore irrational: it is a perfect example of *motivated reasoning*, in which people seem to form beliefs based on what they desire to believe rather than what their evidence or reasons tell them. For the most part, it looks like what I called earlier a we-tricked-ourselves approach. In Jost's eyes, even members of disadvantaged groups have such a need for cognitive and emotional stability that they will accept the very status quo that disadvantages them rather than entertain the possibility that they live in a terribly unjust world.

How plausible is system justification theory? That might depend in part on whether you accept the moral and political vision of the contemporary world which it takes for granted. Even if you do, though, system justification theory might seem to fail empirically to capture the beliefs of members of advantaged and disadvantaged groups. The evidence for this theory emphasizes implicit measures, and implicit bias – which we'll consider in a later chapter – has been challenged on both methodological and empirical grounds.[11] Even if we accept

implicit bias, experimental evidence is by no means univocally in line with system justification theory; for instance, both men and women have been tested as implicitly favoring women – the so-called women-are-wonderful effect[12] – with women's favoritism toward other women more powerful than men's preference. The notion that members of disadvantaged groups will become ambivalent toward their own group, while members of advantaged groups will strongly identify with theirs, also doesn't seem to be borne out by the data; a 2018 survey found that levels of in-group bias among black, Hispanic, and Asian respondents were similar to such levels among non-liberal white respondents, while liberal white respondents were the only group to exhibit out-group favoritism[13] – the opposite of Jost's prediction. To my mind, Jost has a poor understanding of the incentives surrounding system-justifying and system-challenging beliefs; he writes: "In many contexts embracing a progressive, system-challenging worldview is a form of 'political deviance'."[14] But I think there are plenty of contemporary contexts in which the opposite is true. Indeed, it is the very fact that justifying the political system seems "deviant" to Jost that structures his entire approach. System justification theory has also been challenged on replicability grounds.[15]

A final, more general idea along the lines of false consciousness and system justification is what's called *status quo bias*. Status quo bias consists of giving value to the current state of affairs simply because it is the current state of affairs. It is empirically testable. In one kind of experiment, two options are presented to two groups of subjects, but in one group neither option is presented as the status quo, while in the other group one option is presented as the status quo; when an option is presented as the status quo, "its popularity increases significantly."[16] Status quo bias may, however, be rational when viewed as a kind of deference to better-informed experts or authorities (though see Chapter 16 on expertise). Jacob M. Nebel lists other good reasons for some cases of status quo bias, including avoidance of the costs of transitioning from one status quo to another, avoidance of uncertainty, and avoidance of the practical costs of deliberation.[17] Nebel argues additionally that what already exists might be of greater intrinsic value just in virtue of existing.[18] We will come back to whether status quo bias might be rational later, in the chapter on history and progress.

The philosophy of ideology and ideology critique

Both false consciousness theory and system justification theory rely on a prior notion of ideology. It is the ideology that somehow causes or constitutes the false consciousness of the individual, and it is the ideology that provides the individual with the intellectual resources to justify the workings of the system. So it is worth making some general comments about the notion of ideology. First, "ideology" has a somewhat unstable meaning and connotation in a way that is shared by plenty of political terms and maybe even "politics" itself (as we

have discussed previously). The term can be used neutrally to mean something like a sort of comprehensive scheme for thinking about politics. Synonyms for ideology in this simpler, neutral sense might include "worldview," "system of thought," or the German *Weltanschauung*. There is also a more complex, pejorative usage, which gives it the feeling of serving a kind of top-down function of obscuring dark truths about society to buttress the powerful. This is the sense in which false consciousness and system justification theories are theories of ideology, and it is also the sense that is behind the philosophical program of "ideology critique." Having addressed psychologists' ideas of ideology, we'll address philosophers' ideas now.

Sally Haslanger gives a useful distinction which can help us get a grasp on which things are ideologies. She writes:

> Oppression and injustice comes in many forms. One distinction worth noting is between oppression that is repressive, that is, forced upon individuals through coercive measures, and oppression that is ideological, that is, enacted unthinkingly or even willingly by the subordinated or privileged. . . . Chattel slavery in the United States was repressive. Gender oppression is, at least in many contexts, ideological: men and women, even men and women with deep commitments to justice, hardly notice their participation in practices that sustain male privilege and power, and even, sometimes, take them to be central to their identities. . . . Racist oppression in the contemporary United States is mostly hybrid: the majority of the racially subordinated participate unwillingly and experience it as repressive, and the racially privileged enact it unthinkingly. In hybrid cases ideology plays a role, but the ideology is not hegemonic and coercion is often employed to keep the subordinated in their place.[19]

For Haslanger, racist or sexist beliefs – or sets of such beliefs with a certain kind of coherence and causal efficacy – are examples of ideologies. For Marxists, sets of capitalist beliefs would likely be the primary examples. Note that, at least as conceived by the people describing them, these ideologies both involve some people being more powerful than others and this being somehow okay or justifiable. In the most thoroughgoing cases, those of false consciousness and system justification, even the people without the power subscribe to the ideologies, and thus there is no political conflict. But there is still something left to explain. This is an important point: even political *agreement* may stand in need of inquiry about political beliefs. This separates a radical approach from a moderate one.

An important philosophical task in the theory of ideology is, just as we did with "politics," to try to figure out what is meant by the word – what an ideology is. Tommie Shelby writes that ideological beliefs have four features: they are widely shared by members of a certain political group; they "form, or are derived from, a prima facie coherent system of thought," that is, the ideology itself; they

"are a part of, or shape, the general outlook and self-conception of many in the relevant group"; and they "have a significant impact on social action and social institutions" (similar to our belief-action connection for political beliefs).[20] Ideologies themselves have further features which can be cashed out in different ways. First, ideologies are partly or largely false, "distorted by illusions" (theorists of ideology prefer for some reason the *active* language of "distortion" rather than the normal language of falsehood); second, ideologies have something to do with political conditions, especially unequal ones, and the "legitimation" of such conditions, which they help effect; third, the acceptance of, or belief in, ideologies can be explained by the "false consciousness" of those who believe it, where Shelby takes false consciousness to mean "unconscious influence of noncognitive motives" like political interests.[21] So, in slogan form, we might say something like this: for Shelby, ideologies are sets of beliefs which legitimate bad social conditions and are held *despite* being false and *because of* their legitimating function, plus people's lack of awareness that that's why they're held.

Sally Haslanger presents ideology as having similar features, but she relates them to each other in a different, more complex way. She writes that "ideology is best understood functionally: ideology functions to stabilize or perpetuate unjust power and domination, and does so through some form of masking or illusion."[22] So a critique of ideology involves two parts: "The epistemic critique of ideology reveals the distortion, occlusion and misrepresentation of the facts. The moral critique concerns the unjust conditions that such illusions and distortions enable."[23] But "the epistemic and justificatory failings" of ideology "can[not] be identified through ordinary epistemic critique" because "in the social domain, shared beliefs can make themselves true."[24] For this reason, the ideology critic must use methods like genealogy to uncover the function of various beliefs and norms.[25] This also raises the possibility, somewhat obscured in Shelby's account, that an *originating explanation* of how an ideology comes to be might be different from a *perpetuating explanation* of how an ideology maintains its hold. If shared beliefs can make themselves true, this offers a perpetuating explanation of their maintenance and also complicates the question of rationality and epistemic "distortion" or "illusion." After all, how can a truth be a distortion?

Here is one example of how a belief might make itself true. Say employers somehow become convinced that members of a certain group are more likely to steal. For that reason, by and large, members of that group end up more likely to be out of work and indigent. The indigent are, in turn, more likely to be desperate for basics like food, clothing, and shelter; also, having less, they have less to lose if arrested or shunned. So they do, in fact, end up stealing from others more frequently, making the original shared belief true – even though it wasn't true when it was first held by employers. Note that this story draws on the connection between belief and action which we've stressed so much. Precisely because

beliefs can generate actions, and actions can change what's true, beliefs can make themselves true. (This idea is sometimes known as a *self-fulfilling prophecy*.) This also reverses the connection between the mind and the world (much like what we saw in Haslanger's theory of ameliorative analysis). We normally think of true beliefs occurring due to this sort of circumstance: the world is a certain way, and we correctly perceive the world, leading us to correctly represent the world in our beliefs. But in a self-fulfilling prophecy, things happen in a different order: we represent the world in a certain way in our beliefs, which lead us to act in a certain way, which causes the world to be a certain way, which we then perceive and take as confirmation of our original beliefs. Thus, the belief fulfills itself.

Why is it important that an ideology is self-fulfilling? I think for a few reasons, which are mostly pragmatic rather than epistemological. First, it can make even a true political belief seem somehow artificial or constructed rather than natural. Because we are used to perceiving the world and this determining our beliefs, it's odd to think that by changing our political beliefs, our ideologies, we might change the political world such that the old beliefs become false and the new beliefs become true. We might not be content to wait around for evidence for or against our political beliefs, knowing that such evidence could be a product of those political beliefs themselves. Second, it can undermine the legitimacy of people or groups in power. In the ideological explanation, power, or the drive for it, is what explains the political beliefs that justify it rather than people's sensible application of their political beliefs explaining why certain people or groups have power. Recall our potential explanations of political conflict: there are disputes over facts, disputes over values, pure disputes over power or resources, and verbal disputes. If the ideological explanation of some political situation or conflict is right, then what looks like a dispute over facts or values might have its etiology in a pure political dispute. Thus, the notion of self-fulfilling prophecy should make us question just why we hold our political beliefs – where they come from; their *genealogies*.

I think this idea is very clever, and it also makes things complicated for epistemologists, since it suggests that we might not always deserve credit for believing the truth. However, I also think there are some reasons to be a little careful positing self-fulfilling prophecies. Take statements like the following:

> I did well because I believed in myself,
> I did poorly because I didn't believe in myself,
> The relationship failed because neither of us believed it could succeed,
> I got so worried about food poisoning that it made me sick to my stomach.

At a certain point this sort of explanation seems a bit too easy, and the way it works almost starts to seem like magic: we think something and it occurs.

Psychologist Lee Jussim has argued that the effects of self-fulfilling prophecies are usually small.[26]

More generally, philosophers and others who think ideologies play a functional role in perpetuating some system must be careful of the risk of theorizing on the basis of what Kieran Healy has called "evil functionalism."[27] Functionalist social theories have it that different social phenomena are often best understood in terms of the role they play in stabilizing the societies of which they're a part. "Evil functionalism" simply adds to this that such stability is largely the perpetuation of unjust conditions, as against potentially good social change. In addition to plausibly underestimating the value of stability, such views and tendencies also inherit the problems of functionalist theories writ large, which are no longer broadly popular.

Ideology and morality

There is a long-standing debate within the theory of ideology (or false consciousness, or whatever) about how it relates to morality, which is sometimes framed as an interpretive question about Marx's perspective on morality. In Haslanger's conception, ideology produces immoral or unjust outcomes, and that's a big part of the problem with them. But there is a different kind of conception, more in line with the realist and materialist schools we've discussed, which has it that our judgments about justice and right and wrong are most likely to themselves be products of the dominant ideology under which we're living, so that when we say an ideology is unjust, we are likely still operating within it rather than genuinely challenging it.[28] Under this kind of perspective, there are simply different interests, and the best outcome for the powerless is not to achieve justice or to avoid distortion but to determine their beliefs and effectuate outcomes in line with their own interests rather than the interests of the powerful. This might leave things a bit difficult, though, when it comes to the question of what to believe about politics if we're trying to rationally arrive at the truth.

We'll talk more about this conception in a few chapters, but if you're interested in this topic, you might enjoy an old debate between linguist and political commentator Noam Chomsky and social theorist Michel Foucault. This debate crystallized the difference between realist political activists and moralist political activists. Here's an exchange I find telling:

> *FOUCAULT:* So it is in the name of a purer justice that you criticise the functioning of justice? . . . If justice is at stake in a struggle, then it is as an instrument of power; it is not in the hope that finally one day, in this or another society, people will be rewarded according to their merits, or punished according to their faults. Rather than thinking of the social struggle in terms of "justice," one has to emphasise justice in terms of the social struggle.

CHOMSKY: Yeah, but surely you believe that your role in the war is a just role, that you are fighting a just war, to bring in a concept from another domain. . . .

FOUCAULT: But I would merely like to reply to your first sentence, in which you said that if you didn't consider the war you make against the police to be just, you wouldn't make it. . . . The proletariat doesn't wage war against the ruling class because it considers such a war to be just. The proletariat makes war with the ruling class because, for the first time in history, it wants to take power. And because it will overthrow the power of the ruling class it considers such a war to be just.

CHOMSKY: Yeah, I don't agree.

FOUCAULT: One makes war to win, not because it is just.

CHOMSKY: I don't, personally, agree with that. For example, if I could convince myself that attainment of power by the proletariat would lead to a terrorist police state, in which freedom and dignity and decent human relations would be destroyed, then I wouldn't want the proletariat to take power. In fact the only reason for wanting any such thing, I believe, is because one thinks, rightly or wrongly, that some fundamental human values will be achieved by that transfer of power.

FOUCAULT: When the proletariat takes power, it may be quite possible that the proletariat will exert towards the classes over which it has just triumphed, a violent, dictatorial and even bloody power. I can't see what objection one could make to this.[29]

In this exchange, Chomsky's commitment to framing political action in moral terms and Foucault's wholesale rejection of moral language and concepts make them ideal representatives of two ways of thinking about politics, whether of the radical or Marxist or of the everyday, election-winning, bill-passing variety.

Conclusion

The theory of ideology, which is generally a radical theory of political irrationality, has been propounded by all sorts of writers, including contemporary social psychologists and philosophers. But there is a lot of debate about just what it amounts to, how it's supposed to work, and what the evidence is that people live under ideologies. It starts by taking for granted a perspective on politics which itself could easily be explained as the result of an ideology. Further, it might face internal contradictions, seeing as its formulation seems to require both the moral and anti-moral perspectives simultaneously. We will see a way to do similar politics without the notion of ideology in a few chapters.

Discussion questions

1. How much inertia do you think there is in your political beliefs and the political beliefs of those around you? Do you think humans are generally motivated to defend the way things actually work at the expense of the way they might work? Do you often simply accept the social arrangements around you as rational without questioning them? Do you think other people do?
2. Do you think your political beliefs are "there" for a purpose? That is, do you think you, or people in general, have been inculcated into various political belief systems so as to help maintain the smooth functioning of society, whether you think that society is good or bad?
3. What ideologies do you think operate in society today? Note that, for any ideology you name, there is likely a great deal of opposition to it. Is the opposition also due to ideologies, or is it due to simple rationality, or something else? Are there ideologies we can't name, which simply surround us like water surrounds fish?[30]

8

THEORIES OF IDENTITY, SIGNALING, AND PARTISANSHIP

If so far we've seen theories claiming that political disagreement (e.g. disagreement between liberals and conservatives) can emerge from personal differences and political agreement (e.g. agreement on the legitimacy of the status quo or the truth of some ideology) can emerge from human commonalities, we have yet to see a theory claiming, perhaps most parsimoniously, that political disagreement can emerge from human commonalities. Most prominent theories of this type, like much of what we've seen so far, revolve around the notion of a political group and an individual's place in it: their "identity" and how they *signal* that identity to others.

Identity-protective cognition

Dan M. Kahan, a law professor, offers a theory of *identity-protective cognition* to explain people's political beliefs, with "identity" meaning some sort of political group membership. Identity-protective cognition is "a tendency to selectively credit and discredit evidence in patterns that reflect people's commitments to competing cultural groups."[1] This theory of political belief is an individual interest-based one; on Kahan's view, "individuals . . . have a large stake – psychically as well as materially – in maintaining the status of, and their personal standing in, affinity groups whose members are bound by their commitment to shared moral understandings."[2] This goal "can generate motivated cognition relating to policy-relevant facts,"[3] just as the goal of allowing oneself to be satisfied and avoiding discord with one's society generated motivated reasoning in system justification theory. The theory of identity-protective cognition is thus a durability theory of political belief. This theory also emphasizes something we ourselves have stressed so

DOI: 10.4324/9781003355274-10

far about political beliefs: both moral beliefs (beliefs about values) and ordinary beliefs (beliefs about facts, excluding moral facts) can be political. It also suggests a causal-hierarchical relationship between the two kinds of political beliefs: culture (including values) is "*cognitively* prior, in the sense that people's perceptions of what the facts *are* is shaped by their values."[4] We assess the facts in such a way as to put minimal cognitive pressure on our values.

An important source of evidence for identity-protective cognition, which will also be important in our later consideration of the epistemology of democracy and of political expertise, is the finding of *motivated numeracy*. In a characteristic motivated numeracy study,[5] a difficult causal inference problem was given to subjects in four groups. In two groups, the question in the problem was whether a skin cream increased or decreased the incidence of skin rashes; in one group, the data provided supported the conclusion that rashes increased, while in the other, the data provided supported the conclusion that rashes decreased. In the other two groups, the question in the problem was whether a proposed gun control law would lead to an increase or decrease in crime; in one group, the data provided supported the conclusion that the law would lead to an increase in crime, while in the other, the data provided supported the conclusion that the law would lead to a decrease. In all four groups, participants were classified according to political party (by self-identification, from "strong Democrat" to "strong Republican"), political ideology (from "very liberal" to "very conservative"), and by numeracy – that is, facility with quantitative data and statistical inferences.

Kahan and his colleagues made two predictions. First, in both groups considering the different depoliticized skin treatment case, high numeracy would be highly correlated with correctly answering the causal inference problem. Second, and in line with identity-protective cognition, "ideological polarization in the gun ban conditions should be most extreme among those highest in numeracy."[6] Both of these predictions were borne out: while high-numeracy liberal Democrats were much more likely than low-numeracy liberal Democrats to give the correct answer to the problem in what Kahan et al. call the "rash increases," "rash decreases," and "crime decreases" groups, they were only slightly more likely to give the correct answer in the "crime increases" condition – a condition in which, the idea is, the data went against their prior commitment to gun control. Similarly, high-numeracy conservative Republicans did better than low-numeracy conservative Republicans in the "rash increases," "rash decreases," and "crime increases" groups, but did only slightly better in the "crime decreases" group. Numeracy, it seems, didn't help much when the facts went against identity; the authors write: "In other words, higher numeracy improved subjects' performance in detecting covariance" – that is, in answering the causal inference problem correctly – "only in the 'gun control' condition in which the correct response was congenial to the subjects' political outlooks."[7] In fact, low-numeracy participants were hardly polarized at all, even on the politicized question. The explanation

given by Kahan and his colleagues is that a certain level of numeracy is required to determine what the "congenial" conclusion even is;

> those low in numeracy will be less likely to succeed in discerning the correct, identity-affirming conclusion that the data in that condition actually support because they are less likely to possess the sorts of science-comprehension capacities that doing so requires.[8]

So the idea is roughly this: when something is politicized, our brains default to trying to figure out the most politically congenial conclusion, and so those with the highest ability to figure out what's politically congenial will end up with the most politicized beliefs. This has a lot of consequences for our views of phenomena like politicization and polarization; for instance, we might start to think of them as mostly affecting elites rather than low-information voters.

As with status quo bias, though, there may be some at least plausibly rational explanations for motivated numeracy. If people take their prior political beliefs to be rational, then they may be rational in relying on them to answer difficult problems of causal inference. And motivated numeracy, like many of the experimentally tested theories we've considered, seems to fail the test of replication.[9] Psychologists Ben Tappin, Gordon Pennycock, and David Rand argue on the basis of further studies that Kahan et al.'s work does not sufficiently distinguish between political identity and prior political beliefs; in their experiments, "cognitive sophistication magnified a direct effect of prior factual beliefs on reasoning" but there was "no evidence that cognitive sophistication magnified th[e] effect [of political group identity on reasoning]."[10] In other words, according to these psychologists, people put their mental energy toward figuring out whether some new contention or implication is consistent with their prior beliefs and not toward protecting their political identity.

Yet to those who take a moderate approach to politics as well as to the study of political beliefs, there is something very appealing about the notion of identity-protective cognition, just as there is something very appealing about the system justification theory to those who take a radical approach. Just as radicals see injustice everywhere and wonder how it can be allowed to continue, positing ideology as an answer, moderates see political conflict everywhere and wonder why people can disagree so confidently and hatefully, positing political identity (often called "tribal" to indicate the associated virulence) as an answer. We will consider some issues related to this view of politics in Chapters 19 and 20 on polarization.

Political signaling

A related, broader idea is what we might call the political signaling approach to political beliefs. Robin Hanson expresses this idea when he writes that "the main

payoffs from [our evolutionary ancestors'] arguing and negotiating behavior was not via influencing the resulting group beliefs and actions, but from how their words and deeds influenced how others thought of them."[11] An example of the signaling approach is the concept of *moral grandstanding*, which is similar to the popular concept of "virtue signaling." In their book on moral grandstanding, Justin Tosi and Brandon Warmke theorize that two elements are necessary and jointly sufficient for some discursive act to count as moral grandstanding: the grandstander must "want to impress others with their moral qualities," and they must "try to satisfy that desire by saying something in public moral discourse."[12] Note that grandstanding can be part of a theory of political belief only insofar as the grandstander is unaware of what they're doing. When someone is aware that they are grandstanding, or when someone is *merely* signaling, we have moved into the realm of political speech and action. This distinction brings us back to the interpretive question about political disputes and the question of sincerity.

In a recent paper, Dan Williams argues that we should expect sincere beliefs to figure into the story about signaling; this is his theory of *socially adaptive beliefs*.[13] He writes:

> We frequently capitulate to [social] incentives, such that the way in which we form beliefs is highly sensitive to the actual or anticipated effects of candidate beliefs on other agents. Given this, many of our systematic departures from epistemic rationality are driven not by irrationality or the use of cost-effective heuristics but rather by well-calibrated rational self-interest: When the demands of truth conflict with social expedience, it can be practically advantageous to jettison the former in favour of the latter.[14]

So we remain rational in the *practical* or *economic* sense of holding beliefs that fit our self-interest, but fail to be rational in the epistemic sense of holding beliefs justified (for instance) by our evidence.

At least two aspects of Williams's story are of interest for our purposes: "Other agents have reliable access to what we believe and frequently reward us when we hold ungrounded beliefs and punish us when we hold reasonable ones."[15] First, other people's access to our beliefs. Such access is mediated through the belief-action connection which we've touched on so many times. But recall that actions emerge holistically from a set of beliefs and desires, and keep in mind that we're always capable of insincerity in our expressions. This makes inferring someone's beliefs from their actions and speech not so easy after all. However, Williams notes that pretense and prevarication have costs of their own and may have few benefits when compared to false but socially adaptive beliefs. Insincerity "typically requires substantial energy and attention and often elicits strong punishment if it is discovered," especially in cases where the truly held belief would elicit punishment anyway, which is what occurs in the typical case of

socially adaptive belief; plus, "if one has little practical incentive to hold true beliefs anyway," which is also what holds in the typical case of socially adaptive belief, "one will also have little practical incentive to consciously deceive."[16] In other words, insincerity and deception can be *difficult*, *risky*, and *unrewarding* compared to wholesale socially adaptive belief.

Second, the question of which beliefs are rewarded. Of course, *sometimes* a belief that is not grounded in evidence will be what's socially appropriate, but sometimes a belief that is grounded in evidence will be what's socially appropriate. Indeed, all sorts of information-generating institutions must be set up so that their workers – people like researchers and journalists, but also people like lawyers and doctors – have incentives in particular to tell the truth, so that they can't be bribed by interested parties and also so they can't simply misinform others out of laziness. So why think this case is the typical one? One answer:[17] because irrationally held beliefs might be harder to dislodge than rationally held ones and thus better signals of long-term group loyalty; a rationally held belief can be dislodged by a good argument or new evidence, but an irrationally held belief of this sort could only be dislodged by some change in group identity. Irrationally held beliefs might also be better signals of someone's other beliefs: if one holds a group-mandated rational belief for rational reasons, they may reject the group-mandated irrational belief for rational reasons as well, whereas if one holds a group-mandated irrational belief for reasons of group loyalty, there's no reason to think they would then go on to reject the group-mandated rational belief.

Here's one thing that might be missing. Williams's account, along with some of the others, seems to suggest that *conformist* beliefs are the ones that tend to be socially rewarded. Yet I think in at least some cases it is precisely the opposite: people benefit from holding (or seeming to hold) beliefs that *don't* conform – as long as they're the right kinds of beliefs and as long as they don't go too far. Believing might be like fashion in this way. Dressing conservatively is a risk-averse strategy. It minimizes the chance of social disapproval, but it also makes it difficult to stand out. Dressing ostentatiously is a high-risk, high-reward strategy. If your outfit seems to people to "work," then you may reap more rewards than the conservative dresser, but if it seems to fail, then you may be subject to all sorts of criticism and ostracism. Thus it is not the case that, for instance, only socially adaptive beliefs are subject to the sorts of patterns in which Williams is interested. That said, this can only be bad news for human rationality, as it just grows the pot of potentially irrational beliefs. And this is true in the political context: even when people deviate from the norm in their group, it needn't be the case that they do so rationally, since it could be that their deviation was motivated by the potential rewards of approval for the previously deviant belief, possibly even shifting the group consensus and gaining themselves a sort of leadership identity within it.

The social constraint account

Sociologists Andrei Boutyline and Stephen Vaisey have authored a paper utilizing the formal technique of network analysis to elicit the structure of systems of political belief. They also offer a theory of political beliefs from partisan identity, but it differs in important respects from the ones we've considered so far. Drawing on a famous paper by Philip Converse,[18] who emphasized the *non*-ideological nature of most people's explanations of their political beliefs and the "cognitive complexity of ideological reasoning," Boutyline and Vaisey characterize the central idea of their *social constraint* theory as "the claim that people use political identity as a heuristic for acquiring further political beliefs via the flow of information from opinion leaders, including politicians, journalists, and activists."[19] In so doing, people "can replace the abstract question of 'what should I believe?' with the social question 'which team am I on?'"[20] This approach to partisanship is related to two topics we'll talk about in the next part: political ignorance and expertise. Note, however, that it shares an essential feature with the identity-protective cognition account: political and intellectual sophistication, or "expertise," is viewed as making one more reliable as a reporter on or guide of the stances of their political group but not more reliable on the political facts. This is essential to the critical perspective of such theories.

Where do the groups come from?

You might be wondering: Just where does this sense of political identity come from? To whom are all of these people signaling, and why? This is a hard question, but some writers have attempted to answer it. In many major contemporary democracies, at least, one potential answer has to do with the political parties. In the United States, there exists a politico-cultural divide, which can be thought of as left/right, blue/red, Democrat/Republican, or urban/rural depending on just what you think the most important variables are. It could be that we are born into that divide, but another possibility is raised by some theorists of this stripe, like Hyrum Lewis and Verlan Lewis[21] and Hrishikesh Joshi.[22] It could be that there are some political beliefs we're really sure about, where we have really strong convictions, which we think are really important, and which lead us to adopt a partisan identity. We might call these our "central" political beliefs. Then, once we find the political party that takes our view on that issue, we adopt its other platforms and stances as our own beliefs and end up with what we might call "peripheral" political beliefs as well. In this sort of vision, the viciousness of political disagreement is almost entirely explained by party competition, the distinctness of the party platforms, and the polarization of elected officials, with just a little bit of explanation needed for how people arrive at different central political beliefs to begin with.

We have talked a great deal about political groups so far, but we should say something about just why the notion is so attractive to all sorts of theorists, philosophers, psychologists, and sociologists. The prominence of group membership in human psychology is a claim made by the *social identity theory* pioneered by Henri Tajfel. Tajfel's experiments showed that under certain circumstances people would favor members of their "group" even if groups were assigned completely arbitrarily as part of the experimental design and that people would take actions that were not optimal for themselves personally to benefit the group or other group members.[23] Jake Harwood writes of the findings associated with social identity theory:

> As we categorize people into groups, we also categorize ourselves into some of those same groups – social identification. Once we "belong" to a group (our "ingroup"), we seek ways to gain positive feelings from that group membership. Viewing the ingroup more positively than other groups ("outgroups") is one way to achieve those positive feelings. Seeking positive distinctiveness for one's ingroup hence becomes an explanation for holding negative beliefs and attitudes about outgroups in our environment, and hence for prejudice and ultimately discrimination.[24]

We saw earlier that some people think the very nature of politics has to do with distinctions among groups. Social identity theory gives an empirical basis for the importance of such distinctions and support for the notion that advancing the status of one's group can be a powerful motivator for action, including distinctly (perhaps definitively) political action, as well as discrimination and polarization.[25]

Conclusion

A popular kind of explanation for our political beliefs is that they result from our identities as members of political groups. According to this explanation, we tend to adopt the beliefs that will give us the most success as group members, the most success *within* our groups,[26] not quite regardless of the evidence but perhaps sometimes against our best evidence. This theory fits neatly into the American political context with its vicious polarization between partisan groups.

Discussion questions

1. Looking over your beliefs, political and otherwise, do you feel there are some that you've been motivated to adopt to fit in? Do you feel that others know what you truly believe? Do you think you're more likely to believe something silly to fit in or to lie about your beliefs to fit in? Do you ever do the

opposite – believe something silly or lie about your beliefs to distinguish yourself?

2. Are your political beliefs part of your identity? How do your political convictions relate to other parts of your identity? How do they compare to, say, your tastes in music, art, film, literature, or food? How do they compare to your beliefs about academic subjects like sociology or philosophy?
3. How good a thinker do you think you are when it comes to political questions? Do you feel you're sharper when thinking about politics or duller? How about other people you know? Do people make their most insightful points in political arguments, or is that when people are at their most stubborn and their most resistant to evidence and reason?

9

THEORIES OF POSITIONALITY, STANDPOINT, AND EXPERIENCE

Above we considered system justification, an outgrowth of the second part of the Marxist tradition in the theory of political belief. Now we should consider the first part, which in Marx was "class consciousness." The contemporary outgrowth of this notion is *standpoint theory*. Natalie Alana Ashton and Robin McKenna write:

> Marx argued that the different social locations of the bourgeoisie and of the proletariat lead them to have different perspectives on economic exchange and the social relations that hold between the two groups. From the perspective of the proletariat, the oppressive nature of these social relations is, or can be made, visible, whilst from the perspective of the bourgeoisie the oppressive nature of these social relations is obscured.[1]

Standpoint theory updates this theory of class consciousness similarly to how system justification theory updates the theory of false consciousness: by expanding it to different putative political divisions, like race and gender; much work in standpoint theory is done by feminist epistemologists. The notion of social reality being "made visible" under a certain standpoint is sometimes called an *epistemic advantage*.[2] (Later, we'll call it *standpoint expertise*.)

In a forthcoming paper, Emily Tilton notes a few ways that this epistemic advantage (what she calls the "inversion thesis," whereby a social disadvantage generates an epistemic advantage) might accrue, which include better access to evidence, better incentives to seek out evidence, better "conceptual resources" for making sense of their evidence, and more openness to alternative hypotheses.[3] It is also sometimes suggested that people with less power are naturally better

DOI: 10.4324/9781003355274-11

off because they must understand the perspective of the powerful to survive as well as their own; Charles Mills wrote: "Often for their very survival, blacks have been forced to become lay anthropologists studying the strange culture, customs, and mindsets of the 'white tribe' that has such frightening power over them."[4] And Liam Kofi Bright has argued that some white Americans' political beliefs result from a psychology of cognitive dissonance that emerges from being simultaneously familiar with and supportive of the *prescriptive* account of race and equality in the United States and haltingly aware of the ways in which the country *actually* falls short of those ideals.[5]

We will think more about this idea of epistemic advantage in the chapter on expertise; for now our focus is a little different. In this chapter I want to ask: To what extent is standpoint theory, also called standpoint epistemology, a full theory of political belief? It would be a very comprehensive theory indeed if it told us what sorts of beliefs members of advantaged groups have and what sorts of beliefs members of disadvantaged groups have, and it is natural to think that experiences of advantage and disadvantage deeply affect our political beliefs. The idea that demographic identity is an important factor in political beliefs certainly seems commonplace in public discussions of politics, and in the United States in particular polling data is often analyzed for how public opinion breaks down on lines of race, ethnicity, sex, and gender. The claim that "demographics is destiny" has been debated in American electoral politics for decades as the nonwhite proportion of the populace has increased and the white share has decreased, because in the United States nonwhite voters tend to support the Democratic Party and white voters tend to support the Republican Party.

However, sophisticated standpoint theorists have avoided going this far, for two reasons. First, standpoint theorists tend to allow that not all members of a certain group have the epistemic advantage that potentially comes from social disadvantage. Rather, the standpoint is an "achievement"; Amia Srinivasan writes that the epistemic advantage accrues to the political group somehow "as a class," but that in the theories of class consciousness and standpoint epistemology

> the proletarian and feminist standpoints, respectively, are to be achieved, and are not something automatically given in virtue of one's status as an oppressed subject under capitalism or patriarchy. Piercing the ideological appearance requires an overcoming of false consciousness and the achievement of revolutionary consciousness, in turn a matter of both political analysis and political action.[6]

The (purported) facts of phenomena like false consciousness and system justification mean that members of disadvantaged groups *also* face a lot of difficulty in getting their political beliefs right. Therefore, the sort of epistemic advantage at issue in standpoint theory is not merely a "perspectival advantage" like the fact

that "disabled people have better evidence regarding the limitations of certain kinds of building design as a result of their embodied experience as disabled people."[7]

Of course, there are independent reasons to think that a standpoint isn't automatic. It is empirically obvious that, for any political group, there is plenty of political disagreement among its members, even about fundamental issues relating to their experiences. And even when experiences are shared, they – or rather beliefs about them – must be accompanied by other empirical premises to generate beliefs about the causes of those experiences, with ethical premises to generate beliefs about the rightness or wrongness and the significance of those experiences and with both ethical and empirical premises to generate beliefs about what ought to be done in light of those experiences.[i] So then, to qualify as a theory of political belief, we might want to hear quite a bit more from standpoint theorists about just when a standpoint is actually going to be achieved.

The second problem is in the other direction and itself has two parts. First, standpoint theorists disagree over whether a standpoint is exclusive to members of the germane disadvantaged group. According to at least some standpoint theorists,[8] by doing the right kind of research and thinking, a white person, for instance, could achieve the standpoint of a black person, or a man could achieve the standpoint of a woman. So standpoint epistemology is not a fully worked out theory of political belief because it is not univocal as to what sorts of political beliefs members of advantaged groups can even *possibly* have. To my mind, there is very little reason for a standpoint theorist to deny that a member of an advantaged group could achieve the same epistemic advantage as a member of a disadvantaged group. Once the standpoint is construed as an achievement which requires certain kinds of work to get it, we can at least imagine any sort of person undertaking that work. But if members of advantaged groups can achieve the standpoint of disadvantaged groups, then it's not clear anymore what standpoint theory is claiming: some members of disadvantaged groups achieve the standpoint and some don't; some members of advantaged groups achieve the standpoint and some don't. Indeed, Charles Mills writes about his theory of "white ignorance" that the so-called white ignorance "will often be shared by nonwhites to a greater or lesser extent because of the power relations and patterns of ideological hegemony involved."[9] (This is very similar to the theories of false consciousness and system justification which we saw in the previous chapter.) The theory, which seemed to be stated so boldly, ends up making virtually no predictions about people's actual political beliefs. This is not meant to suggest that making such predictions is necessarily, by and large, the *goal* of standpoint theorists. Rather, they are usually more interested in establishing the epistemic

i Thanks to Elizabeth Barnes for this distinction, which she draws in a manuscript still in progress.

advantage of the standpoint. The bolder theory is more commonly found in lay political discussions and in less careful statements in academic work.

If standpoint theory is not quite a full theory of political belief, it certainly functions as a sort of internal political strategy for some political groups, though a controversial one.[10] The idea of social location, position, or standpoint is often used as a kind of meta-consideration in some political circles to determine who even has the standing to speak on an issue – which "voices" should be "centered" and stuff like that. It would be convenient indeed if it were to turn out that the people who we might think deserve to be heard as matters of equity and justice were also due deference as a matter of epistemology. We will discuss this element of standpoint theory a bit more in the chapter on expertise.

Discussion questions

1. What experiences have you had that might give you insight into particular political issues? What is the nature of that insight, and are there aspects of your beliefs about those issues that don't come from those insight-generating experiences?
2. What do you think are the benefits and dangers of trying to predict people's political views based on their membership in demographic groups and of deferring heavily to certain members of certain groups on certain issues?

10

THEORIES OF TUNNELING AND IDEATIONAL DETERMINISM

I want to make a brief mention of the unique theory of political belief offered by Jeffrey Friedman in his 2020 book *Power Without Knowledge*, which he called *ideational determinism*. Ideational determinism combines elements of difficulty and durability, since it says both that the modern political world is too complex for us to understand and that we use irrational, or prerational, heuristics to simplify it for ourselves. Its main claims can be better told as a kind of story.

The modern world is intractably complex – too complex for any of us to understand it directly and completely without some cognitive shortcuts. In Friedman's words, "complexity entailed the need to *interpret* society," and "interpretations might be unreliable."[1] Friedman called those shortcuts "stereotypes," following the journalist and political commentator Walter Lippmann, who wrote:

> The real environment is altogether too big, too complex, and too fleeting for direct acquaintance. We are not equipped to deal with so much subtlety, so much variety, so many permutations and combinations. And although we have to act in that environment, we have to reconstruct it on a simpler model before we can manage with it.[2]

Lippmann continued rather poetically: "To traverse the world men must have maps of the world."[3] Without such stereotypes, according to Friedman and Lippmann, adults would see the world as philosopher and psychologist William James had claimed that babies do, as "one great blooming, buzzing confusion."[4] Without these cognitive shortcuts, all the reality of the modern world, which we seek to manage toward good ends through political action, would make no sense to us. Perhaps irrationally, we don't always realize we're using such shortcuts;

DOI: 10.4324/9781003355274-12

instead, the world seems so simple that "a rational observer could use common sense or intuition to determine the causes of the problems, the likely solutions, and the likely costs of those solution."[5] But this appearance is a result of the imposition of stereotypes.

Friedman supports his claim about the complexity of the modern world in part by reference to the idea of *unintended consequences*. Unintended consequences are just what they sound like: if we institute a policy with the intent that it have some effect, the unintended consequences are all the other effects it has. In the most extreme case, the unintended consequence is the opposite of the intended one. Friedman gives plenty of examples of unintended consequences;[6] here are two: drug and alcohol prohibitions sometimes generate underground criminal syndicates more destructive than the drugs and alcohol that were prohibited, and the Affordable Care Act may have induced some employers to stop hiring and to reduce some employees' hours to avoid responsibilities for providing health insurance under the law. Raising the possibility of unintended consequences is a standard way of opposing some change to the status quo. Albert O. Hirschman incorrectly classifies the idea of unintended consequences as a "reactionary narrative,"[7] but there are plenty of examples of liberal and progressive arguments which raise the possibility of unintended consequences, like the notion that banning abortion will not lower the rate of abortion but will make having an abortion more dangerous. Friedman's conclusion is that "misguided policy missiles can land anywhere," meaning that we truly must survey the entirety of society to be sure our policies will have salutary effects.[8]

Friedman contends, following Lippmann again, that the use of heuristics to reduce all this complexity is necessarily unreliable:

> A stereotype makes part of society legible by mentally predisposing us to see aspects of the otherwise formless chaos as objects connected by causal relationships. A stereotype also highlights whichever pseudo-features of society seem to confirm the accuracy of the stereotype.[9]

This gives an explanation for *confirmation bias*, the human tendency (which we'll consider later) to seek out and give more weight to observations that confirm views we already hold than observations that disconfirm them. This is why ideational *determinism* is what I call a "tunneling theory": because once a stereotype or heuristic takes hold of us, we seem, Friedman thinks, to get kind of stuck in interpreting the world with it. (We saw that some theorists of ideology also posit that ideology has a self-justifying, effectively circular character.) But just one more part of Friedman's argument is required to really ensure that this is bad news for our political beliefs.

The final element of Friedman's argument is the charge that our choice of stereotypes to interpret the world "precedes the use of reason," that it "imposes

a certain character on the data of our senses before the data reach the intelligence."[10] If this were not the case, then we could still make some sort of judgment about which stereotype is best; as it is, our need for cognitive shortcuts generates "an *unavoidable* source of epistemic partiality"[11] – with "partiality" here meaning, I think, a kind of arbitrariness which isn't itself intellectually justifiable and which we have no reason to favor over anyone else's set of cognitive shortcuts. In the language of analytic philosophy, we might say that Friedman and Lippmann believe that "perception is theory-laden." I don't have the space or wherewithal to evaluate this claim in the philosophy of perception.

Friedman's theory is itself rather complicated, and he often makes multiple arguments for the same point or draws out multiple consequences. For instance, he argues that our need to use cognitive shortcuts also makes us unpredictable to policy experts, which is itself yet another source of complexity that results from our original encounter with a complex world.[12] But the main thrust is as above: we must use prerational heuristics to structure our perceptions of an intractably complex world, and those heuristics are self-confirming, meaning that our initial choice of heuristic is an explanation of many of our political beliefs, and for the disagreements we have with others who chose differently. However, people's choices of stereotypes and shortcuts remain unexplained in Friedman's theory; they are sort of arbitrary choices of foundations of political belief. So something seems to be left out, which – even if Friedman is right about stereotypes, complexity, and so on – would need to be explained by some further theory, perhaps one of the ones we've considered already in this text.

Discussion question

1. What are some cognitive shortcuts you use to simplify the complex modern world? How might these heuristics be related to the political beliefs you've ended up with?

11

MINIMALIST AND ELIMINATIVIST THEORIES

Here's an assumption we've held throughout the text so far: people actually *have* political beliefs, and those beliefs are what spur their political actions. In fact, there's a larger assumption at work in virtually everything we've discussed so far: understanding how people form their political beliefs will help us understand politics, and in particular, according to most of the writers we've seen, understanding the irrational influences on people's political beliefs can help us understand some unique dysfunctions of the political realm (though of course these writers may disagree on just what those dysfunctions are, itself a political belief of sorts). A few writers contest this, however. This is the realist or materialist tradition we've mentioned a few times. One way to think about this tradition is as a tradition of finding explanations for actions that don't involve sincere belief – recall the discussion of sincerity from the chapter on identity, signaling, and partisanship. In particular, for each theory we've already discussed, we can mock up a new theory which has all the same concepts, but which is pointed toward explaining people's political *actions*, but *not* their beliefs. As we go on we'll see why these writers prefer that kind of theory, but one reason that's especially relevant for our purposes is that it allows us to keep thinking that people's political belief-forming processes remain relatively rational. If people respond rationally to evidence in determining their beliefs and respond rationally to incentives in determining their actions, then that might be a comforting view for those of us (on some days this includes me) who want to maintain a broadly rationalistic and optimistic picture of human cognition.

DOI: 10.4324/9781003355274-13

Political expressivism

Michael Hannon has argued that what looks like political disagreement, at least in present-day United States, is often just "expressive discourse." Hannon makes three claims about the apparent extreme differences in opinions between the two major groups of American voters: first, those differences are smaller than the differences in attitudes, especially negative attitudes, held about one group by members of the other, what's sometimes called "affective polarization"; second, those attitudes often lead people to "deliberately misreport their beliefs as a way to *express their attitudes*," either to "cheerlead" their side when they have no reason to be truthful to a pollster or out of the assumption their side is right when they don't know much about the issue; and third, voters report themselves as having strong policy views and guiding ideologies when in fact they lack "robust" or "stable" political beliefs.[1]

Hannon musters a number of different sources of evidence for his thesis, which he calls "political expressivism." A theory of genuine political beliefs fails to explain some survey data while political expressivism succeeds in explaining those data. Survey respondents sometimes give truly absurd replies to polling questions. For instance, when asked to compare a photo of Barack Obama's 2009 inauguration to a photo of Donald Trump's 2017 inauguration, where the first photo obviously had more people in it, Trump voters were seven times as likely as Clinton voters to answer that the second photo had more people in it; Hannon asks: "Do these people really *believe* there are more people in the obviously half-empty photo?"[2] For basic perception to go so wrong in a political case would require very odd psychological mechanisms, and Hannon thinks his account does better by positing that people are simply lying to pollsters as a way to express their support for Trump. More generally, partisans from different parties respond more similarly to each other on survey questions when offered monetary incentives for right answers, and they become almost indistinguishable when incentives are also offered for admitting when they don't know an answer.[3] (How does this relate to our consideration of the relationship between belief and action? Why might we think that people's speech is a less reliable guide to their beliefs than their betting behavior?)

Hannon thinks political expressivism can explain a lot of other data about political beliefs as well. Some examples he gives include the following: the fact that people's apparent political beliefs track those of political parties; the fact that disagreement can persist even when evidence is "unequivocal"; the "backfire effect" in which efforts to correct a false belief can lead to that belief seeming to be held more, not less, strongly; political ignorance on factual matters; inconsistent or contradictory political beliefs; and, in general, the poor quality of political debates and the seeming impossibility of resolving them.[4] The most extreme form of Hannon's thesis of political expressivism has it that there is

nothing "underneath" the expressions. Our political behavior is not driven by political beliefs but by political identities, just as cheering or booing during a sports game is driven by one's team allegiances.

The best recent example I can come up with, of partisanship driving political expression with relatively fragile beliefs, is the early party-line response to the coronavirus pandemic. This early consensus was captured nicely by journalist Matthew Zeitlin on Twitter (now X); on February 1, 2020, he wrote:

> [T]his is just a superficial read of twitter, but response to coronavirus is on on [sic] the same politico-cultural lines as everything else[.] [C]ollege-educated liberals: quarantines are ineffective, the flu is more dangerous, relax[.] [C]onservatives: freak out, travel ban, quarantine.[5]

Zeitlin was completely right about the early partisan response to the pandemic and how it broke down by party lines. The later partisan response was, of course, completely the reverse of what he saw at the beginning of that February. But this just goes to show how fragile our political beliefs can be, and this fragility is nicely explained by political expressivism: they're more fragile than normal beliefs because they're not beliefs at all.

Political expressivism is a lot like the identity-protective cognition theory but for speech rather than action. A good question for discussion is how political expressivism might explain the motivated numeracy result from that theory.

Practical ideology

If Hannon's view is identity-protective cognition for speech or action rather than belief, then Olúfẹ́mi O. Táíwò's view is system justification, false consciousness, or ideology theory for speech or action rather than belief. This is what we called earlier a nobody-was-tricked approach. Táíwò writes that we should conceive of ideology as at least potentially resembling the story of the emperor with no clothes.[6] What Táíwò means is that the expressive and goal-directed activities people undertake in public actions are not necessarily straightforwardly reflective of their private beliefs. So politics is not best understood epistemically, and thus any examination of ideology should be "practice first."[7] One example Táíwò gives, which I will interpret perhaps a bit freely here, is of "the waves of disrespectful discourse and slurs that preceded the Rwandan genocide"; these did not *persuade* anyone that the Tutsis were insects but rather were "action-engendering" in the sense of providing information about "how to act," that is, about "incentives" and "schedules of social behavior."[8] Táíwò has redirected the process that seemed to form the irrational "dehumanizing" belief, which in some conceptions involves an actual confusion about who is a member of the human species, toward a rational belief with practical consequences having to do

with which actions will be punished and which actions will be rewarded. People whose interests are served by the dominant ideology act as though they believe it precisely because it serves their interests, and people whose interests are not served by it act as though they believe it because they could face severe consequences if they don't. But we needn't posit that members of either group actually *do* believe the dominant ideology.

Another example of this type is Megan Hyska's analysis of hard propaganda. Cases of hard propaganda involve the dissemination of "crude, heavy-handed, or preposterous" messages.[9] The goal of hard propaganda communicating some claim is not to persuade the viewer, listener, etc. of the truth of that claim but to persuade them of the power of the propagandist:

> [T]he state . . . aims to demonstrate its power, and so intimidate dissidents into silence. . . . One function of hard propaganda is then to convey the belief that just as the propagandist (in the above cases, the state) is powerful enough to dominate a media environment, they are powerful enough to crush resistance.[10]

Hard propaganda is a counterexample to views that hold that propaganda compels those exposed to it through irrational mechanisms.[11] This is because when exposed to hard propaganda, the belief I come to hold is in fact based on evidence – for instance, "[w]here I come to believe that the regime is powerful, it is because I have seen *evidence* of this – the expensive and logistically demanding domination of a media environment, for instance."[12] Then the message spreads not because it convinces people that it is true but because it convinces people that something bad might befall them if they do not spread it. Note, however, how different this is from what we saw, for example, Williams say about socially adaptive beliefs. He was concerned with the idea that something bad might befall people if they do not *believe* what they're not supposed to. These practical theorists are concerned with the idea that something bad might befall people if they do not *say* what they're supposed to, no matter what they believe.

Why I'm not an expressivist

I think it's really good that minimalist and expressivist theories are part of the conversation about political beliefs. And they probably do capture some phenomena of political behavior in certain cases. However, I'd find it really difficult to be a wholesale minimalist or expressivist myself – to take this view about *every* person and *every* issue. First, plenty of people respond well to reasoned arguments about politics. They take their time to consider objections to their political beliefs and either change those beliefs or come up with counterarguments in return. Some people even develop doubts of their own accord and

start searching out objective data or contrasting viewpoints to consider. But if apparent political beliefs were merely an illusory artifact of political expression, all this would be a bit mystifying. Second, political beliefs carry risks which might be hard to explain through an expressivist lens. While Táíwò is right that political behavior can sometimes be explained by self-interest rather than sincerity, political behavior sometimes goes against the individual interest of the actor, as when someone dissents against an orthodoxy with a risk of ostracism, protests with a risk of arrest or physical harm, and so on. In fact, we've seen that the very notion of "interest" is contested in politics, because many theorists think the causally efficacious interests are those of the group rather than the individual. Third, some arguments against expressivism in other areas also work against political expressivism.[13] For instance, political beliefs don't just occur on their own but can (and even characteristically do) function as parts of arguments about political action. We might say things like: "If taxes ought to be raised, then we should vote for Democrats." But if "taxes ought to be raised" and "we should vote for Democrats" are just expressions of raw attitudes, not genuine claims, then such a statement could never be right, because there would be no conflict in holding different *emotional* attitudes about those things. The conflict arises only if the attitude in question is a genuine belief, credence, or something similar.

Discussion questions

1. Is it possible you or some of the people with whom you discuss politics don't actually have political beliefs? What other explanations for political actions, not involving political beliefs, could there be?
2. What do you think about the whole approach of questioning people's sincerity in expressing their political beliefs?
3. Are there things you'd be willing to say but unwilling to bet on? Why might someone think betting is a good guide to our beliefs? Is there a more general principle at work about seeing our beliefs tested somehow? (We'll discuss this in later chapters.)

PART III
Sources

12

DEBUNKING AND RATIONALIZING POLITICAL EPISTEMOLOGY

In the previous part we considered a slew of theories about the formation of political beliefs. In this chapter we'll consider a type of argument that can bridge the gap between an agent's belief that they've formed a subset of their beliefs in a certain way and the rationality of maintaining those beliefs. What philosophical principles are at issue when evaluating our beliefs on the basis of what's caused them? In philosophy, this is called the theory of *debunking arguments*; a debunking argument is just an argument that features some empirical premise about the cause or explanation of a belief and some epistemological premise about rationality, with the conclusion that the belief is irrational or should be abandoned or something similar. In addition to the theories of belief considered earlier, we'll take a quick look at other kinds of biases that may affect our beliefs. It turns out that, as intuitively irrational as some of these belief-forming mechanisms might seem, it is very difficult to state the epistemological principles by which we could conclude that our beliefs would be irrational or that we should abandon them. And yet it is even more difficult to deny that in some cases, knowledge of how a belief was formed can render that belief irrational to maintain. After we've looked at debunking arguments, we'll look at a quick overview of ways to argue that people's political beliefs are rational after all.

Irrelevant causes, convenient beliefs, biased beliefs, induced beliefs, evolved beliefs

In the movie *Beginners*, the main character draws a t-shirt that says: "My personality was created by someone else and all I got was this stupid t-shirt." Debunking arguments express a similar worry: that our beliefs were in a sense generated

DOI: 10.4324/9781003355274-15

by "something else" – something other than our rational evaluation of the evidence. Most of the political theories we considered earlier are examples of that sort of "something else," but here are a few more general categories.

In the most general case, we might worry that our beliefs are caused by something *irrelevant to* whether or not they're true. A common example comes from the political philosopher G. A. Cohen, who writes that when he was studying to become a philosopher, a certain intellectually crucial notion tended to be accepted by philosophers who had studied at Oxford but rejected by philosophers who had studied at Harvard.[1] Likely most members of both groups of philosophers thought they were deciding rationally based on the evidence. But an Oxford philosopher who accepted this notion, or a Harvard philosopher who rejected it, might rationally wonder: Have I *really* judged the evidence rightly? The pattern of acceptance and rejection might make them think they would have judged it differently if they had gone to a different school. But the school one philosopher happens to attend is irrelevant to the truth of philosophical notions. Irrelevant causes are often brought up with phrases like "You just believe that because. . . . "[2] For instance, "You just believe that because you attended Oxford," "You just believe that because you watch a certain news channel," "You just believe that because you grew up in a certain state," and so on.

A subclass of irrelevantly caused beliefs is *convenient beliefs*. "Convenience" isn't a philosophical notion, but it has something to do with things going smoothly. So imagine I hold the belief that I'm a charming conversationalist and pleasant to be around and that I find this belief comforting in some way. Now I attend a cocktail party and notice that people keep making excuses to stop talking to me. If I come up with a bunch of explanations of this trend that have nothing to do with whether or not I'm charming or pleasant to be around, we might think those explanations are rather convenient. As above, someone might say: "You just believe that because you don't want to consider the possibility that you're not as charming as you think."

Very close to the notion of convenience is the notion of *bias*. In fact, most claims of convenience can be expressed with the lay notion of bias; in the previous example, for instance, we could say that I'm biased in favor of beliefs that construe me as charming and pleasant. Philosophers differ on the nature of biases; some take them to be beliefs themselves while others take them to be attitudes, processes, dispositions, or traits.[3] Philosophers also disagree on the relationship between bias and rationality and warn against attributions of bias. We'll talk about bias a bit more very soon.

Another kind of irrelevant influence can result in what I'll call an *induced belief*. There are varying degrees of strength by which a belief could be induced. A really strongly induced belief might be sort of pushed on someone by something akin to brainwashing. Somewhat weaker, but much more common, is really serious rote learning in childhood, and weaker than that is learning "by osmosis"

in childhood – norms and beliefs that just sort of seemed to be in the air that we might have picked up in our early years. Some of our moral, religious, and (yes) political beliefs might be induced in these ways. After childhood examples come weaker adulthood examples, like those in which the social groups we become members of or the news channels we watch influence our beliefs. Cohen's example of Oxford and Harvard is likely of this type. A common feature here is that there are usually actually existing situations in which people took the other belief-forming avenue. For instance, people of one religion who think they likely believe the dogmas of that religion because of the household they grew up in have people who grew up in households of a different religion to whom they can compare themselves. A general soft kind of induced belief is at issue in what Nathan Ballantyne calls the *problem of historical variability*, which is that there are propositions we accept (or reject) which we would reject (or accept) "if [our] background[s] had differed in certain respects."[4]

As just one example of a debunking argument along these lines using considerations we raised in Part II, Hrishikesh Joshi has argued that many political beliefs are socially adaptive in the sense of Williams from the previous chapter. In particular, Joshi thinks that we are very likely to form socially adaptive beliefs

> when three conditions are present: (i) the costs to the individual of being wrong are negligible, (ii) the beliefs fall under sufficiently intense social scrutiny, and (iii) the evidential landscape relevant to the beliefs is sufficiently complex so as to make easy verification difficult to come by.[5]

He writes that these conditions are in turn likely to obtain with regard to what he calls *creedal beliefs*, or "moral, political, religious, sectarian, or ideological assumptions that serve to bind communities together."[6] Note that (ii) is a consideration we saw in Chapter 8 and (iii) is a consideration we saw in Chapter 10. (i), on the other hand, is a consideration we saw to an extent in Chapter 11, when we discussed people's betting behavior being different from their speaking behavior. This idea, roughly that people are much worse at rationally forming beliefs when nothing practical is at stake, is often broached in lay conversations and likely has a pedigree in the psychology of belief[7] but does not make its way into epistemology very often.

Finally, a philosophically important type of irrelevantly influenced belief is the *evolved belief*, which features in evolutionary debunking arguments that target moral beliefs.[8] In an evolutionary debunking argument, we explain moral beliefs by reference to evolution, but then claim that there is no reason to think that evolution would have selected for *true* moral beliefs, which challenges not just the moral beliefs some individual might have but the entire edifice of moral realism. Since moral beliefs are often political beliefs as well, this challenge

might be important for our purposes, too, but we will leave it aside for debunking arguments that are more specifically applicable to politics.

It's tempting to respond to debunking arguments by saying: "Maybe my belief was caused by some irrelevant influence. So I got lucky. I'm still right!"[9] However, it's obvious that in many cases of irrelevant influences this answer would be absurd. Say I was induced to believe something through hypnosis or that my parents flipped a coin to decide whether to brainwash me about a certain subject. In such cases I could not plausibly just hope that I had been lucky. Rationally, I would have to decrease my confidence in the target belief. To believe rationally, we need evidence that we *actually are* lucky.

One *can* be both lucky and rational, though. For instance, imagine that your parents rolled a die when you were born. For each number from one to five, they had assigned a different cult belief system into which they'd indoctrinate you if that number was rolled. But in point of fact they rolled a six and decided not to indoctrinate you at all, letting you perceive the world in a natural way and make up your own mind (let's say rationally) about things. In this case, the fact that you end up with the beliefs you do is a matter of luck, but it's hard to see how the indoctrination that *didn't* occur could prevent those beliefs from being rational. Roger White gives a similar example:

> I will toss a coin now. If it lands Heads I'm going to come and kidnap you and force you to join a cult that teaches that President Obama is a disguised alien bent on colonizing the Earth from Mars. If it lands Tails I'll leave you alone. Your future political beliefs depend on how this coin lands. You will be lucky indeed if you end up believing something remotely true. Phew! It landed Tails. (If it had landed Heads I couldn't make my philosophical point). Your subsequent political judgments have not been impugned in the slightest by my experiment.[10]

So, clearly, the fact that I form or retain a belief merely by luck is not necessarily a problem.

Debunking arguments against political beliefs can be constructed based on all our theories of political belief.

> You just believe that because you have a certain personality type.
>
> You just believe that because you're stuck in a certain ideology.
>
> You just believe that because it's socially appropriate in your group to believe it.
>
> You just believe that because you choose to simplify the world with a heuristic that supports it.

There may be two exceptions. One is from the theories of positionality. On these, we might think that the marginalized standpoint is actually in touch with the

facts, and thus not susceptible to debunking. The other is from the minimalist and eliminativist theories. If I don't have political beliefs at all, then you can't say I have irrational political beliefs, either. Of course, if I'm also going around *acting like* I have political beliefs, something might still strike you as off about my beliefs or how they relate to my behavior. Regardless, we'll now look at another source of debunking: biases.

Bias and rationality

Talk of bias in human judgment is ubiquitous today, with people constantly accusing each other of bias and confessing to their own biases. Spotting bias is considered essential to critical thinking, and people are eager to frame all sorts of things in terms of bias, even when it doesn't make much sense (e.g., "This is my favorite song, but I really like this artist, so I'm biased"). In a survey of the history of the psychology of bias, though, Thomas Gilovich and Dale Griffin explain that this emphasis on bias followed a period of emphasis on *rational choice*. In their telling, the rational choice framework assumes that people make sophisticated calculations about probability and utility in their judgments and that their mistakes are "unsystematic."[11] In the wake of experimental and theoretical challenges to the rational choice approach, Daniel Kahneman and Amos Tversky developed a different research program, called the *heuristics and biases approach*. Gilovich and Griffin write that "Kahneman and Tversky were convinced that the processes of intuitive judgment were not merely simpler than rational models demanded, but were categorically different in kind."[12] The heuristics and biases approach works by identifying heuristics that enable quick, intuitive judgments and then associating them with biases or "departures from the normative rational theory that served as markers or signatures of the underlying heuristics."[13] So the essential contention of this program is that we use shortcuts to make our reasoning quicker and that these shortcuts can be detected by looking at the ways in which we get things wrong. Such shortcuts, it's often said, render our thinking irrational.

Some psychologists think that among our biases is a bias about whether we and other people are biased. We seem to have a "bias bias" or a "bias blind spot" which "results in the conviction that one's own judgments are less susceptible to bias than the judgments of others."[14] One posited cause of this bias includes the fact that we rely on evidence from introspection in our own case while relying on third-personal evidence when evaluating others; another is just the desire to think of oneself as unbiased.[15]

Cognitive biases have been taken to explain political disputes. For instance, in a 2019 *Vox* article, psychologist Brian Resnick detailed "nine essential lessons from psychology to understand the Trump era," all of which involve us reasoning in bad ways.[16] In addition to the biases inherent in the theories of political belief considered earlier, one of the most commonly cited political biases is

confirmation or *myside bias* – really a pair of biases: roughly, we like evidence that favors what we already believe, and we dislike evidence that disfavors what we already believe (or want to believe).[17]

The heuristics and biases research program has been challenged generally by those who disagree about whether the heuristics that it's found are in fact irrational.[18] In an online article, Hanno Sauer notes that, like most of the social psychology we've looked at so far, the heuristics and biases program has been beset by failures of replication.[19] And philosophers have argued that some purported biases are in fact rational; we've seen this with the proposed status quo bias already. Kevin Dorst has argued that confirmation and myside biases can result from rational processes of selective scrutiny.[20] More generally, it can be rational to discount evidence that runs counter to our beliefs somewhat, just as I might doubt a weather report that says it's raining if I couldn't hear rain on my window. As Neil Levy writes, updating our beliefs based on evidence should be "sensitive not only to [an] estimate of how likely the belief is, given the evidence, but also how likely the evidence is, given the belief."[21]

In his book *The Bias That Divides Us*, Keith Stanovich takes a roughly similar stance on these kinds of cases.[22] However, many instances of myside bias don't have this structure. Stanovich gives the examples of "rulings on the adequacy of challenged ballots in elections [being] infected by a partisan bias" and of

> subjects who believed they had worked more on [a] task tend[ing] to think they should be paid more[, with] subjects who believed they had worked *less* on the common task tend[ing] to think both workers should be paid the same.[23]

These aren't cases of people sticking to their guns but people forming beliefs in line with their interests, similar to socially adaptive belief.

Another class of widely cited biases is identity biases: racism, sexism, and so on. These are frequently studied and explained through the research program of implicit bias. But as we mentioned earlier, the implicit bias program has its flaws as well,[24] and the relationship between implicit bias and explicit belief or action, such as might impact politics, is contested.[25] Psychologist Lee Jussim has challenged more broadly the notion that people tend to stereotype each other in inaccurate ways, arguing instead for the thesis of *stereotype accuracy*.[26] Note that the claim that stereotypes are accurate says nothing about the causes of the conditions behind that accuracy. Take the stereotype that in heterosexual single-income marriages, husbands are more likely to work than wives. This stereotype is accurate – it corresponds to a fact – but the explanation of that fact needn't be some deep or innate difference between men and women. Rather, it could be sexism itself, perhaps through a self-fulfilling prophecy, which causes this stereotype based on sex to be true.

The upshot is that research into biases and pervasive human irrationality is hotly contested. The results themselves are up for dispute, and so are their

implications for whether humans are irrational. We'll leave aside empirical questions for now and focus on philosophical machinery.

Cudgels

We haven't been too strict about the form of arguments in this textbook, preferring to focus on other things. But it's important for debunking arguments, because philosophers haven't yet agreed on the precise nature of the bridge principles necessary for debunking arguments. A debunking argument includes an empirical premise describing the way a belief was formed (and picking out some feature of it like irrelevance, bias, convenience, and so on) and a philosophical premise holding that that way of forming beliefs is irrational. I call this philosophical premise the *cudgel* of the debunking argument.[i] There are multiple cudgels: an irrelevance-cudgel, a bias-cudgel, an evolution-cudgel, and so on. The formal statement of a cudgel might be something like "If a belief is caused by X, then it's irrational," where X might be "bias," for example.

One kind of argument in favor of cudgels comes from a principle of *sensitivity*.[27] Principles of sensitivity have it that a belief that some proposition is true can be rational only when it's sensitive to the truth or falsity of that proposition, and someone's belief that a proposition is true is sensitive to the truth or falsity of a proposition if and only if were that proposition false, they would not believe it. So, for instance, my belief that it's dark outside is sensitive to the actual facts about the darkness, since in looking out my window I would be able to tell if it were not dark outside. These principles give a clear explanation of some cases of debunking. For instance, if I was raised to believe a certain dogma, I might think that, even if that dogma were false, I would still believe it; if I were highly motivated to think of myself as charming and pleasant, I might think that, even if I were awkward and unpleasant, I would still believe myself to be otherwise.

Unfortunately, sensitivity has a lot of problems. There are always *some* situations – like the classic skeptical scenarios of dreaming, illusion, one's turning out to be merely a brain in a vat, and so on – in which one would believe what one now rationally believes even if it were false. Further, I take it[28] that there are beliefs that are rational or justified but false. For instance, having bought milk yesterday, and having checked my refrigerator overnight to make sure the milk was still there, I might rationally believe upon waking that there was milk in my refrigerator, even if, for the first time in my life, a thief had broken into my house simply for the purpose of stealing my milk and had, against all odds, not awoken me in doing so. Obviously, such beliefs aren't sensitive in the way the principle

i This comes from an old, rather silly name I tried giving to one such premise in a previous version of this chapter, the *causation undermines justification lemma*.

of sensitivity requires; they *are* false and they *are* believed, so it can't be the case that, were they false, they wouldn't be believed.

Taking sensitivity too seriously also gets some cases wrong. Thomas Kelly gives such a case:

> A parent watches her young child playing normally. The parent can plainly see – and thus, knows – that her child is alive and well, just as anyone else who is viewing the same scene can know the same proposition. However, the parent's belief that the child is alive and well is insensitive: if the child wasn't alive and well, the parent would still believe this, because the parent is so deeply invested in it being true that the child is well, and her desires would ensure that she believes accordingly. If credible evidence began to emerge that the child wasn't alive and well, this would trigger psychological mechanisms that would lead the parent to dismiss that evidence or explain it away so as to allow for the retention of the desired belief. Indeed, these psychological mechanisms would be efficacious in ensuring that the relevant belief continues to be held even if the evidence against that belief became very strong.[29]

This is a bit like a Frankfurt case (see Chapter 17). Sensitivity is stated in terms of what would have happened otherwise, but the "otherwise" might not be sufficiently symmetric to the situation that actually did occur to bear on it. Here, intuitively, the parent rationally gets things right in the actual case even though she would irrationally get things wrong if things had been different.

Sensitivity might also face a self-defeat objection. Self-defeat objections arise when some thesis would imply that it itself is false or otherwise problematic. For instance, a famous principle holding that only empirically verifiable statements are meaningful may be meaningless by its own criterion, as argued by Carl Hempel among others. A nihilist view that nothing is true would undermine itself, since if it were true, it would not be true. Amia Srinivasan has similarly argued that using principles of sensitivity to power debunking arguments is self-defeating, since a belief in such principles is itself susceptible to such arguments.[30]

White, who also argued against sensitivity justifications for debunking arguments, considers an *explanatory* principle that might justify our debunking cudgels. An explanatory principle would have it that someone's belief that some proposition p is true can be justified only if they think the best explanation of why they believe that p makes some reference to the fact that p is, in fact, true. Again, the best explanation of why I think it's dark outside is that I looked outside and saw that it was dark. Similarly, the best explanation of why the person raised to believe some dogma is that they were raised to believe that dogma, not that the dogma is true. But White thinks such a principle makes a mess of our rational predictions about the future. I think that it will no longer be dark in, say,

twelve hours. But time moves forward, not backward, so the fact that it will no longer be dark in twelve hours can't explain why I believe that it will no longer be dark in twelve hours. So under the explanatory principle, my belief that it will no longer be dark in twelve hours is not justified, an absurd result.

There's good reason to think *some* kind of cudgel must work, as a bridge to judgments about irrationality. But cudgels, and the debunking arguments they power, also face direct challenges. One worry is avoiding skepticism. All beliefs have *some* irrelevant causes; White writes: "Any belief . . . can be traced back to innumerable causes from far and wide. Most of these seem obviously irrelevant to the question of whether [the belief is true]."[31] If our cudgels are too strict, we might have to abandon all or nearly all of our beliefs, embracing skepticism, which is generally taken to be a bad result.

White also objects that challenges from debunking arguments might boil down to challenges from disagreement, which we'll consider in the next chapter. Even in cases where there's no actual person who disagrees, White thinks that debunking arguments might show up *potential* disagreement; for example, had Harvard not existed, Cohen could have somehow imagined it.[32] And he suggests that the following two cases would lead to the same epistemological result:

Correlation: Equal numbers of students go to Oxford and Harvard. Those Oxford graduates believe p: Harvard grads believe $\sim p$.

No Correlation: Half of all philosophy graduates believe p, the rest believe $\sim p$. But opinions are randomly distributed with respect to grad schools.[33]

However, White has missed something here which shows that the structure of debunking is often not quite the same as the structure of disagreement. The problem comes from his using the same proportion in both cases. Now consider the following cases:

Different Size Correlation: The twenty thousand students at State think their team will beat Tech in the big game. The five thousand students at Tech think their team will beat State in the big game.

Different Size No Correlation: Sixteen thousand students at State and four thousand students at Tech think State will win the big game. Four thousand students at State and one thousand students at Tech think Tech will win the big game.

Here there is an obvious difference between cases. In Different Size No Correlation, there's a widespread consensus, independent of school attendance, that State has the advantage. In Different Size Correlation, there's no such thing, and the difference is easily explained. So debunking and disagreement work differently "by the numbers," and this plausibly comes from the fact that debunking

shows up *dependence* in views. We'll talk about disagreement by the numbers and independent judgment in the next section and again in the next chapter.

Out-group homogeneity bias

Social psychologists began investigating out-group homogeneity effects in the late 1970s and early 1980s.[34] Bernadette Park and Myron Rothbart described "the out-group homogeneity hypothesis" in 1982 as "predict[ing] that (a) the perception of variability within a group is influenced by one's status as an in-group or out-group member, and that (b) the effect should be symmetrical for both parties of the in-group-out-group dichotomy,"[35] while a 2007 meta-analysis described the effect as simply the "tendency for persons to perceive more variability among ingroup members than among outgroup members."[36] So the out-group homogeneity effect is basically this: that agents generally perceive ingroups as containing more among-member variation than outgroups do. Once we have ideas of independent judgment in hand, we can speculate on how out-group homogeneity might affect our evaluations of the judgments of other groups and their members. However, note that out-group homogeneity also makes debunkings more plausible. If we think everyone in the out-group agrees about everything, then we can posit a unitary explanation for all their beliefs and provide a debunking argument based on that explanation. If instead we conceive of the out-group as being fractious and having a lot of internal conflict (as most groups are), then we'll have to explain our opponents' views more piecemeal, and some of them may be easily debunked while some of them may be harder to contend with. Thus, looking for internal rifts in outgroups may be, among other things, a good counterstrategy to polarization, a widespread phenomenon we'll consider in Chapter 19.

Rationalizing strategies

We've seen that some philosophers and psychologists take a rationalizing rather than debunking approach to political beliefs. There are a few common strategies for this, some of which we've seen and some of which we'll continue to see. One is what I call *rationalizing by truth*, the strategy of arguing that since some way of forming beliefs or some intellectual phenomenon does sometimes get one to the truth, it can't always be irrational. We'll discuss examples of this in the chapters on polarization and conspiracy theories, but the main problem with this strategy is that while truth and rationality are connected, the connection is complicated, and true beliefs can be irrational, so that one must do more than show that a process sometimes reaches the truth to show that that process is not irrational.

Another is what I call *rationalizing by non-ideality*, the strategy of arguing that although some way of forming beliefs might not be rationally ideal, it makes

sense given the limited resources that human cognition can draw on for gathering and evaluating evidence. Regina Rini argues in this way about partisanship, and Endre Begby argues in this way about prejudice.[37] But the non-ideality strategy goes too far. Once we decide that rationality is relative to human capacities, *anything* that seems empirically to be a near-universal feature of human cognition will come out looking rational. But that amounts to stripping the question "Are human psychological practices and mechanisms rational?" of any real content. Even if it's correct, the proposal that the question of rationality means something different than we might think is not bound to be satisfying.

A strategy we've seen is what I call *rationalizing by redirection*, the strategy of arguing that the beliefs challenged by a debunking argument are not actually the beliefs held by the agent. Perhaps the agent is acting insincerely or hiding their beliefs for some reason. Though it's not their main focus, this strategy can be associated with the minimalists, eliminativists, and expressivists about political beliefs and is explicitly avowed by Megan Hyska in her work on propaganda.[38]

Another common strategy is *rationalizing to the group*, the strategy of arguing that although beliefs might be formed irrationally on the individual level, the sum of a bunch of irrationality is itself rational. A trio of philosophers of science calls this the Independence Thesis: "that epistemically rational individuals might form epistemically irrational groups and that, conversely, rational groups might be composed of irrational individuals."[39] Robert E. Goodin argues that a cluster of individual biases can be group-rational, in that when someone who is biased against a certain conclusion endorses it anyway, we can be really certain of that conclusion.[40] Some social institutions might work despite irrationality or even be helped out by it; for instance, we might think that if two opposing attorneys both believe strongly in their side's case, the judge or jury will end up hearing the best possible arguments on both sides. We will see the tension between individual and group rationality come out in our discussion of aggregate epistemic mechanisms like markets and majority votes.

The most extended rationalization project in philosophy comes from Kevin Dorst.[41] He writes that it is a defense of polarization in particular as rational, but what's really important in what he's defending is that the polarization is caused *predictably*, meaning that he is really arguing that it can be rational to be debunkable. His motivating example is interesting not just for his argument but as an example of what sorts of political beliefs can be debunked:

> I haven't seen Becca in a decade. I don't know what she thinks about Trump, or Medicare for All, or defunding the police.
>
> But I can guess.
>
> Becca and I grew up in a small Midwestern town. Cows, cornfields, and college football. Both of us were moderate in our politics; she a touch more conservative than I – but it hardly mattered, and we hardly noticed.

> After graduation, we went our separate ways. I, to a liberal university in a Midwestern city, and then to graduate school on the East Coast. She, to a conservative community college, and then to settle down in rural Missouri.
>
> I – of course – became increasingly liberal. I came to believe that gender roles are oppressive, that racism is systemic, and that our national myths let the powerful paper over the past.
>
> And Becca?
>
> You and I can both guess how her story differs. She's probably more concerned by shifting gender norms than by the long roots of sexism; more worried by rioters in Portland than by police shootings in Ferguson; and more convinced of America's greatness than of its deep flaws.
>
> In short: we started with similar opinions, set out on different life trajectories, and, 10 years down the line, we deeply disagree. . . .
>
> [W]hat's rarely emphasized is that this polarization is predictable: people setting out on different life trajectories can see all this coming. When Becca and I said goodbye in the summer of 2010, we both suspected that we wouldn't be coming back. That when we met again, our disagreements would be larger. That we'd understand each other less, trust each other less, like each other less.
>
> And we were right.[42]

Dorst is writing about a situation where he and his friend are both subject to a debunking argument: *You only believe that because of where you went to college and where you moved after that*. Dorst's central claim is that this can be rational, and he uses a few theoretical notions to make his case.

Dorst focuses on a certain higher-order conception of ambiguous evidence under which some piece of evidence e is ambiguous just when there exists some proposition p such that for any value x, the probability that the probability that p is true given e equals x is less than 1.[43] Dorst suggests that the real-world political environment includes much evidence of this sort and makes another kind of non-ideal argument: ambiguous evidence can polarize at least in part because of limitations of our cognition and in particular of cognitive search. Dorst illustrates this using a word-search task.[44] Word-search tasks involve filling in blanks to make words; Dorst uses the examples of "P _ A _ E T," which can be completed as "PLANET" and "P _ G _ E R," which cannot be completed at all. There is an asymmetry in how we respond to these examples: we are quite certain that the first example resolves, but even if we suspect that the second example will not resolve, it's hard to feel completely sure.[45] When I ran through such examples as a class exercise with a group of undergraduates, they reported being 100 percent certain that the examples that in fact resolved would resolve but only 80 percent certain that the examples that did not resolve would not resolve. Dorst thinks that political information is much like this sort of thing. In a polarized information

and media environment, the evidence we receive in favor of our political side is unambiguous while the evidence we receive against it is ambiguous. Thus, we focus on the clear evidence and believe more and more that our side is right. Dorst brings to bear an enormous theoretical apparatus, and I recommend his work, which I still am puzzling over in places, to anyone reading this book. However, I think it has some problems.

First, Dorst's explanation as given doesn't quite account for the strangest fact about the situation as he describes it. He thinks that in predictable polarization, the following four individuals can be rational *and* all recognize each other as rational: the two moderates before they go their separate ways; the liberal, at some later time, having gone one way; and the conservative, at some later time, having gone another way. But to recognize each other as rational puts certain pressures on these individuals. The liberal and conservative, if they recognize each other as rational, must think the other one is forming their beliefs based on good evidence. But maybe the presence of rationally compelling evidence for the opposite view *should* change their own view (see the next chapter for more on this).

Second, and relatedly, Dorst's system entails undesirable failures of reflection principles and questionable results about evidence. A reflection principle concerns the relationship a rational agent has with their known future beliefs or credences; in particular, reflection principles say that rationality demands deference to known future beliefs or credences. Intuitively, as Adam Elga has discussed,[46] reflection principles fail in cases where one's information or one's reasoning gets worse with time – forgetfulness, drunkenness, mental decline, and so on. However, Dorst explicitly argues that an agent A can be rational in, for instance, believing p at some time t even knowing that at some future time t^* (a) they will have more evidence about p, (b) that additional evidence will be good evidence, and (c) they will rationally disbelieve p. But this seems strange enough to think that something must have gone wrong. In fact, Dorst goes even past this and asserts that at time t^* there can be a second rational agent B who at time t believed all the same things (say p and q) as A on the same evidence as A, that between t and t^* B can get their own good evidence from an alternative source, that A and B can end up at time t^* with opposite belief-pairs (say, A disbelieves p but believes q, and B believes p but disbelieves q) and that all participants at all times can regard all other participants as being rational and having good evidence but facing no rational pressure to change their views.

Third, Dorst seems also to rely on a very strong non-ideality thesis and to sometimes work with a conception of rationality that has both epistemic and pragmatic elements. The agents he has in mind have limited time and limited cognitive resources (the management of which seems like a matter for practical rather than epistemic rationality), they're forgetful, they can't bring themselves to compare their own actual evidence with anyone else's, and they're unskilled

in recognizing which arguments are good and which are bad. Of course, in many ways, this picture fits most of us actual humans and certainly myself. But it is again strong enough that we should wonder to what extent allowing in so much non-ideality renders the discussion unresponsive to the core rationality question.

Finally, I think Dorst errs in saying that his picture of predictable political beliefs as rational will help people in a polarized society see each other more kindly, a claim he makes in several places. When it comes to debunking, the alternative to both sides being rational is both sides being irrational; there's symmetry either way, and in the latter case nobody has the right to simply stick to their guns.

Conclusion

Lots of philosophers, psychologists, and political commentators are interested in the question of how rational our political beliefs are. Debunking arguments can be based on broad theories of how we form our political beliefs or on individual biases that impinge on our decision-making. At the extreme, if such arguments are viable, we might be in the position of being rationally required to be political skeptics.[47] However, it is hard to find the right philosophical principle to power debunking arguments; among other problems, such principles might themselves entail skepticism about far more than politics. Countering debunking strategies are rationalizing strategies, but those strategies too seem unwieldy to me. The question of whether our psychological makeup dooms us to have irrational political beliefs remains, as far as I can tell, very unsettled at the time of this writing.

Discussion questions

1. Do you feel your reasoning is biased in general? What about when it comes to politics in particular? What sorts of strategies do you use to reduce your biases? What sorts of strategies do you use to try to detect other people's biases?
2. What are some beliefs of yours that can be traced to causes that might be irrelevant to the truth of the beliefs? What do you think impels you to maintain these sorts of beliefs? Have you ever attempted to abandon these beliefs?
3. Try Dorst's exercise on ambiguous evidence and cognitive search and see if the results are the same as the ones I got.

13

SKEPTICISM FROM WIDESPREAD DISAGREEMENT

You might think that when you disagree with just one friend or family member across the dinner table, there's not much pressure to change your view. There's a whole world out there beyond the dispute you're having. Maybe everyone else outside that room agrees with you. But what about when you know what people outside that room think? What about when you have access to polls or votes that seem to indicate the consensus view or seem to indicate that there is no consensus view? Or what about when the room is larger – perhaps when a few hundred people watch or even participate in the conversation, and you get a sense of their opinions as well?

I've mentioned that political skepticism – the view that no political beliefs are rational or justified, that we don't or even can't know anything about politics – is an overarching concern of this text. Here is a very roughly stated argument from political division to political skepticism: Lots of people disagree with you about politics. If lots of people disagree with you about something, then you should think that there's a pretty good chance you're wrong about it – some of these people might know just as well as you. And if there's a pretty good chance that you're wrong about something, then you should abandon your quite possibly wrong belief. Therefore, you should abandon your beliefs about politics.

The premises of this argument probably seem pretty broad. And we'll challenge them one by one in just a moment. But for now, let's see by analogy why this argument might seem initially plausible. Imagine you're one of eighty passengers on a bus headed to a big concert. The driver asks which direction to go to get to the concert venue. You say that it's east. But the person sitting next to you says that it's west. In the seat behind you, one person says that it's east, but the other says that it's west. This goes on for each pair of seats in all twenty rows

DOI: 10.4324/9781003355274-16

on the bus. By the time you get to the eightieth person, you might start to feel pretty uncertain about your answer, and you might start to feel rational in that uncertainty. At the very least, you might start wondering: Why did I think it was east to begin with? Why do all these other people think it's west? Where did this disagreement come from? Are there outside sources we can consult? Your belief is now in question where it wasn't before.

First premise: lots of people disagree with you about politics

Why did the above situation seem analogous to politics? It was, after all, one of our starting points way back in the introduction: people disagree over politics; there is a wide range of politics beliefs out there. What could be more obvious about politics? But recall that we also mentioned that there were other perspectives on political disagreement. Apparent political disagreement could be an artifact of merely verbal differences; we could seem to disagree but just be using words differently. Or apparent political disagreement could be a kind of strategy in political struggles with other sources. It could be that people in general are just trying to win and that many people's political expressions are not honest. The appearance of disagreement then could be caused by any number of things.

Recall Michael Hannon's political expressivism, considered a couple of chapters ago. Can Hannon's extreme view undermine our sense that lots of people disagree with us about politics? There are a few reasons to doubt that it can. First, Hannon's argument seems like it might be just as sufficient to cast doubt on whether *we* believe what we think we believe about politics. The argument does not target everyone else's apparent political beliefs *but not our own*; it targets everyone's, period. So we might replace one form of doubt about our political beliefs – doubt based on the ubiquity of political disagreement – with another form of doubt – doubt based on doubt that we have our political beliefs as beliefs to begin with. To the extent that we make an exception for ourselves, that we say "Hey, Hannon might be right about some people, but he's wrong about me," there's little reason to think that there aren't people on the other side about whom that can be said as well. Nothing in Hannon's argument or evidence is partisan, although on the other hand nothing about it says definitively that both sides of contemporary politics are equally empty of political beliefs. Finally, Hannon's argument is based on survey data and empirical psychology; it doesn't seem to make a universal claim about the very possibility of holding political beliefs. All it can really show is that certain things might be explained if a lot of people don't actually have political beliefs. But this doesn't mean that *nobody* has political beliefs. And in that case, yet again, it would remain to be seen what the distribution of *actual* political beliefs looks like.

In popular commentary, pundits often take the view that the opposing side has few real supporters and that one's own side's policies actually have massive popular backing. Even writers whose ideologies are outside of the American

mainstream, like libertarian and socialist commentators, often make this sort of move. It is impossible to consider all the various strange hypotheses that might be put forward for why, if that were the case, American political support seems always to be more or less evenly split between two parties. If it is, there is a lot to explain when it comes to, for instance, voting behavior. For now, we will move away from political science and back to the second premise.

Second premise: disagreement means you might be wrong – others might know better

The second premise brings us back to social epistemology. One of the core questions of social epistemology is what's rational to do when we encounter disagreement, and one of the core questions of political epistemology is what's rational to do when our group encounters a group that disagrees with ours. To begin to answer the first question we need a notion of the relative epistemic standing of individuals, and to begin to answer the second question we need a notion of the relative epistemic standing of groups. We will hold off on the second question until just a bit later.

We make judgments about the relative epistemic standing of individuals all the time. We listen to all sorts of teachers and experts without necessarily questioning their reasoning or their command of the facts. When it comes to certain other people, we don't listen to them at all, because we have no trust in their reliability as potential knowers. Then there are some people who fall near us in their epistemic abilities in some domain, the people social epistemologists call our *epistemic peers*. In the most stringent definition, an agent's epistemic peer is someone who has the same evidence as that agent and an equally good capacity to evaluate that evidence as that agent. More loosely, we might think that the strength of one's evidence and the strength of one's capacity to evaluate evidence can be sort of added up and that an epistemic peer is someone with the same sum of those two things, what we might call a kind of epistemic rating. Looser still, we might think of people with similar though not strictly equivalent epistemic ratings as epistemic peers, and we might think of peerhood from the perspective of the reasoner in question, so that one reasoner ought to treat another as an epistemic peer when their best evidence indicates that they likely have roughly similar combined ratings when it comes to evidence-possession and evidence-evaluation.

The bus example above should show that in at least some cases, when lots of people disagree with you about something, it can mean that you might be wrong about it. But are our political beliefs like that? Maybe the people we disagree with about politics aren't our epistemic peers. There are probably lots of ways to try to establish this sort of claim, but here are three potential strategies. First, we might find ways to argue that the people we disagree with about politics aren't our epistemic peers in a very general sense. Perhaps we could show through

surveys or tests or some other mechanism that they just don't know many facts or aren't very good at analyzing the facts they do know. Second, as suggested by Adam Elga, we might think that nobody could be our epistemic peer if they disagree with us consistently in domains of value, like ethics and politics, because in those domains "one's reasoning about the disputed issue is tangled up with one's reasoning about many other matters,"[1] so that from our perspective they must be getting quite a lot of things wrong.

I don't think Elga's response makes sense in light of our previous discussions of the nature of political belief and the sources of political conflict. Recall that any belief can become a political belief if it becomes enmeshed in political conflict in the right way and that several different kinds of apparent disagreement can cause political conflict: verbal disputes, factual disputes, and disputes about value. On the first of these there is no real disagreement at base, and on the third of these, Elga's analysis may well hold. But what about political disagreements that have factual disputes at their core? Two people might disagree about gun control because they disagree over whether it will save lives; they might disagree over raising the minimum wage because they disagree over whether it will increase unemployment; they might disagree over universal healthcare because they disagree over whether it will lead to nightmarish waiting periods. Of course, in some cases, people who argue over those kinds of facts may be using them only to sugarcoat their real views, but in other cases that's really what's at issue, and differing empirical assessments and predictions lead to differing political views.

A final strategy is to give what's sometimes called an "error theory" of the opposing group's views. An error theory is, roughly, a theory of how or why someone else has made an error. Lay political epistemology is rife with error theories, and we'll go on to examine a variety of kinds of error theories as the text goes on: the other side reads the wrong news sources; the other side believes propaganda; the other side is subject to distorting biases; the other side is caught up in a crazy ideology; and so on. The idea is that if I have a plausible error theory about your belief, it justifies me seeing you as less than an epistemic peer. However, Jonathan Matheson has noted that in at least some cases, providing an error theory for another agent's belief does not undermine their status as one's peer, because one may be fully aware that the other agent can also come up with an error theory for one's own belief.[2] The point of the error theory was to be "symmetry-breaking" in such a way as to give one side a distinct epistemic advantage. But if there are objectively equally plausible error theories of both parties' beliefs, then the situation remains symmetric. Even if the other side's error theory doesn't seem plausible, they'll disagree with you about that – which you'll need another error theory to explain!

The upshot is that it's difficult, though not impossible, for a member of one political group to cast the opposing political group as systematically epistemically inferior in a way that doesn't just circle back to the political disagreement itself. And it is hard to say just what would be unique about political disagreement

that would distinguish it from other examples of disagreement where we do feel quickly that we might be wrong.

Third premise: if you might be wrong, you should abandon your belief

Fair enough; but then what should we do when we encounter an epistemic peer? In the jargon of philosophers, we have two options: we can *reconcile* our views with theirs, or we can remain *steadfast*. To remain steadfast means to keep our previous belief and our previous level of certainty in our belief – our previous *credence*. To reconcile means to move closer to the other side's belief. Theories on which we should always reconcile our views are called *conciliationist* while theories on which we have no such rational obligation are called *steadfaster*. There are also some more nuanced views on which conciliation is *sometimes* or *to some extent* rational, which we'll note at the end of the chapter, but our main concern is with arguments for and against conciliationism.

One source of support for conciliationist views is thought experiments like the bus example above. It certainly seems from such examples that evidence about other people disagreeing with us can rationally undermine our beliefs sometimes – that we ought to be less certain after encountering such evidence. Consider two more:[3]

> Suppose that five of us go out to dinner. It's time to pay the check, so the question we're interested in is how much we each owe. We can all see the bill total clearly, we all agree to give a 20 percent tip, and we further agree to split the whole cost evenly, not worrying over who asked for imported water, or skipped desert, or drank more of the wine. I do the math in my head and become highly confident that our shares are $43 each. Meanwhile, my friend does the math in her head and becomes highly confident that our shares are $45 each.[4]

> You and a friend are to judge the same contest, a race between Horse A and Horse B. Initially, you think that your friend is as good as you at judging such races. In other words, you think that in case of disagreement about the race, the two of you are equally likely to be mistaken. The race is run, and the two of you form independent judgments. As it happens, you become confident that Horse A won, and your friend becomes equally confident that Horse B won.[5]

In each of these cases, the first from David Christensen and the second from Adam Elga, it seems intuitive that we should become less sure of our view after learning of the disagreement. In fact, both Christensen and Elga suggest that what's rational in each case is to give *equal weight* to the friend's assessment. The reason is that, assuming we have already decided that the friend is an epistemic peer, it would be odd to give our own opinion more weight than theirs.

That would amount to a kind of dogmatism under which we could reason from the mere fact that we sometimes disagree with our epistemic peers to the conclusion that they aren't our epistemic peers at all.[6]

Other types of intuitions support conciliationism and the equal weight view more specifically. Ordinarily, when someone we more or less trust tells us something, we give what they say some weight, and it's hard to see why we should stop doing so if we already have an opinion in the matter. As Bryan Frances and Jonathan Matheson put it:

> Learning that a reliable inquirer has come to believe a proposition gives you a reason to believe that proposition as well. The existence of such a reason does not seem to depend upon whether you already have a belief about that proposition.[7]

Frances and Matheson also note that there are analogies between cases of competing reasoners and other cases of competing evidence. They are particularly fond of a thermometer case, in which two people have equally reliable thermometers, and both of them know the thermometers are equally reliable, but the thermometers happen to disagree about the temperature on some particular day. It would be irrational, they say, for each person to simply trust their own thermometer because it happens to be theirs, despite thinking that the other thermometer is equally likely to be right. So, too, should it go with human reasoners, which are input–output devices much like thermometers. Our "readings" of the world, our beliefs and credences, don't rationally merit any more attachment in the face of disagreement than the readings of one particular thermometer over another. These two intuitions can be combined. Say you and a friend are trying to get somewhere in a new city; you each ask a different passerby, but the two give different advice. It would be incredibly strange for each of you to simply trust the one individual you asked rather than pooling the evidence and remaining uncertain.

We should be clear about just what it means to give equal weight to two views. Take the case of the share of the check. The fact that I think our shares are $43 each and my friend thinks our shares are $45 each does not mean that in giving the two equal weight I should come to believe that our shares are $44 each. That would be odd indeed, since as far as we know, nobody believes that our shares come to $44 each. Rather, it means that I should come to have something like 50 percent confidence in the proposition that our shares are $43 each and 50 percent confidence in the proposition that our shares are $45 each. More generally, if we are epistemic peers and before learning of our disagreement my credence in some proposition is x and your credence in that proposition is y, then after learning of our disagreement, the equal weight view says that my credence should rationally be $(x+y)/2$.

There are some complications for the equal weight view. First, there is the question of how to generalize it to people who aren't quite our epistemic peers, especially given that we might not have any full epistemic peers in reality – it might be that on any given question we might give every other reasoner either an advantage or a disadvantage relative to ourselves. Second, and relatedly, there is the question of what happens if we encounter multiple epistemic peers in a sequence. Say S believes A and B to be S's epistemic peers, and say there's some proposition *p* that A and S both believe but B doesn't. First say S talks to A first and B second; then in giving all views equal weight, S sticks with *p* after talking to A, but then concludes that it's a fifty–fifty proposition after talking to B. But say instead that S talks to B first and A second; then in giving all views equal weight, S thinks that *p* is a fifty–fifty proposition after talking to B but then thinks of it as 75 percent likely after talking to A. A principle called *commutativity of evidence* holds that the order in which we get our evidence shouldn't determine our beliefs, and so this situation seems to mean that there's something wrong with the equal weight view.[8] The right solution is to say that epistemic peerhood is fragile and so that after conciliating with one peer I am actually better off than the other peer who hasn't conciliated with them. This makes sense – conciliationism holds that conciliation is rational, after all, and rationality means improving our beliefs – and it fits with the ideas of epistemic democracy we'll discuss soon.

One particular counterargument to conciliationism has seen a lot of discussion in the academic literature. This is the counterargument from what's called "self-defeat." As we've discussed, an epistemological principle is self-defeating if it somehow implies its opposite; that would mean that we couldn't adopt it without logical inconsistency. The self-defeat problem for conciliationism, and for the equal weight view in particular, is relatively straightforward. Lots of philosophers disagree about whether we should reconcile our beliefs with the beliefs of epistemic peers with whom we disagree. Assuming that those philosophers are our epistemic peers, to reconcile our beliefs with theirs would mean to doubt, and perhaps abandon, our belief in conciliationism itself. Adam Elga has argued that beliefs in principles like conciliationism are exempt from their own demands,[9] but this seems objectionably arbitrary: Why should those beliefs be different from any others?

Finally, the very availability of this argument for political skepticism is sometimes seen as a problem for conciliationist views of disagreement. We might think it is simply *obvious* that we needn't abandon our political beliefs (or other controversial beliefs) just because some people disagree with us, so that the argument *must* have a wrong premise somewhere. This way of reversing an argument because you're certain the conclusion must be false is sometimes captured in the phrase "One person's *modus ponens* is another person's *modus tollens*." Yet we will see more arguments for political skepticism throughout the book. To an extent it is a conclusion that can be reached cumulatively.

There are other philosophers who take more nuanced epistemological positions than the equal weight view and the steadfast view. For instance, Tom Kelly has written that whether it's rational to alter our belief in response to disagreement can depend on how strong our initial evidence for the belief is and how strong the evidence that comes from the disagreement is.[10] For Kelly, if we have a lot of initial evidence for the belief, the disagreement shouldn't matter as much, even if it comes from a peer. And Jennifer Lackey has written that whether it's rational to alter our belief in response to disagreement can depend on how strongly we hold the belief and what sort of information we think the person disagreeing with us has access to.[11] For Lackey, if we have a strongly held initial belief and suspect that we and our epistemic peer have asymmetric information, we needn't conciliate with them. It's natural to think that these views will be less conducive than conciliationist ones to the political skepticism entertained in this chapter but more conducive than steadfast ones. Just how they would shake out when it comes to political beliefs depends on how strong our political beliefs are to begin with, how much evidence we think we have, and what sort of evidence we think our epistemic peers on the other side might have. However, the existence of such views is not some sort of epistemological license to stop taking political disagreement seriously.

Conclusion

The skeptical argument from widespread disagreement has it that (a) many people disagree with us about politics, (b) this says something negative about the epistemic status of our political beliefs (perhaps only if we know it, but we all do), and (c) this negative epistemic status means that we should abandon those political beliefs. There are some reasons for doubt about all of these premises, but there are also some genuine considerations in their favor.

Discussion questions

1. Think about different domains of belief, like mathematics, music, sports, and everyday things like the weather, what time an event is scheduled for, and so on. How do you deal with disagreement about those topics? Now think about your political beliefs. How do you deal with disagreement about political beliefs?
2. Do you think there is a "third way" with regard to political disagreement, where we continue to respect those we disagree with but don't abandon our political beliefs?
3. This question anticipates the last chapter of the book, but, in addition to a diversity of belief in contemporary times, there have also been many people in the past who had different political beliefs than we do. How should we react to those disagreements?

14

DEMOCRACY

If a situation in which there is widespread political disagreement should produce skepticism, then perhaps a situation in which there is widespread political agreement should produce confidence. This is the insight behind the theory of epistemic democracy, which arguably shares some features with conciliationism.[1] If the people around us are our peers, and a majority of them have a certain belief, as expressed in a vote or a poll, then we seem to have reason to adopt that belief as well.

Aggregative epistemic democracy

The thinker most closely associated with epistemic democracy is the Marquis de Condorcet, a French Enlightenment philosopher and mathematician. In a simple application of probability theory and the so-called law of large numbers, Condorcet, in his famous *jury theorem*, proved that majority votes are incredibly reliable. The theorem makes a few rigid assumptions, which have been loosened in recent work in political science. The assumptions are that votes are *binary*, between two options; that voters are *independent*, their votes uncorrelated with one another; that voters are *competent*, meaning here that in any binary vote where there is a right option and a wrong option, they are more than 50 percent likely to know which option is right; and that voters are *sincere*, meaning that they vote in line with what they know or believe.

Deductive proofs of Condorcet's jury theorem can be found online easily. I prefer to show it through intuition. If voters are competent, sincere, and independent from one another, and voting on a binary choice, then if we start with two voters, they are more likely to be right than to be wrong (though they may be most likely

DOI: 10.4324/9781003355274-17

to produce a tie). Then if we add two more, they are by the same token more likely to move a wrong vote to a tie or a right vote than they are to move a right vote to a tie or a wrong vote, meaning that the likelihood of the vote being right goes up and the likelihood of the vote being wrong goes down. This happens every time we add two voters. For instance, say each voter is 60 percent likely to vote the right way on some question. Then, given independence, the probability of any two voters getting it right is $.6^2 = .36$, and the probability of two voters getting it wrong is $.4^2 = .16$, giving a .48 probability of a tie. So when we move to four voters, the probability of getting it right is $(.36 \times .48) + (.48 \times .36) + (.36 \times .36) = .4752$, and the probability of getting it wrong is $(.16 \times .48) + (.48 \times .16) + (.16 \times .16) = .1792$. This trend continues asymptotically: as we increase the number of competent, sincere, independent voters, we approach a probability of 1 that they get the right answer.

We will see some general challenges about political sincerity later in the text (Hannon's theory above might be viewed as one), and we won't worry too much about choices being binary. For now, let's focus on the notions of independence and competence, with just a little bit of sincerity to boot.

Earlier, we talked about social epistemology studying conditions on the epistemic standing of individuals and of groups. Well, a level of competence summarizes the epistemic standing of an individual, and a level of independence determines how competent a group is given the competence of its individuals. Imagine a hundred of your friends believe a certain thing. This gives you a really good reason to believe it, too. But then you find out that your friends all believe it because they read it in the same book. Now it might seem that, even though you have a hundred friends with that belief, you're really only dealing with one other reasoner: the person who wrote the book. When we reconcile our belief to those other people's beliefs, doing what Frances and Matheson call "disagreement by the numbers,"[2] it seems as though we must take that into account. Not all philosophers completely agree. Jennifer Lackey suspects that one person's judgments are never completely dependent on someone else's – for instance, we probably always do some sort of initial check that a claim isn't outrageous – so that numbers always matter in disagreement.[3] Regardless, we can still have concepts of relative dependence or independence, and of statistical independence, or lack of correlation, the notion that's operative in Condorcet's jury theorem.[4]

How is a group's competence determined by the independence of its members? Take an extreme case in which all voters vote the same way on every issue. Then the reliability of the majority vote is no better and no worse than the reliability of each member. The easiest toy case to see this in is a case with three voters who each get two out of three votes correct. Consider a situation of perfect correlation among voters:

	Issue 1	*Issue 2*	*Issue 3*
Voter A	Correct	Correct	Incorrect
Voter B	Correct	Correct	Incorrect
Voter C	Correct	Correct	Incorrect
Majority	Correct	Correct	Incorrect

Now compare that to a situation of perfect independence:

	Issue 1	*Issue 2*	*Issue 3*
Voter A	Correct	Correct	Incorrect
Voter B	Correct	Incorrect	Correct
Voter C	Incorrect	Correct	Correct
Majority	Correct	Correct	Correct

The advantage for majority votes of the case of perfect independence is clear.

There is an even simpler version of independent majority votes turning out epistemically well in the Miracle of Aggregation, which Hélène Landemore discusses in her book *Democratic Reason*.[5] The Miracle of Aggregation goes something like this. Imagine that on any issue there are people who know the right answer and people who don't. If the people who don't know the answer end up voting in a random, uncorrelated way, then their votes can be expected to cancel each other out, and the people who know, no matter how relatively few of them there are, can be expected to dominate the vote and drive it to the correct answer. But the assumption of randomness is too strong to really countenance.

How independent are people's political beliefs in the contemporary world? Many of the kinds of error theories for opposing political beliefs we discussed in the previous part are available precisely because different people's beliefs seem to be correlated. As we saw Hannon noting earlier, people's professed political beliefs often follow party platforms. Hannon was saying that to undermine the notion that people followed robust ideologies; but even if they do follow robust ideologies, that'll be a source of correlation, too. And as Hrishikesh Joshi has emphasized,[6] there are significant correlations among political beliefs, as well: we can predict what someone might think about gun control from what they think about immigration, for instance, and so on. Correlations might also emerge from voters having similar psychological mechanisms, heuristics, and biases. And it seems obvious that many people are also influenced in their political beliefs by the news they watch or read, the company they happen to keep, and so on. Even when these are good influences, there will be a cost to independence.

Independence can also be ruined through "belief cascades," in which a sequence of people defer (perhaps rationally) to the beliefs of just one or a few, making it so that a seemingly large group has the epistemic value of just one or a few people's thought processes.

In cases where these are malign influences, the failure of independence will also be a failure of competence. In fact, just as there must be *some* correlation among voters to produce competence,[7] there must be some correlation among voters to produce genuine failures of competence (in which voters do worse than chance; the case of randomness may be acceptable because of the Miracle of Aggregation described earlier). In other words, the most interesting kind of failure case for Condorcet's jury theorem is that of systematic biases among voters.

We considered some types of biases and ideologies above. Now let's consider a clever argument that Bryan Caplan gestures at in his book *The Myth of the Rational Voter*. Caplan notes that since elections are almost never decided by one vote, the expected marginal impact of a single vote is almost zero.[8] So it is likely irrational even to bother voting,[9] and it is particularly irrational to bother learning a lot about politics with the intent of voting correctly. So voters are either irrational or rationally ignorant. Either way, it's bad news for their competence. As I say, this argument is pretty clever. Here's an objection: it seems to make a kind of conflation, that we've already discussed, between rationality in action and rationality in belief. We could be rational in our beliefs despite spending what is, given our goals, an irrational amount of time and effort gathering information to make sure they're correct. However, this objection goes only so far. In some cases, when you know someone is irrationally obsessed with a topic, you trust them more, not less. Someone who is irrationally obsessed with learning about the public transit systems of some foreign country will probably be a great source of information about them. But other types of informational obsession seem more likely to produce false beliefs. Conspiracy theorists, whom we'll discuss in a later chapter, also seem to have this sort of irrational obsession, as do all sorts of fanatics and ideologues.

Some writers have argued that the most knowledgeable voters tend to be the most partisan;[10] if we think that partisanship produces irrationality (see earlier), this would mean that allegedly irrational levels of political information-seeking do in fact correlate with irrational beliefs. Jeffrey Friedman writes that this puts us into a dilemma:

> On the one hand, lacking comprehensive cognitive structures, such as ideological "belief systems," with which to understand politics, most people learn distressingly little about it. On the other hand, a spiral of conviction seems to make it difficult for the highly informed few to see any aspects of politics but those that confirm the cognitive structures that organize their political perceptions.[11]

Attempting to answer this dilemma, Adam Gibbons has argued that while well-informed voters look more partisan if the threshold for being well-informed involves only rather minimal acquaintance with political facts, the very best-informed voters, distinguished by a much higher threshold of knowledge, may end up looking less partisan and hence, maybe, more rational.[12]

When it comes to the average voter's competence, many writers have taken a dismal view.[13] Ilya Somin writes that "the sheer depth of most individual voters' ignorance may be shocking to readers not familiar with the research."[14] Jason Brennan speculates that "perhaps this is because when it comes to voting, knowledge and rationality do not pay, while ignorance and irrationality go unpunished"[15] – a theme we've seen before and will see again. They give the following sorts of examples: before the 2014 election, less than half of survey respondents knew which party controlled the House, and less than half knew which party controlled the Senate;[16] a third of respondents believed that foreign aid was a larger government expenditure than Social Security, transportation, and interest on debt, even though it was the smallest, with Social Security alone costing the government about seventeen times as much;[17] "majorities are ignorant of such basic aspects of the U.S. political system as who has the power to declare war, the respective functions of the three branches of government, and who controls monetary policy";[18] "during election years, most citizens cannot identify any congressional candidates in their district";[19] and so on.

The relationship between this phenomenon, often called *political ignorance*, and competence in voting is often taken to be straightforward: How could people who know nothing about what they're voting *on* somehow vote in a reliably competent way? But keep in mind what competence means: getting more than 50 percent of votes right. Robert E. Goodin and Kai Spiekermann suggest that "sheer ignorance, however, provides no reason to think that people would be worse than random on average (at least given a large number of people)."[20] The idea here is something like the idea from the Miracle of Aggregation: individuals who know nothing could go one way or the other, but there's little reason to expect that the *group as a whole* would *systematically* go one way or the other. But we might think that ignorant people are more likely to fall back on intuitive judgment, which is more likely to be biased, and biased judgment *would* be non-random. As before, the assumption that the votes of the ignorant will be randomly distributed seems like a prohibitively strong one.

What about sincerity? Again, for now we'll consider just one counterargument: strategic voting or strategic behavior more generally. In popular commentary it is often assumed that people vote strategically, so that differences in voting patterns among socioeconomic groups, for instance, can be explained in terms of those different groups' differing interests. Some disputants in the debate about epistemic democracy seem skeptical that political participation tends to be strategic. We saw that Caplan thinks voting is never a strategically rational

way to pursue one's interests, since one vote can't make enough of a difference to justify the cost of voting. And Friedman writes that, rather, "the consensus among political psychologists and survey researchers is that most citizens tend to participate in politics (when they participate at all) 'sociotropically,' not self-interestedly,"[21] where sociotropic political participation is that in which people work "for what they think will serve the economic [and now other] interests of everyone, or the majority, or those who most need help, in their society."[22] Similarly, Landemore talks about the "overwhelming evidence in favor of non-self-interested voting."[23] This view is not universal, though. Brian Kogelmann writes that "the most plausible theory of voter motivation is the *expressive theory*,"[24] a theory much like Hannon's theory of political expressivism. Note that although there is a distinction between voting a certain way because the outcome or its consequences will satisfy one's desires and voting a certain way because *voting that way itself* satisfies some desire, both are inconsistent with the sociotropic theory. Kogelmann also notes that there are other ways that a vote can be neither self-interested nor sociotropic.[25] For instance, a racist voter might vote against their interests to support a racist policy.

Goodin and Spiekermann give a few reasons to think that sociotropic political behavior is a relatively natural result. First, "voters are limited in terms of information and cognitive abilities – much too limited to engage in complex strategic calculations," so that we shouldn't expect them to vote for a self-interest they haven't been able to assess.[26] Therefore, voters are more likely to "fall back on the default rule, 'Vote sincerely'," for a variety of reasons like valuing honesty, expressing genuine viewpoints, feeling solidarity with their political community, and avoiding getting caught lying.[27] Second, they suggest that empirical evidence shows that people talk sincerely about politics during group deliberation.[28] Both of these explanations, however, have some potential problems. First, it isn't always a complex matter to assess one's interests in voting. Often, different candidates appeal directly to different communities. Even when they don't, humans have strong social instincts about who to trust which can direct them. And while deliberation may induce sincerity, it also can reduce independence. The Condorcet assumptions must be treated as a whole.

Remember what we discussed at the beginning of this chapter: the project of epistemic democracy is somewhat different from ours in this text. Epistemic democrats are generally interested in justifying democratic political systems on epistemic grounds. We are interested in if we can, and how we might, justify our political beliefs, and this book is being written at a time when many political systems are democratic. Like most sources of evidence about our political beliefs, it seems that polls and votes can be very reliable in some circumstances and very unreliable in others. What is interesting about research into epistemic democracy is that it helps point us toward what are perhaps the main conditions governing that reliability: sincerity, competence, and my favorite, independence.

Deliberative democracy

There is another line of reasoning about democracies making good decisions: the deliberative tradition. Deliberation in this context means, roughly, talking in a group, using reasons and arguments to try to figure out the right course of action. It is the group equivalent of what might happen internally when you try to come to a decision about what you will do. Deliberation thus sits nicely with our account earlier in the text of just what political beliefs are. While the aggregative tradition in epistemic democracy prizes independence and might be seen as relying on the anonymity of the secret ballot, the deliberative tradition prizes interdependence and the benefits of working through ideas with others.

The basic idea of deliberative democracy seems pretty intuitive. When we think through ideas and possible courses of action on our own, we tend to come up with a better understanding of them. Of course, we do sometimes experience phenomena like obsessiveness and overthinking, but these are probably the exception. The norm is to feel that we've improved our grasp of a situation by cogitating about it. In fact, when we have all the relevant evidence, this seems like the only way to improve our grasp of a situation and develop our plans about the future. You surely have deliberated internally about decisions both major (where to go to college, where to live) and minor (what to eat, what movie to see).

From our perspective, the question is to what extent we should trust or defer to the deliberations of groups of people on political issues. But that raises a question almost immediately: What does it mean to defer to a group? The majority-vote mechanism is clear enough. But who speaks for a group when it's done deliberating? Landemore notes that many theorists, like Joshua Cohen and Jurgen Habermas, require group consensus or unanimous approval.[29] We might think that this is rarely achieved when it comes to contentious political issues. On the other hand, that might be all the more reason to defer to deliberating bodies on the rare occasions when it is achieved. What's more, even in the presence of persistent disagreement, deliberation can still play a clarifying epistemic role, in that it can make it clear just what the disagreement is about, which can aid in understanding and in political negotiation.[30] Thus, even when consensus and unanimity are not reached, we might be able to trust deliberating groups when it comes to the logical structure of the political issues at hand or maybe how different ideas relate to each other.

When we want to defer to a group that's deliberated on some issue, though, we're probably not party to the internal dynamics of that group. One particular bad internal dynamic that Landemore considers is what she calls "reputational influence."[31] Reputational influence can be thought of as a kind of failure of sincerity – a kind of strategic expression. Even in normal social contexts, people feel all sorts of pressures to stay silent or to say things they don't really believe because of what they have to lose or gain from assenting to or dissenting

from various claims. One potential result of these pressures is self-censorship on controversial issues,[32] which can eliminate some perspectives from deliberation which might end up heard in anonymous settings, like secret ballots or suggestion boxes. From within a deliberating group where it's obvious what the social pressures are, one can and ought to adjust one's beliefs to account for other people's likely silence on certain topics,[33] but from outside it may be hard to tell just what those issues are and thus hard to tell on which issues one should defer. And on charged issues people may actively misrepresent their views by "grandstanding" or "virtue signaling" to gain moral approbation.[34]

Equally interesting are what Landemore calls "informational" influences.[35] These influences are often related to polarization, which will be the topic of a later chapter. In some situations, people form what Cass Sunstein calls "deliberating enclaves" or deliberating groups of "like-minded people."[36] This is especially likely to happen in contemporary American society: political divides correlate highly with geographical ones, especially the urban/rural divide, and people's politics can also be predicted by what sorts of hobbies they engage in or even where they buy their coffee.[37] In such groups, Sunstein claims (based on empirical studies), people don't stick to their guns, as a steadfaster about disagreement might hope, and they don't meet in the middle, as a conciliationist about disagreement might hope. Rather, they obey what he names the *law of group polarization*: "members of a deliberating group predictably move toward a more extreme point in the direction indicated by the members' predeliberation tendencies."[38] In fact, in some cases, *most* members of the deliberating enclave ended up with views that were more extreme than those of *any* member prior to deliberation.

Sunstein emphasizes that in the relevant studies, these shifts are evaluated on the basis of anonymous statements, meaning that this dynamic is likely distinct from the social pressure dynamic.[39] We will look at this in more detail later, but on the face of it, such predictable shifts seem like they must be irrational, especially if group members are aware that they are participating in some sort of unrepresentative enclave. It is also worthwhile to think about just what is meant by "extreme" here. Sunstein writes that "the predeliberation median is the best predictor of the direction of the shift" occasioned by enclave deliberation.[40] But just what is meant by this "median"? We talk about there being two sides to political issues, but in the past and perhaps again in the future, the two major positions represented in American politics could be viewed as on the same side, as against what would now seem to be a more extreme viewpoint. The truth of the law of the group polarization, if it does turn out to be true, thus raises as many questions as it answers: Is there some *objective* "median" on issues on which people might deliberate? Based on what kind of measure? How can we find it?

There is much more to say about deliberative democracy. Various writers have explained how deliberative models might address worries about internal influences and power imbalances, and some researchers have done experiments testing how well these and other models work in a controlled setting. For our purposes, what's most important is recognition that in a real-world setting, the appearance of robust and transparent deliberation may be illusory, and understanding what actually goes on in a deliberative setting is crucial for an epistemic agent who is trying to decide whether to defer to the conclusions reached by the deliberating group.

The dilemma of democracy

Democracy has within it a troubling dilemma, a kind of combination of the epistemology of democracy and democratic political theory, which has been touched upon recently in different ways by Hrishikesh Joshi and Robert Talisse.[41] Conciliation with others' viewpoints, and perhaps eventual deference to the judgment of the whole, seems to be part of the ethos of stable, functioning democracies. We can't simply think that we are right about everything and that no one else has anything valuable to contribute; otherwise, our political systems would make little sense, and (at least according to Talisse) we would cease to be good actors within them. However, things also don't seem quite to work if everyone just defers to the judgment of others or of the collective. Why is that?

Remember how we showed the workings of aggregative epistemic democracy in the context of Condorcet's jury theorem. Every time we add two people, if those people are competent and if they are independent of each other and of the people who have already voted, they improve the competence of the group as measured by a majority vote. And the reliability of the majority vote trends toward 100 percent if the voters are even minimally competent, as long as there are enough of them. This potentially justifies deference toward the majority vote. But say people know how everyone before them has voted. Then, if it's right that we should defer to the majority, then they will all simply vote the way most people so far have. But this will destroy independence, meaning that the reliability of the majority vote does not trend toward 100 percent anymore but is merely halted at the stage at which people began deferring to the majority. This can lead to "belief cascades," which we'll discuss in later chapters and which are also a problem for deferring to tradition (which we'll also discuss).

Thus, there is a dilemma of epistemic democracy as well as democratic political practice related to the difference between individual and group competence. Deferring to others might be what's most rational for me as an individual, but it might prevent me from making epistemic contributions to the group.

Conclusion

There are many ways of putting together multiple people's beliefs or reasons for their beliefs to come to a sort of collective or group perspective. The reliability of these methods will often depend both on the mechanism of aggregation or deliberation and the internal structure of the group. Votes, polls, town halls, and markets all give us some evidence about what's true, but these institutions also have features that can distort or obscure the truth. It is hard enough to know what to do when these institutions reach a clear verdict; in contemporary politics they rarely seem to, and so in contemporary politics we might be tempted, by the same token, to think that we should be skeptical about politics as well, abandoning most or even all political beliefs because of the controversy associated with them.

Discussion questions

1. Does anything strike you as undesirable about ignoring the views of the majority? Does anything strike you as undesirable about following them?
2. Just how sincere, competent, and independent do you think most people are when it comes to figuring out their political beliefs? Most theories of political belief we've considered so far would say: not very. Do you think epistemic democracy can survive their objections?
3. How might out-group homogeneity bias relate to epistemic democracy? Think of the independence condition on the jury theorem in particular.

15

DECENTRALIZATION

Voting is one way of aggregating people's preferences or beliefs, and Condorcet gave his theorem about the epistemic advantages of majority votes. Another way of aggregating preferences and beliefs is through markets, and Friedrich Hayek is credited with the major argument concerning the epistemic advantages of decentralized price mechanisms. The idea of the epistemic power of markets more generally has been related in two major ways (that we'll see) to the project of epistemic democracy. One is deliberative, the notion of the marketplace of ideas. The other is aggregative again, the system of prediction markets.

The knowledge argument and market epistemology

Hayek's knowledge argument is viewed as central to economics and to the "central planning" or "economic calculation" debate of the mid-twentieth century, which was basically a debate over whether a planned economy would be more efficient than a decentralized one. This debate is related though maybe not equivalent to the dispute between capitalist and socialist economics, and Hayek's argument is often interpreted as an enormous victory in favor of decentralized economies and maybe in favor of capitalist economies too.[1] For our purposes, however, it is most interesting as a piece of social epistemology, a context in which it is viewed somewhat less often.

Hayek wrote that

> the knowledge of the circumstances of which we must make use never exists in concentrated or integrated form, but solely as the dispersed bits of incomplete and frequently contradictory knowledge which all the separate

DOI: 10.4324/9781003355274-18

> individuals possess. The economic problem of society . . . is a problem of the utilization of knowledge not given to anyone in its totality.[2]

So the central question of economics of his time was an epistemological question about

> the relative importance of the different kinds of knowledge; those more likely to be at the disposal of particular individuals and those which we should with greater confidence expect to find in the possession of an authority made up of suitably chosen experts.[3]

Hayek writes that individuals will be able to contribute "knowledge of the particular circumstances of time and place."[4]

> It is with respect to this that practically every individual has some advantage over all others in that he possesses unique information of which beneficial use might be made, but of which use can be made only if the decisions depending on it are left to him or are made with his active cooperation. We need to remember only how much we have to learn in any occupation after we have completed our theoretical training, how big a part of our working life we spend learning particular jobs, and how valuable an asset in all walks of life is knowledge of people, of local conditions, and special circumstances. To know of and put to use a machine not fully employed, or somebody's skill which could be better utilized, or to be aware of a surplus stock which can be drawn upon during an interruption of supplies, is socially quite as useful as the knowledge of better alternative techniques. And the shipper who earns his living from using otherwise empty or half-filled journeys of tramp-steamers, or the estate agent whose whole knowledge is almost exclusively one of temporary opportunities, or the *arbitrageur* who gains from local differences of commodity prices, are all performing eminently useful functions based on special knowledge of circumstances of the fleeting moment not known to others.[5]

Such practical knowledge, Hayek writes, is crucially important in making economies efficient but is too specific and too mundane for expert planners to understand. John Danaher writes, in interpreting Hayek, that this knowledge base is often "tacit," or "very difficult to express and communicate," and "diverse."[6] Because many people possess some small piece of this body of knowledge, but nobody has much of it, this is a reason to think that a decentralized economy will be more efficient. For that reason, economics "must show how a solution is produced by the interactions of people each of whom possesses only partial knowledge."[7] Danaher suggests that Hayek also has in mind people's subjective

valuations of various goods, which are likely only known to themselves and only implicitly but which are revealed through their decisions to buy and sell at certain prices, and the new knowledge generated by market pressures.

The mechanism for all this distributed information is the price system. Hayek gives an example:

> Assume that somewhere in the world a new opportunity for the use of some raw material, say tin, has arisen, or that one of the sources of supply of tin has been eliminated. It does not matter for our purpose – and it is very significant that it does not matter – which of these two causes has made tin more scarce. All that the users of tin need to know is that some of the tin they used to consume is now more profitably employed elsewhere, and that in consequence they must economize tin. There is no need for the great majority of them even to know where the more urgent need has arisen, or in favor of what other needs they ought to husband the supply. If only some of them know directly of the new demand, and switch resources over to it, and if the people who are aware of the new gap thus created in turn fill it from still other sources, the effect will rapidly spread throughout the whole economic system and influence not only all the uses of tin, but also those of its substitutes and the substitutes of these substitutes, the supply of all the things made of tin, and their substitutes, and so on; and all this without the great majority of those instrumental in bringing about these substitutions knowing anything at all about the original cause of these changes. The whole acts as one market, not because any of its members survey the whole field, but because their limited individual fields of vision sufficiently overlap so that through many intermediaries the relevant information is communicated to all. The mere fact that there is one price for any commodity – or rather that local prices are connected in a manner determined by the cost of transport, etc. – brings about the solution which (it is just conceptually possible) might have been arrived at by one single mind possessing all the information which is in fact dispersed among all the people involved in the process.[8]

Hayek's vision, then, is that new circumstances and new knowledge are quickly integrated into the distributed "knowledge" of the market through the price mechanism by individuals whose actions are felt throughout the whole. When the need for tin changes, the price changes, and information about the value of tin is gained elsewhere even if information about tin itself is not. Much like Condorcet's jury theorem, then, Hayek's knowledge argument aims to show how the partial competence of independent individuals can generate a much more robust competence at the level of a group or whole.

There is one important thing I want to note about the price mechanism and markets. Remember how we talked about reputational influence? In deliberation,

we worried, an idea that turns out to actually have a lot of value might be unpopular, and so people might keep their mouths shut about it in fear of losing social status or being attacked in some way. Would that ever happen in a market governed *only* by the price mechanism, that someone might think something is valuable but not provide that information for fear of repercussions from people who disagree? No; in fact, the opposite is likely. In a market, when other people don't see the value in something, it can be bought cheaply. So the person who sees its value, if they're right, can make the most of it. In other words, in a market, the more you disagree with others, the *cheaper* it is to express your views through buying and selling; the more conventional your views are about what's valuable and useful, the *more expensive* it is to add that information to the market. Of course, market "expression" is also limited by how much money one has to begin with. But this problem is the *least* concerning when those without resources *diverge* from the views of those with resources – a genuine epistemic institutional virtue.

Of course, not everyone is convinced by Hayek's arguments, and markets are said to fail in all sorts of characteristic ways: public goods, negative externalities, monopolies, and different kinds of inequalities. We'll eschew a full examination of these objections, discussing them only insofar as they're relevant to the topic that interests us here, the notion of a marketplace of ideas.

Free speech and the marketplace of ideas

The epistemic virtues of markets, and according to some economists and politicians the virtues of completely free, *laissez-faire* economic systems in particular, have led some people to think that the best collective practice for improving our political beliefs is something like a market but for political ideas. The idea that competition rather than censorship of ideas has the best epistemic consequences was present in many historically important defenses of free speech. John Milton, in his *Areopagitica*, wrote: "Let [Truth] and Falshood grapple; who ever knew Truth put to the wors, in a free and open encounter. Her confuting is the best and surest suppressing."[9] The metaphor of a marketplace of ideas itself comes from opinions in United States Supreme Court cases. In a dissent in *Abrams v. United States*, Justice Oliver Wendell Holmes argued against the constitutionality of the Sedition Act of 1918, writing that upon reflection, people

> may come to believe even more than they believe the very foundations of their own conduct that the ultimate good desired is better reached by free trade in ideas . . . The best test of truth is the power of the thought to get itself accepted in the competition of the market, and that truth is the only ground upon which their wishes safely can be carried out.[10]

And in a concurrence in *United States v. Rumely*, Justice William Douglas argued that it was unconstitutional to require a publisher of political tracts to disclose the names of his customers, writing that "this publisher bids for the minds of men in the market place of ideas."[11] This analogy, then, provides an epistemic argument in favor of free speech: it helps us get at the truth by seeing ideas compete with one another, as products compete in a market.

The philosopher most associated with the marketplace of ideas thesis is John Stuart Mill, who in his work *On Liberty* gave several justifications for free speech and arguments against suppression of dissent. In Mill's view, there are three possible scenarios facing received opinions as against those views that might be censored:

> First, if any opinion is compelled to silence, that opinion may, for aught we can certainly know, be true. To deny this is to assume our own infallibility. Secondly, though the silenced opinion be an error, it may, and very commonly does, contain a portion of truth; and since the general or prevailing opinion on any subject is rarely or never the whole truth, it is only by the collision of adverse opinions that the remainder of the truth has any chance of being supplied. Thirdly, even if the received opinion be not only true, but the whole truth; unless it is suffered to be, and actually is, vigorously and earnestly contested, it will, by most of those who receive it, be held in the manner of a prejudice, with little comprehension or feeling of its rational grounds. And not only this, but, fourthly, the meaning of the doctrine itself will be in danger of being lost, or enfeebled, and deprived of its vital effect on the character and conduct: the dogma becoming a mere formal profession, inefficacious for good, but cumbering the ground, and preventing the growth of any real and heartfelt conviction, from reason or personal experience.[12]

First, suppressed ideas might be wholly correct, and censors might be removing the truth from public discourse, whether intentionally from nefariousness or some other goal or unintentionally from ignorance or rational error. But free speech likely leaves us epistemically better off even if the ideas that would otherwise be suppressed are wholly false. Even if the censors' beliefs are all correct, they will not be fully justified, nor held in a rational way, if they are not subject to challenge. And in that kind of case, it's likely that those beliefs will stop being held genuinely and will become a kind of formality or ritual divorced from people's deeper convictions. Finally, what's still "commoner," according to Mill, than either of these situations is that the would-be censors and the would-be censored each have some share of the truth, which – in line with the deliberative tradition considered earlier – must each be broached in discussion if the full truth is to be appreciated.

A lot of what Mill said sounds pretty reasonable to a lot of people. But in lived political reality, we worry about the consequences of speech, even given the assurance that the truth may eventually come out. False speech might impugn someone's character, leaving them harmed irreparably, and so we have defamation law. False speech might lead people to make bad purchases, and so we have false advertising law. Even true speech might be a kind of stealing, so we have intellectual property law and, in some contexts, rules against plagiarism. More contentious are recent pushes, especially regarding online speech, to censor so-called disinformation and hate speech. These, too, are sometimes thought to have potential consequences that are too unbearable to justify full freedom of speech.

Alvin Goldman and James C. Cox have challenged the marketplace of ideas analogy on various grounds. More specifically, they challenge the claim as they formulate it: "More total truth possession will be achieved if speech is regulated only by free-market mechanisms rather than by other forms of regulation."[13] Their objections could be classified into three categories. First, it's not clear that ideas generate a marketplace in the way that someone like Hayek might have discussed in his analysis of the price mechanism. Second, if ideas do generate a marketplace, it's not clear that what is efficiently maximized by that marketplace is an epistemic good like truth or rationality. Third, there must be something wrong with Mill's arguments, because we accept all sorts of censorship and restrictions on speech in many contexts, like some that we discussed in the last paragraph, and in others we definitely don't think free discourse maximizes truth. Let's take a look at some of these objections.

Do ideas really have a marketplace at all? Recall that for Hayek the price mechanism was crucial for individuals who wanted to disseminate their portion of the truth through buying and selling products. Goldman and Cox ask: What are the products in the marketplace of ideas, and what would it mean to buy and sell them?[14] If the products are the ideas themselves, and the conceit is that the true ideas will win out over the false ones in the marketplace, some problems arise. Markets optimize for best products according to consumers' subjective preferences. Unless we are certain that consumers in the marketplace of ideas prefer true ideas to false ones to begin with, we should worry that they will instead "buy" those ideas that flatter their preconceptions, or that are simple to understand, or that paint the world in a comforting light, or are the most entertaining, or anything else.[15] And the most successful products are not of the overall highest quality but the most efficient given prices of both inputs and outputs;[16] in fact, it is easy to think of examples of industries in which producing low-quality goods is more lucrative than producing high-quality goods – actually, it's hard to think of industries in which that's false! Finally, it is hard to see where the price mechanism enters into the so-called marketplace of ideas at all. When we consider an idea we do not necessarily spend money on it; when we promulgate an idea we do not necessarily receive money for it.

What if there is such a thing as a marketplace of ideas? Should it be free? Goldman and Cox nod at, without endorsing, an objection based on "inequalities of resources or opportunities in the marketplace."[17] This sort of objection is raised to all sorts of markets, not just markets in speech. Inequalities of all sorts may lead to, in Owen Fiss's words, "a public debate that is dominated, and thus constrained, by the same forces that dominate social structure, not a debate that is 'uninhibited, robust, and wide-open'";[18] similarly, Cass Sunstein argues that unequal social relations may undermine the goal of hearing diverse views.[19] An argument vaguely along these lines was also part of Stanley Fish's case that "there's no such thing as free speech."[20] We might think that this sort of worry is similar to the worries we had about reputational influence in deliberation more generally. To improve deliberation, we might want to make rules about everyone getting a chance or against threats or social ostracism. Or, even if we prefer not to make such rules, we might allow that they could improve things on epistemic grounds, so that the *unfettered* marketplace of ideas is no longer what we consider the most efficient.

There are ways in which markets are more classically taken to fail. One really clever argument from Goldman and Cox is that at least some theoretical models of efficient markets assume perfect information for market participants – that is, all buyers and sellers know all prices and details of the products on the market.[21] However, were participants in the marketplace of ideas to have perfect information about the "products" of ideas, statements, and so on, they would already have to have a lot of the knowledge that the marketplace of ideas is supposed itself to provide – possibly even knowledge of which ideas are truth and which are false. So the model on which the marketplace of ideas gets us the truth might require that we know the truth already, rendering it superfluous. However, note that some arguments for markets don't require perfect information, and in fact that the argument we saw from Hayek was based on the *lack* of perfect information. And the risk of market failure is not a devastating objection either, since non-market actors like states have failure modes of their own.[22]

Goldman and Cox also relate the marketplace of ideas to notions of public goods and externalities in economics.[23] This point has to do with who benefits from the truth and who is hurt by falsehood. Markets face problems when the answer is something like "everyone, but no one in particular." This kind of dynamic is typified by famous examples in which individually rational people form an irrational whole, like the Tragedy of the Commons and the Prisoner's Dilemma. These bear some similarities to the Dilemma of Democracy we considered earlier.

Goldman and Cox also adduce several more examples of venues for ideas in which speech isn't wholly free and in which we don't seem to think it should be.[24] In addition to the ones we've already considered, they point out that peer review in academic journals and rules of evidence in courts of law limit free

expression in ways that we think do not hamper but in fact aid in the search for truth and further that there are plenty of contexts in which free talk seems to lead to falsehood, like rumor and gossip, superstition, urban legends, and so forth. These examples could be considered in more detail. For instance, when it comes to legal procedure, the epistemic goal is not rational belief, knowledge, or truth but some specifically delineated threshold of evidence, like "preponderance of the evidence" or "evidence beyond a reasonable doubt." Further, courts sometimes wholly remove evidence, as in the exclusionary rule, a practice with practical but not epistemic justification.

Despite all these objections, something about Mill's point still seems to ring true. As we'll discuss in a later chapter, many people alive today would aver that humanity has made scientific, moral, and social progress over time. This progress has been unpredictable and has often upended received wisdom. If certainty in one's infallibility really is the precondition for advocating censorship, as Mill seemed to suggest, then advocates of censorship must think that such progress is over or that they know exactly the path it will take forward. They must think that we've finally gotten things right. But if everyone in the past who has had that thought has been wrong, why think we've done any better? In the philosophy of science, this kind of argument is called the "pessimistic meta-induction" and generally credited to Larry Laudan.[25] However, this kind of reply might be self-defeating, too. After all, if nobody is ever justified in thinking they've finally gotten things right, then all of the epistemic arguments in favor of free speech, which are meant to help guide us toward the truth, might be beside the point, too. We might be left adrift in a sea of skepticism by such reasoning.

Prediction markets

What if there were a market mechanism, separate from deliberation and free expression, in which ideas could be in some sense bought and sold? We see something like that in systems called *prediction markets*, and Robin Hanson has proposed a form of governance called *futarchy* based around these systems.[26] As usual, our question will not be whether futarchy is a good form of governance but how reliable prediction markets are and whether we should defer to them.

A prediction market in one sense allows you to make bets on the outcomes of events, just as you might bet on the outcome of a soccer game or a horse race. It is more like a marketplace than a bookie, though, in that your bet is represented as a share in the outcome. Say A and B make an even-odds bet about whether it will rain tomorrow, A saying that it will and B saying that it won't. In a prediction market system, this might be operationalized as A and B each paying fifty cents for a token, with A's token being worth a dollar in the event of rain and nothing in the event of clear skies and B's token the reverse. Now say S comes along, and say S is really sure that it will rain, intending to buy plenty of tokens representing

that outcome and offering seventy-five cents each for them. If B is around, the system could create a new token representing rain and a new token representing clear skies, charging S seventy-five cents for the first and charging B twenty-five cents for the second; but S could also buy A's token for seventy-five cents, leaving A with no bet on the table and a twenty-five cent profit.

How can a prediction market become a form of governance? Well, say that instead of representing the outcome of rain, the token somehow represents the outcome of some policy that we're considering implementing. Perhaps C has a token representing the outcome that the unemployment rate will go down after the implementation of policy X, and D has a token representing the outcome that the unemployment rate will stay the same or go up after the implementation of policy X. If C's token is worth much more, then that represents a kind of market consensus that the unemployment rate going down is the more likely eventuality.

In Hanson's futarchy we "vote on values" but "bet on beliefs." We can't make bets about whether lower unemployment is good or bad, because it's not the sort of thing that can be resolved empirically. (Actually, genuine empirical resolution to everyone's satisfaction may be difficult to achieve in plenty of cases, because, as we've discussed, political disagreement runs very deep.) This is because of issues of moral epistemology which we have already broached. So we vote on whether lower unemployment is good or bad and bet on whether policy X will lower unemployment. If we vote that it is good and bet that policy X will achieve it, then we implement it, and the market can be resolved by watching to see what happens next.

The explanation Hanson gives for what he calls the "robust and consistent empirical track record" of successful prediction markets is in part based on incentives to make good bets and to avoid making bad bets: "stronger accuracy incentives tend to reduce cognitive biases, those who think they know more tend to trade more, and specialists are paid to eliminate any biases they find."[27] This means that there must be something real at stake in the bets for a futarchy to work: that people who get things right a lot should get something they want, and people who get things wrong a lot should somehow lose. As in other market settings, this helps to eliminate strategic discourse; we saw earlier that the financial incentive in a market is against conformity and in favor of true information. Recall the connection between action and belief, and Hannon's note that partisans' beliefs seem to moderate when offered monetary incentives. The incentive reveals the genuine belief. And Landemore finds another nice feature in futarchies:

> Information markets also work because they are responsive to the intensity of individual beliefs. Unlike what happens in polls or in voting – where each individual is given only one voice – here the more convinced you are, the more you should be willing to bet. By allowing for a differentiated weight given to more or less firmly grounded beliefs, information markets generally end up aggregating the most reliable information.[28]

In the language we've been using in this text, then, in prediction (or "information") markets, we are not just betting on beliefs but betting on credences. This provides more information than a single up-or-down vote would.

Landemore notes some potential objections to futarchy.[29] Not all questions may have answers partial knowledge of which is distributed to a large number of rational actors who can aggregate that knowledge into a whole through a market mechanism; in some cases, the knowers may be concentrated in a small group. There may be moral concerns about betting on certain events, like terrorist attacks or, say, pandemics, and people's civic virtue might be "corrupted" by integrating the profit motive into their political lives. And prediction markets might be antidemocratic, dominated either by an informed elite (who could profit most easily) or by the already rich (who in futarchies could buy political outcomes through bets, whether rational or irrational). Some institutional designs could help ameliorate some of these problems: in one design, the only funds that could be used to bet in the prediction market would be a certain universal allowance and money gained by betting in the prediction market itself.

Note that there is a dilemma for prediction markets a bit like the Dilemma of Democracy: the more reliable we take the markets to be, the less we think we can gain an advantage within them and the less reliability we'll be able to add by making our own bets. Say you believe the prediction market has rationally and efficiently integrated all information about the events listed therein, so that all the shares are at, in some sense, the correct price; then it can't be rational for you to buy or sell any shares or trade with the other investors. Even more generally, Robert Aumann argues in a classic paper that people who start out from common "priors" and take each other to be rational must end up agreeing,[30] a result which bears some similarities to conciliationism (which can be interpreted as holding that if two people disagree, they can't both be rational and epistemic peers). A related "no trade theorem" establishes that in an idealized market, an agent can never expect even private information to have failed to be integrated by the market, so that they can never expect to get an advantage out of any trade with anyone.[31] But of course this might be taken to establish rather that markets do not operate under idealized conditions.

Conclusion

Philosophers, economists, and others have often been exciting by potential advantages of decentralized epistemic systems, like marketplaces of ideas, prediction markets, and so forth. The notion that freedom in realms like trade or speech might be practically optimal buttresses the arguments of those who take freedom to be a fundamental moral good, and the spontaneous, self-organizing nature of these systems fits nicely into a broader intellectual history in which bottom-up processes like evolution took the place of top-down creation. But market

models don't always match up with the real world, and plenty of social institutions still use top-down rules which are intuitively appropriate. Like epistemic democrats, market theorists have to contend with the reality of political beliefs in developing and justifying their theory.

Discussion questions

1. Do you buy the knowledge argument? Think through Hayek's account of the price system. Does it seem like a better way than central planning to determine prices? Do you agree that no one individual can possess all the information the system integrates?
2. What is your first reaction to the idea of running society using prediction markets? Do you think they could be trusted? Do you think people make better decisions when money is on the line?
3. What do you think of the notion of the marketplace of ideas? Should speech be free? What about the counterexamples raised by Goldman and Cox? There are likely rules about class participation in your course. Don't those rules help discourse proceed productively?

16

EXPERTISE

So far we've discussed various difficulties we might face in forming our political beliefs rationally and reliably. It might be irrational to have strong political beliefs given how many people would disagree with them. Our political beliefs might emerge from distorting ideologies, or be subject to various cognitive biases that lead us away from the truth, or be caused by the strategic evolution of partisan platforms. We'll see that we might base our beliefs only on information from inside an echo chamber, or on propaganda, or on other people's moral grandstanding.

All these issues might prompt you to think: Why keep trying myself? Since I'll likely fail in one of these many ways, why not just trust the experts about politics? We trust expert plumbers about pipes, expert mathematicians about theorems, expert dentists about teeth. When we take a vacation to a place we've never been, we take its residents to be experts relative to ourselves when we ask for directions and recommendations. Consultation with and deference to experts is a normal part of our lives and one that we could hardly do without in the modern world.

In this chapter we'll consider the surprisingly tough philosophical questions about the nature of expertise. We'll also take a look at a proposed specific form of identity-derived expertise about political questions, which I call "standpoint expertise," and at the distinction between expertise when it comes to claims of fact and expertise when it comes to claims of value. We'll conclude by broaching the issues of whether political expertise is possible and whether it would be a good idea to try to form our political beliefs by looking for political experts.

DOI: 10.4324/9781003355274-19

Seven questions about expertise

I organize the questions about expertise under seven headings. The first is the definition question. Just what does it mean to be an expert? Of course, there are people who as a matter of social status or educational credentials are commonly taken to be experts. But for all such people there remains the possibility that they are not actually experts and don't know any more than anyone else.

The second question is the domain question. How do we split up knowledge or potential knowledge into the sorts of domains in which someone could potentially develop and demonstrate expertise? Say someone gives me a convincing answer about agricultural policy in Lithuania. Should I take them to be an expert about policy? About agriculture? About Lithuania? The third question is the possibility question: about whether it is actually possible to achieve expertise in some delineated domain.

Take aesthetics, for instance, or something like art or food criticism. Some might think that aesthetics, or art or food more specifically, boil down only to matters of taste, so that there's no real truth of the matter when it comes to which items are good and bad. But if that's the case, then depending on our notion of expertise, it might be impossible for there to be experts about such things.

Perhaps most philosophically interesting is the identification question. Say we want to look for experts; how do we find them? Of course, this, too, will depend on the definition question. But say an expert is just someone who gets a lot of things right. Then, to find them, we ourselves will need to know a lot of right answers, to test the experts against. But then we seem to be taking ourselves to be the experts, even though the whole situation is based around the fact that we are novices who don't know what to think and are looking for experts to teach us. This expert identification puzzle, also known as the novice-expert problem, has caused philosophers a great deal of consternation.

The deference question is about what to do once we've found the experts. In what sense, and to what degree, is it rational for us to substitute expert judgment for our own? This will likely depend in part on some of our answers to the other questions, like the definition and domain questions. But even when those questions are answered, some complexities remain. While most people might accept some piece of advice from a doctor, others would reject it, and a third group might seek a second opinion. And that brings us to the disagreement question: what to make of the fact that experts often disagree. Any theory on which we should fully and automatically take on a given expert's view on some topic as our own must contend with the fact that we might expect to find differing opinions elsewhere, even among other experts.

The final question, the correspondence question, is about how well the group of people, who some definition of expertise picks out as experts, matches up with

the group of people who are socially or professionally approved as experts – that is, about whether society identifies experts correctly.

Philosophers are highly divided on these questions. In a recent paper, Neil Levy argues that "doing one's own research" is "likely to undermine knowledge."[1] But as we'll see, such independent research may be unavoidable even on *a priori* grounds. So it is a good thing that many other philosophers, as well as some parts of lay common sense, think that it is desirable as well.[2]

The definition question

As mentioned earlier, we can start by taking care to distinguish between social and epistemological senses of expertise. In many cases, someone's meeting or failing to fulfill certain criteria for being an expert in the social sense can be *evidence for* whether they are an expert in the epistemological sense. (We will think more about that when we examine the identification question.) But one thing's being evidence for the other does not make it the same as the other. In the social sense, someone who has graduated from law school and passed the bar is an expert in the law. But the epistemological sense of expertise concerns why those criteria might be useful.

The simplest definition in this sense is that an expert is anyone whose beliefs in a certain domain meet a certain threshold for truth over falsity. This threshold will probably differ from domain to domain. Being able to predict the results of 95 percent of professional sporting contests might suffice to make one an expert – that might be enough to guarantee that one could make a living by gambling on those contests, for instance. But a 95 percent success rate might be too low to qualify as expertise in other fields. An airline, for example, should be expected to get you and your luggage safely to your destination on far more than 95 percent of occasions. Intuitively, though, having true beliefs is a somewhat weak condition for expertise. We tend to think that experts will also have good reasons for their beliefs; their beliefs will be justified, or constitute knowledge, or have some other desirable epistemic property. Experts might also *understand* their domains better; they might be able to see connections between different subdomains, how it all fits together.

The threshold view of expertise fits well with a broader theory of epistemology on which what justifies our beliefs is their source in a reliable method of belief-formation.[3] But on other theories, what justifies our beliefs is our possessing and reflecting on the right reasons that favor them. This has led Jennifer Lackey to propose that we view experts not as authorities but as advisors of sorts.[4] In particular, Lackey rejects the "preemption view" of expert authority, on which whatever reasons one might have for believing or disbelieving something in the domain of the expert, one simply ought to put them aside on learning of the expert's view and adopts that view as one's own.[5]

In our discussion of disagreement, we noted that in the strict sense, one's epistemic peer would be someone who possessed the same evidence as oneself, along with an equal ability to reason through it. Similarly, James Joyce notes that an interlocutor's expertise relative to a novice can consist either in greater information, what he calls "database expertise," or in greater reasoning ability, what he calls "analyst expertise."[6] Each of these is compatible with both the authority and the advisory view of expertise.

The domain question

We can think of the domain question in a few ways. One is that we expect human expertise to come in chunks, from big chunks like physics, mathematics, and zoology to small topics like the biography and letters of a single obscure nineteenth-century novelist. Then the domain question might be just how to split up those chunks. Another is that we often encounter interlocutors who have demonstrated reliability with regard to some propositions and maybe failed to demonstrate reliability with regard to some others. Then the domain question might be just how to predict for some new proposition how reliable they will prove to be. A third is that when an interlocutor is extraordinarily reliable with regard to some set of propositions, we look for an explanation of that reliability. The explanation could be some very specific kind of experience, but it could also be a broader competence with regard to some domain. Then the domain question might be just how to construct such explanations.

Here's an example I often discuss. At a party I was once introduced as a philosopher to a young couple, a man and a woman, and the man asked me: "Oh, there's this philosophical question I've always wondered about. Why do bad things happen to good people?" As it turns out, I have no idea why bad things happen to good people. I'm not sure many philosophers study this at all, but I certainly don't. But there is a sense in which you might expect a philosopher to be able to say something about this. It's a deep question, and deep questions are kind of the purview of philosophy. On the other hand, when I'm on campus or at an academic conference I don't really get treated as though I have broad expertise in deep questions at all or even in academic philosophy. Rather, in that context I'm a specialist, an epistemologist, and often more specifically an epistemologist of politics. The domain in which I've garnered expertise can be broadened or narrowed depending on context, and in particular depending on who else is around me, and what sorts of expertise they have, too.

The domain question may sound abstruse, but it's of great importance to our everyday decision-making and to political epistemology as a theoretical endeavor. Political candidates often come to us with prior experience which they say qualifies them to make the right choices in politics: doctors, lawyers, executives, military officers. Judging their claims to qualification requires a sense of

how broad the domain of their expertise likely is. In other instances, people seem to want to shift the boundaries of their domain to seem authoritative in another. In the early days of the coronavirus pandemic, during the Black Lives Matter protests galvanized by the death of George Floyd at the hands of Minneapolis police, public health researchers Julia Marcus and Gregg Gonsalves argued that racism is a public health matter[7] and thus claimed a level of authority regarding the legitimacy and likely consequences of those protests. We'll think more about the fact that they also claimed a dimension of moral expertise in a later section.

These sorts of issues bring us back to the question of what politics is, what domain we're discussing when we talk about politics, political beliefs, and political knowledge. What would a political expert be? All sorts of people claim political expertise: historians of mass movements, theorists who think they have the best theory of the ideal state, policy wonks who know the costs and benefits of everything under the sun, and campaign managers and legislative directors who know how to get bills passed and candidates elected.

The possibility question

The possibility question: Given some domain, is it possible for there to be experts? There are two reasons it might not be. First, there might not be any actual truths, what philosophers sometimes call "facts of the matter," about that domain. Consider food preferences. Many people find it natural to think that whether one prefers apples to oranges, or chocolate ice cream to vanilla ice cream, is just a matter of variations in taste. (But food critics and professional chefs might disagree.) If so, then there is nothing to be an expert *about*: there aren't any true propositions in the domain about which one could be an authority, and there aren't any relevant reasons that consultation with an advisor could help one appreciate. Anything goes – or what goes is close enough to just anything that expertise is impossible.

But expertise might also be impossible if there are truths that we aren't in a position to know. For instance, there are truths about the future of my life: whether I'll find a spouse and have children, whether I'll live a long life or fall ill young, whether I'll be happy or sad, and so on. People like astrologers, fortune tellers, and tarot card readers advertise themselves as potential experts on the future of my life. But nobody is in a position to know most such truths ahead of time – or at least so many people think. Assuming that such knowledge is impossible, expertise, at least of the authoritative sort, is likely impossible as well.

The identification question

Now we turn to the identification question, often called the novice-expert problem, including by Alvin Goldman, who is largely responsible for posing it and explaining its salience and some first-pass possible solutions.[8] With the

distinction between social experts and experts in the philosophical sense in hand, we can rule out some easy methods of identifying genuine epistemic experts – looking to titles and credentials only guarantees social expertise. So what else can we do?

Let's make sure we understand the problem in its full generality. When we look for experts to defer to in some domain, it's because we are novices: we don't know much about that domain. The question is how we can then figure out who the experts are. Regardless of whether we adopt the conception of the expert as an authority or the conception of the expert as an advisor, to have good evidence that someone is an expert, we must judge how much they know. But to do that, we must take ourselves to know a bunch of facts about the relevant domain. But if we did know those facts, by which we can judge the expertise of others, we wouldn't be novices anymore. In fact, at least under the authority conception of expertise, we would be the experts. This is the novice-expert problem: how a novice in some domain can, while remaining a novice, identify the experts in that domain.

There is one specific kind of cheery case worth noting: that of testable prediction. Say I want to identify the expert meteorologists. I might not right away know the propositions in the relevant domain – say, propositions about what the weather will be like in the future. But I can gradually get external, observable information about which I'm quite confident – say, in the normal way, living from day to day and seeing what the weather is like. Then I can see how well various meteorologists fared at predicting that weather. Here there seems to be little problem, since in observing the weather in the normal way I don't take myself to be an expert meteorologist. This is one example of the first of five types of evidence that Goldman proposes that we use to identify experts, which he calls a track record. The best-case scenario for expert identification is certainly one where track records of potential experts are available to and assessable by the novice.

Goldman's other four criteria are a bit dicier. First, he suggests that we can identify experts by looking to other experts and to "meta-experts" who are tasked with identifying experts in a variety of areas. Such meta-experts are likely to be the sorts of credentialing bodies that bestow upon people what we've been calling social expertise. But we've made sure to stress the difference between social and epistemological expertise. This suggests that these meta-experts might not be much help. And as for the judgment of other experts, well, we'll need to have identified those experts in the first place!

Goldman also suggests that we can use public debates and arguments among would-be experts to judge which of them are really experts. But it seems to me that the novice-expert problem simply recurs as we think about how a novice in some domain might try to evaluate a debate among would-be experts in that domain. How can the novice judge which points should actually be compelling? Some experts may seem to more fluently or convincingly cite reasons for their

views, but how can the novice be sure that those actually are reasons for those views and not just pieces of rhetoric? Moreover, how can the novice be sure that the apparent winner of the debate is not just a good debater, given that debating itself is a skill at which one can become more expert? Indeed, these kinds of questions about experts and rhetoricians have troubled philosophy for as long as it's been around – they feature heavily in Plato's dialogues like the *Gorgias* and the *Sophist*, in which his teacher Socrates opposes those who use clever rhetoric to mask their ignorance of a topic and make themselves seem like experts.

Possibly the most interesting criterion Goldman proposes is that of interests and biases. With an expert, as with any interlocutor, we should try to identify conflicts of interest and sources of bias that could lead either to forming a belief in an unreliable way or to lying about a true belief and propagating a false one. A judge whose spouse is the lead attorney for one of the parties in a case, a doctor who receives a payoff from a pharmaceutical company to prescribe one medication and discourage patients from taking another, a referee who has made an enormous bet on the home team – all of these potential experts are affected by conflicts of interest and sources of possible bias. Yet here, too, there are complications. Say a dentist chooses to become a spokesperson for a particular brand of toothpaste specifically because they think that brand of toothpaste is the best on the market. Then the seeming conflict of interest is no conflict at all: they are expressing their truly held belief. The same may be said for researchers who receive money from think tanks or other organizations with obvious ideological alignments. In many cases, the researchers' opinions come first, and the funding comes second. In this direction, the interest is not necessarily a problem, since it follows on from the genuine belief. So the sword of interests and biases must be wielded only deftly so as not to rule out too much.

The deference question

The deference question gets to the heart of expertise. After all, the main reason we look for experts is to find someone to whose judgment we can defer. Just when and how we ought to defer to other people's judgment? As a kind of trust, deference is a normal part of human social and epistemic practice. If someone has been to a restaurant and you haven't, you trust them when they tell you that the sandwiches are more expensive than the soups. If someone has been to a city and you haven't, you trust them when they tell you that the buses run only every hour. In general, we trust others when they tell us details of their lives that we don't have access to, as well as everyday information that they might simply have spent the time to learn. We can think of deference to experts as an extension of that sort of trust. Of course, we also have a countervailing intuition, especially in the realm of politics, that people should retain some sort of intellectual autonomy and come to their own conclusions. Justin Tiehen has recently argued that

we value, sometimes rightly, the products of our own intellectual labor more than beliefs we gain merely through trust or deference, even when those beliefs are true.[9] I will leave such anti-deferential considerations aside for now, however.

Lackey's distinction between authority and advisor emerges from concerns about deference. She considers an argument from Joseph Raz,[10] which she calls the Track Record Argument. Raz asks us to imagine that we've identified a domain on which we're a novice and someone else is an expert. For a certain subset of our beliefs about this domain we defer our judgment entirely to the expert, and for another subset we retain our independently formed opinions. Raz says that we can infer immediately just from the fact that we are novices and they are experts that we will have a higher proportion of true beliefs in the subset for which we've deferred our judgment entirely: that higher proportion is just what it means that the expert was an expert relative to us. Therefore, the goal of belief being to get at truth or to possess knowledge, we should defer on the whole set of beliefs in the domain and do better overall. As illustration, Lackey cites Linda Zagzebski,[11] who compares two approaches, one of which happens to be the approach of humans and the other of which happens to be the approach of rats. In an experiment where a green light flashes three-quarters of the time and a red light flashes one-quarter of the time, many humans predict three-quarters of the time that the next flash will be green and one-quarter of the time that the next flash will be red. Rats, on the other hand, choose an option that is associated with the green light on all occasions. So rats get things right 75 percent of the time while humans do worse, since although they understand the probabilities, they have no insight into which occasions will be the low-probability, red-light events. Zagzebski thinks that following the Track Record Argument to its conclusion of total expert deference is just a matter of being like the rats rather than like the humans.

Lackey argues otherwise: in the case of expert judgment, we *may* have some insight that helps us distinguish between the occasions on which we should defer and the occasions on which we should retain our independent views. In particular, while some of the expert's pronouncements may seem worth following, others may strike us as "outrageous or morally perverse."[12] In these outrageous pronouncement cases, we might have really good reason to think we're in the low-probability situation. This can be put in probabilistic terms. If we give the expert a 90 percent chance of being right in any given pronouncement, but prior the expert pronouncing some proposition p we were 99 percent certain that p was false, then we ought to remain more on the side of thinking that p is false even after the pronouncement. Lackey notes that in the strong "preemption" formulation of the authority conception of expertise, it's not clear that we can even recognize outrageous pronouncements as such. That's because the preemption view has it that we should simply put aside our prior beliefs, and our reasons for them, when we hear the expert pronouncement. Those beliefs and reasons

having been put aside, or "screened off," it's not clear that we would even retain the intellectual resources for remembering which pronouncements would have struck us as outrageous.

It sounds like Lackey has the better of the deference question here, and that we should retain our independent thinking, treating an expert as an advisor rather than as an authority, so as not to be pushed into believing anything outrageous. But there is something of a complication, which can be seen by reflecting on the domain question. Recall that one way of thinking of a domain of potential expertise was as an explanation of how someone got a lot of similar things correct. This explanation starts to feel rather strange if, in addition to an extraordinarily high number of correct beliefs in a domain, the expert also has a few that even a novice could recognize as simply outrageous. What would be the etiology of such outrageous beliefs from an otherwise highly reliable person? On what grounds would the expert have adopted them? There is nothing logically contradictory about the kind of situation Lackey considers, but there is definitely something strange about it, and thus it may be doubtful whether we can learn anything too general from it about the nature of expertise and the rational requirements of deference to experts.

The disagreement question

The expert disagreement question is a lot like the expert identification question. One situation in which the expert identification question arises is where two people claim to be experts, but disagree, and the novice must figure out which one is the genuine expert. The expert disagreement question concerns the fact that sometimes there are two genuine experts who disagree. The novice must figure out what to make of this scenario.

One thing that philosophers worry about when it comes to expert disagreement is that, as we saw in a previous chapter, rationality may require that when two peers disagree, they move their beliefs toward each other – what we called conciliation. Cases of expert disagreement may seem like cases in which neither expert has done this, and thus, if we subscribe to conciliationism, in which neither expert is being rational about their belief. This might give us a reason to trust *neither* expert when it comes to that belief, which would undermine our proposed strategy of finding and trusting experts about politics.

An interesting kind of scenario often arises with real-world experts but has received little attention that I know of in social epistemology thus far. Different groups of experts, from different domains, may each lay claim in different ways to authority over the same proposition or set of propositions. Consider nature/nurture-style debates; without resolving such debates, we cannot know whether to defer more to, say, biologists or sociologists when it comes to questions of human behavior. There are similar questions about political outcomes; political

scientists, economists, psychologists, sociologists, and philosophers all think they have really important things to say about politics, but both these claims and the background theories that generate them are in conflict with each other, leaving the novice with a debate they must resolve themselves regardless of the presence of experts.

There's another way in which expert disagreement might mean that trusting the experts leaves us no better off than we started. If we are hesitant to trust ourselves with the responsibility of forming our political beliefs because we suspect we have various cognitive biases that may affect our thinking, then we ought to be at least a little hesitant to trust ourselves to identify experts and pick among disagreeing experts ourselves, too. We may expect that we'll engage in what Gabriele Contessa calls "expert shopping"[13] – trusting more the experts who say what we already thought, or what we were hoping they would say, and so on. In sum, if we are bad at forming our beliefs ourselves, we may also be bad at forming our beliefs based on expert judgment, because we must begin by judging for ourselves who the experts are and by judging among them.

The correspondence question

The correspondence question may be the most pressing question there is about experts in the political context. Unfortunately, it is also the least tractable by philosophers. To answer how the socially approved experts match up with the genuine experts in the epistemological sense, we must go out into the real world to do a very broad version of the kind of investigation Goldman recommends.

The classic investigation of that sort is Philip Tetlock's book *Expert Political Judgment*, which paints a rather dismal picture of social expert performance at making political predictions. Tetlock sets out to investigate what he calls "radical skepticism" about social experts (henceforth just "experts" for the purposes of this section), which he takes to consist of six hypotheses.[14] The first is a "debunking hypothesis," which has it that expert performance will be similar to that of chimpanzees and unsophisticated algorithms in experimental settings. The second is a "diminishing marginal returns" hypothesis, which has it that even good expert performance doesn't far exceed that of a reasonably well-informed novice or "dilettante." The third is a "fifteen minutes of fame hypothesis," which has it that good expert performance is often confined to a small domain or to a certain historical moment. The fourth is a "hot air hypothesis," which has it that expertise leads to greater confidence and more elaborate rationalizations rather than better predictions. The fifth is a "seduction hypothesis," which similarly has it that recognition for one's expertise both produces and selects for overconfidence. The sixth is a "sustainable illusion hypothesis," which has it that poor expert performance will by and large never undermine public confidence in experts. I don't have the space to explain Tetlock's experimental methodology here, but

his experiments found significant support for all six hypotheses, with the fifteen minutes of fame hypothesis coming out as maybe the least strongly supported. Tetlock does not think, however, that these results mean that there aren't experts in the epistemological sense at all. Rather, the best predictors have more general knowledge and broad cognitive skills that aid in making good predictions. Examining such individuals became the subject of another book coauthored by Tetlock.[15]

Tetlock began his study of expert political judgment in the fallout from the Iraq War, an event which he and many others thought cast doubt on the level of deference we can rationally give to social experts. Other such notable events have certainly occurred since then, including the housing market and financial crisis of 2008, the election of Donald Trump in 2016, and the coronavirus pandemic in 2020. To my mind, one event, really more of an ongoing scenario in academia, sticks out as important to take account of, what's sometimes called the replication crisis. Stuart Ritchie explains much about this crisis in his book *Science Fictions*.[16] It is a crucial part of scientific practice that data be replicable: that others be able to produce the same results that we have. If a result in one lab couldn't be replicated in some other lab, that's reason to doubt that it really represents some scientific truth; it could be the result of some unrecognized circumstances, some accident or mistake, some statistical manipulation, or worst of all, some fraud. But it turned out that many, many scientific results were failing to replicate. The replication crisis had perhaps its most notorious effects on psychology, and as Ritchie and others have documented, it took down many results, both classic and trendy, from priming effects to the "power pose" to the Stanford Prison Experiment. Humorously, a finding that prediction markets could be used to determine which studies would replicate and which wouldn't itself failed to replicate.[17]

My favorite analysis of the sort of study toppled by the replication crisis comes from a blog post by Daniel Lakens.[18] In a famous study, judges seemed to give favorable rulings to defendants around two-thirds of the time; that ratio drops to zero as lunchtime approaches, rising all the way back up after lunch. Lakens explains: "The idea is that their mental resources deplete over time, and they stop thinking carefully about their decision – until having a bite replenishes their resources." But he avers:

> I think we should dismiss this finding, simply because it is impossible. When we interpret how impossibly large the effect size is, anyone with even a modest understanding of psychology should be able to conclude that it is impossible that this data pattern is caused by a psychological mechanism. . . . If hunger had an effect on our mental resources of this magnitude, our society would fall into minor chaos every day at 11:45. Or at the very least, our society would have organized itself around this incredibly strong effect of

> mental depletion. Just like manufacturers take size differences between men and women into account when producing items such as golf clubs or watches, we would stop teaching in the time before lunch, doctors would not schedule surgery, and driving before lunch would be illegal. If a psychological effect is this big, we don't need to discover it and publish it in a scientific journal – you would already know it exists.

Lakens here gives us great insight not just into the replication crisis but into the relationship between expert judgment and commonsense judgment. Novices would have done well, as Lackey argues, to maintain a sense of which psychological results were simply outrageous, and to categorize the hungry judges study that way. Moreover, experts would have done well to maintain their common sense, and to make it part of their expertise, developing the sort of well-rounded, flexible cognitive quiver with which Tetlock thinks the truly reliable observers are armed. All of this is in tension with a very strict distinction between experts and novices and with an epistemological model which places deference to experts front and center.

Moral expertise

We have emphasized throughout the text thus far that the most distinctively political beliefs have a normative element – often a specifically ethical element which concerns what "we," some polity, ought collectively to do. Expertise about ethics, usually called "moral expertise," presents philosophical conundrums of its own, related to some of the questions we've considered thus far. First, on the possibility question, there are metaethical questions about the nature of morality. For philosophers who take anti-realist metaethical views, on which talk of moral right and wrong is purely emotive, outright false, or even meaningless, it might be hard to countenance the idea of a moral expert – such philosophers seem to think there are no truths that the expert could know better than the novice. We won't consider that worry here, since it takes us too deep into the realm of metaphysics. However, there are also troubles about moral expertise concerning both deference and identification.

Regarding deference, it has been argued that substituting someone else's moral judgment for one's own is objectionably "fishy."[19] Unlike our ordinary, nonmoral beliefs, our moral judgments might have something to do with our character or our responsibility for our own actions. Laura Callahan gives an example of the sort of moral deference case that seems fishy:

> Bill doesn't know whether exclusionary detailing (lying by omission) generally counts as morally wrong. So he consults a person he trusts to be competent in judging these kinds of moral issues. The moral expert tells him it is

> generally wrong to lie by omission, which doesn't sound totally crazy. So Bill accepts this.[20]

Philosophers have struggled to say just what is problematic about deferring in this sort of case. And there is an easy rationale that suggests that such moral deference is not problematic: if we identify a moral expert, then deferring to them will help us to know what we ought to do, thus making it more likely that we do what we ought to; and it is hard to see how it could be wrong to make it more likely that we do the right thing. Moreover, our judgments of the morality of the situations we ourselves are involved in and the actions we ourselves might take are as likely as virtually any judgments to be clouded by motivations, interests, and biases, which is why we often consult others for advice about what to do. So I'm reluctant to spend too much time on the vague complaint about fishiness.

There is a different kind of problem with identification. Sarah McGrath argues that there is nothing like a track record in the case of moral expertise, and so we can never identify moral experts.[21] (Recall that a track record was the most important source of expert identification for Goldman.) McGrath cites Alexander Nehamas, who traces this claim to Plato's depiction of Socrates:

> [I]n the case of ethical experts, it is not clear that we can recognize the experts independently of the fact that we find their views and their reasons for them – their reasons for living as they do – convincing. But to find such reasons convincing is already to follow them. . . . This is one of the most crucial and paradoxical consequences of Socratic ethics: only one good human being can recognize another.[22]

Whereas to check on a doctor's expertise we might check whether their prior patients have gone well, and to check on a lawyer's expertise we might check whether their prior clients have won or lost in court, it seems that we lack "independent access" to or an "independent check," as McGrath puts it, on facts about the moral domain, so that our efforts at identification of moral experts are doomed.

But are things as bad as McGrath thinks? Here are two potential sources of an independent check. The first, discussed by David Enoch,[23] is reflection – in two senses. Say, I find that, in every case where you and I are both confronted with a moral dilemma, you arrive instantly at an answer that takes me several weeks to settle on. Then I might think that I have performed an independent check on your moral expertise by using my weeks of reasoning to check your answer. If the beliefs I settle on after a few weeks tend to be relatively stable, and I'm very confident that those weeks of thinking through the issues improves my ability to get them right, then I might reasonably think that I've identified you as a moral expert relative to me, since you have that improved ability right away. We can

think of this as an application of a "reflection principle" in formal epistemology, which holds that if we know what we'll believe in the future, we should always defer to that future belief. This principle fails in cases of loss of information and loss of reasoning capacity,[24] but ordinary thinking about moral dilemmas and scenarios over time doesn't seem to involve any such loss. So there is one kind of independent check: matching my eventual conclusions about moral matters with another person's immediate conclusions.

There is another kind of source of such an independent check, which is the kind of information we gain by actually performing an action and seeing how we react. This idea is quite a bit like the counterargument from reflection. Sometimes we do things without realizing or without quite believing they're wrong, only to feel embarrassed, ashamed, or regretful afterwards. Such feelings of shame and regret don't always accompany wrongdoing, but it's plausible that when they do, they constitute a source of evidence for us about the morality of our action. So then, like above, say that I've found someone whose recommendations about my actions match up with my feelings of regret and shame, so that I never feel regretful or ashamed about actions they encourage, and often or always feel regretful or ashamed about actions they discourage. Then I might be able to reliably get the same evidence from them that I get from performing the action but without having to risk the possibility that it is wrong.

Political expertise

We have already seen Tetlock claiming that socially approved political experts are generally not genuine experts. But is political expertise even possible? Recall Jeffrey Friedman's theory of ideational determinism,[25] which I called a "tunneling" theory. One consequence of ideational determinism is that political expertise is not possible. Recall that in broad strokes, Friedman claims that the modern social world is too complex for expertise about it to be possible. We have to rely on heuristics to work through the vast repository of interrelated information, and such heuristics end up determining what we believe. So, for Friedman, political experts are experts, at best, at applying the heuristics they've ended up with, but this doesn't make them particularly reliable.

Is Friedman painting with too broad a brush? Recall that we saw in the first chapter that *any* belief can be a political belief under the right circumstances, if subject to the political kind of controversy and positioned as a premise in practical reasoning toward political conclusions. So we might think that this negative view of political expertise leads to a negative view of expertise about anything at all. But even expertise in distinctly social and political domains doesn't seem intuitively impossible. We accept that there are sociologists, economists, national security experts, and so on. The use of heuristics may help rather than harm their reasoning. Political expertise may be possible after all.

Standpoint expertise

Another type of theory that has consequences for our thinking about political expertise is the theory of positionality and standpoint. From this kind of theory we can extract a thesis of *standpoint expertise* – roughly, that marginalized people are experts in the domain of issues relating to their marginalization. If this is the case, then we owe marginalized people whatever sort of deference we have determined in our examination of the deference question. Let's call that further conclusion *standpoint deference* or the standpoint deference thesis.

People who don't like the idea of identity-based deference sometimes try to resist the standpoint deference thesis obliquely, without attacking standpoint expertise itself. For instance, political writer Matt Bruenig, in his essay "What Does Identitarian Deference Require?",[26] argues roughly that standpoint deference requires that we take ourselves to be experts on who is marginalized, that standpoint deference ignores the fact that there is significant disagreement on political issues among members of any marginalized group, and that standpoint deference ignores the fact that majorities of different marginalized groups may disagree with one another. But now that we have considered our seven questions about expertise, we can see that Bruenig's concerns are simply versions of the identification question and the expert disagreement question. So they do little to challenge standpoint deference uniquely, since such worries attach to any form of desired expert deference.

If we want to understand standpoint deference, we must understand standpoint expertise; if we want to challenge standpoint deference, we must challenge standpoint expertise. So: What is appealing about the standpoint expertise thesis? One initial thought is that members of marginalized groups might understand how it feels, or know what it's like, to occupy a certain social position, and that that's understanding and knowledge that people in different social positions could not acquire without listening to members of those groups – and perhaps could only partially acquire even when they do listen. Yet knowing what something is like does not suffice for political expertise. I know what it's like to attend college and to take out loans to attend law and then graduate school. It would be odd to say that this makes me an expert on education and student loan policy. To understand such things I would need a fuller perspective. For instance, I might need to know research on educational outcomes, or I might need to understand trade-offs, like what other areas would lose out if more money were to be invested in education and what the effects of that loss would be. Further, even some important aspects of "knowing what it's like" may not be accessible to marginalized people.[27] And Lidal Dror has noted that "there are informative experiences about the workings of social marginalization that members of *privileged* groups disproportionately have," so that there is little reason to think that, *in principle*, members of marginalized groups have an epistemic advantage over them.[28]

Standpoint theorists themselves give a variety of justifications for the standpoint expertise thesis.[29] They claim that some or all members of marginalized groups – generally, those who have "achieved" the relevant standpoint – will have better evidence, better concepts for making sense of evidence, more relevant experiences, greater openness to alternative hypotheses, a greater interest in justice, and more epistemic virtues like humility. However, these all seem like different explanations for some empirical phenomenon of standpoint expertise. Such a phenomenon is not in fact empirically attested; as Jessica Flanigan puts it, "there is no generalizable evidence that oppression per se makes people better epistemic agents."[30] We might expect a standpoint expertise thesis to be supported by survey data like the data we saw aimed against epistemic democracy in a previous chapter, but such surveys are rarely brought to bear on this question. The often-cited limitation that a genuinely expert standpoint must be "achieved" by members of marginalized groups, often by "consciousness-raising," seems to require that the scales be tilted in favor of certain perspectives, often politically radical ones. This again suggests that standpoint deference does not truly put us in the position of the novice looking for an expert but requires that we begin by taking ourselves to know many substantial political truths.

Finally, let's recall from Chapter 8 one way in particular of establishing the standpoint expertise thesis: that powerless people must understand the powerful people's perspective to survive, what Dror calls "the 'insider-outsider' framework."[31] One way this might be cashed out is to say the following: two perspectives are better than one, and that's the epistemic advantage and the source of the expertise. But it's not always the case that the person who understands both perspectives is epistemically better off. A conspiracy theorist might understand the mainstream narrative better than advocates of the mainstream narrative understand the conspiracy theory, but that doesn't say much. We saw that moral foundations theory holds that conservatives in some sense understand liberal moral principles better than liberals understand conservative moral principles, but that doesn't make conservatives right. This "double vision" approach does not establish expertise, because it does not establish that the most important criterion of expertise, that of a reliable track record on questions in the relevant domain, has been met.

Conclusion

Many philosophers and other writers think we should wholly defer to the judgments of experts to form accurate political beliefs, and many others think such deference is impossible or undesirable for a whole host of reasons. Neither of these extreme perspectives seems to be completely borne out by careful reasoning about questions of expertise. Rather, deference to experts, like every other

aspect of our political judgment, requires weighing a great number of factors, both theoretical and concrete. It is not a route that a completely unknowing reasoner can reliably take in forming political beliefs, but it may be one tool that a savvy reasoner can use in limited situations to improve the beliefs they do have.

Discussion questions

1. What sorts of people do you treat as experts in nonpolitical areas? What sorts of evidence do you draw on to justify deferring to them?
2. What sorts of people, if any, do you treat as experts about politics? Do you think there are any political experts who systematically disagree with your political beliefs?
3. Is there anyone to whom you defer on moral matters – anyone to whom you simply listen if they tell you what's right and wrong in some situation? Do you feel there's something "fishy" about, in essence, letting others direct your actions in this way?

17

THE ETHICS OF BELIEF

Epistemology is a normative discipline. What that means is that it makes evaluations as to when things have gone right or wrong, when one situation is better than another, when agents are built in the right or wrong sorts of ways, or so on. Sometimes, people believe what they shouldn't or fail to believe what they should. We might call these moments of irrationality. We can compare this with a discipline like ethics. Ethics tells us that agents sometimes act in right or wrong ways and that some situations have more value than others. Ethics does not tell us how people *will* act, only how they *should*. Similarly, epistemology is less concerned with how people *do* reason and what they *do* believe and more concerned with how they *should* reason and what they *ought to* believe. We can contrast this with a discipline like physics. A deterministic physical law tells us how things will go. A statistical physical law tells us how things tend to go, or what the distribution of the ways things end up will be if we have a sufficient number of situations with the same starting place, or something like that. There is no question of praising a baseball if its arc is perfectly described by Newtonian mechanics. Nor would there be a question of blaming a baseball if Newtonian mechanics failed to describe its arc.

In the course of disagreements over political beliefs – some more than others – one or both disputants will sometimes offer the following kind of argument as a reason for the other to change their mind: the political belief you hold is *evil*, or at least morally wrong somehow; it reflects on your character, and you *ought* not to hold it, in an ethically robust sense. The point of this chapter is to see if we can make sense of this idea. First, we'll look at a "Big Argument" which aims to undermine the whole idea of morally evaluating beliefs. Both of its premises seem plausible, but they've both been challenged, too; even if we can't figure

DOI: 10.4324/9781003355274-20

out exactly what's wrong about them, the Big Argument might go too far, undermining evaluations of beliefs based on rationality as well. Then we'll look at some more recent ideas, including the notion of right and wrong kinds of reasons, and "encroachment" of practical and moral stakes into the thresholds for rationality, justification, and knowledge when it comes to epistemic evaluations of beliefs.

The Big Argument is pretty simple: our beliefs are not up to us, and moral evaluations of some person can be rightly made only on the basis of what's up to that person, so we can't be morally evaluated on the basis of our beliefs, or in other words we can't be responsible for them. The first premise, that proper moral evaluations concern only what's up to the person being evaluated, is the *ought implies can principle* or just *ought-implies-can* for short. The second premise, that our beliefs are not up to us, is called *doxastic involuntarism*. "Doxastic" is a word that derives from ancient Greek (like "philosophy" and many other words in philosophy), and it means something like "relating to beliefs." Both of these premises are famous, deep, and controversial philosophical views with implications ranging far outside the study of political beliefs.

First premise of the Big Argument: ought implies can

Why think that ought implies can? Many important philosophers throughout history, like Immanuel Kant and G. E. Moore, have thought so,[1] and contemporary philosophers have tried to justify this thought. Bart Streumer argues that ought-implies-can is the only thing that can justify or explain certain moral judgments.[2] Horrible things happened in the past, for instance. What explains the fact that we don't have a responsibility to change those things? The fact that we *can't* go back and change them. David Copp has offered two more arguments in favor of ought-implies-can, both of which concern the nature of moral evaluations.[3] First, Copp says, it would be *unfair* if morality required that an agent do something that that agent is actually incapable of doing, and the idea that morality itself could be unfair is unintelligible, so morality must not require that an agent do something they're incapable of doing. I like this argument, but some would object that morality is in fact unfair, since some argue that one's actions can be better or worse merely on the basis of luck.[4] Second, Copp says, the "point" of morality is to pick out a best decision among a list of possibilities. It would thus be pointless for those actions we can't perform to even be on the list; so ought implies can. Ought-implies-can explains some other things about our moral practices, too, like the fact that we often morally soothe people who feel guilt by telling them there was nothing they could have done. And it figures in a very important argument related to free will; some take the thesis that we have no free will to mean that nothing we do is up to us, which on some understandings of ought-implies-can would mean that we aren't responsible for anything we do, either – that moral responsibility is a kind of myth.

Wesley Buckwalter has recently challenged the view that ought implies can. Buckwalter draws on a wide range of evidence in a cumulative way, aiming to topple ought-implies-can not with a single counterexample but with an argument about what sorts of principles best explain our moral judgments. First, Buckwalter says, ought-implies-can makes it impossible to explain why we often apologize or make excuses for not doing something we couldn't have done; he gives the example of a car accident in bad weather conditions, conditions for which the driver had no responsibility and which left the driver with no control over their vehicle, writing that the driver is still expected to make an apology under such circumstances.[5] Ought-implies-can also makes it possible for people to get out of moral obligations by making them impossible to fulfill, a counterintuitive result;[6] similarly, it would render null certain promises, namely one's promises to do things one cannot (promises one literally cannot keep), which intuitively do generate obligations and can be "genuine" promises.[7] And ought-implies-can means that there are no moral dilemmas – situations in which one is obligated both to do some action A and to do some action A*, despite the fact that they are unable to do both simultaneously, as when a lifeguard can only save one of two drowning swimmers.[8] Certain roles like those of parents or teachers, Buckwalter writes, also come with "oughts," which do not depend on ability – there are just certain things parents and teachers ought to do in virtue of being parents and teachers, and this isn't relative to the abilities of any particular parent or teacher.[9] And a person can coherently complain that they have been prevented from doing what they ought to have done; Buckwalter gives the example of someone who has been tasked with picking up broken glass on a path but who is tied to a tree by mischievous children.[10]

In an article in favor of ought-implies-can, Peter B. M. Vranas considers some further cases that might weigh against it. We might think there are obligations to feel a certain way, even though our feelings, like our beliefs, might be outside our active control; for instance, we might think that if someone does us a favor, we're obligated to feel gratitude, even though what we feel isn't up to us.[11] Vranas moreover notes that we have practices of blaming people with addictions or certain kinds of psychological compulsions for bad acts they do even though we might not think that behavior that satiates an addiction – kleptomania is a good example Vranas mentions here – is really under the control of the addict.[12] Vranas also considers arguments from "Frankfurt cases."[13] Named after Harry Frankfurt, the philosopher who developed them, Frankfurt cases involve agents who seem to have made a choice but could not have acted otherwise. Take a case like this:

> An aspiring monopolist hires an assassin to murder his competitor. Unbeknownst to the assassin, he also hires a neuroscientist to install a small chip in the assassin's brain. If the assassin decides not to commit the murder, the chip will manipulate her neurons, synapses, and so on to override his decision and

> cause her to in fact carry out the assassination after all. The assassin has not been made aware of this contingency plan. In the end, the assassin decides to commit the murder, and the chip does not have to kick in.[14]

In such a case, it seems that it wasn't up to the assassin whether or not she committed the murder. However, it still seems that she is morally responsible. So there is a conundrum for ought-implies-can here too: if ought implies can, we don't seem to be able to express our intuitive view about the assassin, which is something like this: she shouldn't have committed the murder.

There isn't space to go through all the potential replies to these objections, and the counters to those replies, and so on. But one particular strategy that applies to quite a few of them bears mentioning. When we seem to want to say "S ought to do X" even when S cannot do X, we might think this is a way of paraphrasing another kind of obligation, which is within the ability of the agent. In the case of apparent obligations for feelings, that might be (as Vranas suggests) the obligation to *behave as though* one feels a certain way. For example, if one has received a favor, one might owe an expression of gratitude rather than a feeling of gratitude. In the case of getting out of our obligations by making them impossible, it might be the past obligation to not have done this. For example, if a professor seems to get out of the obligation to teach a class by boarding a plane ten minutes before the class begins, and their dean says "That professor ought to be in that classroom right now!", we might take the dean's real meaning to be that the professor ought not to have boarded the plane. As for the Frankfurt case we considered earlier, we might think that the assassin has the responsibility to make the decision not to commit the murder, or to try to make that decision, even though that attempted decision won't effectuate any real-world difference in outcome. I encourage readers to decide how satisfying this strategy seems to them and how many of the objections to ought-implies-can it can handle.

Second premise of the Big Argument: doxastic involuntarism

Why think that we can't control our beliefs? Let's start with the intuitive case; Jonathan Bennett explains that "most of those who think about the question have found it incoherent or absurd to suppose anyone should acquire a belief, just like that, simply because he wanted to, as though acquiring a belief were like raising one's arm."[15] You can't screw up your eyes and try really hard and end up with the belief that the sky is red. Is that because you have so much evidence against it? No: pick some foreign country; you also can't screw up your eyes and try really hard and end up with the belief that it's raining there, a proposition about which you have no evidence either way.

However, there is also an intuitive case against doxastic involuntarism – that is, for the opposite thesis that we can sometimes control our beliefs, what's called doxastic voluntarism. First, to some philosophers it has seemed that we

choose our beliefs when we choose to stop deliberating or inquiring into some matter. We can't simply choose against evidence, but we choose to believe when we make the decision to take the evidence we have as sufficient for forming and maintaining a belief. Second, people do sometimes report feeling as though they have chosen to believe things that make them more comfortable. A parent might say: "I don't know what's happening at that party, but I'm choosing to believe that it's completely innocent." Or a friend of someone accused of a crime might say: "The evidence hasn't all been evaluated yet, but I'm choosing to stick with my friend and believe their story." A student might say: "The test was a blur, but I'm choosing to believe that I did just fine on it, so that I can go about my day." People even say things like: "I don't know if there is an afterlife, but I'm choosing to believe there is, because the alternative is so discomforting." Third, there is the matter of indirect control of belief. Though I might not be able to simply scrunch up my eyes and form a new belief, I might be able to cause myself to form a certain belief by choosing where I get my evidence. If I know that the library in this town has only books that argue for some proposition and I know that the library in the next town over has only books that argue against that proposition, then I might be able to cause myself to believe that proposition (or not) by only reading the books in this library (or the other).

Some classic arguments for doxastic involuntarism come from Bernard Williams.[16] Two of them are what Selim Berker calls the "before the fact" and "after the fact" arguments;[17] I think of them as the "aiming at truth" and "knowing about control" arguments. They have some similarities; they're both based on the fact that beliefs "aim at truth" – a philosophical slogan that essentially means that we give up our beliefs when they turn out false, and think we've done well when they turn out true. The first argument says that since beliefs aim at truth, we can't knowingly form them by mechanisms unrelated to the truth, like choosing them by an act of will. Nishi Shah has objected to this argument on the grounds that it infers too much from the mere fact that beliefs have an aim.[18] Lots of other activities have aims; Shah gives the example of cake-baking. Baking a cake, Shah says, requires mixing ingredients, and so it can be said that we're not baking a cake if we aren't at some point mixing ingredients. Still, we mix ingredients voluntarily when we bake a cake. Similarly, even if we accept that some mental state would not count as a belief if it weren't aimed at the truth, we could deny that choosing a belief isn't itself a way to aim at the truth. But I think we might also doubt that we can't knowingly form beliefs that don't go against the aim of belief. There's no reason to think we couldn't do this through indirect means, as in the library example above.

The second argument says that since beliefs aim at truth, if we were ever to realize that we'd formed our beliefs through choosing them by an act of will, we would immediately lose those beliefs and further that if we were able to choose our beliefs that way, we'd know we had that capacity. However, as Bennett points out, this does not count in favor of doxastic involuntarism. Here's an

analogy: I wake up whenever I realize I'm dreaming, but that doesn't mean I'm never in a dream. Or a closer analogy: whenever we realize a belief is false, we immediately lose that belief, but that doesn't mean we never believe falsehoods. So both of Williams's arguments face very tough challenges.

The idea that belief aims at truth seems related to the representationalist picture of belief considered in an earlier chapter. We also considered a dispositionalist theory of belief, and Carl Ginet has made an argument for doxastic voluntarism based on dispositionalism.[19] Ginet writes that we can decide to count on something's being the case, and that when one counts on something's being the case – that is, when one "stakes something" on something's being the case and also "adopt[s] a dismissive or complacent attitude toward the possibility of losing what one has staked"[20] – that is enough for believing it to be the case. Ginet gives a series of examples; here's just one, which bears some similarities to cases discussed earlier:

> Before Sam left for his office this morning, Sue asked him to bring from his office, when he comes back, a particular book that she needs to use in preparing for her lecture the next day. Later Sue wonders whether Sam will remember to bring the book. She recalls that he has sometimes, though not often, forgotten such things, but, given the inconvenience of getting in touch with him and interrupting his work and the thought that her continuing to wonder whether he'll remember it will make her anxious all day, she decides to stop fretting and believe that he will remember to bring it.[21]

This example makes sense from the dispositionalist perspective. And it also seems to accord with the way we often act. However, I think it is also easy to see why the doxastic involuntarists don't take it to qualify as the sort of thing they want to call a belief. It really doesn't seem that Sue is imposing on herself a representation of the world in which Sam will bring the book. Rather, she is sort of trying to find a way not to think about it, because every way she has of inquiring further about it is somehow annoying. That inquiring further is annoying, however, doesn't seem like evidence for a new belief.

We will continue to consider cases like this later in the chapter, when we talk about the relationship between our actions and what's rational or justified for us to believe. First, though, we'll talk about two potential odd consequences of the Big Argument.

First odd consequence of the Big Argument: moral blame for moral beliefs

Say we aren't morally to blame for our beliefs, whether they turn out true or false. As we've discussed several times, some of our beliefs are moral beliefs – beliefs about what's right and wrong. If we're not to blame when those beliefs

go wrong, we might be in an odd position: we might start to find it hard to blame anyone for anything at all! We often don't blame people for wrong actions that are morally right from the perspective of their beliefs, as long as they can't be blamed for those beliefs. For instance, say that, unbeknownst to me, I live in a world of magic, and in this world of magic, whenever I use my toaster, a small animal experiences some pain. Say I have no way of knowing this and that, indeed, I have no reason to believe it. Then it seems that I cannot be blamed for the animal's pain. I am causally responsible for it but not morally responsible for it. But if I cannot be blamed for any of my beliefs, the conclusion of the Big Argument, then the following case is similar. Say that I know perfectly well that my toaster causes some small animal to experience pain and that rather I don't believe that causing small animals to experience pain is bad. If the conclusion of the Big Argument is correct, then it might also seem that I cannot be blamed for the animal's pain in this situation, either.

There are probably a whole bunch of ways to deal with this conundrum. One, the "volitionist" response, is to accept the conclusion: when we act in line with our moral beliefs (for which, by the conclusion of the Big Argument, we can't be blamed), we can't be blamed for wrongdoing. Does that mean we can *never* be blamed for wrongdoing? No, we can still be blamed when we act *against* our genuinely held moral beliefs, exhibiting what's sometimes called moral incontinence, weakness of will, or akrasia.[22] However, one worry about this response is that it will excuse a lot of intuitively bad actors, like "true believers" in incredibly evil causes.

Another response posits that there is an "asymmetry" between factual beliefs and moral beliefs; Clayton Littlejohn, for instance, writes that "[n]on-culpable factual mistake and ignorance will typically excuse the agent's behavior, but non-culpable evaluative mistake and ignorance will typically not."[23] In other words, we might simply carve out different spaces for the relationship between mistaken beliefs and blameworthy actions based on the type of belief. This response might be criticized for seeming *ad hoc*. If we accept that we should in general be excused for acting on the basis of wrong beliefs when we can't be blamed for holding them, it is not immediately clear what non-arbitrary rationale we have for denying that we should be excused for acting on the basis of beliefs that are not themselves blameworthy, when all that sets those beliefs apart is that they're moral beliefs.

Second odd consequence of the Big Argument: other kinds of responsibility

The Big Argument is often posed not just as an argument about whether we can blame people in a moral sense for holding certain beliefs, but whether we can blame them in any sense whatsoever, including the sense of rationality operative within the kind of normative epistemology that we're taking for granted in this

text.[24] After all, the very fact that rationality is a normative concept means that it contains a kind of "ought" in it. So when we call someone's belief irrational, we're sort of saying that they should not hold it. But to say that, by ought-implies-can, means that they could decide to stop holding it; but by doxastic involuntarism, they could not decide that. So we might think that the Big Argument causes us trouble when extended from morality to rationality.

One strategy for getting out of this jam is to suggest that ought-implies-can is true for the "ought" of morality but not the "ought" of rationality. Richard Feldman has argued that rationality and epistemic "responsibility" are closer to contractual obligations or to the obligations incurred by occupying a certain role,[25] a kind of "ought" we considered earlier as potentially having robust moral force. However, I myself think ought-implies-can is a feature of normativity in general, not just of morality; it explains, for instance, why we never ought to both believe and not believe something.

How do we narrow ought-implies-can for the specific normative realm of epistemology, or epistemic rationality, without destroying it completely? We can take the normativity of games as an example. In any chess position, there is a list of moves that are allowed by the rules of chess, of which one or more are best. If white cannot castle in a certain position, then it is no good to say "you ought to have castled here"; if black cannot queen a pawn in a certain position, then it is no good to say "you ought to have promoted your pawn here." So ought implies can for chess but only for a very weak form of "can" which emerges from the formal structure of the game. Similarly, the "ought" of rational credence, belief, inference, updating, and so on is, I think, generated by the formal structure of a representational system which has credences and beliefs and performs inferences and updates. But I don't have the space to expand much on that thought here.

Ought-implies-can relates to the structure of rationality in a different way which merits a quick mention. There are two broad camps in philosophy when it comes to what rationality requires: the camp that thinks it's very strict (the "ideal theorists") and the camp that thinks it's relatively lenient (the "non-ideal theorists"). One of the arguments the non-ideal theorists marshal in favor of their position is based in ought-implies-can.[26] Ideal theories of rationality seem to demand that we do or have things like perfect recall, knowledge of all logically provable mathematical theorems, and instant updates to our set of beliefs based on new information. But humans are not built to be able to do or have these sorts of things, and since ought implies can, that means we can't be blamed for failing to do or have them. Instead, epistemology should be sensitive to our cognitive limitations and should give recommendations that human agents can actually follow, given the oddities and imperfections of human psychology.

Jennifer Rose Carr argues that this objection to ideal theorists is misguided.[27] She says something a bit like this: the whole background idea of non-ideal epistemology is that epistemology should be concerned with substantively guiding

reasoners given their cognitive limitations. But first, different people have different cognitive limitations. For instance, I might be cognitively limited in being able to do A but not B, and you might be cognitively limited in being able to do B but not A. But then if non-ideal epistemology is correct in its interpretation of the epistemic "ought" and the epistemic "can," there should be something like a different epistemology for you than there is for me, which would seem to strip the field of its subject matter entirely. In theory, non-ideal epistemologists could potentially escape this objection by taking something like a "lowest common denominator" approach and saying that epistemology should be sensitive to the cognitive limitations that *any* human has. But in all likelihood, for virtually any cognitive capacity, it is compromised or absent in at least one human being. So there would be nothing left for epistemology to guide us to do at all; non-ideal epistemology would simply no longer be a normative field. And second, we might want epistemology to be sensitive to some cognitive limitations but not others, regardless of whether or not they're universal. For instance, even if limited memory and the use of stereotypes are both universal among humans, we might think limited memory should be excused, but the use of stereotypes should not.[28] But if non-ideal epistemology takes this tack, then it ends up violating its own interpretation of the ought-implies-can principle in a way that seems arbitrary when viewed in light of its own theoretical motivations.

Another worry: right and wrong kinds of reasons

So much for the Big Argument. Here's another worry about the relationship between morality and belief. Even if beliefs can be morally wrong, we might think that shouldn't matter to epistemologists, because ethical reasons are not the sorts of reasons for belief that affect intellectual rationality. This general kind of concern is known as the *wrong kind of reason problem*. Two of the philosophers who brought attention to this problem are Justin D'Arms and Daniel Jacobson. In their paper on the "moralistic fallacy," they note that moral reasons for or against things like chess moves, jokes, or musical scores often will not count for or against those things from the perspective of chess, humor, or music.[29]

To see why we might think there are distinct areas of normativity which we judge by distinct criteria, consider a thought experiment. Imagine you are playing a chess game with someone. Suddenly, this person tells you, in a way you find convincing, that if your next move is not exceedingly poor, an ally of theirs will steal your rare stamp collection. You had planned to sell this collection right after the game had concluded and to use the money for some very good charitable cause. So it might seem to you that the morally right thing to do, overall, is to play a move that is exceedingly poor by chess standards. The standards of morality and chess come apart in this way.

Here's a case from Alex Worsnip, which he says is of a sort that is "frequently discussed": "You have strong (but not utterly infallible) evidence that the game starts at 3pm. But I, an eccentric millionaire, offer you a bribe: if you can avoid believing that the game starts at 3pm, I will give you $1,000,000."[30] Note that this gives me a really good reason to avoid believing that the game starts at 3pm: I could really use a million dollars. The chance of acquiring a great deal of money is a reason to do all sorts of things, like buying a lottery ticket, taking on a job, writing a novel, making an investment, and so on. So we might think that the offer of money shifts the balance of my reasons in favor of not believing that there will be a soccer game, despite the fact that the evidence favors believing that there will be a soccer game. But the shift is not relevant to judgments about the rationality of my beliefs from the perspective of epistemology. The wrong kind of reasons problem is explaining just what it means that such a perspective, distinct from the perspective of my practical interests or desires, exists.

Mark Schroeder has noted that the wrong kind of reasons problem is very general: "Wrong Kind of Reasons issues arise in any domain which is governed by a standard of correctness" – for instance, "tying a knot, executing [chess] endgames, and setting the table."[31] Just as we can conceive of someone offering you money to believe something unreasonable, we can conceive of someone offering you money to tie a knot in the wrong way. If you take the money, though the knot-tying you perform might be a good action prudentially or morally, it will still be a bad knot.

One approach has been to distinguish between reasons that seem to apply to an action, attitude, or state itself and reasons that apply somehow to the content or object of an action, attitude, or state.[32] For instance, Derek Parfit wrote that "state-given reasons," or reasons that "would make it better if we had some belief or desire," are reasons of the wrong kind (or not reasons at all).[33] However, Schroeder and Pamela Hieronymi have separately challenged these kinds of accounts for being insufficiently general. Hieronymi notes that attitudes like "supposing, imagining, and remembering" can have state-given (or "attitude-related" in Christian Piller's terminology) reasons of the right kind.[34] So the problem remains tricky.

As far as the potential conflict between moral reasons and epistemic reasons is concerned, Selim Berker has argued that there is a crucial difference between practical reasons and epistemic reasons.[35] Reasons for action balance against each other *permissively* while reasons for belief balance against each other *prohibitively*. That is, if you have equal reason to perform or to not perform some action, it is both rational to perform the action and rational to not perform the action. However, if you have equal reason to believe or to not believe some proposition, it is neither rational to believe the proposition is true nor rational to believe the proposition is false. What is rational in the balanced epistemic case is to suspend judgment – to withhold belief. Berker asks us to consider a

case where an agent has equal reasons, though reasons of the wrong kind, to believe and to not believe some proposition. For instance, I might have been offered a great deal of money to believe that the soccer game occurs at 3pm, and I might have been offered the same amount of money to believe that the soccer game does not occur at 3pm. Berker notes that in such a case, the balancing is permissive, not prohibitive. He favors what he calls a "Simple Explanation" for this fact:

> What people call "practical reasons for belief" are really practical reasons to perform a certain action, namely the action of bringing it about that one believes some proposition. Since they are just practical reasons for action, like all practical reasons for action they exhibit permissive balancing.[36]

I find this explanation compelling.

Another route to the ethics of belief: encroachment theories

Some philosophers have argued that practical and moral factors can affect evaluations from the perspective of epistemology even if practical and moral reasons don't count within epistemology. How could that be? Through a posited mechanism called *encroachment*, separated into *pragmatic encroachment* and *moral encroachment*, which occurs when pragmatic or moral factors shift the threshold of evidence, justification, reasons, or whatever else necessary to make a belief rational. These outside factors don't count as reasons, but they can nonetheless affect what we ought to believe.

Blake Roeber distinguishes between two kinds of arguments for pragmatic encroachment.[37] One is the *intuition-based argument*. The canonical presentation of this type of argument comes from Jason Stanley,[38] adapting a thought experiment posed by Keith DeRose in another context.[39] Stanley asks us to consider two people hoping to deposit their paychecks but looking at a long line at a bank on a Friday afternoon. They each remember the bank being open on Saturday several weeks prior. In one case, which Stanley calls "Low Stakes," the person has no impending need of the money. In the other case, which Stanley calls "High Stakes," the person has bills due and no savings and needs the money by, say, Sunday. We should assume for the purposes of argument that the cases are otherwise fully similar. Stanley then argues that the person in the Low Stakes case seems to *know* that the bank will be open on Saturday while the person in the High Stakes case seems *not* to know that the bank will be open on Saturday. The only possible explanation is the difference in their practical interests, that is, the difference in the urgency of their needs for the money.

Roeber has noted, however, that these cases can be reinterpreted to count against pragmatic encroachment. While it might seem sensible to say that the

person in the High Stakes case doesn't know that the bank will be open, it seems just as sensible to say that they *do* know but that they ought to play it safe and double-check.[40] And Alex Worsnip has given a clever argument about these sorts of cases.[41] In essence, he asks: What distinguishes these cases from the bribe case considered earlier? The presence of a practical factor, money, seems to Stanley to be enough to justify suspending belief about whether the bank was open. But in the bribe case, the presence of the practical factor didn't seem like it could rationally affect our beliefs at all. Worsnip surveys possible candidate explanations for why this difference exists and finds them all wanting, concluding that there might be no way to believe in pragmatic encroachment while allowing that practical reasons for belief are reasons of the wrong kind.

Another kind of argument for pragmatic encroachment is the *principle-based argument*, the existence of the phenomenon of pragmatic encroachment is inferred from general principles relating rational action and justified belief. We have already seen that there is a deep connection between action and belief, which was expressed at a first pass in the idea that actions are rationalized by reference to belief-desire pairs. These principles go into more detail. For instance, John Hawthorne and Jason Stanley hold that one knows something only if one can rationally act on it.[42] This is sometimes called the *knowledge-action principle*. And Jeremy Fantl and Matthew McGrath hold that one is justified in believing something if and only if one can rationally act on the belief.[43] This is sometimes called the *justification-action principle*. We can also imagine a *rationality-action principle* along similar lines – in fact, we saw Ginet giving a principle like that above, in line with dispositionalism about belief. But then we can construct an argument like this: one is rational in believing *p* only if one can act as though *p* is the case; whether one can act as though *p* is the case depends on practical factors (like the risks of being wrong); and so whether one is rational in believing *p* depends on practical factors, too.

There are several ways of attacking these kinds of principles. First, we might think they're just counterintuitive. In the context of making assertions, and the norms governing assertions, Jennifer Lackey[44] has discussed the fact that we are sometimes allowed or even obligated to act against our beliefs, for example if we occupy a certain role (her examples involve a teacher, a juror, and a doctor). It might be that we ought to defer to professional communities, their consensuses or norms, when we occupy certain roles, and act based on those rather than our own beliefs, and this is especially the case if we have doubts about the rationality of our own beliefs.

Roeber offers a different kind of objection, what I call the *disjunctive answer objection*.[45] Say I am presented with a bunch of sentences, and I will receive some prize simply for picking a true sentence from this bunch. One sentence says: "The capital of England is London." Another sentence says: "The capital of England is London or the capital of France is Paris." It's irrational for me to pick the first sentence, since the first cannot be true without the second being

true while the second can be true without the first. So it is irrational for me to act on that proposition, and if the principles in question were right, that would mean that I was not justified or rational in believing it or that I did not know it. But, Roeber says, it is highly intuitive that I can in fact retain my states of justification, rationality, and knowledge despite being in this test situation. Roeber's objection is quite fecund. We could turn it into an objection to the effect that the knowledge-action principle entails skepticism, since any proposition whatsoever can be put into a disjunction with some other proposition. Or we could imagine a scenario with three options: "The capital of England is London," "The capital of France is Paris," and "The capital of England is London or the capital of France is Paris." In this last case, the principles will suggest that we are justified or rational in believing, or that we know, the disjunction despite being justified or rational in believing, or knowing, neither of its (logically unrelated) disjuncts, which is at the very least an odd result.

I think these principles, or at least the ones that are formulated as biconditionals rather than mere conditionals, can also be attacked by a *simultaneity objection*. Put sketchily, the idea of such an objection is as follows. At some moment, we might have several different actions which might be based on a belief in the same proposition. Based on the costs and benefits of certain eventualities, it might be the case that some of these actions are rational while others aren't. For instance, say I believe, but am not certain, that it will not rain, and say I am hosting two events tonight, one which will be attended only by locals and the other which will be attended by folks from out of town. Then it might be rational for me to cancel the latter but not the former event, because an event being rained out is a much greater cost to folks who have traveled from out of town than it is to locals. If the principles in question are true, this would mean that I simultaneously know and don't know, or am justified and unjustified in believing, that it won't rain. But that is just a contradiction. So the principles come out looking wrong.

It is pretty simple to extend this argument that practical factors can matter for knowledge or justification to an argument that moral factors matter for knowledge or justification by adjusting the cases to make the relevant interests belong to a party other than the epistemic agent.[46] Some go further. Rima Basu and Mark Schroeder have argued that beliefs themselves can be morally wrong (a topic we considered earlier) and that because of encroachment, morally wrong beliefs can have higher thresholds of justification.[47] James Fritz calls the former, simple kind of encroachment from moral factors *moderate moral encroachment* and the latter kind of encroachment from the morality of beliefs themselves *radical moral encroachment*.[48] Here is the sort of case of radical moral encroachment Basu and Schroeder have in mind, as given by Renee Bolinger.[49]

> The night before he is to be presented with the Presidential Medal of Freedom, John Hope Franklin hosts a celebratory dinner party at the Cosmos

> Club. All the other black men in the club are uniformed attendants. While walking through the club, a woman sees him, calls him over, presents her coat check ticket and orders him to bring her coat.

Basu and Schroeder think the woman's belief that John Hope Franklin is an attendant is morally wrong, based as it is on the fact that he is black. And they think morally wrong beliefs have moral stakes, which gives them a higher threshold of necessary justification (through encroachment), making this one irrational.

Fritz has argued that radical moral encroachment cannot avoid the wrong kind of reason problem,[50] and he and Elizabeth Jackson have argued that if there is radical moral encroachment on beliefs, there must also be encroachment on our levels of confidence in our beliefs, which all participants in the debate agree would be objectionable.[51]

However, even if there is radical moral encroachment, there is one really strong inference Schroeder takes from the proposed existence of that phenomenon that is really interesting for our purposes here. Schroeder argues that through radical moral encroachment, moral and epistemic evaluations of beliefs are unified. Any belief that would wrong someone will have a high enough threshold that to believe it would be irrational.[52] In other words, we are never in the position of being epistemically required, or perhaps even epistemically permitted, to believe something when that belief would be morally wrong. In a paper of my own, I argue that if the accounts of how beliefs can be morally wrong proposed by Basu and Schroeder are correct, this conclusion of Schroeder's must be incorrect.[53] Basu and Schroeder's accounts of what makes beliefs wrong all revolve around something in the reasoning process, about the sorts of evidence that should be admitted into forming certain kinds of beliefs about people. But it can never be rational – that is, epistemically praiseworthy – to exclude reasons or evidence from our belief-forming process, and so avoiding the beliefs that are on their accounts immoral would require an agent to be irrational. As with Worsnip's criticisms of pragmatic encroachment, there may be another forthcoming account which would solve this problem, but the existing explanations do not.

Political irrationality from encroachment

In this chapter we've talked a lot about morality and epistemology but not too much about political beliefs. The most obvious relationship between the topics is that the sorts of beliefs we might find abhorrent are often political in some sense or another. We might think it is evil to believe a certain conspiracy theory about members of the opposing political party or evil to disbelieve the advice given by members of our own party. Sometimes, it is said to be evil to deal in stereotypes about members of political groups (or other groups – such stereotypes are probably always political).

However, there is another potential relationship between encroachment theses and political beliefs, in that encroachment theses might be able to generate yet another argument for political skepticism, or at least for the irrationality, broadly speaking, of political beliefs. We can pose this argument as a dilemma. Either there are significant moral and practical stakes to our political beliefs or there are not. If there are significant moral and practical stakes to our political beliefs, and if encroachment theses are correct, then the threshold of evidence we need to justify political beliefs is very high, and it is hard to see how we could meet it on controversial political propositions. But if there are no significant moral and practical stakes to our political beliefs, then we are behaving irrationally in spending a lot of time researching and debating those beliefs, at least if we are not political figures with real political power ourselves. (Bryan Caplan makes something like this latter argument.)[54] This might not mean our beliefs aren't justified once we have the evidence, but then again, it might cast doubt on those beliefs as well: if our epistemic interest in politics is so irrational, it might undermine our faith in our own abilities to form rational judgments even once we have the evidence. We might simply expect ourselves to have some sort of bias.

Whether political beliefs have hefty practical and moral stakes is itself an interesting question. In a large democracy, our votes make little difference. We might think that they reflect something about our characters, but then again, the same moral system can generate different political beliefs given differences in empirical evidence.[55] I leave the reader to reflect on that at their leisure.

Conclusion

A deep argument, what I've called the Big Argument, suggests that we can never be morally praised or blamed for our beliefs, but this argument's two premises, ought-implies-can and doxastic involuntarism, are each subject to critique, and philosophical debate about them has not been entirely conclusive; further, the conclusion of the Big Argument generates some perplexities about whether we can be praised and blamed even for our actions, given that our actions are rationalized by moral beliefs and about the nature of rationality and epistemic responsibility. Even if the Big Argument is correct, encroachment phenomena may be sufficient to link epistemology and morality, but theories of pragmatic and moral encroachment are themselves open to a lot of counterarguments, and they may undermine the rationality of our political beliefs wholesale.

Discussion questions

1. Do you think some beliefs are moral and some beliefs are immoral? What sorts of beliefs do you have in mind? In particular, do you have political beliefs in mind?

2. Do you think our beliefs should change based on the moral or practical stakes of acting on them? How else might we think about how to integrate moral or practical stakes into our decision-making?
3. Do you feel you have control over your beliefs? Is there anything you feel you've actively made yourself believe or disbelieve? Can we distinguish believing from more clearly active processes like attending to certain facts, helping oneself or deceiving oneself, and so on?
4. How high do you think the stakes of our political beliefs are, as individuals? Do you think people act as though political beliefs are more important than they actually are, less important than they actually are, or just precisely as important as they actually are?

18

THE EPISTEMOLOGY OF LIBERALISM

In discussing ideology we saw that there was a complex way to think about it, involving some sort of top-down function of obscuring the dark truths about society, and a simple way to think about it, as a sort of comprehensive scheme for thinking about politics. Synonyms for ideology in the second, simple sense might include "worldview," "system of thought," or the German *Weltanschauung*. In this chapter we're going to think about a set of ideas that is an ideology in this simple sense, and that is maybe the dominant ideology, in that sense, of our times: the ideology of liberalism.

Like most topics we've considered in this text, whole books have been written about liberalism. For our purposes, it's enough to think of liberalism as committed to claims like the following: people should have whatever worldviews they see fit; institutions should be neutral about values and ideologies; personal autonomy and freedom are fundamental values; the main, perhaps the only, reason to restrict one person's freedom is to prevent them from restricting another person's. Liberalism is thus the hands-off ideology *par excellence*, at least according to some of its supporters. According to others, and according to some critics, liberalism can't remain completely hands-off, or it'll be destroyed by competing ideologies; and according to still other critics, liberalism isn't hands-off at all but is just as insistent on enforcing support for itself as other ideologies, like fascism and communism.

These sorts of worries about whether liberalism can be hands-off are part of the underlying concern of this chapter. They are sometimes popularized through Karl Popper's *paradox of tolerance*. Popper wrote that

> unlimited tolerance must lead to the disappearance of tolerance. If we extend unlimited tolerance even to those who are intolerant, if we are not prepared to

DOI: 10.4324/9781003355274-21

> defend a tolerant society against the onslaught of the intolerant, then the tolerant will be destroyed, and tolerance with them. . . . We should therefore claim, in the name of tolerance, the *right* not to tolerate the intolerant.[1]

Often presented online in cartoon form, this passage has been used by all sorts of American political actors to justify illiberal actions, despite the fact that the passage also included the following caveat:

> I do not imply, for instance, that we should always suppress the utterance of intolerant philosophies; as long as we can counter them by rational argument and keep them in check by public opinion, suppression would certainly be most unwise.[2]

Despite its misapplication, it makes an important theoretical point about whether liberalism can really survive if it remains neutral between itself and illiberal ideologies.

Criticisms of liberal neutrality

Kevin Vallier has distinguished between three criticisms of liberal neutrality: that it's incoherent or impossible, that it's infeasible or impractical, and that it's undesirable or unattractive.[3] He gives the following example:

> Imagine you're in a challenging marriage. You disagree with your spouse about everything: how to raise the kids, spend money, and so on. Imagine that, despite all your efforts, you can't agree. But you don't want a divorce because your life together is special. Divorce would be messy and expensive, and neither of you would be happier. So you decide to see a marriage therapist.
>
> The therapist helps you see that your marital disputes aren't going anywhere. She says that you must find ways to live together before conflict destroys your marriage and depletes your shared love. The therapist offers compromises. Neither of you likes her suggestions, but you admit that they will heal your pain. You're disappointed. You mourn the marriage you dreamed of, but you continue to love your spouse.
>
> Like you, the marriage therapist is another human being and has her own concerns. She has probably taken sides privately. But she distances herself from her personal view so that she can do her job. The therapist's job is still coherent and feasible. She draws on the human capacity to take the perspective of others to bring you two together. Therapists are not alone. If you live with other people, you take other [perspectives] all the time, if you're a decent person. In a fight, even if you think you are right, you frequently yield to others to sustain the relationship. It is not easy to get along with other humans, and we often fail. But it is still appropriate and even inspiring to try.

> The ideal of liberal neutrality expands this logic to the exercise of state power. The liberal state is not a therapist, of course. But it preserves functioning relationships between citizens and between the public and the state and prevents conflicts from erupting into violence. To do so, political officials do not take sides between our competing moral visions. At least within limits.[4]

Vallier's point is a clear one. There are cases where neutrality is obviously possible, feasible, and desirable. And some of those cases look a lot like what we should want from the state. Given that it's part of the state's job to maintain internal unity and avoid violent internal strife, it must be neutral.

Critics would, I think, object to this characterization. First, they might say that the marriage therapist is not a neutral party. The goal of the therapy is to save the marriage; it begins with a value judgment. Similarly, they might say, even a liberal state must make value judgments, deciding what to pursue and what to try to avoid. The judgment about internal strife is one such value judgment.

Some critics go further. They say that neutrality is impossible in the sense that a chosen "neutral" policy will always have certain effects compared to various kinds of non-neutral policies. The liberal neutralists are picking those effects in as non-neutral a way as they would be if they were picking the effects of the non-neutral policies.

I think these criticisms are pretty easy to answer. Neutrality is only neutrality *vis-a-vis* some specific set of choices and some specific goal. If a father tells his children that they can pick which restaurant to eat dinner at, he is neutral with regard to the question of restaurant but not with regard to the question of food entirely; he may remain committed to the family eating dinner *somewhere*, and he may cross the possibility of eating from a dumpster, for instance, off his list. This means that neutrality can come in degrees, and a critic might fairly wonder just how much non-neutrality it takes for something to no longer count as neutral; but the answer isn't "any non-neutrality at all." Nor does it make sense, if, say, two of the children vote to eat pizza and the third child votes to eat tacos, to say that the father has somehow picked pizza by letting his children vote; neutrality is a feature of the process, not the result.

David Beaver and Jason Stanley argue that liberal neutrality is impossible in a different way. They take neutrality to be "an ideal in deliberation," that of "speak[ing] without bias or taking sides."[5] Then a fully neutral deliberative space is one in which "participants in a discussion solely focus on exchanging reasons."[6] This is impossible, for Beaver and Stanley, "because utterances of words are moves in speech practices that have various resonances beyond the contents of those words."[7] So the idea is that neutrality is a matter of using words in a certain way, without connotation or resonance, which on their view is impossible. However, it's hard to see what the relationship is between Beaver and Stanley's notion of neutrality in language and the political liberal's notion of neutrality.

What about the idea that neutrality is undesirable? There's a clear sense in which this might be true. Perhaps there is some objectively best restaurant to which the father could take his children, and the father knows this. Perhaps one of the spouses is in the wrong and the other is in the right, and the marriage therapist knows this. Then it might be undesirable for the parties who don't know what the right choice is to be in the position of choosing. But there are a lot of caveats to make here. First, opposing liberalism on these grounds requires being confident that the process that replaces it will feature a decision-maker who consistently knows what's right and wrong in this way. We might think that it's rare to find such a person. Second, we might have reason to think that a neutral process won't just be *random* – that it will find the right answer if there's one to be found. We considered arguments for this kind of claim in regard to democracies and markets, for instance. Third, as Vallier notes, although the father or the marriage therapist might know what's right and wrong, there might be bad consequences of them imposing their view. It might be that the children or the spouses are better off figuring this out for themselves even if the decision-maker knows, or it might be that imposing a view causes some sort of conflict or violence, as Vallier worried it might. Fourth, we might think that it is a good in itself for a decision not to be imposed. Liberalism is often accused of being a form of proceduralism, but we might think that liberal neutrality and its freedoms are substantive goods. (Note that freedom itself is a kind of neutrality. If someone leaves you free to either do or not do X, then they are remaining practically neutral on the question of whether to do X.)

Liberalism and self-defeat; the burdens of judgment and the overlapping consensus

The argument that liberalism is self-defeating is related to the argument that liberal neutrality is impossible. Where one argument worries that an internal issue in liberalism means that it can't survive, the other accuses liberalism of being only falsely neutral, explaining its survival. If we accept these arguments together, then either classical liberalism isn't actually liberal; that is, it enforces the worldview of liberalism in a non-neutral way on people and institutions and demands that people conform to the ideology, making it *logically* self-defeating; or classical liberalism isn't actually capable of defending itself or establishing a stable society; that is, it demands that people and institutions stand by neutrally as liberal principles are attacked or eroded, making it *practically* self-defeating.

These concerns are also closely related to the concerns that animated John Rawls in his book *Political Liberalism*. Rawls asks: "How is it possible for there to exist over time a just and stable society of free and equal citizens, who remain profoundly divided by reasonable religious, philosophical, and moral doctrines?"[8] We can interpret those questions in our terms: How can a society

remain liberal while also being just (avoiding logical self-defeat, that is, maintaining liberal principles) and stable (avoiding practical self-defeat)?

Rawls's answer relies on the notion of the *burdens of judgment*. The burdens of judgment are the features of moral and political evaluation that make such evaluation difficult and thus that explain the possibility of reasonable disagreement. In essence, Rawls is talking about all the sorts of things we've already talked about and will continue to talk about in this book and drawing out a lesson in political theory from them. He counts among the burdens of judgment (I'm paraphrasing) that: (a) evidence may be complex and contain conflicts, (b) people may disagree about how important different considerations are, (c) the concepts we use to navigate political situations are vague, (d) people's experiences shape their views, and so the variety of human experiences leads to a variety of human opinions, (e) competing values may be hard to compare, or even incommensurable, and (f) not every consideration can be feasibly taken into account in our social, political, and legal institutions.[9] A different textbook on political epistemology could be written with one section on each of these considerations; they are certainly real issues.

According to Rawls, reasonable citizens will recognize these burdens and therefore their own fallibility and the possibility of reasonable disagreement; thus, "they endorse some form of liberty of conscience and freedom of thought. It is unreasonable of us to use political power, should we possess it, or share it with others, to repress comprehensive views that are not unreasonable."[10] Note that while "burdens of judgment" is an epistemological concept, "reasonable" here is not meant in an epistemological sense but rather in the sense of *fairness*, which is related to his broader political philosophy.[11] For Rawls, controversially, reasonableness also requires *publicity*, meaning that a reasonable point in political deliberation is one that others are in some sense able to accept. (Some object that this unacceptably rules out religious conviction.)[12] Since, therefore, all reasonable ideologies will respect some basic liberal principles, liberalism needn't impose itself on them. Rather, liberalism can emerge through an "overlapping consensus," in which all reasonable ideologies agree on those principles.[13] Importantly, each ideology will endorse liberalism "from its own point of view"[14] from within the values of its own ideology.

Rawls's argument is clever, and it's been extraordinarily influential. However, there's at least one troubling counterargument available. In a review of *Political Liberalism*, Michael Sandel notes that we can agree with Rawls that every reasonable ideology will make space for liberal values without agreeing that those values will always outweigh other values native to the ideologies themselves.[15] Indeed, this is the kind of point that Rawls's own discussion of the burdens of judgment should bring out. Say I adhere to some political ideology which includes liberal principles and other principles; those principles will sometimes conflict with one another, and liberalism, hands-off ideology that it is, can't be in

the position of saying that its own dictates should always win out. So although Rawls's argument, if we accept it, can establish that reasonable ideologies will all *include* liberal values, it can't establish that a liberal society will necessarily remain stable, if that stability is premised on those values never being superseded by other, competing values.

Conclusion

The epistemology of liberalism, such as it is, is based in the same kind of neutral, hands-off approach to disputes over facts and values in which liberal conceptions of free speech and free markets are based. In this chapter we mostly focused on Rawls, but plausibly the conciliationist view of disagreement might also be considered part of liberal epistemology. The link between liberalism and epistemology deserves further exploration, especially as the critiques of liberal politics often accompany critiques of liberal epistemological desiderata, like objectivity and neutrality.

Discussion questions

1. Do you think you can remain neutral on some question if you're in a position of power that's somehow related to that question? Do you think you should?
2. How much reasonable disagreement do you think there is about politics? Do you agree with Rawls about the burdens of judgment?
3. Rawls thinks the only admissible reasons in public deliberation should be reasons that, in some sense, ought to matter to other discussants. Is he right? Why or why not?

PART IV
Dynamics

19

POLARIZATION AS SORTING

Polarization is surely the most important dynamic in contemporary American politics. The split – between left and right, blue and red, Democrat and Republican – defines our political culture, and many commentators and researchers think that it's been getting worse for decades. A 2006 academic article about presidential politics held: "George Bush in 2004 was the most polarizing presidential candidate in modern political history."[1] But surely the negativity of the 2016 presidential election, and of the political season following it and perhaps continuing now, dwarfed that of 2004.

Types of polarization, broadly

But polarization is "said in many ways": there are a lot of things we mean by the term. A lot of what's bad about polarization is traceable to *affective polarization*. That's "polarization" of people's feelings: widening gaps between their views of their own side (typically measured by their views about one major political party) and the other side (typically measured by their views about the other political party). The rise in affective polarization is mostly explained by *negative partisanship*, the negative view of the other side, rather than positive partisanship, the positive view of one's own side. In a *Politico* article on negative partisanship coauthored by Alan Abramowitz, also one of the coauthors of the 2006 paper mentioned earlier, we find the following data about the 2016 election:

> Neither Trump nor Clinton was especially well-liked. Data from a Pew survey conducted before the 2016 national conventions found that both candidates

DOI: 10.4324/9781003355274-23

> received mediocre ratings from supporters of their own party – and record low ratings from members of the opposing party. On a "feeling thermometer" scale of zero to 100 degrees, Clinton received an average rating of 12 degrees from Republicans, while Trump received an average rating of 11 degrees from Democrats. In fact, 68 percent of Democrats rated Trump at zero, and 59 percent of Republicans rated Clinton at zero – an extraordinary reading with no modern precedent.[2]

This atmosphere of mutual hostility and hatred is so tangible that I have simply assumed familiarity with it in the text thus far.

There's another aspect of polarization, which Robert Talisse calls *political saturation*.[3] Political saturation can be thought of as ubiquitous politicization, perhaps in the sense of politicization we developed earlier in the book. Talisse thinks of political saturation as involving several steps. First, "social spaces are growing increasingly politically homogeneous," as have "sources of news, information, and entertainment."[4] You can make a really good guess about someone's politics from what they watch on television, what books or newspapers they read, or even what restaurants they go to. This sorting also occurs on geographical lines.[5] We expect to know something about what aspects of American culture someone participates in when we know their zip code – the divide between "blue states" and "red states" is part of this, but even more important is the divide between urban voters and rural voters, a relatively recent phenomenon. Our social spaces are not just sorted, though; according to Talisse, we also treat that sorting as sort of sensible, so that when people engage in an activity that has been "sorted" along political lines, we can ascribe political sentiments to them:

> More and more behavior is now commonly treated as at once an indicator and an expression of one's political allegiances. These days, political significance is commonly ascribed to behavior of all kinds, everything from buying groceries and rooting for a sports team to residing in a given neighborhood, enjoying a particular style of music, or watching one late-night television program rather than another.[6]

We tend to make predictions about people in line with our expectations about their political beliefs based on sorting.

This brings us naturally to the kind of polarization that's most important for us here, in a book about political beliefs: *belief polarization*. Like polarization more broadly, belief polarization is a plural phenomenon. A few years ago, a large group of authors developed a list of ten ways of modeling political polarization, which are all somewhat independent of one another.[7] We don't have to look at all

ten of them, though. They fall roughly into two kinds of measures: those that are like the negative partisanship above, in that they measure how far apart political groups are, and those that are like political saturation above, in that they measure how homogeneous political groups are. These are two different kinds of ideas, and they each may be related to different kinds of social problems that emerge from polarization. But in line with the theme of our text, we'll be most interested in questions about the rationality of polarized political beliefs and what we should do if we discover that our political beliefs are polarized.

So this chapter will consider questions about political beliefs that have been sorted to increasingly resemble those of co-partisans, and the following chapter will consider questions about political beliefs that have become more extreme, that have moved away from some posited political center.

Belief polarization as sorting

What does the polarization of *political beliefs* in the sense of *sorting* mean? At least two things. First, consider a case in which fifty people have opinions about what the income tax rate should be: the first person thinks it should be 1 percent, the second percent thinks it should be 2 percent, and so on. Then, after some time or after some event transpires, the distribution of beliefs in this group of fifty people changes. Now the first twenty-five people (from the original fifty) think the tax rate should be 15 percent and the last twenty-five people think the tax rate should be 40 percent. The beliefs have sorted in the sense of coalescing toward two poles; we can increasingly predict, from what *side* someone is on, what their *exact* belief is. Interestingly, philosophers have yet to say much about polarization in this single-issue sense of sorting.

We'll start instead with multi-issue sorting and then try to come back and say some things that are relevant to both kinds of sorting. Multi-issue sorting occurs much like political saturation: when we can increasingly predict, from one political belief someone holds, what their other political beliefs are. Compare these two situations or populaces (just examples I've come up with myself):

Populace 1

	Abortion restrictions	*Gun restrictions*	*Tax cuts*	*Border control*
Person A	Against	Against	In favor	In favor
Person B	Against	In favor	Against	In favor
Person C	In favor	Against	Against	Against
Person D	In favor	In favor	In favor	Against

Populace 2

	Abortion restrictions	*Gun restrictions*	*Tax cuts*	*Border control*
Person A	In favor	Against	In favor	In favor
Person B	In favor	Against	In favor	In favor
Person C	Against	In favor	Against	Against
Person D	Against	In favor	Against	Against

Here, populace 2 – which looks a bit like the two major contemporary American political parties – is far more sorted than populace 1. In populace 1, you have four distinct distributions of views, with no real correlations among them. In populace 2, you have two pairs of people who agree about everything and disagree about everything with the other pair. This huge increase in correlation among views is what we mean by sorting. Note that sorting also reduces statistical independence in the sense required by the Condorcet jury theorem. In general, on the assumption that a population is made up of individually competent people, the more sorted it is, the less reliable its majority votes will be.

Some of the theories we've considered so far explain sorting pretty well. The identity and partisanship theories explain why people would have political beliefs that track the platforms of their political groups, and Hannon's expressivist theory explains why people would *seem* to have political beliefs that track the platforms of their political groups. Theories of personality type also do a good job at explaining sorting. If personality categories determine political beliefs, and there are a really limited number of personality categories, then there should be a limited number of political-belief distributions. Highly partisan writers in particular seem to like the idea that there are two types of people in the world – good people and bad people – and that the good people flock to the good political beliefs because they like good things while the bad people flock to the bad political beliefs because they like bad things. In his pop-political book *Why We're Polarized*, pundit Ezra Klein writes that "sorting has made the Democrats into a coalition of difference and driven Republicans further into sameness."[8] But of course when it comes to political *beliefs*, sorting is always a matter of sameness.

Sorting and orthogonality

Is sorting rational? You might think that it's hard to come up with a general answer to that question. Surely it depends on the beliefs of each side and how they're sorted! But Hrishikesh Joshi has argued that there are some inherently irrational characteristics of sorting. His argument is complex, but it goes a bit

like this.[9] Joshi argues that many political beliefs are *orthogonal* to each other: accepting a certain view about issue X doesn't put any *rational* pressure on us to accept a certain view about issue Y and vice versa. He then argues that orthogonality renders sorting irrational.

First, orthogonality. We've seen that genuine political disputes are likely to be attributable either to disputes over facts or to disputes over values. People like Haidt and Klein, and indeed most partisans, would probably say that orthogonality is unlikely because there are broad background moral beliefs that affect many different object-level political beliefs: things like "the most in need should be given priority," "people should by and large be free," and so on. In response to this, Joshi says, rightly, that to generate a political stance, moral principles aren't enough; they must be conjoined with empirical premises, and in political beliefs, empirical beliefs are controversial: "Many of the issues with respect to which the political camps disagree rest on contentious empirical claims. Hence, a high-level moral principle or theory . . . is not going to by itself yield [political beliefs] as corollaries."[10] He gives the following examples: there's disagreement over whether raising the minimum wage actually helps unskilled workers, and there's disagreement over whether affirmative action actually helps its recipients.[11] However, I think Joshi hasn't done quite enough here. He's shown that it's *possible* to agree on moral principles but disagree on empirical claims and hence disagree when it comes to political beliefs. But we'll see that what he's concerned with is, as we were earlier in the text, what *actually causes* the political disagreements, by and large. Joshi also doesn't spend much time on the possibility that there is some overriding *empirical* dispute that explains all the different partisan disagreements. For instance, partisan political differences might be explained by different beliefs about empirical issues such as how humans respond to incentives, how humans deal with change and difference, how well humans thrive in unstructured societies as opposed to hierarchical ones, and other questions about human nature such as whether humans are ultimately good and trustworthy.[12]

In their book *The Myth of Left and Right*, Hyrum Lewis and Verlan Lewis survey a number of possible explanations of partisan disagreement of this sort. They find them wanting in part because they are interested in changes to partisan stances *over time*, and even when one is able to construct an explanation of partisan differences at one moment in time, that explanation often falls apart when one considers past views of the parties or political sides. A lot of the explanations they consider are related to the theories of political differences we've looked at so far; they include

> change versus preservation, . . . arrogance vs. humility, autonomy vs. control, . . . big vs. small government, . . . chaos vs. order, collectivism vs. individualism, compassion vs. greed, complex vs. simple, . . . equality vs. hierarchy, equality vs. liberty,[13]

and a couple dozen more. They offer takedowns for each of these too-simple dichotomies, though at times they make the same questionable move: they explain how *in principle* the proposed dichotomy might be reversed, so that for example a left-wing view is on the side of "hierarchy" while a right-wing view is on the side of "equality," but they don't show that this is the *psychological reality* of partisans. A further reason to be doubtful of such ideological stories is that social scientific research casts doubt on the idea that the average voter understands and applies ideologies;[14] we've already mentioned Philip Converse's classic paper[15] examining this issue.

Note that accepting Joshi's orthogonality thesis about political beliefs means rejecting almost all of the explanations of political beliefs we considered in this text's second part. The only remaining contenders would be the identity, signaling, and partisanship theories and the minimalist theories.

The next step is to show that orthogonality renders sorting irrational. Many of our beliefs are orthogonal to each other; without a debunking argument behind them, this doesn't show anything about whether they're irrational. I believe that it's dark outside as I write this, I believe that Paris is the capital of France, and I believe that two and two make four; those beliefs have nothing to do with one another, but seeing that they have nothing to do with one another isn't about to make me doubt any of them.

Joshi needs an extra step, which is to think about *our opponents*' political beliefs as well as our own. This isn't from the perspective of conciliation or the philosophy of disagreement but the perspective of *explanation*. Why does my side hold the beliefs it does while my opponent's side holds the beliefs it does? Crucially, Joshi says that it won't do to say that we know better, for reasons related to what we saw in discussing epistemic democracy above: even if we do know better, so that we have a good explanation of why *we* get everything *right*, we still need an explanation of why *they* get *everything* wrong rather than, say, getting some issues wrong and some issues right (which might happen in a universe of unsorted political beliefs).[16] And crucially, if the orthogonality thesis is right, we need not one but many explanations – something like what we called "error theories" of disagreement earlier in the text. So that's the rub: if the orthogonality thesis is correct, then there's a lot that needs explaining. However, note that some of the theories of political belief we considered, like the system justification theories, *do* aim to offer explanations of lockstep but wrong political beliefs.

Joshi's argument has a final step of sorts, which is to say that he has a better explanation, most naturally framed as a debunking argument, of the sorting of political beliefs in this manner.

> The phenomenon of belief polarization across a population raises the prima facie worry that social and historical factors, as opposed to truth-tracking belief forming mechanisms, explain at least some of the predominant beliefs

> among the groups in question. The natural explanation for why a typical conservative or liberal holds the political views she does would appeal to facts about her social circles and, relatedly, the sorts of news outlets she pays attention to.[17]

Again, this is one of the identity, signaling, and partisanship-style theories we saw earlier in the text. So, in the end, this is another kind of debunking argument based on that theory, with polarization as sorting offering a new line of empirical support for it.

Sorting and bubbling

According to some, one way that people sort themselves in contemporary politics, often as a result of technology, is into *bubbles*. The basic idea of a bubble can be taken as an extension of the idea of political homogeneity in Talisse's account of political saturation; a bubble is a homogeneous social space large enough that its occupants don't really escape it, at least not to encounter the political views of those outside. One jumping-off point is the prediction with which Cass Sunstein begins his 2001 book *Republic.com*:[18]

> It is some time in the future. Technology has greatly increased people's ability to "filter" what they want to read, see, and hear. General interest newspapers and magazines are largely a thing of the past. The same is true of broadcasters. The idea of choosing "channel 4" or instead "channel 7" seems positively quaint. With the aid of a television or computer screen, and the Internet, you are able to design your own newspapers and magazines. Having dispensed with broadcasters, you can choose your own video programming, with movies, game shows, sports, shopping, and news of your choice. You mix and match.
>
> You need not come across topics and views that you have not sought out. Without any difficulty, you are able to see exactly what you want to see, no more and no less. . . .
>
> Perhaps you have no interest at all in "news." Maybe you find "news" impossibly boring. If so, you need not see it at all. Maybe you select programs and stories involving only music and weather. Or perhaps you are more specialized still, emphasizing opera, or Beethoven, or the Rolling Stones, or modern dance, or some subset of one or more of the above. If you are interested in politics, you may want to restrict yourself to certain points of view, by hearing only from people you like. In designing your preferred newspaper, you choose among conservatives, moderates, liberals, vegetarians, the religious right, and socialists. You have your favorite columnists; perhaps you want to hear from them, and from no one else. If so, that is entirely feasible with a simple "point and click." Or perhaps you are interested in only a few

> topics. If you believe that the most serious problem is gun control, or global warming, or lung cancer, you might spend most of your time reading about that problem, if you wish from the point of view that you like best.
>
> Of course everyone else has the same freedom that you do. Many people choose to avoid news altogether. Many people restrict themselves to their own preferred points of view – liberals watching and reading mostly or only liberals; moderates, moderates; conservatives, conservatives; neo-Nazis, neo-Nazis.[19]

This prediction was still entirely speculative when Sunstein made it, but as filter algorithms improved, people quickly began to think that "some time in the future" had arrived. In 2011, Eli Pariser wrote a book called *The Filter Bubble*; the ten intervening years had seen the rise of Facebook, Twitter, and YouTube and the transition of Netflix from a mail-based service to a streaming service, for instance, though TikTok was still some years away. Pariser opens the book with a quote from media theorist Marshall McLuhan: "We shape our tools, and therefore our tools shape us" – a point illustrated by Pariser's opening salvo, a story about Google's introduction of "personalized search":

> Starting that morning [of December 4, 2009], Google would use fifty-seven *signals* – everything from where you were logging in from to what browser you were using to what you had searched before – to make guesses about who you were and what kinds of sites you'd like. Even if you were logged out, it would customize its results, showing you the pages it predicted you were most likely to click on. . . . In other words, there is no standard Google anymore.[20]

Google develops a sense of who I am based on what I search for, and then my future searches are shaped based on that sense, and go on to further develop it, and so on. The technology itself locks me in or "sorts" me into a certain universe of results.

Philosopher C. Thi Nguyen has worried a bit about two kinds of bubbling effects: *epistemic bubbles* and *echo chambers*. Roughly, both epistemic bubbles and echo chambers are situations in which perspectives at odds with some epistemic agent's perspective are somehow missing from that agent's experience, but in an epistemic bubble, "relevant voices have been excluded by omission," whereas in an echo chamber, "relevant voices have been actively discredited."[21] Of course, much depends on what we mean by "relevant voices." On a broad reading of that phrase, we're in an epistemic bubble with regard to pretty much everything we believe.[22]

Nguyen considers the sort of question we're interested in in this text: Are agents in epistemic bubbles and echo chambers rational? The epistemic bubble case depends a bit on one's background assumptions in highly abstract

epistemology, but my instinct is to say that it's not irrational to be in an epistemic bubble until you realize it, at which point it's rational to downgrade your bubbled or sorted beliefs. Echo chambers are a harder case. How do we distinguish, for instance, the (presumably good) way the peer review system in academic publishing intentionally keeps some ideas out of academic journals from the (presumably bad) functioning of an echo chamber? In addition to this example, Nguyen gives the examples of "medical board exams and university hiring practices."[23] He worries about the broader possibility that people in echo chambers, presented with what looks like evidence that people outside are untrustworthy, might be rationally but inextricably "epistemically trapped."[24]

Nguyen's suggested solution – as far as the question of rationality goes – traces from philosopher Thomas Kelly, who also presents it in the context of belief polarization.[25] You might recall the principle of *commutativity of evidence*, which holds that it shouldn't rationally matter what order we get our evidence in. For someone in an echo chamber, in which some evidence discredits other evidence, order

> matters very much. . . . [S]ince [they] encountered the echo chambers and assimilated its beliefs first, their use of background beliefs to vet new sources will lead them to continually increase their trust in the echo chamber and their distrust of outsiders.[26]

In other words, the person in the echo chamber is being irrational because they have violated the commutativity principle in discarding new evidence on the basis of their perspective from within the chamber, which was determined by old evidence. However, I'm not sure I'm convinced by this idea. If order doesn't matter, then it also looks like people *outside of* echo chambers might be irrational as well simply because, if they had received their evidence in a different order, they might themselves be in an echo chamber after all. At the very least, the notion that "order doesn't matter" needs to be stated more precisely. For instance, rationality might require that we balance *all* possible orders in which evidence could've been received.

Endre Begby has considered a similar phenomenon, that of "evidential preemption." Evidential preemption occurs when one epistemic agent prepares another epistemic agent for evidence they might receive in the future, most typically when that evidence is the testimony of a third epistemic agent, and warns them that it will be misleading in some way.[27] Though Begby thinks there is in some sense something bad about this, he writes that as a non-ideal theorist of epistemology (see the chapter on the ethics of belief), he doesn't think that it's irrational: "While it is true that I might be in possession of more information at the later time, I may also foresee that my ability to make correct use of that information will be significantly impaired."[28] This is similar to his strategy regarding prejudice, which we've seen.[29]

Even the echo chamber analysis might be too generous to some partisans. In a 2010 study, Brendan Nyhan and Jason Reifler found that "corrections frequently fail to reduce misperceptions among the targeted ideological group," and also noted "several instances of a 'backfire effect' in which corrections actually increase misperceptions among the group in question."[30] If corrections actually backfire, then partisans would be doing something worse than discounting new evidence because they're in an echo chamber. They'd be becoming *more certain* of their views despite encountering evidence *against* them – which just must be irrational. We could come up with lots of possible explanations for such a seemingly paradoxical backfire effect based on our theories of political belief. For instance, if political beliefs are really a matter of group identity, then the existence of contrary viewpoints might induce a feeling of challenge or threat, meaning that the partisan must become that much more extreme in their adherence to the group consensus. However, the following decade has established that the backfire effect is probably quite rare if not nonexistent.[31] Indeed, partisans seem to respond pretty rationally to factual corrections in experimental contexts.[32]

There are empirical doubts about epistemic bubbles and echo chambers themselves, too. A British study of the prevalence of echo chambers and filter bubbles suggests that in most countries, "cross-cutting exposure," in which people are exposed to multiple viewpoints, is more the norm than echo chambers; the authors write:

> Across a range of different countries, including the highly polarised United States, several cross-platform studies – both those reliant on survey data and those reliant on passive tracking data – have found that few people occupy politically partisan online news echo chambers. . . . Even in the United States, researchers have long found that echo chambers are smaller and less prevalent than commonly assumed.[33]

Of course, small echo chambers can also make a big difference to political reality if they can generate extremist groups (see later). However, cross-cutting exposure isn't necessarily the greatest news, either, as the authors note that "cross-cutting exposure, at least on social media, also seems to be able to increase polarisation, at least among political partisans."[34] This is another sign that the idea of bubbling remains relatively intuitive and underdeveloped, both theoretically and empirically.

Sorting, charity, and humility

In an earlier chapter we talked about interpretive charity. Some commentators take some version of interpretive charity to be an important tool against polarization; for example, in their book *The Coddling of the American Mind*, Greg

Lukianoff and Jonathan Haidt include charitable interpretation as part of their recommendation about how to "explicitly reject the untruth of us versus them: [that is, that] life is a battle between good people and evil people."[35] However, recall that charitable interpretation can work only through maximizing another person's true attributed beliefs "by our own lights" – to be charitable, we interpret someone else so as to maximize the number of beliefs on which they agree with us. Being too charitable, then, actually risks *exacerbating* epistemic bubbles and echo chambers, in Nguyen's sense, because it risks interpreting fundamental disagreement which might puncture a bubble as something else. This is especially the case because partisans are likely to think of political beliefs as very central and obvious, so that interpreting someone as disagreeing with them about politics would be very uncharitable indeed.

For example, take the slogan that we should "abolish the police." This is a position that at least some people claim to hold. Now imagine someone named Tom, who begins from the position that the notion of abolishing the police is simply ludicrous. Tom might interpret the slogan "abolish the police" to mean that we should change the current form of law enforcement, or fire large groups of police officers due to misconduct, or retrain all officers in different kinds of dispute resolution techniques. Those would all be charitable interpretations by Tom's lights. So he avoids encountering the true view.

Similarly, having too much epistemic humility, or overapplying conciliationism about disagreement, can lead to the maintenance of epistemic bubbles and echo chambers. If you're *in* an epistemic bubble or echo chamber, you're going to be mostly talking to people who hold the same views on politics. If you are constantly modifying your views to take into account the views of your interlocutors, as humility dictates and as the thesis of conciliationism might have it, then any disagreement you have with the group will be quickly tamped down as a result of that modification. For example, take Sally, who is in an echo chamber all of whose members believe in a certain conspiracy theory. In the dark of night, as she's about to fall asleep, Sally keeps coming up with trenchant, devastating objections to this conspiracy theory. But when she wakes up, she encounters all the other true believers, and she thinks: "I can't possibly be right and all of them wrong; that's so arrogant." Staying humble, Sally loses faith in her objections and ends up continuing to believe the conspiracy theory. This is an unfortunate but not necessarily irrational effect. It gives yet another example of the ways in which rationality does not necessarily help us escape from bad epistemic situations.

Sorting and understanding

A related danger of sorting, and related bubbling phenomena, is that we lose the ability to understand others – something we might think we're responsible

for doing, for one reason or another. Concern over this sort of danger was part of what fueled Haidt to develop his moral foundations theory. As we saw when we covered that theory, it can be taken to imply that conservatives understand liberals better than liberals understand conservatives, since conservatives "have" liberal moral foundations and liberals don't "have" conservative moral foundations (though what it means to "have" a foundation was unclear to us, given that the relationship between politics and moral foundations was only partial and gradual). And this is how Haidt interprets his own work on moral psychology. He writes:

> In a study I did with Jesse Graham and Brian Nosek, we tested how well liberals and conservatives could understand each other. We asked more than two thousand American visitors to fill out the Moral Foundations Questionnaire. One-third of the time they were asked to fill it out normally, answering as themselves. One-third of the time they were asked to fill it out as they think a "typical liberal" would respond. One-third of the time they were asked to fill it out as a "typical conservative" would respond. This design allowed us to examine the stereotypes that each side held about the other. More important, it allowed us to assess how accurate they were by comparing people's expectations about "typical" partisans to the actual responses from partisans on the left and the right. Who was best able to pretend to be the other?
>
> The results were clear and consistent. Moderates and conservatives were most accurate in their predictions, whether they were pretending to be liberals or conservatives. Liberals were the least accurate, especially those who described themselves as "very liberal." The biggest errors in the whole study came when liberals answered the Care and Fairness questions while pretending to be conservatives.[36]

Haidt's explanation of *why* liberals get conservatives wrong is especially interesting: "If you don't see that Reagan [for instance] is pursuing positive values of Loyalty, Authority, and Sanctity, you almost have to conclude that Republicans see no positive value in Care and Fairness."[37] In other words, if you don't understand what conservatives – who (in Haidt's theory) share the liberal values – have to weigh those values *against*, you end up thinking that conservatives don't give them any weight at all. Imagine you've read some research that says a very delicious dessert is unfortunately very unhealthy. Now imagine I offer you some of that dessert, knowing nothing about that research. Imagine, finally, that you decline my offer. I might come to the conclusion that you don't think the dessert is very delicious; if you did, why would you turn it down? Because you "have" a consideration in mind that I don't have, my interpretation of your action, and my attribution to you of values and views, comes out wrong.

This answers a question we saw in examining Hanno Sauer's criticisms of moral foundations theory earlier: What does it mean to "appreciate" a moral

foundation without granting it normative authority oneself? One thing it means is to understand that it factors into the decision-making of those who possess it, so that they must weigh the shared moral foundations against it. But while that gets us past the position where we must infer that the opposing side doesn't value the shared foundations *at all*, it might not get us much further than that. Say again that I really value my delicious dessert, and now say that you don't think it's delicious. Now say that someone is screaming "Help, help!" outside, but I'm just sitting here calmly eating. Once you understand how delicious I think the dessert is, you might understand that I could plausibly still care *somewhat* about helping someone who is screaming outside, because the value of eating my dessert could simply have outweighed the value of helping a stranger. But that won't suffice as a full defense of my actions in your eyes. I *shouldn't* care so much about the dessert; I should care *far more than I do*, relative to my other concerns, about helping the stranger. Similarly, as long as liberals don't themselves accept conservative moral foundations, it seems fully rational for them to say that conservatives don't care *enough* about the shared moral foundations.

It would be very easy to say simply that because conservatives seem to understand liberals better than liberals understand conservatives, there's some sort of burden on liberals of considering the alternative viewpoint that conservatives don't bear. But that assumes a kind of underlying symmetry that liberals needn't accept exists. Say I believe in a ridiculous conspiracy theory about some event and you believe in the mainstream view about that event. Because there is so much information out there on the mainstream view – perhaps it's taught in schools – I have a great understanding of it while you have no understanding of the conspiracy theory. But this alone doesn't generate some responsibility in you for understanding the conspiracy theory. Even in a situation without ridiculousness, say of a scientist who produces a new and correct theory of some phenomenon and a scientist who understands the appeal of that view but holds to their incorrect older theory, we might not think the person who understands only one side of things is necessarily doing anything wrong nor that the person who understands both sides is necessarily doing something right. If liberals really believe they are *right* about morality and about politics, then they will similarly believe that there are far more important things than mutual understanding to deal with. And our theory shouldn't bake in the assumption that liberals aren't right!

Discussion questions

1. What feelings do you have about people who disagree with you about politics?
2. What activities do you engage in that you think are related to your political identity, even if they aren't political themselves? What political beliefs might someone conclude you have from details like where you live, where

you shop, and what entertainment you consume? Would they be right? Do you draw these sorts of inferences about others? Are you generally right?

3. Do you think your political beliefs are in a bubble or an echo chamber?
4. How well do you think you understand the motivations and arguments of someone on the other political side? Bryan Caplan has called contests where people on opposite political sides attempt to mimic each other "ideological Turing tests."[38] Do you think these tests are a good idea? How well do you think you would do? What do you think they show?

20

POLARIZATION AS EXTREMISM

The notion of extremism is likely a bit simpler than sorting. Society is polarizing in the sense of extremism if, by some measure, views on each side become more and more extreme over time, whether or not they're sorted. (You might even just need one side to become more extreme to call it polarization.) Imagine again that we do two surveys of the same fifty people about the income tax rate. Imagine that the first survey gives us just the result with which we ended the polarization-as-sorting example: twenty-five people think the income tax rate should be 15 percent and twenty-five people think the income tax rate should be 40 percent. After some time has passed, we do a second survey. This time, we end up with the result that ten people think the income tax rate should be 5 percent, ten people think the income tax rate should be 10 percent, ten people think the income tax rate should be 20 percent, ten people think the income tax rate should be 35 percent, and ten people think the income tax rate should be 50 percent. The range of views has gotten wider, and we might think that has made things more extreme.

Or has it? Maybe things aren't so simple. How have we picked the *center* of the income tax rate debate? Maybe there's a sense in which the *correct* view is that the income tax rate should be 5 percent or 50 percent. If that's the case, things might have gotten no more (or only a little more) extreme relative to the *true* center in this example. This shows an inherent difficulty with the notion of extremism: there doesn't seem to be an easy way to decide what's extreme and what's normal before the fact without introducing one's own substantive political beliefs into the measurement, something that we've seen is intellectually very dangerous.

DOI: 10.4324/9781003355274-24

But we can still explain what extremism is in a relative rather than absolute way. Taking an initial distribution of views (say the half-and-half distribution of views on the income tax rate at 15% and 40% that we discussed earlier), we can identify what its most extreme points are, count how many people are there, identify which points are more extreme than any in that distribution, and call the average (or median, or mode, or whatever) the *center*. Then we have our polarization as extremism in the previous example: the new most extreme points are more extreme than the old most extreme *by its measure*. Another way to become more extreme is by having more *extremists* – say by moving from a distribution in which one person thought the income tax rate should be 5 percent, twenty-four thought it should be 15 percent, twenty-four thought it should be 40 percent, and one person thought it should be 50 percent to a distribution in which twenty-four people thought the income tax rate should be 5 percent, one person thought it should be 15 percent, one person thought it should be 40 percent, and twenty-four people thought it should be 50 percent. There, the actual views represented are the same, but they're represented at different rates.[1]

The arbitrariness of the "center" might lead us to think that polarization as extremism is not necessarily irrational. Don't you need to know whether the center is right or wrong before you know whether moving away from it is necessarily irrational? In this vein, Neil Levy writes:

> There seems no a priori reason to think that the truth is more likely to lie in the middle of a group of deliberators, prior to their sharing their opinions with one another . . . than at the extremes. . . . [T]here's no reason to think [some] normative claim should be rejected because it was initially held only by a minority of the deliberators. Everything depends on the composition of the group and the distribution of opinion within it. Extreme opinions about race and gender were more accurate than more moderate opinions in the antebellum United States, for instance.[2]

This is obviously right in one sense and obviously wrong in another sense. It's obviously right in the sense we've examined so far: the "center" easily could be wrong. However, first, it's not clear that there isn't an *a priori* reason to think that the truth should lie in the middle. That's precisely the result that the conciliationist theory of disagreement and the related thesis of epistemic democracy gave us when we looked at them in prior chapters.

Furthermore, when we evaluate the rationality of extremism, we are not necessarily asking whether it gets us to the truth, but whether it does so by the right process of reasoning. That a practice or mechanism results in some true beliefs gives us only limited information about its rationality. Stopped clocks give the right time twice a day. "Always heads" calls a coin flip right half the time. But trusting a stopped clock is not rational even when the clock is right, and believing a coin will flip heads is not rational even when it does come up heads.

Similarly, polarization as extremism cannot be called rational merely in virtue of sometimes bringing us closer to the truth. Indeed, many versions of belief polarization as extremism can't *always* bring us closer to the truth, since in these versions the idea of polarization is that for each group that goes in one direction there's a group that goes in the other direction. We'll see a similar problem with an argument in the next chapter: that conspiracy theories sometimes turn out to be true does not mean that believing conspiracy theories sometimes turns out to be rational. These are both examples of "rationalizing by truth," which we considered in Chapter 12.

Levy seems to be aware of this potential line of criticism, writing: "Of course, it might be the case that group polarization leads to more accurate beliefs in this or other cases only by chance," but he follows this up by saying that group polarization "might nevertheless be epistemically objectionable."[3] This lays bare the problem. A belief-forming mechanism that results in truth only by chance is *ipso facto* epistemically objectionable, not *nevertheless* epistemically objectionable. A belief-forming strategy or mechanism that relies on luck is not itself reliable, and it confers no warrant. We cannot vindicate even *some* instances of polarization and conspiracy theorizing on the grounds that *some* instances of polarization and conspiracy theorizing turn out to generate true beliefs, because the rationality of these practices in some specific instance depends on their reliability in general.

So, to make a long story short, if we want to figure out whether polarization as extremism is or can be rational, we must start by figuring out how it happens.

Group ("enclave") polarization

One influential theory about how polarization as extremism happens comes from Cass Sunstein. Sunstein considers situations of people in *enclaves*, which are just deliberating groups that include only people with a certain level of initial agreement, and finds that extremism can result when people deliberate in such enclaves. Note that enclaves are not as restrictive even as epistemic bubbles. We discussed this briefly in the chapter on democracy; let's look at it a bit more closely now.

Sunstein draws on a wealth of examples, both social-scientific experiments and "natural experiments" in which enclaves form and their effects are measured without the need of a controlled setting. Interestingly, Sunstein's "law of group polarization" holds outside of political contexts and thus has the capacity to explain them quite reductively. Here are some of Sunstein's examples:

> [W]e conducted a[n] experiment, involving about 3,000 jury-eligible citizens and 500 deliberating juries, each with six people. . . . People read about a personal injury case, including the arguments made by both sides. They were also asked to record, in advance of deliberation, an individual "punishment

judgment," . . . on a scale of 0 to 8, where . . . 0 indicated that the defendant should not be punished at all, and 8 indicated After the individual judgments were recorded, jurors were sorted into six-person groups and asked to deliberate to reach a unanimous "punishment verdict." You might predict (as we did) that people would compromise and that the verdicts of juries would be the median of punishment judgments of jurors. But your prediction would be badly wrong.

Instead, the effect of deliberation was to create both a *severity shift* for high-punishment jurors and a *leniency shift* for low-punishment jurors. When the median judgment of individual jurors was 4 or higher on the 8-point scale, the jury's verdict ended up *higher* than that median judgment. Consider, for example, a case involving a man who nearly drowned on a defectively constructed yacht. Jurors tended to be outraged by the idea of a defectively built yacht, and groups were significantly more outraged than their median members. High levels of outrage and severe punitive judgments became higher and more severe as a result of group interactions.

But when the median judgment of individual jurors was below 4, the jury's verdict was typically *below* that median judgment. Consider a case involving a shopper who was injured in a fall when an escalator suddenly stopped. Individual jurors were not greatly bothered by the incident, seeing it as a genuine accident rather than a case of serious wrongdoing. In such cases, juries were more lenient than individual jurors. Here, then, is a lesson about what happens when people discuss wrongdoing. If group members are upset, they will probably get more upset after talking to each other. If group members think that what happened is not a big deal, they will usually think that what happened is basically nothing after a period of discussion.[4]

With respect to many decisions, members of deliberating groups became significantly more disposed to take risks after a brief period of collective discussion. On the basis of such evidence, it became standard to believe that deliberation produced a systematic "risky shift." For a significant period, the major consequence of group discussion, it was thought, was to produce that risky shift – a thought that would bear on many parts of social life, because groups are often asked to decide whether to take a gamble or, instead, to take precautions.

But later studies drew this conclusion into serious question. They even raised the question whether culture, rather than group dynamics, is responsible for the risky shift. On many of the same questions on which Americans displayed a risky shift, Taiwanese subjects showed a "cautious shift." On most of the topics just listed, deliberation led citizens of Taiwan to become significantly less risk-inclined than they were before they started to talk. Nor was the cautious shift limited to the Taiwanese. Among Americans, deliberation sometimes produced a cautious shift as well, as risk-averse people became

more reluctant to take certain risks after they talked with one another. There are two major examples of cautious shifts: the decision whether to marry (!) and the decision whether to board a plane despite severe abdominal pain, possibly requiring medical attention. In these cases, the members of deliberating groups moved toward greater caution.

At first glance, it seemed hard to reconcile these competing findings, but the reconciliation turned out to be simple: *The predeliberation median is the best predictor of the direction of the shift*. When group members are disposed toward risk-taking, a risky shift is observed. When members are disposed toward caution, a cautious shift is observed. It follows that the striking difference between American and Taiwanese subjects is not a product of any cultural difference in how people behave in groups. It results from a difference in the predeliberation medians of the Americans and the Taiwanese on the key questions. When Americans show a predeliberation median in favor of caution, discussion moves them toward greater caution; the same is true of Taiwanese. When American groups show a risky shift, and Taiwanese a cautious shift, it is simply because of a difference in their initial inclinations. Thus the risky shift and the cautious shift are both subsumed under the general rubric of group polarization.[5]

I reproduced these examples at full length because I think they're very fully striking and because Sunstein does a great job of explaining them himself. One fascinating thing that emerges from them is that it might seem as though there *must* be some objective "center" at work based on these examples. The center is precisely the point from which group polarization radiates outward. If a deliberating group subject to the law of group polarization which shared some opinion would move, say, to the "left" after deliberating, then that is a "left"-wing opinion; if the group would move to the "right" after deliberating, then it's a right-wing opinion; and if it wouldn't move at all after deliberating, then the opinion is in the center. Of course, Sunstein admits readily that not *all* deliberating groups will be subject to the law of group polarization; in particular, people who "feel really strongly" before deliberation are unlikely to be moved,[6] and groups deliberating about "eureka problems," problems with answers that can be quickly recognized as correct once suggested, like crossword puzzles, won't polarize.[7]

The shift in favor of the predeliberation "direction" could still be rational if, for instance, extreme group members were being taken as experts, and other members were deferring to them. But this isn't the case; the more directionally extreme a member is to begin with, the more they shift when the group polarizes.[8] It is hard to come up with a rational explanation of this change in beliefs. But it might still be possible to come up with a rational explanation of the way *some* group members' beliefs change.

Sunstein thinks a few factors explain enclave polarization. The most amenable to a rationalizing treatment is that enclave polarization can supply deliberators with new information, but that this new information tends to be one-sided.[9] Members of the enclave encounter new evidence and arguments in favor of their own side far more than they encounter new evidence and arguments in favor of the other side, and they may also encounter counterarguments to some of their concerns about their own side. This can justify the fact that deliberating members do not simply average out their views in line with conciliationism. Conciliationism is a theory of what we should do in response to the *mere fact* of disagreement or of new evidence about other people's beliefs more generally. But in a deliberating group, we also hear *why* the other people have their beliefs. If there are ten of us deliberating, and we all believe some proposition *p* somewhat hesitantly, but we all believe that proposition for different reasons, then we might come out of deliberation all having ten different reasons to believe *p*, which might make us very certain indeed. Similarly, hesitant members might not even be sure that their evidence does support their view in the way they think, and thus even discussion with people who possess only the same evidence or arguments they do can provide corroboration, and corroboration of a moderate position on one side might make an extreme position on that side more plausible, too.[10]

Less rational, but in line with some of our theories of political belief, are explanations in terms of reputation, social comparison, and identity. If I conceive of myself as *the kind of person* who is really strongly on one side of a political issue but then end up deliberating in a group where I actually seem like a moderate on that issue, I may revise my view to be more extreme to maintain that self-conception.[11] And deliberating groups which feature "a shared identity and a high degree of solidarity" exhibit more polarization than other groups.[12]

Finally, one that's a bit unclear: Sunstein suggests that deliberation can reveal to people what their "hidden" beliefs and desires truly are.[13] People may self-censor or even deceive themselves about what they really think about some political issue because it is too extreme. This is what seems to happen in "preference falsification," a phenomenon of false responding to opinion polls based on the social desirability of certain answers.[14] But once placed in a deliberating enclave, these opinions that might have seemed unacceptable may become acceptable or even popular. In this view, enclave polarization merely elicits our true thoughts about politics rather than irrationally creating new beliefs.

Asymmetric polarization?

The theory of polarization as extremism being caused by enclave deliberation and other such bubbling effects raises the possibility of *asymmetric polarization* in which one political side is more polarized than the other. After all, one side could have more insular enclaves or bubbles than the other side. In particular,

some political scientists argue that in the United States, Republicans are polarizing and Democrats are not.[15] While it is important that we not simply assume that polarization is symmetrical, the methods often used to establish this particular thesis are highly questionable. In particular, it is usually justified using DW-NOMINATE, a statistical tool that tracks politicians' voting patterns. However, voting patterns aren't necessarily the same as ideologies, don't necessarily reflect broader changes over time, don't necessarily reflect polarization in society as a whole, and don't necessarily reflect choices by legislative leaders as to what's even up for a vote.[16] As with other social-scientific tools we've examined so far in the text, it's always important to be clear about what *exactly* is being measured, so that we can compare it to our description of the phenomenon we're interested in and see to what extent the measurement will be informative.

Theory of the Overton window?

Our discussion of extremism has naturally involved some talk about a political "center" and about which political beliefs are acceptable and unacceptable. This "window" of political acceptability is often called the "Overton window," after Joseph P. Overton, who theorized that politicians are limited in their scope: "they generally only pursue policies that are widely accepted throughout society as legitimate policy options."[17] But Overton's main lesson was that the rest of us are not so limited; we can *shift* the Overton window, making what was previously unthinkable part of the mainstream. As Derek Robertson puts it (in describing the theory),

> Ring the bell loudly for your idea, no matter how unpopular, and back it up with plenty of research and evidence, so the thinking went. Today's fringe theory can become tomorrow's conventional wisdom by the shifting of the finely tuned gears that move popular opinion.[18]

In recent American politics this has often been expressed through the concept of "normalization"; some who accused the candidate and then President Donald Trump of being dangerously politically extreme also accused others of not doing enough to emphasize how extreme Trump was in talking about him, thus "normalizing" him.[19] (Note that this was a rather strange, new usage of the term; to "normalize" a situation would previously have meant to *render it* normal, not to *treat it as* normal. I suspect this usage is related to the top-down, emphasis-on-language view of political beliefs that has come up quite a bit in this text.)

Little philosophical background has been developed to support the Overton window, and the whole concept of "normal" faces some of the same difficulties as the concept of "center," I think. Some writers have tried to explain what cognitive processes people undertake in "normalizing" unethical behavior, but the

strategies they mention – denying that they're responsible, denying that anyone has been injured, and so on – don't seem different from the strategies people would use if they *actually weren't* behaving unethically.[20]

Another way of thinking about "normalization" might be drawn from the work of Philip Tetlock on moral psychology. Tetlock considers the fact that in most moral systems, some values are considered sacred and not subject to trade-offs or calculations, while others are best dealt with by precisely that sort of economic reasoning, as what Tetlock calls "secular" values.[21] (This can be usefully compared with the backdrop of Haidt's moral dumbfounding experiments.) Consider that, during the coronavirus pandemic, hospitals deciding which patients to treat in which order and with what resources – what's known as *triage* – garnered a great deal of outrage. Outraged people thought there was something deeply wrong about making such decisions about other people's lives, because life is (for lack of a better word) sacred. Moral taboos about sacred values as related to beliefs may also explain some views on the ethics of belief; for instance, Tetlock considers "forbidden base rates" or taboo statistical generalizations about demographic groups.[22] These are precisely the sorts of generalizations that are targeted by the philosophical literature on the ethics of belief, as we saw a few chapters ago. The notion of sacredness and taboo may provide one possible account of what is meant by "normalization" and of *some* cases of shifting the Overton window. It may be that a political discourse seems to treat certain ideas as taboo, as violations of sacred values, but then suddenly subjects them to analysis and critique through the "secular" lens of trade-offs and calculations, perhaps in the political realm through policy analysis and punditry. Thus, "normalization" might not mean acceptance but rather discussion through a "normal" frame, which takes previously inviolable values as subject to compromise. It will be hard, though, to differentiate this from normal social changes in values over time.

Normalization may also have something to do with attention. What's normal, even when it's not ideal, can be taken for granted as part of the architecture of personal and social life; it's normal that my bus is sometimes late, that I sometimes feel tired even after a good night's sleep, and so on. So shifting the Overton window and "normalizing" something might be taken to mean changing social norms such that something that formerly would have grabbed one's attention as initially outrageous now no longer strikes one as particularly remarkable. This has some similarities to the analysis in terms of sacredness.

The reverse of broadening the Overton window is narrowing it; the reverse of "normalizing" is stigmatizing or, perhaps, problematizing. We will discuss this more in the chapter on political beliefs and the philosophy of history, but one mechanism for narrowing the Overton window is broadening the meaning of pejorative terms, something we discussed in the chapter on verbal disputes. In most contexts in contemporary American politics, terms like "fascist" and "communist" communicate a great deal of disapproval. Over time, the meanings of

these terms seem to expand, especially when the bad characteristics they denote become rarer. Scientists studying "prevalence-induced concept change" note that

> when the "signal" a person is searching for becomes rare, the person naturally responds by broadening his or her definition of the signal – and therefore continues to find it even when it is not there. . . . For example, when blue dots become rare, participants start calling purple dots blue, and when threatening faces become rare, participants start calling neutral faces threatening.[23]

More concretely, psychologist Nick Haslam has written about "concept creep" in his discipline:

> Many of psychology's concepts have undergone semantic shifts in recent years. These conceptual changes follow a consistent trend. Concepts that refer to the negative aspects of human experience and behavior have expanded their meanings so that they now encompass a much broader range of phenomena than before. . . . The concepts of abuse, bullying, trauma, mental disorder, addiction, and prejudice . . . illustrate these historical changes.[24]

Note, however, that when one Overton window narrows, another broadens: expanding our notions of what's taboo in one area may contract them in another, for instance when an expanded notion of harm results in greater openness to violating liberty to protect people. This and other aspects of the Overton window deserve more careful study.

Epistemic blowback

What I call *epistemic blowback* occurs when bad arguments in favor of some political claim convince us to reject that claim. In general, we shouldn't care too much about the *existence* of bad arguments in favor of some claim. Take the following argument: Everything is blue, therefore the sky is blue. That's a bad argument – an unsound one; it has a false premise. But the fact that we can formulate it doesn't mean that we should doubt that the sky is blue. More broadly, an argument is just a series of propositions, with one set apart as the conclusion. There are infinitely many propositions, and so there are infinitely many bad "arguments" we can construct for any conclusion.

Epistemic blowback doesn't occur because of the existence of bad arguments but because of their *prevalence* and *popularity*. When partisans of a certain party or adherents of a certain view give us arguments in favor of that party or view, we figure that they're either honestly expressing the reasons they have for their own political beliefs or providing the reasons they think are most likely to convince us. If the arguments are bad, then, we have grounds for thinking either that their

own beliefs are unfounded or that, from our perspective, there are going to be no convincing reasons for sharing their political beliefs. Either way, the view is weakened.

Epistemic blowback is plausibly related to some of the phenomena we've considered so far. In polarized contexts, people are likely to be unaware of the weaknesses of arguments made on their side. Plus, a bad argument might be as good a signal of one's partisan loyalty or membership as a good argument, or even better – so if we buy a signaling-based theory of partisanship, it's easy to understand why someone might choose to express bad arguments. In fact, Robert Talisse has suggested that working to recognize the flaws of our own arguments, and our own side's arguments more generally, is a good antidote to problems of polarization.[25]

When a viewpoint gets too dominant, people get lazy in their support for it, thinking that they can take its truth, and people's acceptance of that truth, for granted. Perhaps they don't investigate it enough to think of opposing arguments, or perhaps they find its truth so obvious that they never even think to ask why they believe it. In philosophy, we think about a lot of ideas like this, leading to some of the classic philosophical puzzles about whether other people exist, whether there's an external world, whether there's such a thing as right and wrong, whether there's such a thing as causation, and so on. Plausibly, in some political communities, some political beliefs will seem just as obviously true as the belief that there's a world outside our own heads. (In fact, many philosophers seem to think that doubt about political beliefs is less appropriate than doubt about these sorts of matters!)

Epistemic blowback plausibly occurs in, and perhaps helps to remediate, this sort of situation. Insular political groups often take their shared political beliefs to be obvious. When political process requires that they convince others to achieve their goals, though, the best arguments they can muster are bad ones. They're not just unconvincing but actively counterproductive: the fact that they don't have better reasons for their political beliefs suggests that there might not be good reasons for those political beliefs.

Epistemic blowback is thus like the flip side of the "Ideological Turing Test" coin. Polarization seems to make it harder for us to understand the other political side. But it might make it harder for us to understand our own side as well. And the less we understand our own side, the less successful we'll be in political argument. If I'm right, then political groups with bad epistemic practices are bound to suffer epistemic blowback sooner or later.

Conclusion

Polarization is the dominant force in American politics, and it comes in many flavors: affective polarization, political saturation, polarization as sorting, polarization as extremism. For the types of polarization related to political belief, there

are fascinating arguments which trace polarization to irrational sources of belief. But there may be rational sources to our polarized beliefs as well. And there are some difficult questions facing the theory of polarization as well – what's the center and what's extreme, what's normal and what's not, and so on. Opponents of polarization must find a way to characterize their core notions without simply taking for granted that political moderates are correct.

Discussion questions

1. Where do you think the center of politics can be found? Does it matter where the center is? If it doesn't, how does that relate to our arguments about disagreement and epistemic democracy? If it does, how can we explain why political change over time might be good, especially in moving societies from truly abhorrent evils like slavery?
2. Try an enclave polarization experiment: pick some claim (doesn't need to be political), have students rate (for instance) how much they would bet on the claim being true or false, then have them deliberate about the claim in groups for a short time. After the deliberation, see what ratings they give to the claim and whether those ratings have obeyed the law of group polarization.

21

CONSPIRACY THEORIES

Few kinds of political belief come in for more criticism than conspiracy theories, and few kinds of political belief are more often used to characterize what's wrong with how we form our beliefs. Conspiracy theories have been related to both types of polarization, as both a cause of sorting and an effect of sorting and as both a cause of extremism and an effect of extremism. They have also been related to ideologies, as propagandistic (see next chapter) tools meant to redirect the attention of the gullible citizen. Yet it's not easy at all to come up with an account of just what it means to be a conspiracy theory, and once an account is in hand it's not easy to say just what is wrong with conspiracy theorizing – what makes it *irrational*, the question to which we keep returning. Of course, many conspiracy theories turn out to be false, but some turn out to be true, and anyway the trick is to figure out how them turning out false is related to them being conspiracy theories. For these reasons, some philosophers are *particularists* about conspiracy theories, meaning that every conspiracy theory has to be evaluated on its own merits and the evidence for and against it – the mere fact that a conspiracy theory *is* a conspiracy theory isn't enough to discount it. For particularists, conspiracy theories are just like any other theories. The opposite position is *generalism*. However, Keith Raymond Harris has sketched a middle ground position, on which something's being a conspiracy theory is *part of* the evidence against that theory, though not enough to undermine it completely.[1]

Separate from questions about the epistemology of conspiracy theorizing are questions about its prevalence and changes in its prevalence over time. Political scientist Joe Uscinski has written about the ubiquity of breathless reporting about the apparent rise in conspiracy theorizing, but he remarks that the very frequency with which this claim is made – he cites examples in prominent publications

DOI: 10.4324/9781003355274-25

from 1964, 1977, 1991, 1994, 2004, 2011, 2013, 2017, and 2018 – seems to undermine the notion that conspiracy theorizing is really seeing a present-day rise.[2] He writes that

> claims about the ebb and flow of conspiracy theories in the public come without any precision. They confuse and conflate the number of conspiracy theories, the number of people who believe those theories, the willingness of people to believe, and the salience of those beliefs. An increase in any one of these would indicate something different, and would have to be measured differently.[3]

With a group of other scholars, he studied the issue, finding "troublingly high percentages" of belief in conspiracy theories but "little supportive evidence" regarding *growth* in such beliefs.[4] Unlike some technological and algorithmic issues we've discussed, then, conspiracy theories are, for better or for worse, a perennial concern for political epistemologists and other theorists of political belief.

What are conspiracy theories?

Uscinski et al.'s study proceeds as though we already know what the conspiracy theories are. But philosophers will worry about the initial categorization. Just what is a conspiracy theory? Say I come up with some new political belief; how do I know whether or not I'm engaging in conspiracy theorizing? At first blush we might think that any belief that posits a conspiracy is a conspiracy theory. But conspiracies often happen in the real world, and some seem unproblematic; and anyway, "conspiracy" itself is tough to define. For example, M. R. X. Dentith and Brian L. Keeley raise the possibility of a "minimal conception of what counts as a conspiracy," which has only three conditions: "There exists or existed some set of agents with a plan, steps have been taken by the agents to minimize public awareness of what they are up to, some end is (or was) desired by the agents."[5] But under this definition a surprise birthday party is a conspiracy. Conspiracies seem to require secrecy, but it's not clear what kind of secrecy they require, and in any event the notion of secrecy is none too clear itself.[6]

In an earlier work, Keeley suggested the following definition: "A conspiracy theory is a proposed explanation of some historical event (or events) in terms of the significant causal agency of a relatively small group of persons – the conspirators – acting in secret."[7] So a conspiracy theory posits that something was caused by conspirators (we'll forget that he said "a relatively small group," which is probably not part of the precondition for being a conspiracy theory).[8] This seems pretty minimal too, but in fact this definition may be too restrictive, slanted toward political conspiracy theories. When I asked a class

of undergraduates for their favorite conspiracy theories, one response was the theory that the rapper Tupac was still alive, and another was the theory that pianist Stevie Wonder was not blind.[9] Neither of these are proposed explanations of historical events; the conspiracy itself is enough.

Keeley offers an expanded concept of *unwarranted conspiracy theories* that have some of the common characteristics of the conspiracy theories we tend not to like. These further characteristics include contradiction of "some received, official, or 'obvious' account"; the attribution of "nefarious true intentions" to the conspirators; and "seek[ing] to tie together seemingly unrelated events."[10] But we might worry that these kinds of conditions make the question of what is a conspiracy theory hopelessly relative, and, also, we might worry that none of these aspects are inherently irrational. Let's look at each in turn before we move on to some other considerations.

Conspiracy theories and official stories

Take official stories. Charles Pigden writes (in 2007, of the beginning of the Iraq War):

> In Iraq the theory that Saddam was in cahoots with Al Qaeda was both inconsistent with the official view and posited evil deeds on the part of the Head of State. Thus in Iraq it was a conspiracy theory in [Keeley's "unwarranted"] sense. Not so in Britain. Thus it was permissible to believe it in Britain but rationally wrong to believe it in Iraq. . . . An epistemic principle that forbids a theory in Baghdad but allows it in London leaves something to be desired.[11]

How can it be, Pigden seems to wonder, that the same theory counts as conspiratorial in one place, but not another – indeed, that *disbelieving* that theory might mark one as a conspiracy theorist in the second place? Yet I'm not so sure how convincing this criticism is. It's normal for people in different places to have different evidence. Imagine there's a place where it virtually never rains, so that I, located somewhere else, have a rational belief that it's *not* raining there right now. But then there's a freak rainstorm, so that the people who are there right now actually observe it raining, making it rational for them to believe that it *is* raining there right now. Why shouldn't this be "forbidden" to me but "allowed" to them, given that our evidence varies? It may be a somewhat different question whether the category of conspiracy theory should be sensitive in this way, but I don't see a clear or immediate reason why not.

What things rest on instead is the evidential value of "official stories." But here we run into some trouble. As David Coady writes, "whether the pervasive skepticism of people and institutions in authority that Keeley warns us against is warranted, depends on the kind of people and institutions in authority at the time and place in question."[12] We might think that people who dissent from the

official story in London are in a different position than people who dissent from the official story in Baghdad. Perhaps people in London have more *reason* to believe the official story than people in Baghdad. But once we talk of reasons to dissent, we are in the realm of simply evaluating how much we can trust other people's testimony and to considerations of expertise that we've already discussed. It seems odd to create a special category merely for distrust that goes too far. Trust in the official story might go too far as well. But the appropriate amount of trust probably depends on the sorts of factors we discussed in the chapter on expertise, like the track record of the person or institution in question.

There's another interesting element of trust here when it comes to institutions. Official stories tend to be sent down from big, blocky-looking buildings into whose inner workings we might have little insight, leading us to rely on track records. But what if we do know how the inner workings of such institutions function? We might care in particular about whether the people in those institutions have the right incentives to produce truthful official stories. Take the media: it would make sense to trust the media if we think truthful reporters tend to be promoted and mistaken or dishonest reporters tend to be chastised or even fired. In such a situation, knowing what we know about human behavior, we would expect media employees to work to produce truthful reporting. This inference from incentives is a bit like what we saw in the chapter on decentralization.

Neil Levy offers some thinking along these lines in characterizing official stories as coming from "properly constituted epistemic authorities," elucidating: "Epistemic authorities are properly constituted to the extent to which they consist in a distributed network of agents, trained in assessing knowledge claims, who make their evidence and processes available to scrutiny, within and beyond the network."[13] To Levy, the presence of many people engaging in inquiry and disseminating some sort of collective perspective makes institutions authoritative. However, we have seen that the presence of multiple epistemic agents produces a highly reliable group only when those agents tend to be both competent and independent of each other. Coady cashes this worry out with a discussion of belief cascades (or "information cascades"), which "can occur when people express their opinions about the answer to a certain question in a publicly observable sequence," which can lead to people later in the sequence simply following people earlier in the sequence – a cascade.[14]

There is also a danger that in characterizing the "official story" as coming from *epistemic* authorities we are missing the point. Conspiracy theories tend to be about the actions of *political* authorities, and we are often, especially when conspiracy theories are on the table, not in the position of determining whether political authorities are also epistemically trustworthy. Of course if we begin by knowing which institutions to trust, we aren't wrong to trust them. But it's not clear that we tend to have access to that sort of information. That situation is arguably part of what produces conspiracy theories.

Conspiracy theories and skepticism; conspiracy theories and numbers

A side worry about doubting the "official story" is that it will lead us into skepticism about far too many things. Keeley writes that conspiracy theories

> throw into doubt the various institutions that have been set up to generate reliable data and evidence. In doing so, they reveal just how large a role trust – in both institutions and individuals – plays in the justification of our beliefs. . . . There is the free press, made up of reporters, editors, and owners who compete to publish "the scoop" before others do. There are governmental agencies charged with investigating incidents, producing data, and publishing findings. And there are, of course, various "free agents" (including the conspiracy theorists themselves) who are members of the public. Inherent in the claim that alleged evidence against a theory should be construed as evidence for that theory is a pervasive skepticism about our public, fact-gathering institutions and the individuals working in them. Thus, as a conspiracy theory matures, attempt after attempt to falsity a conspiracy theory appears to succeed, and this apparent success must be explained as the nefarious work of the conspirators. As a result of this process, an initial claim that a small group of people is conspiring gives way to claims of larger and larger conspiracies.[15]

There are a few different ways of interpreting Keeley's criticism here. First, we might interpret Keeley as saying that the problem is that *believing* a conspiracy somehow has the intellectual implication that one rationally must abandon many of one's other beliefs as well, risking skepticism. Second, we might interpret Keeley as saying that the problem is that in *inquiring about* a conspiracy that one believes, one encounters a lot of disagreement and counterevidence, meaning that maintaining the belief requires abandoning an increasing number of our other beliefs.

Relevant to the first interpretation is an argument made by Pigden, who focuses on the real existence of conspiracies in history, such as any number of coups and assassinations:

> Is [never believing conspiracy theories] a sensible belief-forming strategy? Obviously not. An epistemic strategy should maximize the chances of truth and minimize the chances of error. But if this strategy had been pursued in the past, many politically important truths would never have come to light. . . . All these theories were once inconsistent with official opinion, though nowadays official opinion has managed to catch up with the facts. Thus it would have been an epistemic mistake to have adopted this strategy in the past.[16]

In fact, this is not a very good argument. Just because an epistemic strategy results in some truths does not mean that it maximizes the chances of truth and

minimizes the chances of error. If I believe that every roll of a die will come up six, I end up believing truly about one-sixth of the time, and I believe more truths than I would if I were to suspend belief about what number to expect when I see someone roll a die. But it is rational to suspend belief even though it means losing out on those true predictions. This is similar to what we saw in the previous chapter when Levy attempted to rationalize extremism by reference to true extreme beliefs.

However, Pigden's emphasis on the existence of true historical conspiracies *does* help in rebutting the argument about skepticism. It is simply difficult to see how believing the truth about conspiracies to assassinate leaders or take over governments would have undermined the remainder of someone's beliefs about the world. And even if it did, that would have been the right result for them. The fact that some conspiracy theories end up being true means that their intellectual implications cannot be the only grounds for rejecting them. If those implications are indeed rationally inferred from true beliefs, then they themselves are rational to believe and hence not undesirable implications.

The second interpretation has a bit more appeal to it. Certainly some agent's belief in a conspiracy theory is not irrational simply because they *might* at some future point encounter contrary evidence in inquiring about that theory. We don't have a rational responsibility to anticipate all the evidence we might encounter. However, there are likely good reasons to discount one's belief in conspiracy theories when the number of people involved seems to go up. We might think that there is some sort of difficulty involved in keeping secrets or dissimulating, so that the more people are involved in a conspiracy, the more likely it is that someone has let the cat out of the bag at some point. We might also think that large conspiracies face the same difficulties as large organizations in general, of being difficult to manage and corral. We might think that purported low-level conspirators are unlikely to be sufficiently incentivized to participate, that they're unlikely to be getting anything out of it that would be worth the risk. Finally, we might think that moral goodness is more common than moral nefariousness, so that it is hard to find many people willing to participate in a nefarious conspiracy, making a large one unlikely.

Conspiracy theories and nefariousness

The inclusion of "nefariousness" in the criteria for unwarranted conspiracy theories might also seem puzzling. First of all, the term itself is strange. Philosophers rarely describe things as nefarious rather than, say, bad, unethical, or immoral. Second, what's meant by "nefariousness" needs further spelling out. Am I a conspiracy theorist if I think a group is acting nefariously from the perspective of *my* moral beliefs? From the perspective of *their* moral beliefs? From the perspective of the *truth* about morality? And there is a question about conspiracy and secrecy themselves – if they are taken to be somehow *inherently* nefarious, then

the nefariousness condition does nothing for us. Anyway, there are likely a lot of conspiracies in which people take themselves, perhaps correctly, to be acting rightly and with good intention – think of a coup that overthrows an evil dictator.

There is, however, a reason to think that in the present day, the attribution of bad intentions is a common part of conspiratorial thinking and a bad one at that. Recall from our discussion of polarization that much of American polarization is asymmetrically negative: many people's main political motivation is antipathy toward the other side. In such an environment, theories that posit nefarious conspiracies undertaken by the other side may seem especially attractive. Thus, nefariousness may be part of a lot of bad conspiracy theories in virtue of being downstream from irrational polarization.[17]

Conspiracy theories and explanation; conspiracy theories and sophistication

A common thread in philosophical critiques of conspiracy theories is that they take the wrong view of which sorts of explanations of events are admissible and likely. In a classic early work on conspiracy theory, Karl Popper wrote that the "conspiracy theory of society" is "the view that whatever happens in society . . . are the results of direct design by some powerful individuals or groups."[18] What conspiracy theorists miss is that "nothing ever comes off exactly as intended."[19] Similarly, Keeley writes that

> conspiracy theories embody a thoroughly outdated world view, a perspective on the meaning of life that was more appropriate in the last century. . . . Conspiracy theorists are, I submit, some of the last believers in an ordered universe. By supposing that current events are under the control of nefarious agents, conspiracy theories entail that such events are *capable of being controlled*.[20]

And Pete Mandik writes that because of difficulty of interpreting actions and attributing belief/desire pairs to other agents, which we discussed long ago near the beginning of the book, "any given attribution of a belief-desire pair is . . . likely to be simply *post hoc*," including the attribution of a conspiratorial intent.[21] What these writers share is that they think there's something wrong with explaining political events in terms of a few individuals' actions and intentions.

Certainly recent intellectual history has many instances of largely replacing centralized theories explaining some feature of the world by reference to a few powerful agents' intentional actions with more sophisticated decentralized theories explaining some feature of the world by reference to many agents' intentional actions or even to random chance. We saw one in an earlier chapter: Hayek's argument against central planning; another one is the theory of evolution. However, we certainly still do explain many events by reference to people's

intentional actions, and the fact that real conspiracies have occurred in history and in the world again suggests that such explanations are sometimes the right ones. It's pretty likely that at least some conspiracy-theory explanations of political events are part of our normal and best practices of inference to the best explanation.[22] Moreover, refusing to countenance the possibility of conspiracy theories may be its own form of irrationality; Coady notes the possibility of "coincidence theorists" who "irrationally reject evidence of conspiracy," as well as of "institutional theorists" who offer explanations in "impersonal institutional terms" and by reference to market forces (again, much like the sorts of explanations that might be inspired by Hayek).[23]

Institutional theorists have a certain sophistication. But how different are their explanations than those of conspiracy theorists? Coady notes a few things.[24] First, explanations of events in terms of institutions do not necessarily remove conspiracies from the picture. The history of the relevant institutions may involve conspiracies at their founding, for instance. And market forces can lead to conspiracies, as in price gouging, which often must be addressed by explicit law. Finally, institutions and markets are made up of individuals, and their sort of overall or collective actions often can't be explained except by reference to the smaller actions of those individuals, so that we never really escape explanation in terms of people's intentions. For Coady, then, it is hard to see why this "systemic" or "structural" approach represents a wholesale *epistemological* improvement on conspiracy theorizing.

Conspiracy theories and ideology

The question of systems and structures brings us back to one type of theory of political belief, the ideological type, which had examples like the false consciousness theory and the system justification theory. In their most brute forms, the ideological theories of political belief are simply conspiracy theories of political belief, because they posit that, through propaganda (see next chapter) and other mechanisms, already-powerful interests organize the political beliefs of the less powerful so as to maintain and increase their own power, a nefarious intent. Thus, it is hard to see how philosophers engaged in ideology critique could oppose conspiracy theorizing. Of course, some such philosophers prefer the more sophisticated, institutional forms of theorizing, like the ones we just saw Coady discussing. For instance, we might think of Haslanger's idea of self-fulfilling belief as buttressing a sophisticated ideological theory that removes the need for top-down explanation. However, Haslanger also practices genealogy, tracing the history of our concepts and social practices to critique them and offer debunking arguments, and just as Coady suggests, the presentation of the *beginning* parts of this history can seem a lot closer to conspiracy theorizing than institutional theorizing. Finally, by explaining ideologies in terms of their "functions" in maintaining an order, ideology critics seem to me to be offering

something close to an intentional explanation anyway. It is hard to avoid the conclusion that ideological theories of political belief are at best superficially sophisticated versions of conspiracy theorizing. The good news for ideology critics is that it's not clear, as we saw in this chapter, that conspiracy theorizing is universally or even generally epistemically irresponsible.

The idea that ideology critique might be a form of conspiracy theorizing is an example of one last kind of point I want to make. It is rare for those who take conspiracy theorizing to be necessarily irrational to address the possibility of a *conspiracy-theoretic dilemma*: a situation in which an epistemic agent is faced with believing either one conspiracy theory or another, with no viable option of believing neither. This is an especially important dilemma since accounts of the spread of conspiracy theories can often seem like conspiracy theories themselves – for instance, they might explain that the interests of powerful actors are served by lots of laypeople believing some conspiracy theory. In rationality, there must be at least one acceptable option; this is a form of the ought-implies-can principle.[25] So if it's possible for an epistemic agent to be faced with a choice between believing one conspiracy theory and believing another conspiracy theory, it must be the case that it is rational for that agent to believe at least one conspiracy theory.

Conclusion

Popular commentary often presents conspiracy theorizing as an urgent problem involving a great deal of irrationality. But both the urgency and the irrationality are debatable. We've seen that it's hard to find a definition of conspiracy theories and, once a definition is in hand, that it's hard to figure out just what is irrational about them. Though arguments have been offered on both sides of the debate about the rationality of conspiracy theories, the confidence with which many commentators wave away conspiracy theories and conspiracy theorists has not been supported thus far by the philosophical analyses of social epistemologists.

Discussion questions

1. Do you believe any conspiracy theories? If so, what attracts you to them? Do they fit the definitions offered in this chapter, or are they counterexamples to those definitions? If not, why not? Do you think there's something wrong in general with conspiracy theorizing?
2. Some opposition to conspiracy theorizing seems to require that we trust almost blindly in the goodness of our leaders and institutions. How do you feel about this?
3. As we saw, explanations of social circumstances that avoid considering the intentions and actions of powerful individuals and instead focus on broader incentives and market forces seem to be more sophisticated. Broadly speaking, in trying to understand politics, do you prefer the former, individualistic kind of reasoning or the latter, structural, systemic kind of reasoning?

22

PROPAGANDA, DEHUMANIZATION, AND GULLIBILITY

We've discussed the ideological theories of political belief. One of the proposed mechanisms of ideological spread and enforcement is *propaganda*. Like "ideology," "propaganda" has both a neutral and a negative meaning. The contemporary philosophical study of propaganda is often focused around language, and in this chapter we'll critically discuss one major book, Jason Stanley's *How Propaganda Works*, and one major idea, the idea of dehumanization, which both take propaganda to have a lot to do with language. However, we might begin by thinking that this focus on language is misplaced. The Wikipedia entry for "propaganda" largely features *images* like these:[1]

DOI: 10.4324/9781003355274-26

The left-hand picture is described as "Poster depicting Winston Churchill as a 'British Bulldog'" while the latter is described as "Poster of the 19th-century Scandinavist movement," a pro-unification movement in the Scandinavian countries.

Propaganda and code

Stanley makes about three major positive claims about the nature of propaganda in *How Propaganda Works*. The first is that propaganda often, and most characteristically, operates by code words and dog whistles. Stanley gives examples like the following: "When the news media connects images of urban Blacks repeatedly with mentions of the term 'welfare,' the term 'welfare' comes to have the not-at-issue content that Blacks are lazy."[2] His idea is that the very *word* "welfare" can become propaganda because it encodes this stereotype. Stanley considers a similar possibility with regard to the word "lady," on which use of this word suggests that "well-dressed white women" are all ladies and ladies are all "submissive and in need of care."[3] Because this content is "not at issue" or "smuggle[d] in,"[4] it is difficult to challenge or even recognize. This not-at-issue content is related to (but not identical with) the phenomenon of *presupposition*, in which there are propositions taken for granted beyond the statement one explicitly expresses. One neat example of this is with verbs like "stop": if I say "China has stopped stockpiling metals," then my statement presupposes that China *used to* stockpile metals.[5] Another kind of example might be something like this: "We haven't yet found Smith's murderer." This presupposes that Smith *has* a murderer and thus that Smith was murdered. A final type of coded language, not discussed as much by Stanley, is a *dog whistle*, a statement which

> allow[s] two plausible interpretations, with one interpretation being a private, coded message targeted for a subset of the general audience, and concealed in such a way that this general audience is unaware of the existence of the second, coded interpretation.[6]

Stanley thinks propaganda largely functions through these coding devices like not-at-issue content, presupposition, and dog whistles. If he's right, then it's important for our understanding of political beliefs. It could mean that many of our political beliefs emerge from aspects of language we don't even recognize.

But there are two difficult challenges for his theory on this front. First, it's not clear that the theory behind these apparent coding devices is solid. As Jason Brennan mentions,[7] little is done to argue for this particular interpretation of the word "welfare," for example. To my mind, the same goes for the other devices. Just as we mentioned in the last chapter regarding ideology theories of political belief, this code-word theory of propaganda, in which evil, more or less subconscious suggestions and inferences are lurking behind normal words and

statements, feels a bit like a conspiracy theory, in which some masters of language are secretly manipulating us to believe things that serve their purposes. The second challenge is that it's not at all clear that propaganda tends to work this way. Take the images above; nothing is coded about them. And as Brian and Samuel Leiter note,[8] much propaganda is not coded at all – offensive content about race, for instance, one of the types of examples Stanley is most interested in, is often asserted explicitly by racists rather than coded, and this is probably even more the case in more racist societies, which certainly will have their share of racist propaganda. So coding is a questionable phenomenon, and its relationship to propaganda is questionable as well.

Propaganda and ideals

Stanley relates propaganda to "ideals."[9] Ideals are something like values, goals, or principles. Stanley differentiates two relationships propaganda might have to an ideal:

> *Supporting Propaganda*: A contribution to public discourse that is presented as an embodiment of certain ideals, yet is of a kind that tends to increase the realization of those very ideals by either emotional or other nonrational means.
>
> *Undermining Propaganda*: A contribution to public discourse that is presented as an embodiment of certain ideals, yet is of a kind that tends to erode those very ideals.[10]

Brennan writes that wartime propaganda (like the bulldog Churchill above) does not bear either of these relationships to ideals.[11] I'm more concerned with "undermining propaganda" in particular. Recall four characterizations we had of political disputes: disagreements over values, disagreements over nonmoral facts, merely apparent disagreements over words, and a final, purer type of dispute in which each party sort of just wants to win. In a disagreement over nonmoral facts, both sides may share values but disagree on what best enacts those values.

Take a debate over affirmative action in which both parties share the value of anti-racism. One disputant might say that implementing an affirmative action policy enacts the value of anti-racism because it redresses historical racist injustices, uplifts people who were marginalized on account of racism, and improves the diversity of institutions, leading people who are a part of them to be less racist. Another disputant might say that implementing an affirmative action policy undermines the value of anti-racism because it discriminates based on race, makes assumptions about people based on race and treats them as "tokens," and may lead to "mismatch" situations in which beneficiaries of the policy are unable to thrive. Of course, it's entirely possible that one of these disputants is *right* and

the other one is *wrong*. The question is whether the one who is wrong is *engaging in propaganda*. Under Stanley's definition, they must be, since whichever side is wrong is advocating for something that in fact tends to erode the ideal of anti-racism. But this is an absurd result: it suggests that good-faith disagreement between people who share values is necessarily propagandistic. Stanley might be on safer ground with an "internalist" approach on which undermining propaganda is taken to be of a kind that the propagandist *knows* tends to erode the ideals it is presented as embodying.

Propaganda and irrationality

We might think that supporting propaganda is more common than undermining propaganda. But Stanley's definition of supporting propaganda requires that it use "emotional or other nonrational means." First, it's not clear that emotional means are necessarily nonrational means. A Scandinavian seeing the friendly pro-unification poster above might feel warmed and thus, rationally, come to understand how much they desire the unification of Scandinavia. Second, there are examples of types of propaganda that don't need nonrational means to be effective. Megan Hyska writes that "hard propaganda" and "propaganda by the deed" are both counterexamples to this "irrationalist" conception of propaganda.[12] Hard propaganda and propaganda by the deed share that they motivate the observer of the propaganda through power – in hard propaganda, the power of the state to transmit an absurd message; in propaganda by the deed, the power of the working class (for instance) to resist the state. Both the intended and actual effects are rational in that the observer is meant to update their beliefs based on a new but also newly accurate understanding of who has what sort of power and to act in line with those new beliefs.

I've focused on Stanley's account because it is a recent, prominent, influential, and philosophical theory of propaganda. Many others have been offered, but as far as I can tell, they tend to be massively overinclusive or sadly obscure. In particular, they either include Stanley's irrationality condition or fail to say anything distinctive about propaganda. Online, Mira Sotirovic has compiled a list of definitive statements about propaganda.[13] Certain features of these stick out: (a) Propaganda is generally aimed to change an audience's attitude or behavior; (b) The *scope* of propaganda is generally large, a "mass" or "collective" or "society"; (c) The *effort* of propaganda is generally extended, long term, or somehow systematic; (d) The *goal* of propaganda is generally political, that is, it's political attitudes or behaviors that are targeted; (e) The *method* of propaganda is generally not an appeal to reason but some sort of manipulation or psychological "technique." There are hard cases for such theories, like a brilliant novel with a political message, an advertisement for candy, or an attractive celebrity's Instagram feed. These conditions could be further refined to address such cases.

Dehumanization

Philosophers have recently taken an interest in *dehumanizing propaganda* and its purported consequence of *dehumanization*. The theory of dehumanization attempts to explain some of the worst human political conduct in history, like genocide. The overarching idea is that humans are social animals who must learn to deal with each other with grace and etiquette and to obey social rules about interpersonal interaction; thus, instances of extreme violence and evil are deviations that stand in need of explanation, and the best explanation of such extreme deviations from our rules about dealing with other humans is that they involve ceasing to see the other humans *as* human and treating them as though we might treat animals or objects.[14] Nobody denies that people who commit atrocities often compare their victims to animals (like "rats, lice, dogs, and cockroaches").[15] But the question is whether people really stop believing their victims are human and start believing that they're animals.

Harriet Over has recently raised a few challenges for this kind of hypothesis. First, she writes, animalistic language can be used to praise as well as denigrate, as in the example of "Lionheart"; more generally, many societies treat certain nonhuman animals quite well, so that lacking humanity is not necessarily a "risk factor" for being victimized.[16] Second, though propaganda against victim groups may sometimes compare them to animals, it will also sometimes describe them in ways that are distinctly human, like as "traitors," or as cunning, malicious, jealous, spiteful, dishonest, or disloyal.[17] Indeed, such states may be precisely why some groups are victimized.[18] Third, and perhaps most obviously, comparison or association is not identity; the dehumanization hypothesis seems to assume that analogies and metaphors are beyond a great number of people.[19] Separately, Kate Manne has challenged the dehumanization hypothesis by suggesting that it overestimates the degree to which humans normally treat each other well.[20]

In response to these sorts of challenges, David Livingstone Smith has recast the theory of dehumanization somewhat, suggesting among other things that dehumanizing propaganda causes atrocities only in tandem with some sort of ulterior motive on the parts of victimizers who hear the propaganda and that it's generally obvious from context which animal comparisons are dehumanizing and which aren't.[21] I'm mostly unconvinced by his responses, in part because they rely upon the idea that dehumanizers see the people they purportedly dehumanize as simultaneously human and nonhuman, an odd contradiction that does not help his case.[22] However, for our study of political beliefs, the really central question is the obvious one. To what extent should we think of people as being so gullible? Can hearing a leader insult another group by comparing them to rats or cockroaches really start to make us think that members of those groups *are* rats and cockroaches and to forget the obvious fact that we're members of a common species?

Gullibility; trust

In his book *Not Born Yesterday*, Hugo Mercier considers the question of human gullibility.[23] His conclusion is that gullibility is easy to overstate. When people's apparent beliefs seem to be the result of being too gullible, Mercier offers alternative explanations of those beliefs. One explanation he gives is that people are only acting as though they believe what we might take them to gullibly believe because such action may be in their interest for other reasons; we've seen this move made by Hannon, Táíwò, and Hyska as well. In other cases, our apparent beliefs are of the sort that doesn't eventuate in action; Mercier notes that many people who seem to have irrationally extreme political beliefs don't take associated extreme political actions, indicating that the beliefs aren't genuinely held (compare again Hannon and the bets). There are also issues on which we appear to have formed beliefs gullibly but which don't bear on our lives at all; Mercier writes that when this happens we may instead be simply enjoying the "mind candy" of dealing with information that seems counterintuitive or clandestine and is therefore sort of fun or naughty to think or talk about.[24] This idea also builds from the belief-action connection.

In cases where we do seem to have genuine beliefs, Mercier stresses that even when they seem gullible they may be based off of sensible or intuitive inferences. One of my favorite examples here is bloodletting. Bloodletting has been a popular medical practice throughout history, and it might be tempting to try to understand its popularity by "trac[ing] it back through the great physicians who had defended it."[25] Indeed, this is a genealogical method that we've seen in the chapter on debunking arguments. But Mercier suggests that bloodletting crops up in all sorts of human societies, meaning that its popularity can't be primarily the result of many people gullibly believing some theory some individual came up with. Rather, there must be something commonly or even universally intuitive about the practice,[26] even if it turns out to be unscientific in the end.

Irrational gullibility is an important part of the theory of propaganda. If most people aren't irrationally gullible, it's hard to see how the worst kinds of propaganda and dehumanization could be as effective as some – those who make propaganda central in their accounts of people's political beliefs – think it is. Interestingly, there are also accounts that hold that people are insufficiently trusting, rather than being too trusting, of authorities and fellow citizens. We saw some hints of this in the chapter on expertise; it is also present in some recent work on trust in politics, like Kevin Vallier's book *Trust in a Polarized Age*. Vallier opens that book by writing that "Americans are finding it harder and harder to trust one another. . . . [O]ur trust in our fellow citizens has fallen dramatically. . . . Trust in government and democracy has fallen steeply as well."[27] We haven't discussed trust in politics, but it is obviously an important notion for social and political epistemology, as one aspect of trust is epistemic: trusting someone in this sense involves believing things on the basis of their word. Vallier

also connects trust to polarization, writing that "falling trust and increasing partisan divergence are mutually reinforcing."[28] But it is difficult to say anything general about just how much we should trust one another.

Conclusion

There's no question that all sorts of political actors produce propaganda and that propaganda sometimes affects other people's beliefs and actions. However, just how propaganda works remains a matter of some controversy, and in particular it's not clear that propaganda must work by the use of code words or by irrational mechanisms. Whether political action and, in particular, horrible atrocities can be explained by people genuinely believing propaganda and, in particular, dehumanizing propaganda depends in part on how gullible we think people are. As we've seen elsewhere in the text, there are reasons to think that people aren't so gullible, meaning that propaganda may not play quite the role some theorists assign to it.

Discussion questions

1. Do you believe political discourse proceeds by the use of code words? Do you think political suggestions are made covertly, whether through texts, images, or something else, which lead people to have a certain political outlook relatively unquestioningly?
2. How gullible do you think people are in general? What are some considerations in favor of the idea that people are too gullible? What are some considerations against it? In terms of our social structure, why might it make sense for humans to be very trusting, and why might it make sense for humans to be very vigilant?

23

NEWS, NARRATIVES, AND RUMORS

Few terms seemed to capture the difficult social-epistemological situation American political commentators took the country to be in starting around 2016 better than "fake news," but in the end, few terms exemplified that situation better than "fake news," too. Now largely dropped for the more technical-sounding but little clearer term "misinformation," the furor around "fake news" prompted a flurry of articles in epistemology and even a few books. But the most commented-upon aspect of "fake news" by philosophers is its vague and shifting definition. Indeed, we'll see that "fake news" is an even slippier concept than the ones, like "politics" and "conspiracy theory," we've wrangled with so far. Many philosophers recommend abandoning it entirely as inappropriate for precise analysis.[1]

In this chapter we'll discuss non-expert sources of political information and concerns of trustworthiness around them, focusing on media sources in general but broadening to look at political testimony in general toward the end. Like many works of political epistemology, this text has emphasized political disagreement and the partisanship of news sources. But we should start out by reminding the reader how much agreement there is in the media and how trustworthy reporting is in general. If – in the present-day American media world at the time of this writing – you see on the news or read in the newspaper that a certain natural disaster has occurred, or that a certain team has won a sporting event, or that a certain treaty has been signed somewhere, you will almost never regret believing what you see or read. So where does the problem arise? Why is trust in the media so low?

Practical problems in this area are similar to the problems we saw regarding expertise. To do a good job in believing a proposition on the basis that it was asserted by a reliable news source, first we need to be able to recognize which

DOI: 10.4324/9781003355274-27

things are reliable news sources and then we need to be able to ensure we know just which are the propositions in regard to which those sources are reliable.

Media trust and objectivity; balanced media diets

Overall, trust in the American media is at a very low point at the time of the writing of this book. Yet trust in the media is also polarized. Linley Sanders, a data journalist at YouGov, wrote the following about a March 2022 poll they did:

> YouGov asked 1,500 Americans where they get their news from and how much they trust a variety of prominent media organizations and news anchors. The poll, conducted from March 26–29, shows that while Americans are more likely to trust than distrust many prominent news sources, there are very few organizations that are trusted by more than a small proportion of Americans on both sides of the political aisle. In fact, the most Americans overall place trust in an organization that rarely covers domestic politics: the Weather Channel (52% of Americans trust it). The Weather Channel is trailed by the U.K. news outlet, BBC (39%), the national public broadcaster, PBS (41%), and The Wall Street Journal (37%).
>
> The most politically polarizing media outlet is CNN, a cable news outfit that has been a frequent target for Donald Trump. Nearly two-thirds of Democrats (66%) rate CNN as trustworthy, compared to 11% of Republicans, a 55-point difference, larger than for any other outlet. The second-largest partisan difference, 49 points, is for news coming from The New York Times: 63% of Democrats trust the outlet, whereas 14% of Republicans do.
>
> Democrats, generally, are more trusting of any mainstream news outlet. The exception is Fox News, which 53% of Republicans trust, compared to 19% of Democrats.[2]

This poll gives a good sense of the outlook of the American public on its news media. On the picture that emerges here, the American media is at its most trustworthy when it's reporting on issues unrelated to American politics. On issues related to American politics, the news networks seem to be interpreted by the public as being partisan, with those on the same (perceived) side as the network giving it a *small* trustworthiness boost and those on the opposite (perceived) side giving it a *large* trustworthiness drop. This difference between the size of the same-side trustworthiness boost and the size of the opposite-side trustworthiness drop is consistent with what we discussed about the largely negative character of affective polarization in America.

Every now and then, charts circulate online categorizing various news sources by left/right alignment and by reliability. I've never seen any evidence that such charts are much more than the off-the-cuff evaluations of a single individual.

Still, a politically informed or even politically sensitive viewer or reader can probably assess the partisan lean of a news source pretty quickly by the tone of its coverage, especially its opinion pieces. Many media outlets separate reporting from opinion, but this separation isn't as strictly enforced as it used to be, and the partisan split in trustworthiness will virtually always track the tendency of a source's opinion page (i.e., an opinion page that tends to feature left-wing or pro-Democratic Party pieces will usually be more trusted by Democrats and less trusted by Republicans, and vice versa for an opinion page that tends to feature right-wing or pro-Republican Party pieces).

Alex Worsnip has suggested that we respond to the bias or partisanship of news sources by diversifying our sources rather than leaning into our preconceptions or attempting to consume only unbiased sources.[3] Regina Rini has argued otherwise. She writes that we can more readily take those we agree with about politics as epistemic peers than those we disagree with, one consequence of which is that we can be rational in trusting media sources on our side and not media sources on the other side.[4] However, Rini hands off responsibility for preventing fake news and other kinds of misinformation to institutional actors like fact-checkers.[5] Having determined that partisanship can be rationally acceptable, she doesn't offer an explanation of why, then, these fact-checkers, even the most rational ones, wouldn't reason in exactly the way she vindicates elsewhere. Fact-checking just *is* doing inquiry and epistemology right; so if partisanship is a way of doing inquiry and epistemology right, fact-checking won't solve the problems caused by partisan news coverage.

There is also a question of partisan lean when it comes to the "mainstream media" as a whole. For instance, a 2020 *Washington Examiner* article suggested that over 90 percent of journalists' political donations in that election year had gone to Democrats.[6] Questions about media bias often lead into debates about objectivity, which are much like the debates about neutrality we considered in the chapter on liberal epistemology. Detractors, including many journalists, argue that objectivity is impossible or that objectivity is undesirable. It is said to be impossible because journalists, like anyone else, enter into work on a story with all sorts of opinions, which affect what they think is likely and what they think is important. This is true, but the fact that objectivity is not an automatic state does not mean that it is an impossible one to achieve. It is also said to be impossible because the very act of choosing what to write and publish about requires making value judgments of the sort that objectivity demands we avoid. This is true too, but it's not clear why someone couldn't be objective in researching and writing a story after having been subjective in choosing to do so. Philosophers of science used to make a similar distinction between the "context of discovery," in which research questions were chosen and hypotheses constructed, and the "context of justification," in which experiments were done and evidence evaluated. Objectivity is finally said to be undesirable because it prevents the news

media from being maximally politically potent in an activist role by taking sides. To me, this doesn't seem undesirable at all, but in any event, the ability of the media to be politically potent is probably parasitic upon its being perceived as objective, so that eschewing objectivity can only be politically efficacious in the short term. These sorts of debates about what sort of attitude journalists should take toward their inquiries are essentially matters of applied epistemology, to the extent that concept makes sense.

Fake news and misinformation

What is "fake news"? The philosophical controversy over the analysis of this term revolves around just what the important fakery is in a fake news story and what the goal of that fakery is.[7] The broadest possible definition comes from Justin McBrayer; for him, "something counts as fake news if it's misleading and yet news," and "fake news is misleading information."[8] Romy Jaster and David Lanius define fake news as "news that lacks truth and truthfulness," that is, news that is misleading and which was authored without the goal of being truthful.[9] Even more strictly, Duncan Pritchard thinks the goal of fake news must be to mislead, so that mere "bullshit," in which a speaker doesn't care whether what they say is true or false, will not count as fake news.[10] For other philosophers, the fakeness of fake news involves its pretense to be real news, in the sense of reporting, not in its pretense to truth or truthfulness. A group of philosophers offers this definition: "Fake news is the broad spread of stories treated by those who spread them as having been produced by standard journalistic practices, but that have not in fact been produced by such practices."[11] Finally, Don Fallis and Kay Mathiesen prefer a definition of fake news as "counterfeit news":

> A story is fake if and only if it is not genuine news, but is presented as genuine news, with the intention and propensity to deceive. . . . Genuine news has been produced by professionally trained reporters, fact checkers, and editors, who are attempting to provide fair and accurate accounts of current events.[12]

Our normal philosophical method, which we used in discussing terms like "politics," was to check such definitions against our intuitions. Potential counterexamples to some of these definitions include stories created to get "clicks" or sell magazines rather than specifically to spread a falsehood, false stories which well-meaning journalists published thinking they were reporting the truth, and satirical stories which are intended to make people laugh but are sometimes interpreted as serious. I'll leave thinking through these sorts of potential counterexamples as an exercise for the reader. There's a different question here: Do we really have strong intuitions about the term "fake news"? It's a recent linguistic innovation, a neologism which was contested almost instantly upon being

invented. It's hard for me to see how it could be the sort of term with which philosophical analysis could see much success.

Another complication has to do with the importance of "standard journalistic practice," "professionally trained reporters," and so on in some of the definitions of "fake news." For an epistemologist, what's important about a journalistic practice or a way of training reporters should be how reliable it is – its relationship to the truth. But if we were at a point where we could characterize which sorts of trainings and practices would reliably lead to the truth, then we could simply define "fake news" as news that's produced in other ways than those. If a newspaper publishes a story of a certain type but the reporter forgets to ask a key figure for comment, which let us imagine is standard procedure for a story of that type, does that make the story fake news? On the other hand, what if an upstart magazine with journalists who haven't received professional training manages to uncover a great story no other publication had found – would that be fake news simply because of the training issue? And there are all kinds of criticisms of standard journalistic practice, like the debate over "objectivity" and bias broached in the previous section and in the chapter on liberalism. For these sorts of reasons, it's hard for me to see why we should care about "fake news" if it's defined merely as against "real news," where the reality of real news is simply the standard operating procedure of a news media whose reliability is already a matter of controversy.

Moreover, the term "fake news" has largely given way in public conversation to the terminology of "misinformation" and "disinformation," so philosophical debate about its meaning seems to no longer be particularly helpful. That said, I don't think those terms are much more perspicacious than "fake news." We have already discussed the concept of propaganda, and of course we all have a very intuitive, commonsense background notion of *lying*. Personally, a cottage industry in constantly creating new terms used to castigate political opponents' speech is not to my liking.

How big a problem is fake news? We already looked at some suggestions that echo chambers and conspiracy theories might not be as big a problem as people think. David Coady suggests the same about fake news and indeed suggests that the term "fake news" itself serves a propagandistic function of undermining and even censoring opposing views.[13] These sorts of concerns are why I opened this chapter by discussing the debate about fake news rather than fake news itself.

Narrative lean

Knowing the partisan lean of a news source is one thing. In some cases, we may know something about what I think of as its *narrative lean*: what sorts of stories it tends to sell and produce, which has something to do with the worldview of its journalists and executives (perhaps an "ideology" or perhaps just heuristics

that simplify the facts, in the way Friedman suggested we all must use), but also something to do with its interpretation of the market: what its journalists and executives think can sell papers and gain readership. To my mind, a great deal of media bias is better conceptualized as a narrative lean than as a partisan lean and as honest confusion rather than deceptive politicking.

In recent years, narrative lean is most easily observed in rather sensitive topics; hopefully my discussion of these topics doesn't upset any readers. I'm thinking in particular of stories like *Rolling Stone*'s 2014 article "A Rape on Campus," an explosive piece which was retracted in its entirety less than five months later, and the reporting on the hate crime hoax perpetrated by actor Jussie Smollett in 2019. Rightly or wrongly, many journalists took for granted that these stories were true as reported, and I think this is because they seemed to these journalists like *the sorts of things that would likely happen* in a world with ubiquitous poor treatment of women and minorities. The combination of this judgment of likelihood, backed by narrative lean, with the activist mindset many contemporary journalists possess, which has to do with wanting to directly improve the state of the world through reporting rather than simply reporting on it, led to a great deal of overeager and unquestioning reporting in these instances. (It doesn't help that newsrooms have slashed budgets for investigative reporting, meaning that a substantial amount of journalism is simply reporting on other people's reporting.) As we'll see later, narrative heuristics may also affect our judgments of the political epistemology of history.

At the extreme, some people theorize that the selection of narratives you consume from the media, like the "stereotypes" that concerned Jeffrey Friedman in his theory of ideational determinism, "dictates your perception of reality."[14] This view, that people's political beliefs result simply from a series of broadcasted constructs interacting with each other and not from any contact with material reality, was for a time popular in postmodern social theory,[15] but I think it's unlikely to fully capture political epistemology. Though journalistic frames may affect our perceptions, bread-and-butter economic issues like unemployment and inflation, as well as issues of basic safety, are easily felt in an unmediated way by people they affect. Our theory of political belief should take this into account.

Political rumors, political testimony, political skepticism

The difficulty of trusting media sources or anyone else when it comes to politics is the source of another argument for political skepticism, due to Blake Roeber.[16] To simplify a bit, Roeber argues that we can't come up with justified political beliefs without trusting others, but that from what we know about political beliefs, we can't really trust others about politics, either, so that we can't come up with justified political beliefs, period. Roeber develops this argument in a book of his own, but what I find most compelling is a single example he uses. Roeber

gives an anecdote in which he tells a friend that he read somewhere that "handgun owners are more likely to shoot someone on accident than use their guns to defend anyone."[17] When pressed, he tries to remember where he first read this claim, only to find no original trustworthy source for it, and indeed finding that it would be a rather difficult claim to make confidently, since there are no reliable statistics about how often handguns are used defensively. This sort of claim is what I now think of as a *political rumor*.

Political rumors operate just like regular rumors. People start saying them and then other people start believing them, and then all of a sudden they feel like common knowledge, even though they appeared almost out of thin air. Despite its high intellectual standards in some contexts, academia is rife with political rumors. In college, for instance, I was taught (in a sort of offhand, did-you-know-that sort of way) that vibrators – yes, those kinds of vibrators; I'm sorry – were used in Victorian medicine to help treat hysteria. But in a 2018 paper, Hallie Lieberman and Eric Schatzberg suggest that this claim is false and that it traces to a single source which provided no real evidence for it. They write that a scholar named Rachel Maines, in a book called *The Technology of Orgasm*,

> argues that Victorian physicians routinely treated female hysteria patients by stimulating them to orgasm using electromechanical vibrators. The vibrator was, according to Maines, a labor-saving technology that replaced the well-established medical practice of clitoral massage for hysteria. She states that physicians did not perceive either the vibrator or manual massage as sexual, because neither method involved vaginal penetration.
>
> This argument has been repeated in dozens of scholarly works and cited with approval in many more. A few scholars have challenged various parts of the book. Yet no scholars have contested her central argument, at least not in the peer-reviewed literature. Her argument even spread to popular culture, appearing in a Broadway play, a feature-length film, several documentaries, and many mainstream books and articles. This once controversial idea has now become an accepted fact.
>
> But there's only one problem with Maines' argument: we could find no evidence that physicians ever used electromechanical vibrators to induce orgasms in female patients as a medical treatment. We examined every source that Maines cites in support of her core claim. None of these sources actually do so. We also discuss other evidence from this era that contradicts key aspects of Maines' argument. This evidence shows that vibrators were indeed used penetratively, and that manual massage of female genitals was never a routine medical treatment for hysteria.[18]

An *Atlantic* article on the controversy quotes Maines as claiming that her theory was only intended as a "hypothesis."[19] But if it's true, that's even worse

news: it means that this sort of intellectual rumor (and this one is definitely a political rumor, given its relationship to sex and gender issues) can spread without anyone even initially making a claim. Political beliefs start to seem like a game of "Telephone." No wonder Roeber thinks we should embrace "political humility" instead.

There are rumors in all areas of life, though. Why think that political rumors are especially dangerous? The reasons hearken back to some of the theories of political belief we discussed earlier and to some other political dynamics. If we think we are especially prone to tribalistic thinking when it comes to politics, then we will also think that we are especially prone to believe nasty or strange rumors, and this effect will be magnified in a context of widespread polarization. If we think we are especially prone to conspiracy theorizing when it comes to politics, well, the rumor is the natural conduit for conspiracy theories: underground, unofficial, hard to trace to a source. If we think politics is intractably complex, so that we have to trust and defer to others when it comes to forming our political beliefs, then we will expect to be especially susceptible to the spread of rumors when thinking about politics, since we must consult others in doing so. Finally, political dissent often leads to social ostracism, so that people often engage in it secretly if at all. In other words, politics is arguably complex and social in a way that often involves conflict, antipathy, and punishment, and this seems like a breeding ground for rumor, insinuation, innuendo, and secrecy. More generally, if we think we are more likely to be irrational about political beliefs than other beliefs, as (we've seen) many people do, then we will expect people to be vulnerable to political rumors in a way they aren't to others.

Conclusion

How bad is the news environment in the United States? There are certainly robust patterns of partisanship in the media. Yet they are not universal, and even from partisan sources we can still get trustworthy information on a wide variety of topics. Some news is certainly false or misleading, as explained by partisan and narrative lean and by the existence of fake news and misinformation, but the extent of these problems is easily overstated. The political epistemology of the news media is likely to continue developing, and I encourage readers to think more about it as they consume media products.

Discussion questions

1. Have you ever been tricked by a fake news source? Have you ever thought that a satirical joke was a real news article? Either way, what sorts of cues do you think you use to determine how trustworthy a news source is? (One of mine: the more ads, and the "sketchier" the ads, the less trustworthy the news.)

2. How partisan do you think the media is? Should a political partisan stick with the news sources that fit their political views, attempt to find unbiased news sources, or try to consume a range of media products?
3. Is there a clear line between opinion writing and news reporting? What are some ways in which a reporter's political beliefs might influence their reporting?
4. Do you think objectivity is possible? Do you think it is desirable?

24

POLITICAL EPISTEMOLOGY AND HISTORY

Some think the two sides that seem always to emerge in politics represent different sides of or perspectives on history. One side thinks we must preserve the past, or even return to it, while another thinks we'll know better in the future. "Conservative" and "progressive" do seem like chronological terms, but I don't think people's views of history determine their views on contemporary politics. However, history does have a relationship with political epistemology that's been heretofore underexplored by professional philosophers. In this chapter I'll commence that exploration. There is a lot of raw material in this connection for future research.

Tradition as "the democracy of the dead"

Remember that our arguments for epistemic democracy saw that as competent and independent voters are added, majority votes become more and more reliable. Now consider the following view on tradition given by G. K. Chesterton:

> Tradition is only democracy extended through time. . . . Tradition may be defined as an extension of the franchise. Tradition means giving votes to the most obscure of all classes, our ancestors. It is the democracy of the dead. Tradition refuses to submit to the small and arrogant oligarchy of those who merely happen to be walking about. . . . Democracy asks us not to neglect a good man's opinion, even if he is our groom [as in stable hand]; tradition asks us not to neglect a good man's opinion, even if he is our father.[1]

Here Chesterton is talking about a whole bunch of justifications for democracy. But it is the epistemic justification for democracy that concerns us in this text.

DOI: 10.4324/9781003355274-28

If we know what people in the past thought, why not find a way to aggregate their views with those of people alive today? Of course, one response is simply to reject the arguments for epistemic democracy considered earlier in the book. Then the argument from democracy to tradition flounders immediately.

Another response might be this: we already *do* include the views of people in the past in our own. In fact, this was part of a debunking argument we considered some chapters ago. Much of our worldview comes from our surroundings, experiences, and education in early life, and much of that is determined by those who raise and teach us, whose worldview in turn came from those before them, and so on. The anti-traditionalist might say: there is no need to add *extra* consideration for the political beliefs of the past into our deliberations or votes, because they affect us so much already.

A third response might be this: in the past, people were not sufficiently competent and independent to do more good than harm when added to our voting pool. This is possible, of course, but I think it's a response to be careful about. Regarding independence, although people in the past might not have been that independent from each other, they will have a lot of differences from people in the present, meaning that including past perspectives could add independence to our group overall. Regarding both independence and competence, there is a question of just what might have changed to make *us* so much more reliable than people in the past. This judgment itself might be liable to be debunked; of course we think the present is so reliable – it's where *we* are! Finally, there is a worry that this kind of reply limits the scope of our support for democracy. Of course we can be trusted with it, but those uncivilized people in the past, they couldn't be trusted with it.

Robert E. Goodin and Kai Spiekermann mention a different kind of worry about respecting tradition. They worry that it can lead to failures of independence, and to belief cascades in particular – what they call "few minds, many mimics."[2] In the ideal situation for epistemic democracy, independence is ensured by secret, anonymous ballots. But including tradition in our epistemic democracy, as the "democracy of the dead," is the opposite of secret and anonymous. In that situation, past votes help determine future ones. And, Goodin and Spiekermann say, there may be cases of "double-counting," as our predecessors may themselves have taken tradition into account.[3] This is the "Burkean paradox":[4] that reliance on tradition now makes the tradition we establish for the future weaker. But note that this paradox is just another version of the dilemma of democracy. The dilemma arises whenever a process reliably aggregates views to which we can either contribute or defer.

Chesterton's fence and Hayek's spontaneous order

Chesterton's fence is another principle from G. K. Chesterton about respecting past ideas. We will have to do some work to interpret it as a principle of political

epistemology in particular, but it is worth reading his original formulation, since it is quite lovely.

> In the matter of reforming things, as distinct from deforming them, there is one plain and simple principle; a principle which will probably be called a paradox. There exists in such a case a certain institution or law; let us say for the sake of simplicity, a fence or gate erected across a road. The more modern type of reformer goes gaily up to it and says, "I don't see the use of this; let us clear it away."
>
> To which the more intelligent type of reformer will do well to answer:
>
> "If you don't see the use of it, I certainly won't let you clear it away. Go away and think. Then, when you can come back and tell me that you do see the use of it, I may allow you to destroy it."
>
> This paradox rests on the most elementary common sense.
>
> The gate or fence did not grow there. It was not set up by somnambulists who built it in their sleep. It is highly improbable that it was put there by escaped lunatics who were for some reason loose in the street. Some person had some reason for thinking it would be a good thing for somebody. And until we know what the reason was, we really cannot judge whether the reason was reasonable.
>
> It is extremely probable that we have overlooked some whole aspect of the question, if something set up by human beings like ourselves seems to be entirely meaningless and mysterious.
>
> There are reformers who get over this difficulty by assuming that all their fathers were fools; but if that be so, we can only say that folly appears to be a hereditary disease.
>
> But the truth is that nobody has any business to destroy a social institution until he has really seen it as an historical institution.
>
> If he knows how it arose, and what purposes it was supposed to serve, he may really be able to say that they were bad purposes, or that they have since become bad purposes, or that they are purposes which are no longer served. But if he simply stares at the thing as a senseless monstrosity that has somehow sprung up in his path, it is he and not the traditionalist who is suffering from an illusion.
>
> We might even say that he is seeing things in a nightmare.[5]

Now what was Chesterton saying here? Though we may have all sorts of advantages over the past based on advances made in the intervening time (scientific advances, moral advances, what have you), people from the past have a certain advantage over us, which is that they (according to Chesterton, at least) likely understand why they did the things they did. This might not seem like much, but of course everything that exists now was made in the past – so it is just a question of how far back we can trace many of our laws and social structures.

The further we go into the past, the greater the extent to which past people are *experts* on those laws and social structures relative to us. They understand why they're there and we don't. We should defer to them to the extent we can on their purposes, but we can retain our independent judgment regarding those purposes themselves.

Chesterton can be both compared to and contrasted with Hayek, who gave a very different kind of argument for traditionalism. Recall from the chapter on decentralization that Hayek thought everyone had only partial knowledge, which market mechanisms could aggregate, thus beating out any individual expert or "central planner." Hayek thought social institutions, the "fences" Chesterton was interested in, also emerged from market mechanisms, developing a "theory of socio-cultural evolution [a]s a generalization of his theory of spontaneous market order."[6] Just as prices developed in markets, large-scale social structures emerge for Hayek from the small choices of tons of people with only partial information. Thus, for a Hayekian, the source of respect for tradition is not that we must look to the past and understand why people established the social structures they established; it is that even *they* do not understand quite how those social structures came about and that we should expect the mechanisms of spontaneous order, like the market, to do better than we *or* they can on our own. Now, those social structures which have stuck around likely serve a good function; this function being likely a kind of mystery to us, given our only partial knowledge of the spontaneous and aggregative order of society, we should, as Chesterton thought, be very hesitant about tearing long-lasting social structures down. But both Chesterton and Hayek are essentially making arguments in political epistemology; they are arguing that the people Chesterton calls "reformers" are generally not in a position to justify their political beliefs, because they don't understand the social structures those beliefs are about.

Of course, there is an extent to which respect for tradition is the norm. Laws don't have expiration dates – although we do sometimes repeal them. Buildings don't erode – although we do sometimes knock them down. We might think that the past plays enough of a role without us taking it into account in determining our political beliefs. Regardless, it's now time to think about the other side of the coin: arguments about progress.

Progress

Here are two pieces of data of differing strengths. First, we contemporary humans – "contemporary" meaning "of late 2022," the time of this writing – take ourselves to have made moral progress over past humans: we take our civilizations to be better, we take our actions to do more good and less wrong, and we take ourselves to know more moral facts. Horrors of the past are regularly used to motivate the idea that there even is such a thing as right and wrong.

Second, but more weakly, we anticipate the continuation of such progress. In particular, we will sometimes make comments like the following (for some action φ): "In the future, φing will be seen as horrific, barbaric, evil, etc.," or even "In the future, φing will be seen as horrific, barbaric, evil, etc., just as slavery, the Holocaust, etc. are seen as evil today." Each of these types of future judgments is, I think, typically offered as though they have at least some evidentiary value when it comes to whether or not φing is right or wrong. That is, they are offered as though a future judgment against φing gives us a reason to believe that it is wrong to φ. Of course, the value of future judgments is defeasible; their evidentiary value could be outweighed by other evidence. Reasons-talk allows this: a reason can be just one among many. But the reason must be rather strong as well. If the future judgment is of the comparative type, for instance, we must take it that the judgment of the future counts for more than the judgment of the past and likely of the present. So when we make future judgments, we take future people to be something like experts relative to present people.

How is the second piece of data related to the first one? First, it's common to take the present to be a midpoint in some sort of story about moral progress. For instance, Peter Singer's book *The Expanding Circle* develops the idea in its title: human society has gradually become concerned with the welfare, rights, or other morally relevant properties of more and more kinds of entities.[7] And Michael Huemer discusses a "pervasive trend toward liberalization of values over human history."[8] So, in this sense, the second piece of data might sort of continue on from the first piece of data.

Moreover, we might need to believe in moral progress to avoid a debunking argument regarding the first piece of data. After all, our confidence in our own time's moral values, as against those of the past, could easily be explained away the same way that one society's preference for their own values over another society's values could be explained away: as a kind of chauvinism.[9] In Whit Stillman's 1990 film *Metropolitan*, two characters have the following exchange:

Tom: Nearly everything Jane Austen wrote is ridiculous from today's perspective.

Audrey: Has it ever occurred to you that today looked at from Jane Austen's perspective would look even worse?[10]

To stick to our thesis regarding moral progress, the first piece of data, we need a way of characterizing the difference between the past and the present as more than merely a difference of "perspective." So the idea that the future will be better than the present, our opinions about which are represented by the second data point, is needed to save from a debunking challenge the idea that the present is better than the past, our opinions about which are represented by the first data point.

We are certainly now experts relative to the past on quite a few things. Mathematics, the sciences, engineering: these have obviously progressed, and there are things with which even laypeople have some passing familiarity now that the greatest scientists of, say, the Enlightenment would have had little conception, like quantum physics, relativity theory, many aspects of computer science, and so on.[11] Is morality like these things? Recall from the chapter on expertise that moral expertise may work a bit differently than other kinds of expertise. In particular, when it comes to identifying moral experts, we seem to have a harder time checking candidates' judgment without recourse to our own prior values. Is there any reason to think that, if we found ourselves in the future, we would think it much better than the past? Our values are at a certain distance from past values; future values will be at a certain distance from our values. For moral realists, it's very likely that one set of values is *more accurate* than the others; but why think it will be the future's? Why not ours?

The idea of moral progress can be thought of like a graph, with time on the *x*-axis and moral knowledge on the *y*-axis. Perhaps there are some moments where the line slants down, but in the long haul it just goes up and up. People rarely consider the possibility of a different view, what we might call a *parabolic* view of history, in which things, after a time of rising action, have peaked, are peaking, or will peak, and this peak is followed by falling action. Given the fact that we have to rely on our own moral beliefs to judge moral experts, it's not clear to me why it's so irrational for an epistemic agent to believe that the peak is roughly *now*, which is likely when their values are the most widely shared.

Why think that progress won't continue? First, there are many interpretations of history. Some will dispute whether or not we really do have a widening moral circle, whether we really are seeing constant liberalization of values, etc. Second, it seems to me that the idea of "perspective" found in the *Metropolitan* exchange can be extended to whole stories about progress. Just as we could tell a whole story of progress which involved many intermediate steps and decision points and a kind of line or lines connecting all of them, the person from the past could tell a whole story of decline which involved similar steps, points, and lines. Moreover, patterns can continue longer than we might like. Consider a meal of pasta with no sauce. The dish gets better and better as you add sauce, but only up to a point. Eventually, there is enough sauce, and adding more makes the dish worse; you can *go too far* in this regard. So it might be if the historical patterns that have fueled moral progress continue. Say, for instance, that determining a certain policy involves balancing concerns of safety with concerns of liberty. One of these values may grow stronger to the point of the ideal balance, after which further growth would be detrimental to the balance.

Another reply is what I call the "many roads objection." It is difficult to predict how the future will view our present practices, which suggests a lot of contingency in that regard; there is no reason to think that the future will arrive on the

one correct "road." A few years ago, I encountered a video on Facebook showing people's heart-wrenching final moments with their beloved but painfully ill pets before a veterinarian sedates and euthanizes those pets. We can imagine at least three future attitudes toward such a video. First, values might not change, and the video might remain heart-wrenching but morally unproblematic. Second, values might change such that euthanizing a pet comes to be seen as wrong, as, for example, a kind of rights violation. Third, values might change such that becoming too emotionally involved with an animal comes to be seen as wrong, as, for example, a way of irresponsibly ignoring the far more important suffering of other living humans. The many roads objection insists that we cannot really know which of these eventualities will actually come to pass. But if in general we cannot know in what direction social norms will evolve, then it seems that by the same token we cannot in general know that future social norms will be an improvement on present social norms, by our own lights. If we really would be able to identify future people as moral experts, then we should be able to know which of the possible "roads" is the one a moral expert would take.

Sometimes it's claimed that moral progress works through "the accumulation of new empirical knowledge."[12] Charles Barzun, in explicating the jurisprudence of former Supreme Court Justice David Souter, gives the example of "[s]ocial attitudes about homosexuality hav[ing] changed dramatically," which he thinks is explained by the fact that "as more and more people have realized that their brothers and sisters and sons and daughters are gay, the more their views about the nature of homosexuality have changed."[13] Similarly, we might think that part of the explanation for Singer's notion of an "expanding circle" comes merely from societal awareness that a wider variety of lifestyles and social conditions do not lead to cataclysmic social upheaval. Such empirical knowledge, however, is necessarily contingent, and in any event has a stopping point; so it's not clear that we should be confident that moral progress will largely continue based on this sort of explanation of it.

Just how do values change? Do they do so rationally? As part of a recent book, Will MacAskill has argued that the abolition of slavery, for instance, was the result of changing moral attitudes, though not the result of inevitable moral progress.[14] But moral attitudes can change for a lot of reasons. A 2019 article in *The Economist* holds that "generational replacement is what shifts opinion."[15] In other words, as an older generation with one set of values dies out, a newer generation with another set of values takes control. And it could be that apparent moral progress is actually the result of technological and scientific progress, because such progress obviates the sorts of threats that in earlier times might have motivated people to act immorally, for instance, by going to war over scarce resources or turning away visitors who might carry unknown pathogens. But if what looks like moral progress does not involve actually convincing anyone that the new view is better by the use of reason, why think that reason favors

the new view at all? And there may be debunkable incentives that lead people to take on new moral perspectives as though they're fashions. We've discussed moral grandstanding and "virtue signaling" a few times already;[16] these result in part from people's desire to be seen as especially morally good. If what looks like moral progress is actually driven by signaling behavior, then we might think of it more like a series of changes in moral fashion than as a serious, consistent improvement in moral beliefs.

Parallelism; how "isms" change their faces

One especially interesting potential mechanism for (what might look like) moral progress is through denials of moral progress. We saw in an earlier chapter that stigmatization can occur through "concept creep," as the definitions of negative concepts seem to expand when fewer and fewer real-world instances match them. Steven Pinker has proposed that there is a "negativity bias" which leads us to see more, and focus more on, negative news, giving the impression that things are getting worse or that things aren't getting better as fast as they actually are.[17] Gregory Mitchell and Philip E. Tetlock have found in psychological studies "a pervasive tendency . . . to see things as getting worse than they really are," and moreover they see this bias as "almost the opposite" of the status quo bias considered earlier in the text.[18]

Examples of concept creep and negativity bias abound. One common umbrella phrase for these things is "changing face" – "the changing face of racism," "the changing face of fascism," and so on. What does it mean for an "ism" to "change its face"? "Isms" aren't agents in history; Cass Sunstein has warned of "the difficulty of demonstrating, and the potential recklessness of claiming, that one or another 'ism' is causally associated with concrete social developments."[19] If "isms" are theories, then they may be individuated by their propositional content. But I think in cases like "racism" and "fascism" (though not the "liberalism" Sunstein is investigating), it makes more sense to think of "isms" as being habits of mind or psychological tics or biases. The "changing face" claim is best interpreted as a claim that some psychological tendency, which would previously have eventuated in now-obviously evil acts, is now eventuating in acts that don't seem as evil but in fact are (evil because, I suppose, they emerge from the same psychological tendency). A lot of philosophical work is required to explain just what an "ism" is such that it could feature in this sort of way in this sort of discourse.

Sometimes people make arguments based on parallels between historical events and present-day politics. They say things like: "X is the modern-day equivalent of Y." One example I've heard: "Supporting Black Lives Matter is the modern-day equivalent of opposing slavery." Another example I've heard:

"Opposing abortion is the modern-day equivalent of opposing slavery." In evaluating arguments like these, it's important to try to figure out just what is taken to be parallel about the situations. In what sense is the *moral* or *political* wrong of the past situation similar to the *moral* or *political* wrong of the present situation? In college, for instance, I worked for an organization that tried to raise awareness about ethnic cleansing in Darfur. Activists around that issue often compared it to the Holocaust, saying: "Never again." There the parallel was meant to be clear: people were killed based on their race or ethnicity in Nazi Germany, and the same thing was happening again in Darfur. The rhetorical force of these arguments, I think, comes from the fact that we are much more certain about what was evil in the past than we are about what is evil now and that we all like to think of ourselves as people who might have done the right thing in the past. These arguments based on historical parallels try to leverage moral certainty about the past to combat moral uncertainty about the present.

Slippery slopes and genetic arguments

As far as I know, philosophers have said little about the relationship between two informal fallacies, the *slippery slope fallacy* and the *genetic fallacy*. Slippery slope arguments say: If you accept X now, it's just a slippery slope to accepting Y in the future; so you shouldn't accept X. They can also be run past-to-present rather than present-to-future: X occurred, and it led to Y, which is bad; so we shouldn't have accepted X. In this form, we might also think of such arguments as "ideas have consequences" arguments.[20] Genetic arguments are genealogical debunking arguments; they say: Y came from X, and X was bad; so you shouldn't accept Y. I think these arguments are mirror images of each other. One says that the cause was bad and therefore we shouldn't like the effect. The other says that the effect was bad, and therefore we shouldn't like the cause. Though these arguments commit fallacies, they're fallacies of the type that in social epistemology we're often okay with, because they rely on non-deductive heuristics which we must evaluate substantively. Unsurprisingly, I find genetic arguments, which judge a purportedly obviously bad past as reflecting poorly on the present or future, to be more popular among progressives (both political progressives and social-epistemological progressives) while I find slippery slope and ideas-have-consequences arguments, which judge a purportedly bad present or future as reflecting poorly on the past, to be more popular among conservatives and traditionalists (again, of both stripes). Though genetic arguments and genealogies interest philosophers, not much general has been said about ideas-have-consequences arguments, which to my mind is an oversight likely caused by political imbalance in the discipline. There is definitely work to be done about the form of these arguments and about what, if anything, they can teach us.

Political exegesis

Another way history relates to political epistemology is through what I call *political exegesis*. This is the interpretation of political events and their relation to various trends and principles. What is an event "about"? During and after the 2016 election, for instance, philosopher Justin Tiehen compiled a Twitter thread of something like a thousand proposed causes of Trump's prominence and victory. The exegesis had great importance for what came after: Were Trump voters racist, or were they angry about an economy that had left them behind? Was the cause of Trump's victory a stifling left-wing political correctness, or was it the enduring appeal of white supremacy? Both the language of "aboutness" in regard to events and the difficulty of puzzling apart various contributory factors are proper objects of reflection on the part of philosophers of political epistemology.

Political exegesis is related to history because our claim that some event is "about" something likely represents a reflection on what came before that event or a prediction about what will come after. Exegesis in this sense is a matter of putting political events in the proper context and subsuming them under some larger, grander kind of explanation. This kind of historical theory-building can affect our political beliefs in ways that might be rational and ways that might be irrational. If we already know all the details of an event, and those details have eventuated in a judgment about that event, it's not always clear how contextualizing that event should change that judgment. On the other hand, thinking of how an event fits into a trend can sometimes point up potential consequences of that event. Political exegesis is also related to the phenomenon of "isms" changing their faces, discussed earlier. A seemingly innocuous event might through parallelism gain virtuous or vicious associations with past events, and it is political exegesis which assigns those associations.

Living documents

One final place where moral progress and the philosophy of history are important is in constitutional interpretation and the philosophy of law. We have already seen that Justice Souter linked his jurisprudence to a theory of moral progress; Chief Justice Earl Warren wrote likewise about "the evolving standards of decency that mark the progress of a maturing society,"[21] and in the recent decision legalizing gay marriage across the United States, Justice Anthony Kennedy wrote for the majority that "new dimensions of freedom [have] become apparent to new generations."[22] Similarly, legal scholar David Strauss, in the introduction to his book *The Living Constitution*, writes that a constitution that does not adapt to changing social mores would end up being "a relic that would keep us from making progress,"[23] and legal scholar Michael Dorf writes that he and others like him "believe that human history is, if not exactly a steady march of

moral progress, at least a process in which, over the long run, society's norms change for the better."[24] On the other side, famous originalist Justice Antonin Scalia expressed skepticism that the passage of time "mark[s] progress . . . as opposed to rot,"[25] and Justice Samuel Alito has asked: "Is it true that our society is inexorably evolving in the direction of greater and greater decency? Who says so, and how did this particular philosophy of history find its way into our fundamental law?"[26]

The doctrine of living constitutionalism has it that social mores now can and should affect our interpretation of laws written in the past; that what the law is now has something to do with social mores now. It makes sense for living constitutionalists to believe in moral progress; little else could explain why their approach is a good one, and in particular if they didn't believe in moral progress, they would be committed to overturning rightly held decisions, an odd stance. As we saw earlier, this is something that comes up frequently in the writings of jurists and legal scholars, but the apparent reliance of jurisprudence on the philosophy of history has not led to a resurgence of interest in the latter field. I think this relationship, and the relationship between political epistemology and the philosophy of history, amply demonstrate the potential intellectual fecundity of such a resurgence.

Conclusion

Conservatives and traditionalists; progressives; even parabolic theorists – there are many views on history, and it seems like they might bear on our political epistemology, as well as on our politics. The past may be a resource for added "minds" that improve our thinking in one way or another or a hindrance that shackles us to old ways of thinking and of doing politics. The future may see the full fruition of our moral values, the continued upward movement of moral progress and improvement in our moral and political beliefs, or it may see a fall in which delicate balances are thrown out of whack and new atrocities, now unthinkable, dwarf the old. Whatever the case may be, there were people at some points in the past, and there will be people at least some points in the future, and so social and political epistemology, informed by the philosophy of history, have something to say about how we should deal with those people's beliefs, just as they have something to say about how we should deal with the beliefs of those living today.

Discussion questions

1. Do you think people's moral beliefs will be better in the future? Why or why not?
2. What are some examples of present practices that might be viewed as morally abhorrent in the future? What are some examples of present practices that

might have been viewed as morally abhorrent in the past? Which judgment do you care about more? Why?

3. Do you think there are things we fail to understand now that were understood in the past? Why did this understanding seem to slip away?
4. Why do you think moral values change? Do you think people are convinced over time? Do you think the side of good tends to win?

CONCLUSION

Our study of political beliefs has taken us to many places. We've seen the perspectives of philosophers, psychologists, political scientists, political theorists, economists, sociologists, historians, pundits, and others. And we've seen how much disagreement there is, not just about politics but about political beliefs: how they're formed, whether they're rational, what we ought to do about it. Some of these disagreements are political disagreements recast as epistemology or psychology, but some of them reflect the genuine difficulty of the subject matter and the complexity of the topic. If you've made it this far, good work. (I'm saying that to myself as much as to you.)

What should you take away from this book? What should you do? There are some lessons we've seen repeated throughout this book. One is that people's motivations for developing their political beliefs are often rather obscure. It's often inappropriate to question such motives within a political dispute rather than responding to the content of someone's speech, but in this "meta" context, I think the politically involved reader should ask themselves just what they get out of it. And don't fall back on slogans like "it's good to be aware of what's going on." Why is it good? Things are going on everywhere, some of them political, some not. Why is it good to be aware of the political things? A related lesson is that engaging with politics, and even with political information, does not necessarily make one more knowledgeable in every important sense. We saw that political information may be associated with high partisanship and with ideological capture. Very well-informed readers should ask themselves whether their sustained engagement with politics is best described as an upwards epistemic trend, where they continue learning new things and improving their understanding of political

DOI: 10.4324/9781003355274-29

reality, or a kind of spiral, where they end up more and more certain of the sorts of political beliefs they started out with.

The most important lesson, I think, is of the value of independent thought. It's rarely rational – rarely intellectually responsible – to mold one's inner life, one's beliefs and credences, to fit the demands of an ever-changing party platform or some bespoke *au courant* ideology. (You'd expect philosophers of all people to understand this, but plenty of philosophers work hard doing just that. At least they tend to be better at it than other people!) Even when thinking for yourself doesn't work out that well for you, it works out well for the group. If you must think about politics, try to avoid simplifying frames, stock narratives, and language and concepts that serve a political purpose. I suppose what I want to say is this: be political as little as you can; either avoid politics, or do politics in the least political way you can manage.

At the end of the day, I'm one of those irritating people who believes in objectivity, neutrality, rationality, and free speech, who thinks we should conciliate when presented with opposing views, who doesn't want to take anyone else's word for anything, and who doesn't want to be put in a box with a neat and tidy ideological label. People like me often get a lot of things wrong, and we look silly doing it. Heck, we look silly when we're right, too. I don't have a solution for the problem of looking silly, but I prefer it to looking like I came off of some intellectual assembly line.

I went back to graduate school in philosophy in large part to study political epistemology. Back when I did, I agreed with many of the people we've encountered in this book that we were in a unique epistemic crisis in American politics, one of bubbles and echo chambers and conspiracy theories and fake news and the competition between various media talking heads who couldn't evaluate evidence if their lives depended on it. But now I'm not so sure. First, I'm not so sure things now are so different from what they've always been. Technology changes, but people, I think, stay the same. Second, I'm not so sure that the political epistemology I see everywhere isn't just politics by another name. Accusing others of being in thrall to ideologies and biases and so forth is a political strategy like any other. Third, I'm not so sure that individual irrationality is what matters. I still get incensed when I see someone making some bonkers claim about politics, especially one that I know they take for granted because of their social milieu, as anyone who's read my book reviews knows. But sometimes I think that I'm looking at the wrong thing and that what matters is how it all comes together on the societal level. And if I'm being honest, on the societal level, I'm no longer sure that the problem of political epistemology is the biggest political problem we have. I wish I could figure out what is.

BIBLIOGRAPHY

Abramowitz, Alan I. and Walter J. Stone (2006). "The Bush Effect: Polarization, Turnout, and Activism in the 2004 Presidential Election." *Presidential Studies Quarterly* 36/2: 141–154.

Abramowitz, Alan I. and Steven Webster (2017). " 'Negative Partisanship' Explains Everything." *Politico*. www.politico.com/magazine/story/2017/09/05/negative-partisanship-explains-everything-215534/.

Adorno, Theodor W., Else Frenkel-Brunswik, Daniel J. Levinson, and R. Nevitt Sanford (1950). *The Authoritarian Personality*. Harper.

Allen, Amy (2016). "Feminist Perspectives on Power." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/entries/feminist-power/.

Alston, William (1988). *Epistemic Justification: Essays in the Theory of Knowledge*. Cornell University Press.

Alston, William (2005). *Beyond "Justification": Dimensions of Epistemic Evaluation*. Cornell University Press.

Althusser, Louis (1970). "Idéologie et appareils idéologiques d'État." *La Pensée* 151.

Anderau, Glenn (2021). "Defining Fake News." *KRITERION* 35/3: 197–215.

Anderson, Luvall and Michael Barnes (2022). "Hate Speech." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/entries/hate-speech.

Arguedas, Amy Ross, Craig T. Robertson, Richard Fletcher, and Rasmus K. Nielsen (2022). "Echo Chambers, Filter Bubbles, and Polarisation: A Literature Review." *Reuters Institute for the Study of Journalism*. https://royalsociety.org/-/media/policy/projects/online-information-environment/oie-echo-chambers.pdf.

Ashton, Natalie Alana and Robin McKenna (2020). "Situating Feminist Epistemology." *Episteme* 17/1: 28–47.

Audi, Robert (2020). *Seeing, Knowing, and Doing*. Oxford University Press.

Audi, Robert and Nicholas Wolterstorff (1997). *Religion in the Public Sphere: The Place of Religious Convictions in Political Debate*. Rowman and Littlefield.

Aumann, Robert J. (1976). "Agreeing to Disagree." *Annals of Statistics* 4/6: 1236–1239.

Austin, J. L. (1962). *How to Do Things with Words*. Harvard University Press.

Aytac, Ugur and Enzo Rossi (forthcoming). "Ideology Critique without Morality: A Radical Realist Approach." *American Political Science Review*.

Ballantyne, Nathan (2012). "The Problem of Historical Variability." In *Disagreement and Skepticism*, Diego Machuca, ed. Routledge: 239–259.
Ballantyne, Nathan (2019). *Knowing Our Limits*. Oxford.
Ballantyne, Nathan and Peter H. Ditto (2021). "Hanlon's Razor." *Midwest Studies in Philosophy* 45: 309–331.
Ballarini, Cristina and Steven Sloman (2017). "Reasons and the 'Motivated Numeracy Effect'." *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*: 1580–1585.
Barzun, Charles (2018). "Justice Souter's Common Law." *Virginia Law Review* 104: 655–726.
Basu, Rima (2018). "Can Beliefs Wrong?" *Philosophical Topics* 46: 1–17.
Basu, Rima (2019). "What We Epistemically Owe to Each Other." *Philosophical Studies* 176: 915–931.
Basu, Rima (2021). "A Tale of Two Doctrines: Moral Encroachment and Doxastic Wronging." In *Applied Epistemology*, Jennifer Lackey, ed. Oxford University Press: 99–118.
Basu, Rima and Mark Schroeder (2019). "Doxastic Wronging." In *Pragmatic Encroachment in Epistemology*, Brian Kim and Matthew McGrath, eds. Routledge.
Beaver, David, Bart Geurts, and Kristie Denlinger (2021). "Presupposition." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/entries/presupposition/.
Beaver, David and Jason Stanley (2021). "Neutrality." *Philosophical Topics* 49/1: 165–185.
Bedard, Paul (2020). "90% of Media Political Donations to Biden, Sanders, AOC, Democrats: Report." *Washington Examiner*. www.washingtonexaminer.com/washington-secrets/90-of-media-political-donations-to-biden-sanders-aoc-democrats-report.
Begby, Endre (2021a). "Evidential Preemption." *Philosophy and Phenomenological Research* 102/3: 515–530.
Begby, Endre (2021b). *Prejudice: A Study in Non-Ideal Epistemology*. Oxford University Press.
Bennett, Jonathan (1990). "Why is Belief Involuntary?" *Analysis* 50: 87–107.
Berker, Selim (2011). "Class Notes." https://scholar.harvard.edu/files/sberker/files/phil244-meeting2-williams.pdf.
Berker, Selim (2018). "A Combinatorial Argument Against Practical Reasons for Belief." *Analytic Philosophy* 59: 427–470.
Bishop, Bill (2008). *The Big Sort: Why the Clustering of Like-Minded America is Tearing Us Apart*. Houghton Mifflin.
Boldry, Jennifer G., Lowell Gaertner, and Jeff Quinn (2007). "Measuring the Measures: A Meta-Analytic Investigation of the Measures of Outgroup Homogeneity." *Group Processes and Intergroup Relations* 10: 157–178.
Bolinger, Renee Jorgensen (2018). "The Rational Impermissibility of Accepting (Some) Racial Generalizations." *Synthese* 197/6: 2415–2431.
Boutyline, Andrei and Stephen Vaisey (2017). "Belief Network Analysis: A Relational Approach to Understanding the Structure of Attitudes." *American Journal of Sociology* 122/5: 1371–1447.
Bramson, Aaron, Patrick Grim, Daniel J. Singer, William J. Berger, Graham Sack, Steven Fisher, Carissa Flocken, and Bennett Holman (2017). "Understanding Polarization: Meanings, Measures, and Model Evaluation." *Philosophy of Science* 84/1: 115–159.
Brennan, Geoffrey and Loren Lomasky (1997). *Democracy and Decision: The Pure Theory of Electoral Preference*. Cambridge University Press.
Brennan, Jason (2016). *Against Democracy*. Princeton University Press.
Brennan, Jason (2017). "Propaganda about Propaganda." *Critical Review* 29/1: 34–48.
Brennan, Jason (2020). "The Ethics and Rationality of Voting." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/entries/voting/#RatiVoti.
Bright, Liam Kofi (forthcoming). "White Psychodrama." *Journal of Political Philosophy*.

Brownstein, Michael (2019). "Implicit Bias." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/entries/implicit-bias.
Bruenig, Matt (2013). "What Does Identitarian Deference Require?" https://mattbruenig.com/2013/02/26/what-does-identitarian-deference-require/.
Buckwalter, Wesley (2020). "Theoretical Motivations of 'Ought Implies Can'." *Philosophia* 48: 83–94.
Callahan, Laura Frances (2018). "Moral Testimony: A Re-Conceived Understanding Explanation." *The Philosophical Quarterly* 68/272: 437–459.
Caplan, Bryan (2007). *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*. Princeton University Press.
Caplan, Bryan (2011). "Ideological Turing Tests." www.econlib.org/archives/2011/06/the_ideological.html.
Cappelen, Herman (2018). *Fixing Language*. Oxford University Press.
Carr, Jennifer Rose (forthcoming). "Why Ideal Epistemology?" *Mind*.
Carter, J. Adam and Robin McKenna (2020). "Skepticism Motivated: On the Skeptical Import of Motivated Reasoning." *Canadian Journal of Philosophy* 50/6: 702–718.
Caruso, Gregg (2018). "Skepticism about Moral Responsibility." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/entries/skepticism-moral-responsibility/.
Case, Spencer (2019). "The Boy Who Inflated the Concept of 'Wolf'." *Quillette*. https://quillette.com/2019/02/14/the-boy-who-inflated-the-concept-of-wolf/.
Cassam, Quassim (2019). *Conspiracy Theories*. Polity.
Chalmers, David (2011). "Verbal Disputes." *Philosophical Review* 120/4: 515–566.
Chen, Angela (2017). "Is A Liberal Bias Hurting Social Psychology?" *Pacific Standard*. https://psmag.com/social-justice/is-a-liberal-bias-hurting-social-psychology.
Chesterton, Gilbert Keith (1908). *Orthodoxy*. John Lane Company.
Chesterton, Gilbert Keith (1929). "The Thing." https://archive.org/stream/G.K.ChestertonTheThing/G.K.Chesterton-The%20Thing_djvu.txt.
Christensen, David (2007). "Epistemology of Disagreement: The Good News." *Philosophical Review* 116/2: 187–217.
Coady, David (2003). "Conspiracy Theories and Official Stories." *International Journal of Applied Philosophy* 17/2: 197–209.
Coady, David (2007). "Are Conspiracy Theorists Irrational?" *Episteme* 4/2: 193–204.
Coady, David (2021). "The Fake News about Fake News." In *The Epistemology of Fake News*, Sven Bernecker, Amy K. Flowerree, and Thomas Grundmann, eds. Oxford: 68–81.
Cohen, G. A. (2000). *If You're an Egalitarian, How Come You're So Rich?* Harvard.
Contessa, Gabriele (forthcoming). "Shopping for Experts." *Synthese*.
Converse, Philip E. (1964). "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, David E. Apter, ed. Free Press: 206–261.
Copp, David (2008). " 'Ought' Implies 'Can' and the Derivation of the Principle of Alternative Possibilities." *Analysis* 68/1: 67–75.
Cottam, Martha L., Elena Mastors, Thomas Preston, and Beth Dietz (2016). *Introduction to Political Psychology* (3rd ed.). Routledge.
Danaher, John (2016). "Understanding Hayek's Knowledge Argument." https://philosophicaldisquisitions.blogspot.com/2016/11/understanding-hayeks-knowledge-argument.html.
Daniels, Norman (2016). "Reflective Equilibrium." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/entries/reflective-equilibrium/.
D'Arms, Justin and Daniel Jacobson (2000). "The Moralistic Fallacy: On the 'Appropriateness' of Emotions." *Philosophy and Phenomenological Research* 61: 65–90.
Davidson, Donald (1963). "Actions, Reasons, and Causes." *Journal of Philosophy* 60: 685–700.
Davidson, Donald (1974). *Inquiries into Truth and Interpretation*. Clarendon.

Dentith, M. R. X. (2016). "When Inferring to a Conspiracy Might Be the Best Explanation." *Social Epistemology* 30/5–6: 572–591.
Dentith, M. R. X. and Brian L. Keeley (2019). "The Applied Epistemology of Conspiracy Theories: An Overview." In *The Routledge Handbook of Applied Epistemology*, David Coady, ed. Routledge.
Dentith, M. R. X. and Martin Orr (2017). "Secrecy and Conspiracy." *Episteme* 15/4: 433–450.
DeRose, Keith (1992). "Contextualism and Knowledge Attributions." *Philosophy & Phenomenological Research* 52: 913–929.
Donahue, T. J. (2007). "What is Politics?" https://papers.ssrn.com/sol3/papers.cfm?abstract_id=963511.
Dorf, Michael (2012). "The Undead Constitution." *Harvard Law Review* 125: 2011–2055.
Dorst, Kevin (2020a). "Reasonably Polarized: Why Politics is More Rational than You Think." *Stranger Apologies*. www.kevindorst.com/stranger_apologies/rp.
Dorst, Kevin (2020b). "Confirmation Bias is Rational." *Psychology Today*. www.psychologytoday.com/us/blog/reasonably-polarized/202010/confirmation-bias-is-rational.
Dorst, Kevin (forthcoming). "Rational Polarization." *Philosophical Review*.
Driver, Julia (2005). "Moralism." *Journal of Applied Philosophy* 22: 137–151.
Dror, Lidal (2023). "Is There an Epistemic Advantage to Being Oppressed?" *Nous* 57: 618–640.
Duarte, José L., Jarret T. Crawford, Charlotta Stern, Jonathan Haidt, Lee Jussim, and Philip E. Tetlock (2015). "Political Diversity Will Improve Social Psychological Science." *Behavioral and Brain Sciences* 38.
Dworkin, Ronald (1996). "Objectivity and Truth: You'd Better Believe It." *Philosophy and Public Affairs* 25/2: 87–139.
Eagly, Alice H. and Antonio Mladinic (1994). "Are People Prejudiced Against Women? Some Answers from Research on Attitudes, Gender Stereotypes, and Judgments of Competence." *European Review of Social Psychology* 5: 1–35.
The Economist (2019). "Societies Change Their Minds Faster than People Do." www.economist.com/graphic-detail/2019/10/31/societies-change-their-minds-faster-than-people-do.
Elga, Adam (2007). "Reflection and Disagreement." *Nous* 41/3: 478–502.
Elga, Adam (2010). "How to Disagree about How to Disagree." In *Disagreement*, Richard Feldman and Ted Warfield, eds. Oxford University Press.
Enoch, David (2014). "A Defense of Moral Deference." *Journal of Philosophy* 111: 229–258.
Falbo, Arianna (forthcoming). "Inquiring Minds Want to Improve." *Australasian Journal of Philosophy*.
Fallis, Don and Kay Mathiesen (forthcoming). "Fake News is Counterfeit News." *Inquiry*.
Fantl, Jeremy and Matthew McGrath (2009). *Knowledge in an Uncertain World*. Oxford University Press.
Feldman, Richard (1988). "Epistemic Obligations." *Philosophical Perspectives* 2: 235–256.
Feldman, Richard (2001). "Voluntary Belief and Epistemic Evaluation." In *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, Matthias Steup, ed. Oxford University Press: 77–89.
Fish, Stanley (1994). *There's No Such Thing as Free Speech – And It's a Good Thing, Too*. Oxford University Press.
Fiss, Owen (1991). "State Activism and State Censorship." *Yale Law Journal* 100/7: 2087–2106.
Flanigan, Jessica (forthcoming). "Credibility and the Standpoint Expectation." *Georgetown Journal of Law and Public Policy*.

Foucault, Michel (1976). *Histoire de la sexualité 1: La volonté de savoir*. Éditions Gallimard.
Frances, Bryan and Jonathan Matheson (2018). "Disagreement." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/archives/win2019/entries/disagreement/.
Frankfurt, Harry (1969). "Alternate Possibilities and Moral Responsibility." *The Journal of Philosophy* 66/23: 829–839.
Friedman, Jeffrey (2006). "Democratic Competence in Normative and Positive Theory: Neglected Implications of 'The Nature of Belief Systems in Mass Publics'." *Critical Review* 18/1–3: i–xliii.
Friedman, Jeffrey (2017). "Trump Voters and Economic Grievances: It's the Media, Stupid." *Niskanen Center*. www.niskanencenter.org/trump-voters-economic-grievances-media-stupid/.
Friedman, Jeffrey (2019). *Power Without Knowledge: A Critique of Technocracy*. Oxford University Press.
Fritz, James (2017). "Pragmatic Encroachment and Moral Encroachment." *Pacific Philosophical Quarterly* 98: 643–661.
Fritz, James (2020). "Moral Encroachment and Reasons of the Wrong Kind." *Philosophical Studies* 177: 3051–3070.
Fritz, James and Elizabeth Jackson (2021). "Belief, Credence, and Moral Encroachment." *Synthese* 199/1–2: 1387–1408.
Gardiner, Georgi (2014). "The Commutativity of Evidence: A Problem for Conciliatory Views of Peer Disagreement." *Episteme* 11/1: 83–95.
Gedeon, Péter (2015). "Spontaneous Order and Social Norms: Hayek's Theory of Socio-Cultural Evolution." *Society and Economy* 37/1: 1–29.
Gerber, Alan S., Gregory A. Huber, David Doherty, and Conor M. Dowling (2012). "The Big Five Personality Traits in the Political Arena." *Annual Review of Political Science* 14: 265–287.
Gibbons, Adam (2022). "Is Epistocracy Irrational?" *Journal of Ethics and Social Philosophy* 21/2: 251–281.
Gibbons, Adam (forthcoming). "Rational Conceptual Conflict and the Implementation Problem." *Inquiry*.
Gigerenzer, Gerd (2008). "Why Heuristics Work." *Perspectives in Psychological Science* 3/1: 20–29.
Gigerenzer, Gerd (2018). "The Bias in Behavioral Economics." *Review of Behavioral Economics* 5/3–4: 303–336.
Gilovich, Thomas and Dale Griffin (2002). "Introduction: Heuristics and Biases: Then and Now." In *Heuristics and Biases: The Psychology of Intuitive Judgment*, Thomas Gilovich, Dale Griffin, and Daniel Kahneman, eds. Cambridge University Press: 1–18.
Ginet, Carl (2001). *Knowledge, Truth, and Duty*. Oxford University Press.
Glüer, Kathrin (2011). *Donald Davidson: A Short Introduction*. Oxford University Press.
Goldberg, Zach (2019). "America's White Saviors." *Tablet*.
Goldman, Alvin (1979). "What is Justified Belief?" In *Justification and Knowledge: New Studies in Epistemology*, George S. Pappas, ed. D. Reidel Publishing Company: 1–25.
Goldman, Alvin (2001). "Experts: Which Ones Should You Trust?" *Philosophy and Phenomenological Research* 63/1: 85–110.
Goldman, Alvin and Bob Beddor (2021). "Reliabilist Epistemology." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/archives/sum2021/entries/reliabilism/.
Goldman, Alvin and James C. Cox (1996). "Speech, Truth, and the Free Market for Ideas." *Legal Theory* 2: 1–32.

Goodin, Robert E. (2006). "The Epistemic Benefits of Multiple Biased Observers." *Episteme* 3/3: 166–174.
Goodin, Robert E. and Kai Spiekermann (2018). *An Epistemic Theory of Democracy*. Oxford. www.tabletmag.com/sections/news/articles/americas-white-saviors.
Graham, Jesse, Jonathan Haidt, and Brian A. Nosek (2009). "Liberals and Conservatives Rely on Different Sets of Moral Foundations." *Journal of Personality and Social Psychology* 96/5: 1029–1046.
Graham, Peter (2000a). "Transferring Knowledge." *Nous* 34/1: 131–152.
Graham, Peter (2000b). "Conveying Information." *Synthese* 123/3: 365–392.
Green, Mitchell (2020). "Speech Acts." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/entries/speech-acts.
Habgood-Coote, Joshua (2019). "Stop Talking About Fake News!" *Inquiry* 62/9–10: 1033–1065.
Haidt, Jonathan (2001). "The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment." *Psychological Review* 108/4: 814–834.
Haidt, Jonathan (2012). *The Righteous Mind: Why Good People are Divided by Politics and Religion*. Pantheon.
Haidt, Jonathan, Fredrick Björklund, and Scott Murphy (2000). "Moral Dumbfounding: When Intuition Finds No Reason." https://polpsy.ca/wp-content/uploads/2019/05/haidt.bjorklund.pdf.
Hannon, Michael (2021). "Disagreement or Badmouthing? The Role of Expressive Discourse in Politics." In *Political Epistemology*, Elizabeth Edenberg and Michael Hannon, eds. Oxford: 297–318.
Hannon, Michael (2022). "Are Knowledgeable Voters Better Voters?" *Politics, Philosophy & Economics* 21/1: 29–54.
Hanson, Robin (2009). "Political Signaling Theories." *Overcoming Bias*. www.overcomingbias.com/2009/05/political-signaling-theories.html.
Hanson, Robin (2013). "Shall We Vote on Values, but Bet on Beliefs?" *Journal of Political Philosophy* 21/2: 151–178.
Harris, Keith Raymond (2022). "Some Problems with Particularism." *Synthese* 200/6: 447.
Harwood, Jake (2020). "Social Identity Theory." *International Encyclopedia of Media Psychology*, Jan Van den Bulck, ed. https://onlinelibrary.wiley.com/doi/book/10.1002/9781119011071.
Haslam, Nick (2016). "Concept Creep: Psychology's Expanding Concepts of Harm and Pathology." *Psychological Inquiry* 27/1: 1–17.
Haslanger, Sally (2000). "Gender and Race: (What) are They? (What) Do We Want Them to Be?" *Nous* 34/1: 31–55.
Haslanger, Sally (2012). *Resisting Reality*. Oxford University Press.
Haslanger, Sally (2017a). *Critical Theory and Practice*. Koninklijke Van Gorcum.
Haslanger, Sally (2017b). "Culture and Critique." *Aristotelian Society Supplemental Volume* 91: 149–173.
Hawthorne, John and Jason Stanley (2008). "Knowledge and Action." *Journal of Philosophy* 105: 571–590.
Hayek, Friedrich (1945). "The Use of Knowledge in Society." *The American Economic Review* 35/4: 519–530.
Henne, Paul, Jennifer Semler, Vladimir Chituc, Felipe De Brigard, and Walter Sinnott-Armstrong (2019). "Against Some Recent Arguments for 'Ought' Implies 'Can'." *Philosophia* 47: 131–139.
Hieronymi, Pamela (2005). "The Wrong Kind of Reason." *Journal of Philosophy* 102: 437–457.
Hirsch, Eli (2005). "Physical-Object Ontology, Verbal Disputes, and Common Sense." *Philosophy and Phenomenological Research* 70/1: 67–97.

Hirschman, Albert O. (1991). *The Rhetoric of Reaction: Perversity, Futility, Jeopardy*. Belknap.
Huang, Haifeng (2018). "The Pathology of Hard Propaganda." *Journal of Politics* 80: 1034–1038.
Huemer, Michael (2016). "A Liberal Realist Answer to Debunking Skeptics: The Empirical Case for Realism." *Philosophical Studies* 173/7: 1983–2010.
Hughes, Nick (2019). "Dilemmic Epistemology." *Synthese* 196/10: 4059–4090.
Hyska, Megan (2023). "Against Irrationalism in the Theory of Propaganda." *Journal of the American Philosophical Association* 9/2: 303–317.
Iyer, Ravi, Spassena Koleva, Jesse Graham, Peter Ditto, and Jonathan Haidt (2012). "Understanding Libertarian Morality: The Psychological Dispositions of Self-Identified Libertarians." *PLoS One* 7/8: e42366.
Jackson, Elizabeth G. (2020). "The Relationship Between Belief and Credence." *Philosophy Compass* 15/6.
Jacobs, Tom (2016). "Donald Trump's Appeal to the Authoritarian Personality." *Pacific Standard*. https://psmag.com/news/trumps-appeal-to-the-authoritarian-personality.
Jacobson, Daniel (2012). "Moral Dumbfounding and Moral Stupefaction." *Oxford Studies in Normative Ethics* 2: 289–316.
James, William (1890). *The Principles of Psychology: Volume 1*. Henry Holt.
Jaster, Romy and David Lanius (2021). "Speaking of Fake News." In *The Epistemology of Fake News*, Sven Bernecker, Amy K. Flowerree, and Thomas Grundmann, eds. Harvard University Press: 19–45.
Joshi, Hrishikesh (2020). "What Are the Chances You're Right About Everything? An Epistemic Challenge for Modern Partisanship." *Politics, Philosophy & Economics* 19/1: 36–61.
Joshi, Hrishikesh (2021). *Why It's OK to Speak Your Mind*. Routledge.
Joshi, Hrishikesh (forthcoming a). "The Epistemic Significance of Social Pressure." *Canadian Journal of Philosophy*.
Joshi, Hrishikesh (forthcoming b). "Debunking Creedal Beliefs." *Synthese*.
Jost, John T. (2017). "A Theory of System Justification." *Psychological Science Agenda*. www.apa.org/science/about/psa/2017/06/system-justification.
Jost, John T. (2020). *A Theory of System Justification*. Harvard University Press.
Joyce, James M. (2007). "Epistemic Deference: The Case of Chance." *Proceedings of the Aristotelian Society* 107: 187–206.
Judd, Charles M. and Bernadette Park (1988). "Out-Group Homogeneity: Judgments of Variability at the Individual and Group Levels." *Journal of Personality and Social Psychology* 54: 778–788.
Jussim, Lee (2012). *Social Perception and Social Reality: Why Accuracy Dominates Bias and Self-Fulfilling Prophecy*. Harvard University Press.
Kahan, Dan M. (2013). "Ideology, Motivated Reasoning, and Cognitive Reflection." *Judgment and Decision Making* 8/4: 407–424.
Kahan, Dan M. (2017a). "Misconceptions, Misinformation, and the Logic of Identity-Protective Cognition." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2973067.
Kahan, Dan M. (2017b). "The Expressive Rationality of Inaccurate Perceptions." *Behavioral and Brain Sciences* 40.
Kahan, Dan M. and Ellen Peters (2017). "Rumors of the 'Nonreplication' of the 'Motivated Numeracy Effect' are Greatly Exaggerated." *Yale Law & Economics Research Paper* 584.
Kahan, Dan M., Ellen Peters, Erica Cantrell Dawson, and Paul Slovic (2017). "Motivated Numeracy and Enlightened Self-Government." *Behavioral Public Policy* 1/1: 54–86.

Kanai, Ryota, Tom Feilden, Colin Firth, and Geraint Rees (2011). "Political Orientations are Correlated with Brain Structure in Young Adults." *Current Biology* 21/8: 677–680.

Kaufman, Scott Barry (2016). "The Personality of Political Correctness." *Scientific American*. https://blogs.scientificamerican.com/beautiful-minds/the-personality-of-political-correctness/.

Keeley, Brian L. (1999). "Of Conspiracy Theories." *Journal of Philosophy* 96/3: 109–126.

Kelly, Thomas (2005). "The Epistemic Significance of Disagreement." *Oxford Studies in Epistemology* 1: 167–196.

Kelly, Thomas (2008). "Disagreement, Dogmatism, and Belief Polarization." *Journal of Philosophy* 105/10: 611–633.

Kelly, Thomas (2010). "Peer Disagreement and Higher-Order Evidence." In *Disagreement*, Richard Feldman and Ted A. Warfield, eds. Oxford University Press: 111–174.

Kelly, Thomas (2022). *Bias: A Philosophical Study*. Oxford University Press.

Klein, Ezra (2020). *Why We're Polarized*. Avid Reader Press.

Kogelmann, Brian (2023). "Finding the Epistocrats." *Episteme* 20/2: 497–512.

Kopko, Kyle C., Sarah McKinnon Bryner, Jeffrey Budziak, Christopher J. Devine, and Steven P. Nawara (2011). "In the Eye of the Beholder? Motivated Reasoning in Disputed Elections." *Political Behavior* 33: 271–290.

Kugler, Matthew, John T. Jost, and Sharareh Noorbaloochi (2014). "Another Look at Moral Foundations Theory: Do Authoritarianism and Social Dominance Orientation Explain Liberal-Conservative Differences in 'Moral' Intuitions?" *Social Justice Research* 27/4: 413–431.

Kukla, Quill R. (2021). "Situated Knowledge, Purity, and Moral Panic." In *Applied Epistemology*, Jennifer Lackey, ed. Oxford University Press: 37–66.

Kuran, Timur (1995). *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Harvard University Press.

Kvalnes, Øyvind and Salvör Nordal (2019). "Normalization of Questionable Behavior: An Ethical Root of the Financial Crisis in Ireland." *Journal of Business Ethics* 159: 761–775.

Lackey, Jennifer (2008). *Learning from Words: Testimony as a Source of Knowledge*. Oxford University Press.

Lackey, Jennifer (2010). "A Justificationist View of Disagreement's Epistemic Significance." In *Social Epistemology*, Adrian Haddock, Alan Millar, and Duncan Pritchard, eds. Oxford University Press: 298–325.

Lackey, Jennifer (2013). "Disagreement and Belief Dependence: Why Numbers Matter." In *The Epistemology of Disagreement: New Essays*, David Christensen and Jennifer Lackey, eds. Oxford University Press: 243–268.

Lackey, Jennifer (2018). "Experts and Peer Disagreement." In *Knowledge, Belief, and God: New Insights in Religious Epistemology*, Matthew A. Benton, John Hawthorne, and Dani Rabinowitz, eds. Oxford University Press: 228–245.

Lakens, Daniel (2017). "Impossibly Hungry Judges." *The 20% Statistician*. https://daniellakens.blogspot.com/2017/07/impossibly-hungry-judges.html.

Landemore, Hélène (2013). *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton University Press.

Laudan, Larry (1981). "A Confutation of Convergent Realism." *Philosophy of Science* 48/1: 19–49.

Leader Maynard, Jonathan and Alex Worsnip (2018). "Is There a Distinctly Political Normativity?" *Ethics* 128: 756–787.

Leftwich, Adrian, ed. (2004). *What is Politics?* Polity Press.

Leiter, Brian and Samuel Leiter (2015). "Not Your Grandfather's Propaganda." *The New Rambler*. https://newramblerreview.com/book-reviews/philosophy/not-your-grandfather-s-philosophy.

Levari, David E., Daniel T. Gilbert, Timothy D. Wilson, Beau Sievers, David M. Amodio, and Thalia Wheatley (2018). "Prevalence-Induced Concept Change in Human Judgment." *Science* 360: 1465–1467.
Levy, Neil (2005). "The Good, the Bad, and the Blameworthy." *Journal of Ethics and Social Philosophy* 1/2: 1–16.
Levy, Neil (2007). "Radically Socialized Knowledge and Conspiracy Theories." *Episteme* 4/2: 181–192.
Levy, Neil (2021a). "Virtue Signalling is Virtuous." *Synthese* 198: 9545–9562.
Levy, Neil (2021b). "Is Myside Bias Irrational? A Biased Review of *The Bias That Divides Us*." *Social Epistemology Review and Reply Collective* 10: 31–38.
Levy, Neil (2022). "Do Your Own Research!" *Synthese* 200/5: 1–19.
Lewis, Hyrum and Verlan Lewis (2022). *The Myth of Left and Right*. Oxford University Press.
Liao, Shen-Yi and Nat Hansen (2023). "'Extremely Racist' and 'Incredibly Sexist': An Empirical Response to the Charge of Conceptual Inflation." *Journal of the American Philosophical Association* 9/1: 72–94.
Lieberman, Hallie and Eric Schatzberg (2018). "A Failure of Academic Quality Control: *The Technology of Orgasm*." *Journal of Positive Sexuality* 4/2: 24–47.
Linville, Patricia W., Gregory W. Fischer, and Peter Salovey (1989). "Perceived Distributions of the Characteristics of In-Group and Out-Group Members: Empirical Evidence and a Computer Simulation." *Journal of Personality and Social Psychology* 57: 165–188.
Lippmann, Walter (1922). *Public Opinion*. Macmillan Publishers.
Littlejohn, Clayton (2014). "The Unity of Reason." In *Epistemic Norms: New Essays on Action, Belief, and Assertion*, Clayton Littlejohn and John Turri, eds. Oxford University Press: 135–152.
Lopez, German (2017). "Trump Won Because of Racial Resentment." *Vox*. www.vox.com/identities/2017/12/15/16781222/trump-racism-economic-anxiety-study.
Loury, Glenn (1994). "Self-Censorship in Public Discourse." *Rationality and Society* 6/4: 428–461.
Lukács, G. (1971). *History and Class Consciousness: Studies in Marxist Dialectics*. MIT Press.
Lukianoff, Greg and Jonathan Haidt (2018). *The Coddling of the American Mind*. Penguin.
MacAskill, William (2022). *What We Owe the Future*. Basic Books.
Machery, Edouard (2022). "Anomalies in Implicit Attitudes Research." *WIREs Cognitive Science* 13/1.
MacWilliams, Matthew (2016). "The One Weird Trait That Predicts Whether You're a Trump Supporter." *Politico*. www.politico.com/magazine/story/2016/01/donald-trump-2016-authoritarian-213533/.
Malka, Ariel, Yphtach Lelkes, and Nissan Holzer (2017). "Rethinking the Rigidity of the Right Model." In *The Politics of Social Psychology*, Jarret T. Crawford and Lee Jussim, eds. Taylor & Francis: 116–135.
Mandik, Pete (2007). "Shit Happens." *Episteme* 4/2: 205–218.
Manne, Kate (2016). "Humanism: A Critique." *Social Theory and Practice* 42/2: 389–415.
Manne, Kate (2017). *Down Girl: The Logic of Misogyny*. Oxford University Press.
Mansbridge, Jane, James Bohman, Simone Chambers, David Estlund, Andreas Føllesdal, Archon Fung, Cristina Lafont, Bernard Manin, and José Luis Martí (2010). "The Place of Self-Interest and the Role of Power in Deliberative Democracy." *Journal of Political Philosophy* 18/1: 64–100.
Marcus, Julia and Gregg Gonsalves (2020). "Public-Health Experts are Not Hypocrites." *The Atlantic*. www.theatlantic.com/ideas/archive/2020/06/public-health-experts-are-not-hypocrites/612853/.

Matheson, Jonathan (2015). "Disagreement and Epistemic Peers." *The Oxford Handbook of Topics in Philosophy*. https://doi.org/10.1093/oxfordhb/9780199935314.013.13.
Mattson, Kevin (2018). "The Trumpian Personality." *Dissent*. www.dissentmagazine.org/article/revisiting-frankfurt-school-authoritarian-personality-trump.
Mayo-Wilson, Conor, Kevin J. S. Zollman, and David Danks (2011). "The Independence Thesis: When Individual and Social Epistemology Diverge." *Philosophy of Science* 78/4: 653–677.
McBrayer, Justin P. (2020). *Beyond Fake News: Finding the Truth in a World of Misinformation*. Routledge.
McGrath, Sarah (2009). "The Puzzle of Pure Moral Deference." *Philosophical Perspectives* 23: 321–344.
Mercier, Hugo (2020). *Not Born Yesterday*. Princeton University Press.
Messick, David M. and Keith P. Sentis (1979). "Fairness and Preference." *Journal of Experimental Social Psychology* 15/4: 418–434.
Meyer, Robinson and Ashley Fetters (2018). "Victorian-Era Orgasms and the Crisis of Peer Review." *The Atlantic*. www.theatlantic.com/health/archive/2018/09/victorian-vibrators-orgasms-doctors/569446/.
Milgrom, Paul and Nancy Stokey (1982). "Information, Trade, and Common Knowledge." *Journal of Economic Theory* 26/1: 17–27.
Mill, John Stuart (1859). "On Liberty." www.econlib.org/library/Mill/mlLbty.html.
Mills, Charles W. (2017). *Black Rights/White Wrongs: The Critique of Racial Liberalism*. Oxford University Press.
Mills, Michael (director) (2010). *Beginners* (film).
Milton, John (1644). "Areopagitica." https://milton.host.dartmouth.edu/reading_room/areopagitica/text.html.
Mitchell, Gregory and Philip E. Tetlock (forthcoming). "Are Progressives in Denial about Progress? Yes, but So is Almost Everyone Else." *Clinical Psychological Science*.
Miton, Helena, Nicolas Claidière, and Hugo Mercier (2015). "Universal Cognitive Mechanisms Explain the Cultural Success of Bloodletting." *Evolution and Human Behavior* 36/4: 303–312.
Moss, Jordan and Peter O'Connor (2020). "The Dark Triad Traits Predict Authoritarian Political Correctness and Alt-Right Attitudes." *Heliyon* 6/7.
Nagel, Thomas (1979). *Mortal Questions*. Cambridge University Press.
Nebel, Jacob M. (2015). "Status Quo Bias, Rationality, and Conservatism about Value." *Ethics* 125: 449–476.
Nehamas, Alexander (1998). *The Art of Living*. University of California Press.
Nelkin, Dana K. (2019). "Moral Luck." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/entries/moral-luck/.
Nguyen, C. T. (2018). "Escape the Echo Chamber." *Aeon*. https://aeon.co/essays/why-its-as-hard-to-escape-an-echo-chamber-as-it-is-to-flee-a-cult.
Nguyen, C. T. (2020). "Echo Chambers and Epistemic Bubbles." *Episteme* 17/2: 141–161.
Nyhan, Brendan (2021). "Why The Backfire Effect Does Not Explain the Durability of Political Misperceptions." *Proceedings of the National Academy of Sciences* 118/15: e1912440117.
Nyhan, Brendan and Jason Reifler (2010). "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior* 32: 303–330.
Ornstein, Norm (2014). "Yes, Polarization is Asymmetric – and Conservatives are Worse." *The Atlantic*. www.theatlantic.com/politics/archive/2014/06/yes-polarization-is-asymmetric-and-conservatives-are-worse/373044/.
Over, Harriet (2021). "Seven Challenges for the Dehumanization Hypothesis." *Perspectives on Psychological Science* 16/1: 3–13.

Parfit, Derek (2011). *On What Matters: Volume 1*. Oxford University Press.
Pariser, Eli (2011). *The Filter Bubble: What The Internet is Hiding from You*. Penguin Press.
Park, Bernadette and Myron Rothbart (1982). "Perception of Out-Group Homogeneity and Levels of Social Categorization." *Journal of Personality and Social Psychology* 42/6: 1051–1068.
Pedersen, Walker S., L. Tugan Muftuler, and Christine L. Larson (2018). "Conservatism and the Neural Circuitry of Threat: Economic Conservatism Predicts Greater Amygdala-BNST Connectivity During Periods of Threat vs. Safety." *Social Cognitive and Affective Neuroscience* 13/1: 43–51.
Pepp, Jessica, Eliot Michaelson, and Rachel Katharine Sterken (2019). "What's New about Fake News?" *Journal of Ethics and Social Philosophy* 16/2: 67–94.
Persson, Emil, David Anderson, Lina Koppel, Daniel Västfjäll, and Gustav Tinghög (2021). "A Preregistered Replication of Motivated Numeracy." *Cognition* 214.
Pigden, Charles (2007). "Conspiracy Theories and the Conventional Wisdom." *Episteme* 4/2: 219–232.
Piller, Christian (2006). "Content-Related and Attitude-Related Reasons for Preferences." *Royal Institute of Philosophy Supplements* 59: 155–181.
Pinker, Steven (1994). "The Game of the Name." *The New York Times*. https://stevenpinker.com/files/pinker/files/1994_04_03_newyorktimes.pdf.
Pinker, Steven (2019). "Correct for the Media's Negativity Bias." *Politico*. http://www.politico.com/interactives/2019/how-to-fix-politics-in-america/misinformation/correct-for-the-medias-negativity-bias/.
Plunkett, David and Timothy Sundell (2013). "Disagreement and the Semantics of Normative and Evaluative Terms." *Philosophers' Imprint* 13/23: 1–37.
Podgorski, Abelard (2017). "Rational Delay." *Philosophers' Imprint* 17.
Popper, Karl (1950/2013). *The Open Society and its Enemies*. Princeton University Press.
Popper, Karl (1963). *Conjectures and Refutations*. Routledge.
Pratto, F., J. Sidanius, L. M. Stallworth, and B. F. Malle (1994). "Social Dominance Orientation: A Personality Variable Predicting Social and Political Attitudes." *Journal of Personality and Social Psychology* 67: 741–763.
Pritchard, Duncan (2021). "Good News, Bad News, Fake News." In *The Epistemology of Fake News*, Sven Bernecker, Amy K. Flowerree, and Thomas Grundmann, eds. Oxford University Press: 46–67.
Putnam, Hilary (2004). *The Collapse of the Fact/Value Dichotomy and Other Essays*. Harvard University Press.
Quattrone, George A. and Edward E. Jones (1980). "The Perception of Variability Within In-Groups and Out-Groups: Implications for the Law of Small Numbers." *Journal of Personality and Social Psychology* 38: 141–152.
Rawls, John (1971). *A Theory of Justice*. Harvard University Press.
Rawls, John (1993). *Political Liberalism*. Columbia University Press.
Raz, Joseph (1988). *The Morality of Freedom*. Clarendon Press.
Raz, Joseph (2009). "Reasons: Explanatory and Normative." In *New Essays on the Explanation of Action*, Constantine Sandis, ed. Palgrave: 184–202.
Resnick, Brian (2019). "9 Essential Lessons from Psychology to Understand the Trump Era." *Vox*. www.vox.com/science-and-health/2018/4/11/16897062/political-psychology-trump-explain-studies-research-science-motivated-reasoning-bias-fake-news.
Restall, Greg (2015). "Arguments Matter, Even If They Come Down to 'Semantics'." *The Conversation*. https://theconversation.com/arguments-matter-even-if-they-come-down-to-semantics-43188.

Rini, Regina (2017). "Fake News and Partisan Epistemology." *Kennedy Institute of Ethics* 27: E43–E64.
Ritchie, Stuart (2020). *Science Fictions*. Metropolitan Books.
Robertson, Derek (2018). "How an Obscure Conservative Theory Became the Trump Era's Go-to Nerd Phrase." *Politico*. www.politico.com/magazine/story/2018/02/25/overton-window-explained-definition-meaning-217010/.
Robin, Corey (2011). *The Reactionary Mind: Conservatism from Edmund Burke to Sarah Palin*. Oxford University Press.
Robinson, Nathan J. (2017). "Politics is a Contest of Domination." *Current Affairs*. www.currentaffairs.org/2017/07/politics-is-a-contest-of-domination.
Roeber, Blake (2018). "The Pragmatic Encroachment Debate." *Nous* 52: 171–195.
Roeber, Blake (2020). "How to Argue for Pragmatic Encroachment." *Synthese* 197: 2649–2664.
Roeber, Blake (manuscript). *Political Humility*.
Rudy-Hiller, Fernando (2018). "The Epistemic Condition for Moral Responsibility." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/entries/moral-responsibility-epistemic/.
Sampson, Eric (manuscript). "Partisan Political Belief Clusters are Not Irrational."
Samuelson, William and Richard Zeckhauser (1988). "Status Quo Bias in Decision Making." *Journal of Risk and Uncertainty* 1: 7–59.
Sandel, Michael (1994). "Political Liberalism." *Harvard Law Review* 107/7: 1765–1794.
Sanders, Linley (2022). "Trust in Media 2022: Where Americans Get Their News and Who They Trust for Information." *YouGovAmerica*. https://today.yougov.com/topics/politics/articles-reports/2022/04/05/trust-media-2022-where-americans-get-news-poll.
Sankaran, Kirun (2020). "What's New in the New Ideology Critique." *Philosophical Studies* 177: 1441–1462.
Sanneh, Kelefa (2019, August 12). "The Fight to Redefine Racism." *The New Yorker*.
Sauer, Hanno (2011). "Social Intuitionism and the Psychology of Moral Reasoning." *Philosophy Compass* 6/10: 708–721.
Sauer, Hanno (2018). *Debunking Arguments in Ethics*. Cambridge University Press.
Sauer, Hanno (2020). "The Reason Wars are Over. Reason Won." https://medium.com/science-and-philosophy/the-reason-wars-are-over-reason-won-89c1f4d648a9.
Saul, Jennifer (2018). "Dogwhistles, Political Manipulation, and Philosophy of Language." In *New Work on Speech Acts*, Daniel W. Harris, Daniel Fogal, and Matt Moss, eds. Oxford University Press: 360–383.
Scalia, Antonin (1997). *A Matter of Interpretation*. Princeton University Press.
Scarantino, Andrea and Ronald de Sousa (2018). "Emotion." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/entries/emotion/.
Schimmack, Ulrich (2021). "Prediction Markets of Replicability." *Replicability-Index*. https://replicationindex.com/2021/05/10/osc-pm/.
Schmitt, Carl (1932). *The Concept of the Political*. Duncker & Humblot.
Schroeder, Mark (2010). "Value and the Right Kind of Reason." *Oxford Studies in Metaethics* 5.
Schroeder, Mark (2018). "When Beliefs Wrong." *Philosophical Topics* 46: 115–127.
Schwarz, Jon (2007). "Democrats and the Iron Law of Institutions." www.tinyrevolution.com/mt/archives/001705.html.
Schwitzgebel, Eric (2019). "Belief." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/entries/belief/.
Shah, Nishi (2002). "Clearing Space for Doxastic Voluntarism." *The Monist* 85/3: 436–445.
Shelby, Tommie (2003). "Ideology, Racism, and Critical Social Theory." *Philosophical Forum* 2: 153–188.

Singal, Jesse (2016). "Why It Took Social Science Years to Correct a Simple Error About 'Psychoticism'." *The Cut*. www.thecut.com/2016/07/why-it-took-social-science-years-to-correct-a-simple-error-about-psychoticism.html.
Singal, Jesse (2017). "Psychology's Favorite Tool for Measuring Racism Isn't Up to the Job." *The Cut*. www.thecut.com/2017/01/psychologys-racism-measuring-tool-isnt-up-to-the-job.html.
Singer, Peter (1981). *The Expanding Circle: Ethics and Sociobiology*. Farrar, Straus, & Giroux.
Skitka, Linda J. and Anthony N. Washburn (2016). "Are Conservatives from Mars and Liberals from Venus? Maybe Not So Much." In *Social Psychology of Political Polarization*, P. Valdesolo and J. Graham, eds. Routledge: 78–101.
Smith, David Livingstone (2020). *On Inhumanity: Dehumanization and How to Resist It*. Oxford.
Somin, Ilya (2016). *Democracy and Political Ignorance* (2nd ed.). Stanford University Press.
Sotola, Lukas K. and Marcus Credé (2022). "On the Predicted Replicability of Two Decades of Experimental Research on System Justification: A Z-Curve Analysis." *European Journal of Social Psychology* 52/5–6: 895–909.
Spaulding, Shannon (2021). "Beliefs and Biases." *Synthese* 199/3–4: 7575–7594.
Sperber, Dan and Hugo Mercier (2017). *The Enigma of Reason*. Harvard University Press.
Srinivasan, Amia (2015). "The Archimedean Urge." *Philosophical Perspectives* 29/1: 325–362.
Srinivasan, Amia (2017). "Genealogy, Epistemology, and Worldmaking." *Proceedings of the Aristotelian Society* 119/2: 127–156.
Srinivasan, Amia (2020). "Radical Externalism." *Philosophical Review* 129/3: 395–431.
Stancil, Will (2023). "Tweet." https://twitter.com/whstancil/status/1612182474892738560.
Stanley, Jason (2005). *Knowledge and Practical Interests*. Oxford University Press.
Stanley, Jason (2011). "The Ways of Silencing." *New York Times*. https://archive.nytimes.com/opinionator.blogs.nytimes.com/2011/06/25/the-ways-of-silencing/.
Stanley, Jason (2015). *How Propaganda Works*. Princeton University Press.
Stanovich, Keith (2021). *The Bias That Divides Us: The Science and Politics of Myside Thinking*. MIT Press.
Stillman, Whit (director) (1990). *Metropolitan* (film).
Strauss, David A. (2010). *The Living Constitution*. Oxford University Press.
Strawson, Galen (2010). "The Depth(s) of the Twentieth Century." *Analysis* 70/4: 607.
Strawson, P. F. (1963). "Carnap's Views on Conceptual Systems Versus Natural Languages in Analytic Philosophy." In *The Philosophy of Rudolf Carnap*, Paul Arthur Schilpp, ed. Open Court: 503–518.
Strawson, P. F. (1974). *Freedom and Resentment and Other Essays*. Routledge.
Street, Sharon (2006). "A Darwinian Dilemma for Realist Theories of Value." *Philosophical Studies* 127/1: 109–166.
Streumer, Bart (2007). "Reasons and Impossibility." *Philosophical Studies* 136: 351–384.
Sunstein, Cass (1993). *Democracy and the Problem of Free Speech*. The Free Press.
Sunstein, Cass (2001). *Republic.com*. Princeton University Press.
Sunstein, Cass (2002). "The Law of Group Polarization." *Journal of Political Philosophy* 10/2: 175–195.
Sunstein, Cass (2009). *Going to Extremes: How Like Minds Unite and Divide*. Oxford University Press.
Sunstein, Cass (2019). "Ismism, or Has Liberalism Ruined Everything?" *Harvard Law Public Working Paper No. 19-19*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3372364.

Swire-Thompson, Briony, Joseph DeGutis, and David Lazer (2020). "Searching for the Backfire Effect: Measurement and Design Considerations." *Journal of Applied Research in Memory and Cognition* 9/3: 286–299.
Táíwò, Olúfẹ́mi O. (2018). "The Empire Has No Clothes." *Disputatio* 10/51: 305–330.
Táíwò, Olúfẹ́mi O. (2020). "Being-in-the-Room Privilege: Elite Capture and Epistemic Deference." *The Philosopher*. www.thephilosopher1923.org/post/being-in-the-room-privilege-elite-capture-and-epistemic-deference.
Tajfel, Henri (1970). "Experiments in Intergroup Discrimination." *Scientific American* 223/5: 96–103.
Talisse, Robert (2019). *Overdoing Democracy*. Oxford University Press.
Talisse, Robert (2021). *Sustaining Democracy*. Oxford University Press.
Tappin, B. M., G. Pennycock, and D. G. Rand (2021). "Rethinking the Link Between Cognitive Sophistication and Politically Motivated Reasoning." *Journal of Experimental Psychology: General* 150/6: 1095–1114.
Taylor, Shelley E., Susan T. Fiske, Nancy L. Etcoff, and Audrey J. Ruderman (1978). "Categorical and Contextual Bases of Person Memory and Stereotyping." *Journal of Personality and Social Psychology* 36: 778–793.
Tetlock, Philip (2003). "Thinking the Unthinkable: Sacred Values and Taboo Cognitions." *Trends in Cognitive Science* 7/7: 320–324.
Tetlock, Philip (2005). *Expert Political Judgment*. Princeton University Press.
Tetlock, Philip and Dan Gardner (2015). *Superforecasting: The Art and Science of Prediction*. Crown Press.
Tiehen, Justin (forthcoming). "The IKEA Effect & the Production of Epistemic Goods." *Philosophical Studies*.
Tilton, Emily (forthcoming). "'That's Above My Paygrade': Woke Excuses for Ignorance." *Philosophers' Imprint*.
Tirrell, Lynne (2013). "Studying Genocide: A Pragmatist Approach to Action-Engendering Discourse." In *Pragmatism, Law, and Language*, Graham Hubbs and Douglas Lind, eds. Routledge.
Tosi, Justin and Brandon Warmke (2016). "Moral Grandstanding." *Philosophy and Public Affairs* 44/3: 197–217.
Tosi, Justin and Brandon Warmke (2020). *Grandstanding: The Use and Abuse of Moral Talk*. Oxford University Press.
Traldi, Oliver (2021). "The Dehumanization Debate." *The Hedgehog Review*. https://hedgehogreview.com/issues/who-do-we-think-we-are/articles/the-dehumanization-debate.
Traldi, Oliver (2022). "The Mismeasure of Ideology." *The Bulwark*. www.thebulwark.com/the-mismeasure-of-ideology/.
Traldi, Oliver (2023). "Uncoordinated Norms of Belief." *Australasian Journal of Philosophy* 101/3: 625–637.
Traldi, Oliver (manuscript a). "Conceptual Engineering and Reasons for Belief."
Traldi, Oliver (manuscript b). "Conciliating Toward Expertise."
Uscinski, Joseph (2019). "Is This the Year of the Conspiracy Theory?" *Arc Digital*. https://medium.com/arc-digital/is-this-the-year-of-the-conspiracy-theory-7d875f7a2e9e.
Uscinski, Joseph, Adam Enders, Casey Klofstad, Michelle Seelig, Hugo Drochon, Kamal Premaratne, and Manohar Murthi (2022). "Have Beliefs in Conspiracy Theories Increased Over Time?" *PLoS One* 17/7: e0270429.
Vallier, Kevin (2020). *Trust in a Polarized Age*. Oxford University Press.
Vallier, Kevin (2022). "Defending Neutrality." *Liberal Currents*. www.liberalcurrents.com/defending-neutrality/.
van Roojen, Mark (2018). "Moral Cognitivism vs. Non-Cognitivism." In *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, ed. https://plato.stanford.edu/entries/moral-cognitivism/.

Vavova, Katia (2018). "Irrelevant Influences." *Philosophy and Phenomenological Research* 96/1: 134–152.
Vayrynen, Pekka (2013). *The Lewd, the Rude and the Nasty: A Study of Thick Concepts in Ethics*. Oxford University Press.
Vermeule, Adrian (2008). *Law and the Limits of Reason*. Oxford University Press.
Vranas, Peter B. M. (2007). "I Ought, Therefore I Can." *Philosophical Studies* 136: 167–216.
Wallace, David Foster (2009). *This is Water*. Little, Brown and Company.
Weaver, Richard M. (1948). *Ideas Have Consequences*. University of Chicago.
Whelan, Frederick G. (1981). "Language and its Abuses in Hobbes' Political Philosophy." *American Political Science Review* 75/1: 59–75.
White, Roger (2010). "You Just Believe That Because. . ." *Philosophical Perspectives* 24: 573–615.
Widerker, David (1991). "Frankfurt on 'Ought Implies Can' and Alternative Possibilities." *Analysis* 51/4: 222–224.
Williams, Bernard (1973). "Deciding to Believe." In *Problems of the Self*. Cambridge: 136–151.
Williams, Dan (2021). "Socially Adaptive Belief." *Mind and Language* 36/3: 333–354.
Wittgenstein, Ludwig (1953). *Philosophical Investigations*. Blackwell Press.
Wood, Thomas and Ethan Porter (2019). "The Elusive Backfire Effect: Mass Attitudes' Steadfast Factual Adherence." *Political Behavior* 41: 135–163.
Worsnip, Alex (2019). "The Obligation to Diversify One's Sources: Against Epistemic Partisanship in the Consumption of News Media." In *Media Ethics: Free Speech and the Requirements of Democracy*, Carl Fox and Joe Saunders, eds. Routledge: 240–264.
Worsnip, Alex (2021). "Can Pragmatists Be Moderate?" *Philosophy and Phenomenological Research* 102/3: 531–558.
Zagzebski, Linda (2012). *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief*. Oxford University Press.
Zaller, John R. (1992). *The Nature and Origins of Mass Opinion*. Cambridge University Press.
Zeitlin, Matthew (2020). "Tweet." https://twitter.com/mattzeitlin/status/1223718846139457538.
Zmigrod, L., P. J. Rentfrow, and T. W. Robbins (2020). "The Partisan Mind: is Extreme Political Partisanship Related to Cognitive Inflexibility?" *Journal of Experimental Psychology: General* 149/3: 407–418.

NOTES

Chapter 1

1 Donahue 2007: 1.
2 Cottam et al. 2016: 1–14.
3 Wittgenstein 1953: §65–116.
4 Donahue 2007: 4.
5 Donahue 2007: 5.
6 Donahue 2007: 7.
7 See Allen 2016.
8 For example, Donahue 2007: 5–10.
9 Donahue 2007: 6, 9.
10 Leftwich 2004: 2.
11 For example, Robinson 2017 ("Politics is a Contest of Domination").
12 Leader Maynard and Worsnip 2018.
13 Sources: www.cnn.com/2005/WORLD/asiapcf/02/20/cnna.clinton.bush/; https://abc7news.com/nancy-pelosi-politics-donald-trump-common-wealth-club/2057215/; https://michiganadvance.com/2019/12/16/slotkin-to-vote-yes-on-impeachment-this-is-bigger-than-politics/; www.heraldmailmedia.com/story/news/2021/04/01/gov-gavin-newsoms-anti-recall-strategy-brand-the-other-side-as-republican/43738617/.

Chapter 2

1 Jackson 2020: §2.2.
2 Jackson 2020: §2.1.
3 Schwitzgebel 2019: §1.
4 Raz 2009.
5 Davidson 1963.
6 Audi 2020: 168.
7 Schwitzgebel 2019: §1.2.
8 See Schwitzgebel: §1.1.2.
9 Strawson 2010.

10 Vayrynen 2013.
11 Putnam 2004.

Chapter 3

1 Donahue 2007: 4; see also Schmitt 1932.
2 Driver 2005: 137–138.
3 I drew this objection from a conversation with Rich Eva, who took the view that purely moral disagreement is not political disagreement.

Chapter 4

1 See, for example, Whelan 1981.
2 Schwarz 2007.
3 Davidson 1974: 197, quoted in Glüer 2011: 112–113.
4 See, for example, Hirsch 2005 for an investigation of similar ideas within metaphysics.
5 Ballantyne and Ditto 2021.
6 Elga 2007: 492–494.
7 Elga 2007.
8 Elga 2007: 492.
9 See van Roojen 2018.
10 See Strawson, P. F. (1963). "Freedom and Resentment." in Strawson 1974: ch1.
11 See Scarantino and de Sousa 2018.
12 In Austin 1962.
13 Stanley 2011.
14 The canonical source is Austin 1962.
15 Adapted from Green 2020: §2.
16 Stanley 2011.

Chapter 5

1 Chalmers 2011: 522.
2 Chalmers 2011: 519.
3 See, for example, Plunkett and Sundell 2013. In fact, metalinguistic negotiation is a broader phenomenon than this and includes other types of considerations about downstream effects of deciding how to use a word; thanks to Adam Gibbons for this point.
4 Restall 2015.
5 Haslanger 2000: 33.
6 Manne 2017: 44.
7 See discussion at Sanneh 2019.
8 Strawson 1963: 505 influentially posed this challenge to Rudolf Carnap's (e.g. 1950) project of explication.
9 Both at Cappelen 2018: 101.
10 Or so I argue in Traldi manuscript a.
11 See, for example, Case 2019, but see Liao and Hansen 2023, arguing that the gradability of these sorts of terms undermines the dilution critique (I have my doubts about their argument).
12 Pinker 1994.
13 Gibbons forthcoming.
14 From the *Cambridge Dictionary of Philosophy* (2015), Robert Audi, ed., under "emotive conjugation".
15 Beaver and Stanley 2021.

Chapter 6

1 See Joshi 2020; Lewis and Lewis 2022.
2 For example, Haidt et al. 2000; Haidt 2001.
3 Sauer 2011.
4 Sauer 2011.
5 Haidt et al. 2000.
6 Sperber and Mercier 2017.
7 Haidt et al. 2000.
8 Haidt et al. 2000.
9 Haidt et al. 2000.
10 Haidt et al. 2000.
11 Graham et al. 2009: 1031.
12 See Iyer et al. 2012.
13 Graham et al. 2009: 1031.
14 Graham et al. 2009: 1031–1040.
15 Joshi 2020.
16 Sauer 2018: 132.
17 Sauer 2018: 132–133.
18 Sauer 2018: 133.
19 Sauer 2018: 136–137.
20 Sauer 2018: 134.
21 Jacobson 2012: 298–304.
22 In Rawls 1971.
23 Daniels 2016.
24 Adorno et al. 1950: 35.
25 Adorno et al. 1950: 37.
26 Jacobs 2016.
27 Adorno et al. 1950: 259–260.
28 For example, Kugler et al. 2014.
29 Pratto et al. 1994: 742.
30 Malka et al. 2017: 117.
31 For example, Kanai et al. 2011; Pedersen et al. 2018.
32 For example, Gerber et al. 2012.
33 MacWilliams 2016.
34 Mattson 2018.
35 Robin 2011: 4.
36 Described in Malka et al. 2017.
37 For example, Lopez 2017.
38 Quoted in Chen 2017.
39 Joshi 2020: 47.
40 For example, Duarte et al. 2015; see also Skitka and Washburn 2016.
41 Described in Singal 2016.
42 See Kaufman 2016.
43 Zmigrod et al. 2020.
44 Moss and O'Connor 2020.

Chapter 7

1 For example, Althusser 1970; Foucault 1976.
2 Jost 2017.
3 Jost 2020: 104.
4 Jost 2020: 104.
5 Jost 2020: 105.

6 Jost 2020: 111.
7 Jost 2020: 119.
8 Jost 2020: 130–132.
9 Jost 2020: 124.
10 Jost 2020: 276.
11 See Machery 2022.
12 Eagly and Mladinic 1994.
13 Goldberg 2019.
14 Jost 2020: 285.
15 Sotola and Crede 2022.
16 Nebel 2015: 451, citing Samuelson and Zeckhauser 1988.
17 Nebel 2015: 454.
18 Nebel 2015: 461–466.
19 Haslanger 2017b: 149–150.
20 Shelby 2003: 158.
21 Shelby 2003: 162–164, 183–184.
22 Haslanger 2017b: 150.
23 Haslanger 2017b: 150.
24 Haslanger 2017b: 150–151, citing Haslanger 2012 and Haslanger 2017a.
25 See Haslanger 2012.
26 Jussim 2012: ch13.
27 In now-deleted tweets.
28 See Aytac and Rossi forthcoming. See also discussion at Sankaran 2020: 1445.
29 Chomsky, Noam and Michel Foucault (1971). "Human Nature: Justice versus Power." http://chomsky.info/1971xxxx/.
30 A metaphor famously from Wallace 2009.

Chapter 8

1 Kahan 2017a: 1.
2 Kahan et al. 2017; see also Kahan 2017b.
3 Kahan 2013: 408.
4 Kahan 2017a: 2.
5 Kahan et al. 2017.
6 Kahan et al. 2017.
7 Kahan et al. 2017.
8 Kahan et al. 2017.
9 Persson et al. 2021; Ballarini and Sloman 2017, though see Kahan and Peters 2017.
10 Tappin et al. 2021.
11 Hanson 2009.
12 Tosi and Warmke 2020: 15.
13 Williams 2021.
14 Williams 2021: 334.
15 Williams 2021: 334.
16 Williams 2021: 339.
17 I've heard this answer broached in conversation many times and don't know if it has a unique originating source.
18 Converse 1964.
19 Boutyline and Vaisey 2017: 1377, citing Zaller 1992: 6.
20 Boutyline and Vaisey 2017: 1378.
21 In Lewis and Lewis 2022.
22 In Joshi 2020.
23 For example, Tajfel 1970.

24 Harwood 2020.
25 For example, Klein 2020.
26 See again Schwarz 2007.

Chapter 9

1 Ashton and McKenna 2020: 32, citing Lukács 1971.
2 For example, Ashton and McKenna 2020: 33.
3 Tilton forthcoming.
4 Mills 2017: 53.
5 Bright forthcoming.
6 Srinivasan 2020: 411–412.
7 Tilton forthcoming, glossing an example from Kukla 2021.
8 See discussion at Tilton forthcoming, which also argues for this thesis.
9 Mills 2017: 57.
10 See, for example, Táíwò 2020 critiquing this way of putting standpoint theory into practice.

Chapter 10

1 Friedman 2019: 82.
2 Lippmann 1922: 16, cited in Friedman 2020: 96.
3 Lippmann 1922: 16.
4 Lippmann 1922: 80, citing James 1890: 488.
5 Friedman 2019: 35.
6 Friedman 2019: 55–57.
7 Hirschman 1991.
8 Friedman 2019: 57.
9 Friedman 2019: 97.
10 Friedman 2019: 97, citing Lippmann 1922: 98.
11 Friedman 2019: 97.
12 Friedman 2019: ch4.

Chapter 11

1 Hannon 2021: 297–298.
2 Hannon 2021: 301.
3 Hannon 2021: 301.
4 Hannon 2021: 305–308.
5 Zeitlin 2020.
6 Táíwò 2018.
7 Táíwò 2018: 306, citing Tirrell 2013.
8 Táíwò 2018: 322.
9 Hyska 2023, citing Huang 2018: 1034.
10 Hyska 2023.
11 For example, Stanley 2015.
12 Hyska 2023.
13 See, for example, van Roojen 2018: §4.

Chapter 12

1 Cohen 2000: 18.
2 White 2010.

3 See discussion on the case of implicit biases in particular at Brownstein 2019: §2, as well as Spaulding 2021.
4 Ballantyne 2012.
5 Joshi forthcoming b.
6 Joshi forthcoming b.
7 It is mentioned frequently in Mercier 2020, for instance.
8 For example, Street 2006.
9 See discussion at Vavova 2018: 138, citing Dworkin 1996: 125.
10 White 2010: 599.
11 Gilovich and Griffin 2002: 1–2.
12 Gilovich and Griffin 2002: 3.
13 Gilovich and Griffin 2002: 3.
14 Ballantyne 2019: 130–131. See also Kelly 2022: ch4.
15 Ballantyne 2019: 131; Kelly 2022: 88–92.
16 Resnick 2019.
17 Stanovich 2021 distinguishes the two (e.g. 4–6) and also defines them more rigorously than I have.
18 See generally the work of Gerd Gigerenzer on this count, for example, Gigerenzer 2008, 2018.
19 Sauer 2020.
20 Dorst 2020b.
21 Levy 2021b: 32. Sauer 2020 makes a similar point.
22 Stanovich 2021: ch2.
23 Stanovich 2021: 12, citing Kopko et al. 2011; Messick and Sentis 1979.
24 See Machery 2022 for an overview.
25 See Singal 2017.
26 Jussim 2012, especially ch10–12, and elsewhere.
27 See, for example, White 2010: 580–582.
28 So do most epistemologists; for dissenters, look to Timothy Williamson, Clayton Littlejohn, and Amia Srinivasan.
29 Kelly 2022: 173.
30 Srinivasan 2015: 330. She also makes this point about other justifications in Srinivasan 2015 and about a safety condition at Srinivasan 2017: 134.
31 White 2010: 574.
32 White 2010: 607–608.
33 White 2010: 577.
34 For example, Taylor et al. 1978; Quattrone and Jones 1980.
35 Park and Rothbart 1982: 1052.
36 Boldry et al. 2007: 157, citing Park and Rothbart 1982, Judd and Park 1988, and Linville et al. 1989.
37 Rini 2017; Begby 2021b.
38 Hyska 2023.
39 Mayo-Wilson et al. 2011: 653.
40 Goodin 2006.
41 Dorst forthcoming.
42 Dorst 2020a.
43 Dorst forthcoming.
44 Dorst forthcoming.
45 This difference is reminiscent of Cass Sunstein describes as “eureka problems,” for example, Sunstein 2009: 46.
46 Elga 2007: 480–481.
47 See Carter and McKenna 2020.

Chapter 13

1 Elga 2007: 492.
2 Matheson 2015.
3 Cited in Matheson 2015.
4 Christensen 2007: 193.
5 Elga 2007: 486.
6 See Elga 2007: 486–487.
7 Frances and Matheson 2018.
8 Gardiner 2014.
9 Elga 2010.
10 See, for example, Kelly 2005; Kelly 2010.
11 See, for example, Lackey 2010.

Chapter 14

1 See Traldi manuscript b.
2 Frances and Matheson 2018.
3 Lackey 2013; see also Lackey 2018.
4 See Goodin and Spiekermann 2018 for more on independence, especially chapter 5.
5 Landemore 2013: 156–160.
6 Joshi 2020; see also the chapters on partisanship theories of politics and polarization as sorting.
7 Correlation produced presumably by their getting reliable information about the world itself; see discussion at Goodin and Spiekermann ch5.
8 Caplan 2007.
9 See also Brennan 2020.
10 See Hannon 2022.
11 Friedman 2006: i, in a kind of exegesis of Converse 1964.
12 Gibbons 2022.
13 See, for example, Friedman 2006; Caplan 2007; Brennan 2016; Somin 2016.
14 Somin 2016: 17.
15 Brennan 2016: 24.
16 Somin 2016: 17.
17 Somin 2016: 18.
18 Somin 2016: 20.
19 Brennan 2016: 25.
20 Goodin and Spiekermann 2018: 53.
21 Friedman 2019: 11.
22 Friedman 2017.
23 Landemore 2013: 154.
24 Kogelmann 2023, citing Brennan and Lomasky 1997: 33.
25 Kogelmann 2023.
26 Goodin and Spiekermann 2018: 49.
27 Goodin and Spiekermann 2018: 49.
28 Goodin and Spiekermann 2018: 50.
29 Landemore 2013: 92.
30 Landemore 2013: 94, following Mansbridge et al. 2010.
31 Landemore 2013: 120.
32 See Loury 1994.
33 See Joshi forthcoming a.
34 Tosi and Warmke 2016, 2020.

35 Landemore 2013: 120.
36 Sunstein 2009: 4.
37 Bishop 2008; see also Talisse 2019.
38 Sunstein 2002: 176.
39 Sunstein 2009: 7.
40 Sunstein 2009: 17.
41 Joshi 2021; Talisse 2021.

Chapter 15

1 Note that some modern socialists are "market socialists" who take themselves to have integrated Hayek's insights.
2 Hayek 1945: 520.
3 Hayek 1945: 521.
4 Hayek 1945: 521–522.
5 Hayek 1945: 522.
6 Danaher 2016.
7 Hayek 1945: 530.
8 Hayek 1945: 526.
9 Milton 1644.
10 250 U.S. 616 (1919): 630.
11 345 U.S. 41 (1953): 56.
12 Mill 1859.
13 Goldman and Cox 1996: 4.
14 Goldman and Cox 1996: 16–17.
15 Goldman and Cox 1996: 17–19.
16 Goldman and Cox 1996: 17.
17 Goldman and Cox 1996: 9.
18 Fiss 1991: 2100, cited at Goldman and Cox 1996: 10.
19 Sunstein 1993, cited at Goldman and Cox 1996: 10.
20 Fish 1994.
21 Goldman and Cox 1996: 19–20.
22 Thanks to Adam Gibbons for pressing these last two points.
23 Goldman and Cox 1996: 23–26.
24 Goldman and Cox 1996: 12–16.
25 For example, Laudan 1981.
26 Hanson 2013.
27 Hanson 2013: 156–157.
28 Landemore 2013: 176.
29 Landemore 2013: 178–183.
30 Aumann 1976.
31 Milgrom and Stokey 1982.

Chapter 16

1 Levy 2022.
2 For example, Tiehen forthcoming.
3 See, for example, Goldman 1979. See also Goldman and Beddor 2021 for an overview of such theories.
4 Lackey 2018.
5 Lackey 2018: 231, citing Zagzebski 2012.
6 Joyce 2007.

7 Marcus and Gonsalves 2020.
8 Goldman 2001.
9 Tiehen forthcoming.
10 Lackey 2018: 232–233, citing Raz 1988: 68–69.
11 Lackey 2018: 233, citing Zagzebski 2012: 115, fn11.
12 Lackey 2018: 235.
13 Contessa forthcoming.
14 Tetlock 2005: 41–44.
15 Tetlock and Gardner 2015.
16 Ritchie 2020.
17 Schimmack 2021.
18 Lakens 2017.
19 The word "fishy" is from Enoch 2014.
20 Callahan 2018.
21 McGrath 2009.
22 Nehamas 1998: 81–82, cited in McGrath 2009: 334–335.
23 Enoch 2014: fn20.
24 See discussion at Elga 2007: 480–481.
25 Friedman 2019.
26 Bruenig 2013.
27 In a manuscript, Sara Chan argues that one's own well-being is not accessible in this way, which undercuts some claims to authority made by disability activists.
28 Dror 2023: 622.
29 Here I draw from Flanigan forthcoming and Tilton forthcoming.
30 Flanigan forthcoming.
31 Dror 2023: 623.

Chapter 17

1 Caruso 2018: §3.3.
2 Streumer 2007, outlined in Henne et al. 2019: 132–133.
3 Copp 2008.
4 Nelkin 2019; see Nagel 1979 for a locus classicus.
5 Buckwalter 2020: 85–86.
6 Buckwalter 2020: 86.
7 Buckwalter 2020: 86–87.
8 Buckwalter 2020: 87.
9 Buckwalter 2020: 87–88.
10 Buckwalter 2020: 88.
11 Vranas 2007: 174–175.
12 Vranas 2007: 182–186.
13 Vranas 2007: 192–196; see also Widerker 1991.
14 Adapted from Frankfurt 1969.
15 Bennett 1990: 89–90.
16 Williams 1973.
17 Berker 2011.
18 Shah 2002: 437–440.
19 Ginet 2001: 63–76.
20 Ginet 2001: 65.
21 Ginet 2001: 64.
22 See, for example, Levy 2005.
23 Littlejohn 2014: 144.

24 Alston 1988, 2005.
25 Feldman 1988; Feldman 2001.
26 See Carr forthcoming; see also Podgorski 2017 on rational delay.
27 Carr forthcoming.
28 Though see Begby 2021b, discussed in an earlier chapter.
29 D'Arms and Jacobson 2000.
30 Worsnip 2021: 531–532.
31 Schroeder 2010.
32 See, for example, Parfit 2011; Piller 2006.
33 Parfit 2011: 420.
34 Hieronymi 2005: 442.
35 Berker 2018.
36 Berker 2018: 432.
37 Roeber 2020.
38 Stanley 2005: 3–4.
39 DeRose 1992: 913.
40 Roeber 2018: 188–190; see also Falbo forthcoming.
41 Worsnip 2021.
42 Hawthorne and Stanley 2008.
43 Fantl and McGrath 2009.
44 Lackey 2008: ch4; see also Graham 2000a, 2000b.
45 Roeber 2018: 177–179.
46 See Fritz 2017.
47 Basu and Schroeder 2019; see also Basu 2018, 2019; Schroeder 2018.
48 Fritz 2017.
49 Bolinger 2018: 2416.
50 Fritz 2020.
51 Fritz and Jackson 2021.
52 Basu and Schroeder 2019; Basu has since rejected this view – personal correspondence, and see Basu 2021.
53 Traldi 2023.
54 Caplan 2007.
55 See Joshi 2020.

Chapter 18

1 Popper 1950/2013: 581.
2 Popper 1950/2013: 581.
3 Vallier 2022.
4 Vallier 2022.
5 Beaver and Stanley 2021: 165.
6 Beaver and Stanley 2021: 166.
7 Beaver and Stanley 2021: 167.
8 Rawls 1993: 4.
9 Rawls 1993: 56–57.
10 Rawls 1993: 61.
11 Rawls 1993: 62.
12 For example, Nicholas Wolterstorff's contributions in Audi and Wolterstorff 1997.
13 Rawls 1993: 133–172.
14 Rawls 1993: 134.
15 Sandel 1994: 1776–1782.

Chapter 19

1 Abramowitz and Stone 2006: 141.
2 Abramowitz and Webster 2017.
3 Talisse 2019: ch3.
4 Talisse 2019: 73.
5 As explored by Bishop 2008.
6 Talisse 2019: 73.
7 Bramson et al. 2017.
8 Klein 2020: ch9.
9 Adapted from Joshi 2020.
10 Joshi 2020: 49.
11 Joshi 2020: 49.
12 See Sampson manuscript.
13 Lewis and Lewis 2022: ch4.
14 See, for example, Somin 2016: 21.
15 Converse 1964.
16 Joshi 2020: 40–41.
17 Joshi 2020: 54.
18 Sunstein 2001. See also Sunstein's follow-ups: *Republic.com 2.0* (2009) and *#Republic* (2018).
19 Sunstein 2001: 3–4.
20 Pariser 2011: introduction.
21 Nguyen 2020: 142. See also Nguyen 2018.
22 Thanks to Nevin Climenhaga for making something like this point.
23 Nguyen 2020: 149.
24 Nguyen 2020: 155.
25 Kelly 2008.
26 Nguyen 2020: 156.
27 Begby 2021a.
28 Begby 2021a: 527.
29 Begby 2021b.
30 Nyhan and Reifler 2010.
31 See Nyhan 2021; see also Swire-Thompson et al. 2020.
32 For example, Wood and Porter 2019.
33 Arguedas et al. 2022.
34 Arguedas et al. 2022, citing Bail et al. 2018.
35 Lukianoff and Haidt 2018: 259.
36 Haidt 2012: 334.
37 Haidt 2012: 335.
38 Caplan 2011.

Chapter 20

1 This is the kind of distinction developed extensively in Bramson et al. 2017.
2 Levy 2021a: 9550–9551.
3 Levy 2021a: 9551.
4 Sunstein 2009: 14–15.
5 Sunstein 2009: 15–16.
6 Sunstein 2009: 12.
7 Sunstein 2009: 46.
8 Sunstein 2009: 40.
9 Sunstein 2009: 22–23.

10 Sunstein 2009: 23–26.
11 Sunstein 2009: 26–30.
12 Sunstein 2009: 42–44.
13 Sunstein 2009: 30–31.
14 See Kuran 1995.
15 For example, Ornstein 2014.
16 As I argue in Traldi 2022.
17 From www.mackinac.org/OvertonWindow.
18 Robertson 2018.
19 See, for example, Alterman, Eric (2016, September 14). " 'Normalizing' Trump." *The Nation*; Homans, Charles (2016, November 1). "The 'Normalization' of Trump, and What Comes After." *The New York Times*; Cole, Teju (2016, November 11). "A Time for Refusal." *The New York Times Magazine* ("In the early hours of November 9, 2016, the winner of the presidential election was declared. . . . All around were the unmistakable signs of normalization in progress."); Yglesias, Matthew (2016, November 30). "The Case for Normalizing Trump." *Vox*.
20 Examples from Kvalnes and Nordal 2019: 764.
21 Tetlock 2003.
22 Tetlock 2003: 321–322.
23 Levari et al. 2018: 465.
24 Haslam 2016.
25 Talisse 2021.

Chapter 21

1 Harris 2022, in particular the distinction between weak and strong particularism.
2 Uscinski 2019.
3 Uscinski 2019.
4 Uscinski et al. 2022.
5 Dentith and Keeley 2019.
6 See Dentith and Orr 2017.
7 Keeley 1999: 116.
8 See discussion at Coady 2003: 198.
9 Thanks to the fall 2021 "Epistemology in Practice" class at Notre Dame for this and so many other helpful discussions.
10 Keeley 1999: 116–117.
11 Pigden 2007: 228–229.
12 Coady 2003: 203.
13 Levy 2007: 188.
14 Coady 2007: 201.
15 Keeley 1999: 121–122.
16 Pigden 2007: 230.
17 See also Cassam 2019, which makes a link between conspiracy theories, extremist propaganda, and cognitive biases.
18 Popper 1963: 341.
19 Popper 1963: 124.
20 Keeley 1999: 123.
21 Mandik 2007: 211.
22 See Dentith 2016.
23 Coady 2007: 196–197.
24 Coady 2007: 197.
25 Though see Hughes 2019.

Chapter 22

1 At http://en.wikipedia.org/wiki/Propaganda.
2 Stanley 2015: 138.
3 Stanley 2015: 205.
4 Stanley 2015: 138.
5 An example from Beaver et al. 2021.
6 Saul 2018: 362, cited by Andreson and Barnes 2022.
7 Brennan 2017: 39.
8 Leiter and Leiter 2015.
9 Stanley 2015: ch2.
10 Stanley 2015: 53.
11 Brennan 2017: 41.
12 Hyska 2023.
13 At "What is Propaganda?" https://publish.illinois.edu/mirasotirovic/whatispropaganda.
14 See, for example, Smith 2020.
15 Over 2021: 3.
16 Over 2021: 5–6, 9–10.
17 Over 2021: 6–8.
18 Over 2021: 8–9.
19 Over 2021: 7.
20 Manne 2016.
21 Smith 2020.
22 See my review at Traldi 2021.
23 Mercier 2020.
24 Mercier 2020: 158.
25 Mercier 2020: 200.
26 See also Miton et al. 2015.
27 Vallier 2020: 2.
28 Vallier 2020: 6.

Chapter 23

1 See, for example, Habgood-Coote 2019; Coady 2021.
2 Sanders 2022.
3 Worsnip 2019.
4 Rini 2017: E51.
5 Rini 2017: E55–E58.
6 Bedard 2020.
7 See Anderau 2021 for a survey of over a dozen proposed definitions.
8 McBrayer 2020: 3–4.
9 Jaster and Lanius 2021: 20.
10 Pritchard 2021: 46.
11 Pepp et al. 2019: 69.
12 Fallis and Mathiesen forthcoming.
13 Coady 2021.
14 Stancil 2023.
15 Maybe in the work of Jean Baudrillard, for instance.
16 Roeber manuscript.
17 Roeber manuscript.
18 Lieberman and Schatzberg 2018: 24–25.
19 Meyer and Fetters 2018.

Chapter 24

1 Chesterton 1908: 84–85.
2 Goodin and Spiekermann 2018: 151–153.
3 Goodin and Spiekermann 2018: 153.
4 Vermeule 2008: 75–77, cited in Goodin and Spiekermann 2018: 154.
5 Chesterton 1929: ch4.
6 Gedeon 2015: 1.
7 Singer 1981.
8 Huemer 2016.
9 See, for example, Ballantyne 2012.
10 Stillman 1990.
11 Though see Laudan 1981 for a kind of reason to think otherwise.
12 *District Attorney's Office for the Third Judicial District v. Osborne*, 557 U.S. 52, 105 (2009) (Souter, J., dissenting), cited in Barzun 2018: 702.
13 Barzun 2018: 703.
14 MacAskill 2022: ch3.
15 *The Economist* 2019.
16 See Tosi and Warmke 2016, 2020.
17 For example, Pinker 2019.
18 Mitchell and Tetlock forthcoming.
19 Sunstein 2019.
20 After Weaver 1948, a book called *Ideas Have Consequences* which makes this argument about medieval nominalism.
21 In *Trop v. Dulles*, 356 U.S. 86, 101 (1958).
22 576 U.S. 644 (2015).
23 Strauss 2010: 14.
24 Dorf 2012: 2045.
25 Scalia 1997: 40–41.
26 *Miller v. Alabama*, 132 S. Ct. 2455, 2487 (2012) (Alito, J., dissenting).

INDEX